Journal of Computational Physics **171,** 632–677 (2001)
doi:10.1006/jcph.2001.6800, available online at http://www.idealibrary.com on IDEAL®

High-Order Two-Way Artificial Boundary Conditions for Nonlinear Wave Propagation with Backscattering

Gadi Fibich* and Semyon Tsynkov*,†,[1]

*School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel; and †ICASE, MS 132C, NASA Langley Research Center, Hampton, Virginia 23681-2199
E-mail: fibich@math.tau.ac.il, tsynkov@math.tau.ac.il

Received August 28, 2000; revised April 13, 2001

When solvinglinear scattering problems, one typically first solves for the impinging wave in the absence of obstacles. Then, using the linear superposition principle, the original problem is reduced to one which involves only the scattered wave (which is driven by the values of the impinging field at the surface of the obstacles). When the original domain is unbounded, special artificial boundary conditions (ABCs) have to be set at the outer (artificial) boundary of the finite computational domain in order to guarantee the reflectionless propagation of waves through this external artificial boundary. The situation becomes conceptually different when the propagation equation is nonlinear. In this case the impinging and scattered waves can no longer be separated, and the problem has to be solved in its entirety. In particular, the boundary on which the incoming field values are prescribed should transmit the given incoming waves in one direction and simultaneously be transparent to all the outgoing waves that travel in the opposite direction. We call such boundary conditions *two-way ABCs*. In the paper, we construct the two-way ABCs for the nonlinear Helmholtz equation, which models a continuous-wave laser beam propagation in a medium with a Kerr nonlinear index of refraction. In this case, the forward propagation of the beam is accompanied by backscattering, i.e., generation of waves in the opposite direction to that of the incoming signal. Our two-way ABCs generate no reflection of the backscattered waves and at the same time impose the correct values of the incoming wave. The ABCs are obtained in the framework of a fourth-order accurate discretization to the Helmholtz operator inside the computational domain. The fourth-order

The work of the first author was partially supported by Grant 97-00127 from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel. The work of the second author was partially supported by the National Aeronautics and Space Administration under NASA Contract NAS1-97046 while in residence at ICASE, NASA Langley Research Center, Hampton, VA, 23681-2199.

[1] Current address: Department of Mathematics, North Carolina State University, Box 8205, Raleigh, NC 27695. E-mail: tsynkov@math.ncsu.edu.

0021-9991/01 $35.00

convergence of our methodology is corroborated experimentally by solving linear model problems. We also present solutions in the nonlinear case using the two-way ABC which, unlike the traditional Dirichlet boundary condition approach, allows for direct calculation of the magnitude of backscattering.

Key Words: artificial boundary conditions (ABCs); two-way ABCs; radiation; Helmholtz equation; nonlinearity; nonparaxiality; fourth-order schemes; self-focusing; backscattering.

1. INTRODUCTION

The Helmholtz equation

$$\Delta E(x_1, \ldots, x_D) + k^2 E = 0, \quad \Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_D^2} \tag{1}$$

models time-harmonic wave propagation in D dimensions. The simplest case is when $k \equiv k_0$, which corresponds to propagation of waves through a homogeneous medium. For example, in optics E is the electric field, $k_0 = \omega_0 n_0/c$ is the wavenumber, ω_0 is frequency, n_0 is the (linear) index of refraction of the medium, and c is the speed of light.

In many applications, one wants to solve Eq. (1) in the presence of an impinging wave field and boundaries, which can be either surfaces of obstacles or interfaces between different media. The source of the impinging wave is prescribed by a relation of the form

$$E_{\text{inc}} = E_{\text{inc}}^0 \quad \text{on } \Sigma_{\text{source}}, \tag{2}$$

where Σ_{source} can, for example, be a point (i.e., a point source) or a plane (i.e., a plane wave), and E_{inc}^0 is given. The physical properties of surfaces and/or interfaces, i.e., how they handle the impinging wave in terms of propagation through and/or reflection, are given by *linear* operator relations of the form

$$\boldsymbol{L}[E] = 0 \quad \text{on } \Sigma_{\text{interface}}. \tag{3}$$

For example, if $\Sigma_{\text{interface}}$ is the surface of a perfect conductor, then (3) reduces to $E = 0$ on $\Sigma_{\text{interface}}$ (total reflection).

Since Eqs. (1)–(3) are linear, one can solve the scattering problem in two sequential stages as follows. The solution is split into two components

$$E = E_{\text{inc}} + E_{\text{scat}}.$$

At the first stage one solves for the incoming wave field E_{inc}, which is the solution of Eq. (1) in $\mathbb{R}^D$ in the absence of any obstacles and/or interfaces, driven by the known source term (2). Typically, one can write this solution explicitly as a superposition of plane and/or spherical waves. Then, at the second stage, one solves for the scattered wave field E_{scat}, which satisfies Eq. (1) with no sources, subject to the boundary condition

$$\boldsymbol{L}[E_{\text{scat}}] = -\boldsymbol{L}[E_{\text{inc}}] \quad \text{on } \Sigma_{\text{interface}},$$

which directly follows from (3). In the process of solving numerically for E_{scat}, one has to replace $\mathbb{R}^D$ with a bounded computational domain. In doing so, one needs to introduce *the artificial boundary conditions* (ABCs), see [34], which make the boundary transparent for outgoing waves and guarantee the solvability of the truncated problem on the finite computational domain, such that the computed solution is close to the original infinite-domain solution.

In addition to the simplest case $k \equiv k_0$, there are numerous applications where the medium is nonhomogeneous, i.e., $k = k(x_1, \ldots, x_D)$. In this case, one may also need to solve for the incoming field E_{inc} numerically (using ABCs), rather than analytically. However, as this problem is linear as well, one can still employ the linear superposition principle and thus solve first for E_{inc} and then for E_{scat}.

In the current study, we consider a more complex case where k also depends on the field intensity, i.e., $k = k(\omega_0, |E|^2)$. For example, the propagation of an intense continuous-wave (cw) laser beam[2] through a Kerr-type medium such as water or silica, is described by Eq. (1) with $k^2 = k_0^2(1 + \epsilon|E|^2)$, where $\epsilon = 4\epsilon_0 c n_2$ and n_2 in the Kerr coefficient (e.g., [4, 21]). In this case, beam propagation is governed by the nonlinear Helmholtz equation (NLH)

$$\Delta E + k^2 E = 0, \quad k^2 = k_0^2(1 + \epsilon|E|^2). \tag{4}$$

We note that the natural definition of the index of refraction is in the frequency domain. In the time domain, the cubic nonlinearity becomes a nonlocal convolution, which, in the case of almost-monochromatic wavepackets, reduces, to leading order, to a cubic nonlinearity; see Section 8.7 and Ref. [8] for more details. Because of the nonlinearity in the equation of propagation (4), the components E_{inc} and E_{scat} can no longer be decoupled as in the linear case. From a numerical point of view, this nonlinear coupling adds a new twist to the construction of the ABCs, since the Kerr medium interface at $z = 0$ is required to transmit E_{inc} in one direction and, at the same time, transmit E_{scat} in the opposite direction. Deriving and implementing this *two-way ABC* in the discrete nonlinear framework is a key emphasis of this study.

2. PHYSICAL MODEL

Although our numerical approach is quite general, in order to motivate the presentation we relate it to a specific physical problem, namely that of an intense laser beam propagating through a nonlinear Kerr medium. The Kerr medium is located in the half-space $z \geq 0$; the directions of increasing and decreasing z are called right and left, respectively; and the wave source in the model is a right-traveling beam, impinging on the Kerr medium at $z = 0$. Therefore, the only physical boundary in the model is the transverse two-dimensional (x, y) plane at $z = 0$. For simplicity, we assume that the input beam is radially symmetric in the transverse plane and denote the transverse coordinate by $r = \sqrt{x^2 + y^2}$. For the purpose of solving the problem numerically, the original unbounded physical domain $0 \leq z < +\infty$ is replaced with the truncated domain $0 \leq z \leq z_{\max}$. Therefore, the left computational boundary at $z = 0$ coincides with the Kerr medium interface, and the right computational boundary at $z = z_{\max}$ is a far-field artificial boundary. The desired behavior of the solution at the boundaries $z = 0$ and $z = z_{\max}$ is described in the remainder of this section.

[2] A cw laser beam is a time-harmonic monochromatic wave, as opposed to, say, pulses and wave packets.

2.1. *Two-Way Propagation of Waves at Media Interface*

At $z = 0$, the electric field E has both incoming and backscattered components. The value of the incoming wave upon entering the nonlinear medium is given by

$$E_{\text{inc}}(r, 0) = E^0_{\text{inc}}(r). \tag{5}$$

The backscattered field at $z = 0$ cannot be known ahead of time, because backscattered waves originate inside the domain and subsequently propagate toward the left boundary (see Section 7.1.1 for more details). The simplest formulation of the problem, which is the one used in this study, is to require that the boundary $z = 0$ be completely transparent to all backscattered waves. Consequently, the two-way ABC at $z = 0$ has to ensure the reflectionless propagation of backscattered waves through the boundary (*a radiation boundary condition*) and at the same time correctly prescribe the incoming signal (5).

The main idea in our implementation of two-way ABCs is the following. The overall solution E is composed of outgoing waves E_{out} that propagate through $z = 0$ with no reflection, and an incoming wave E_{inc}, which is prescribed at $z = 0$. Assume that outgoing waves E_{out} satisfy a homogeneous boundary condition $\boldsymbol{B}E_{\text{out}} = 0$, where $\boldsymbol{B}$ denotes a linear nonreflecting boundary operator. The incoming wave E_{inc} cannot satisfy this boundary condition, because otherwise the solution E would not be unique. Therefore, if we apply the operator $\boldsymbol{B}$ to $E = E_{\text{out}} + E_{\text{inc}}$, we obtain $\boldsymbol{B}E = \boldsymbol{B}E_{\text{out}} + \boldsymbol{B}E_{\text{inc}} = \boldsymbol{B}E_{\text{inc}} \neq 0$. Since the incoming wave E_{inc} is given [see (5)], we see that the nonhomogeneous relation $\boldsymbol{B}E = \boldsymbol{B}E_{\text{inc}}$ can serve as a two-way ABC with the desired properties. In the present study the implementation of this idea is carried out through linearization in an iterative framework with subsequent analysis of the linear problem in Fourier space. A detailed description of the procedure is given in Section 4 for the continuous formulation of the problem, and in Section 6 for the discrete formulation of the problem.

Our construction of a two-way ABC is based on the assumption of reflectionless propagation of all left-going waves, as well as the incident right-going wave, through $z = 0$. A more accurate physical model should, of course, include reflections from the media interface $z = 0$, because the linear index of refraction can be discontinuous across this interface (see Section 8.2 for more details). These reflections can result in different values of the incoming wave field on two sides of the interface, i.e., $E_{\text{inc}}(r, -0) \neq E_{\text{inc}}(r, +0)$. In the current study we disregard this effect, which can be interpreted as either considering E^0_{inc} of (5) to be the part of the incoming wave that has already been transmitted past the $z = 0$ interface, or assuming continuity of the wavenumber across the interface. Similarly, we neglect the reflection of backscattered waves by the media interface at $z = 0$. In Section 8.2 we briefly comment on how one can incorporate a reflecting interface (i.e., discontinuity in k at $z = 0$) in the methodology that we are building. In fact, we consider this one of the future extensions of our current work.

2.2. *Behavior as $z \to +\infty$*

Basically, as $z \to +\infty$, we require that E have no left-propagating components. In this study we assume that at large distances propagation is diffraction-dominated and the field amplitude decays to zero, i.e., $\lim_{z\to\infty} \max_{0\le r<\infty} |E(r, z)| = 0$, so that

$$\lim_{z\to+\infty} k^2 = k_0^2.$$

Therefore, at large z's the solution is a linear superposition of right-traveling waves.

As mentioned, the simulation is carried out on a truncated domain $0 \le z \le z_{\max}$ (see Sections 5 and 6). Therefore, the desired behavior of the solution as $z \to +\infty$ has to be captured by a far-field ABC at the artificial boundary $z = z_{\max}$. This boundary condition should guarantee the reflectionless propagation of all the waves traveling toward $z = +\infty$.

3. PARAXIAL APPROXIMATION

Most research on wave propagation in a Kerr medium has been carried out in the framework of the nonlinear Schrödinger equation (NLS), rather than the NLH. We now briefly describe how one derives the NLS from the NLH and quote some results on wave propagation in the NLS model. For more information on NLS theory, see, e.g., [10, 21, 31, 32].

For reasons that will become clear, we consider the NLH in $\mathbb{R}^D$ with a general power-law nonlinearity

$$\Delta E + k^2 E = 0, \quad k^2 = k_0^2(1 + \epsilon |E|^{2\sigma}).$$

We denote the axial coordinate by $z := x_D$ and assume radial symmetry in the transverse plane of the first $D-1$ coordinates, i.e.,

$$E = E(r, z), \quad r = \sqrt{x_1^2 + \cdots + x_{D-1}^2}.$$

We also separate the slowly varying envelope ψ from the fast oscillations, and we introduce the nondimensional variables

$$E = (r_0 k_0 \sqrt{\epsilon})^{-1/\sigma} \exp(i k_0 z) \psi(\tilde{r}, \tilde{z}), \quad \tilde{r} = \frac{r}{r_0}, \quad \tilde{z} = \frac{z}{2L_{DF}},$$

where r_0 is the initial beam width and $L_{DF} = k_0 r_0^2$ is the diffraction length. After dropping the tildes, the equation for the amplitude ψ, in nondimensional form, is given by

$$\gamma_{\text{np}} \psi_{zz} + i\psi_z + \Delta_\perp \psi + |\psi|^{2\sigma} \psi = 0,$$

where the transverse Laplacian is

$$\Delta_\perp = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_{D-1}^2} = \frac{\partial^2}{\partial r^2} + \frac{D-2}{r} \frac{\partial}{\partial r},$$

and

$$\gamma_{\text{np}} = \left(\frac{1}{2 r_0 k_0} \right)^2.$$

In typical physical setups the beam width r_0 is much larger than the wavelength λ, which implies that $0 < \gamma_{\text{np}} \ll 1$ (or, equivalently, in dimensional variables, that $\psi_{zz} \ll k_0 \psi_z$). Therefore, it is customary to employ the *paraxial approximation*, i.e., neglect the $\gamma_{\text{np}} \psi_{zz}$ term. In that case, NLH reduces to the nonlinear Schrödinger equation (NLS):

$$i\psi_z + \Delta_\perp \psi + |\psi|^{2\sigma} \psi = 0. \tag{6a}$$

The NLS is an evolution equation where z plays the role of "time" and the initial condition is given at $z = 0$ for all r:

$$\psi(r, 0) = E^0_{\text{inc}}(r). \tag{6b}$$

Therefore, under the paraxial approximation one approximates a boundary-value problem for the NLH with an initial-value problem for the NLS. Since the NLS accounts only for the forward-propagating wave, backscattering effects are neglected in this model. The question arises, therefore, whether and how the results of the NLS model remain valid at the NLH level, or alternatively, how these results are affected by backscattering. As of yet, almost no rigorous studies of these issues have been conducted. We therefore hope that the current study, which focuses primarily on developing a computational methodology for solving the NLH, will provide a means for numerically comparing the NLH and the NLS in the future.

Let us now proceed with describing some specific results in the NLS model which are interesting to look at in the framework of the NLH.

3.1. *Critical Self-Focusing—Arrest of Collapse*

We recall that the focusing NLS (6a) is called *subcritical*, *critical*, or *supercritical*, when $\sigma(D-1)$ is less than, equal to, or greater than 2, respectively. It is known that the solutions of both critical and supercritical NLS can actually develop singularities, i.e., blow up, at a finite z. There is, however, a marked difference between these two cases, as near the singularity nonlinearity dominates over diffraction in the supercritical case, while they are of the same magnitude in the critical case. As a result, unlike the supercritical case, singularity formation in the critical NLS is highly sensitive to perturbations, which can arrest the blowup even when they are small [10, 11]. In this paper we focus on the critical case, which corresponds to the physical self-focusing ($\sigma = 1$ and $D - 1 = 2$). In that case, solutions of the NLS can become singular (i.e., blow up) after finite propagation distance, provided that their initial power (L^2 norm) is above a certain threshold N_c, which is called *the critical power*.

The observation that the paraxial approximation breaks down near the singularity has been already noted by Kelley, in his celebrated paper on self-focusing [16]. Feit and Fleck [7] were the first to demonstrate that nonparaxiality of the beam can arrest the blowup, by showing numerically that initial conditions that lead to singularity formation in the NLS result in focusing–defocusing oscillations in the NLH. In these simulations, however, they did not solve a true boundary-value problem for the NLH. Instead, they solved an initial-value problem for a "modified" NLH that describes the right-going wave only (while introducing several additional assumptions along the way). Akhmediev and collaborators [1, 2] analyzed an initial-value problem for a different "modified" NLH; their numerical simulations also suggested that nonparaxiality arrests the singularity formation. Both numerical approaches ([7] and [1, 2]), however, failed to fully account for the effect of backscattering. Fibich [9] applied asymptotic analysis to derive an ODE in z for self-focusing in the presence of small nonparaxiality. His analysis suggests that nonparaxiality indeed arrests the singularity formation, resulting instead in decaying focusing–defocusing oscillations. However, backscattering effects were neglected in this asymptotic analysis.

Since there are no singularities in nature (i.e., the laser beam continues to propagate beyond the NLS blowup point), a natural question is whether initial conditions that lead to

blowup in the NLS correspond to global solutions of the corresponding NLH. To the best of our knowledge, the very issue of the solvability of the NLH still remains unresolved, including the critical case $\sigma(D-1)=2$. Therefore, we are interested in solving numerically the critical NLH *as a true boundary-value problem*, in order to address this question. Another issue of interest in the critical case is calculating the amount of power which is backscattered for beams which do not blow up in the NLS model. We note that at present, there are no data coming from either analysis or numerical simulations, on the actual extent of backscattering, besides the general notion that it should be small.

In order to simplify the calculations, we consider the critical NLH with $D=2$ and $\sigma=2$, i.e.,

$$\left[\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial r^2}\right]E(z,r)+k^2E=0,\quad k^2=k_0^2(1+\epsilon|E|^4), \tag{7}$$

which corresponds to the critical NLS

$$i\psi_z+\psi_{rr}+|\psi|^4\psi=0. \tag{8}$$

Based on the insight gained from NLS theory, we can expect that the results for the critical NLH with $D=2$ and $\sigma=2$ would also apply for the critical NLH with $D=3$ and $\sigma=1$.

4. NONLINEAR ITERATION APPROACH

In this section we use a continuous formulation to outline and motivate the iterative numerical approach that we adopt in this study for solving the foregoing nonlinear wave propagation problem. The actual derivation, however, is done completely at the discrete level in Sections 5 and 6.

We are interested in solving the NLH (7) in the half-space $z\geq 0$, subject to boundary condition (5) for the incoming field, decay in the transverse direction

$$\lim_{r\to\infty}E(r,z)=0,$$

and radiation conditions at $z=0$ and $z=+\infty$ for the outgoing waves, as discussed in Sections 2.1 and 2.2. We build the iteration algorithm as follows. First, we define the *linear* version of the problem as

$$\boldsymbol{L}_F[E]=0, \tag{9}$$

where

$$\boldsymbol{L}_F=\left[\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial r^2}\right]+k_0^2(1+\epsilon F(r,z)), \tag{10}$$

$F(r,z)$ is a given function, and E satisfies the same boundary conditions as in the nonlinear problem. Then, we find the solution of the nonlinear problem (7) using the iterations

$$\boldsymbol{L}_{F^{(n)}}\left[E^{(n+1)}\right]=0,\quad F^n=\left|E^{(n)}\right|^4\quad \text{for } n=0,1,2,\ldots,\mathcal{N}, \tag{11}$$

with the initial guess $E^{(0)}(r, z) \equiv 0$. Since there is no rigorous theory that guarantees the convergence of algorithm (11), our simulations (see Section 7) serve as a numerical test for the convergence of these iterations. In Section 8.3 we briefly discuss alternative approaches to the nonlinear iterations.

We note that iterative approaches have been used previously in numerical simulations of problems over infinite domains, although in completely different settings; see, e.g., [12, 13, 22].

4.1. *Iterative Solution of the Variable-Coefficient Linear Equation*

In general, one can use any linear Helmholtz solver to solve equation (11) with respect to $E^{(n+1)}$ while keeping $F^{(n)}$ frozen. In this study we solve (11) also iteratively as

$$\boldsymbol{L}_0\left[E^{(m+1)}\right] = -\epsilon k_0^2 F^{(n)} \cdot E^{(m)} \quad \text{for } m = 0, 1, 2, \ldots, \mathcal{M}(n), \tag{12}$$

where

$$\boldsymbol{L}_0 = \left[\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial r^2}\right] + k_0^2.$$

Note that the function $F^{(n)}$ does not change in the course of the iterations (12).

By rewriting formula (12) in the form

$$E^{(m+1)} = \boldsymbol{L}_0^{-1}\left[-\epsilon k_0^2 F^{(n)} \cdot E^{(m)}\right],$$

we see that it formally corresponds to the standard fixed point iteration scheme. Therefore, these iterations are more likely to converge when the right-hand side (RHS) is small. We note that this occurs when $\epsilon F^{(n)} \ll 1$, i.e., when the nonlinearity in the NLH is weak ($k^2 \approx k_0^2$). We can expect this to be the case in physical self-focusing for the following reason. The Kerr coefficient of the medium n_2 is so small that even for intense laser beams, upon entering the nonlinear medium, $\epsilon|E^0_{\text{inc}}|^2 \ll 1$. In the framework of the NLS model, if the initial beam power is above the threshold for collapse, the nonlinear contribution to the index of refraction $\epsilon|E|^2$ [see (4)] would eventually become comparable to the linear one n_0. However, the asymptotic analysis in [9] suggests that nonparaxiality arrests self-focusing when $\epsilon|E|^2 \ll 1$. As a result, $k^2 \approx k_0^2$ for all $z \geq 0$.

4.2. *Direct Solution of the Constant-Coefficient Linear Equation*

At each iteration of the inner loop (12), we solve a linear constant-coefficient equation of the form

$$\boldsymbol{L}_0 E = \Phi(r, z), \tag{13a}$$

where the RHS Φ is given by

$$\Phi = -\epsilon k_0^2 F^{(n)} \cdot E^{(m)}. \tag{13b}$$

Equation (13a), with Φ given by (13b) and subject to the boundary conditions discussed above, is solved in the following way. We use Fourier decomposition in the transverse

direction for the solution E, the RHS Φ, and boundary data $E^0_{\text{inc}}(r)$:

$$\begin{aligned} E(r,z) &= \sum_l u^l(z)\cos(lr), \\ \Phi(r,z) &= \sum_l f^l(z)\cos(lr), \\ E^0_{\text{inc}}(r) &= \sum_l u^{0,l}_{\text{inc}}\cos(lr). \end{aligned} \tag{14}$$

Because of the orthogonality of the Fourier modes, the l-th Fourier mode $u^l(z)$ of $E(r,z)$ satisfies the ordinary differential equation

$$u^l_{zz}(z) + k_l^2 u^l(z) = f^l(z), \quad k_l^2 = k_0^2 - l^2, \tag{15}$$

subject to the Dirichlet condition for the right-going wave at $z = 0$ [cf. (5)],

$$u^l_{\text{inc}}(0) = u^{0,l}_{\text{inc}}; \tag{16}$$

a radiation condition for the left-going wave at $z = 0$; and a radiation condition at $z = +\infty$. It is at this level, i.e., after the separation of variables, that we implement the two-way ABC at $z = 0$ and the radiation boundary condition at $z = +\infty$. For that, we use the concept of the *one-way Helmholtz equations*.[3]

4.2.1. One-way Helmholtz equations and the radiation principle. Equation (15) admits two linearly independent eigenfunctions: $u^{(1)} = e^{i\sqrt{k_l^2}z}$ and $u^{(2)} = e^{-i\sqrt{k_l^2}z}$. When $k_l^2 > 0$, $u^{(1)} = e^{i|k_l|z}$ is the right-propagating wave and $u^{(2)} = e^{-i|k_l|z}$ is the left-propagating wave, whereas when $k_l^2 < 0$, $u^{(1)} = e^{-|k_l|z}$ is the right-decaying (evanescent) wave and $u^{(2)} = e^{|k_l|z}$ is the left-decaying (evanescent) wave. Therefore, the one-way Helmholtz equations that each admits only one of the two eigenfunctions while prohibiting the other one are

$$u_z - i\sqrt{k_l^2}u = 0, \tag{17a}$$

$$u_z + i\sqrt{k_l^2}u = 0. \tag{17b}$$

Equation (17a) corresponds to the right-traveling or right-evanescent wave $u^{(1)}$, and Eq. (17b) corresponds to the left-traveling or left-evanescent wave $u^{(2)}$.

As mentioned at the end of Section 2.2, for the purpose of numerical solution we truncate the infinite domain $[0, +\infty)$ in z and reduce it to the finite interval $[0, z_{\max}]$. The one-way Helmholtz equations (17) can be used as boundary conditions for Eq. (15) on the interval $[0, z_{\max}]$. Indeed, if we want to make sure that near both edges of the interval $[0, z_{\max}]$ the solution is only composed of outgoing waves, then we need to use relation (17a) as the boundary condition at $z = z_{\max}$ and relation (17b) as the boundary condition at $z = 0$:

$$u_z - i\sqrt{k_l^2}u = 0 \quad \text{at } z = z_{\max}, \tag{18a}$$

$$u_z + i\sqrt{k_l^2}u = 0 \quad \text{at } z = 0. \tag{18b}$$

[3] The term "one-way wave equation" was apparently first introduced by Trefethen and Halpern in [33].

Clearly, as the boundary conditions (18a) and (18b) each eliminate one of the two eigenfunctions $u^{(1)}$ and $u^{(2)}$, the homogeneous version of Eq. (15) on $[0, z_{\max}]$ (i.e., when $f^l \equiv 0$) with these two boundary conditions is only satisfied by the trivial solution. Consequently, the nonhomogeneous Eq. (15) with boundary conditions (18) is uniquely solvable for any RHS f concentrated on the interval $[0, z_{\max}]$. From the standpoint of physics, the resulting solution is only composed of waves due to sources located inside $[0, z_{\max}]$, which radiate to the right and to the left, and contains no incoming waves from sources outside this interval. A solution of this type is said to satisfy the *radiation principle*.

4.2.2. Adding the incoming power. As mentioned above, for the particular problem that we are studying we also need to prescribe the incoming wave at $z = 0$, i.e., complement the radiation boundary condition (18b) for the left-traveling waves at $z = 0$ with a Dirichlet boundary condition (16) for the given right-traveling wave, which altogether will yield the two-way ABC. In the continuous framework, this can be done as follows. The incoming wave (16) gives rise to a solution of the form $u^{0,l}_{\text{inc}} e^{i\sqrt{k_l^2}z}$. Substituting this expression into the one-way Helmholtz equation (17b), we arrive at the following inhomogeneous relation

$$u_z + i\sqrt{k_l^2}u = 2i\sqrt{k_l^2}e^{i\sqrt{k_l^2}z}u^{0,l}_{\text{inc}}. \tag{19}$$

As in the case of any inhomogeneous linear differential equation, the general solution to Eq. (19) can be written as a sum of the general solution u_{H} to the corresponding homogeneous equation (17b) and a particular solution u_{p} to the actual nonhomogeneous Eq. (19):

$$u = u_{\text{H}} + u_{\text{p}}.$$

We may pick the particular solution as the one generated by the incoming wave, $u_{\text{p}} = u^{0,l}_{\text{inc}} e^{i\sqrt{k_l^2}z}$; and the general solution to (17b) is obviously given by $u_{\text{H}} = \text{const} \cdot e^{-i\sqrt{k_l^2}z}$.

4.2.3. Obtaining the overall solution. In order to add the incoming power to the radiation solution, we replace the homogeneous boundary condition (18b) with relation (19) interpreted as a boundary condition at the left edge of the interval:

$$u_z + i\sqrt{k_l^2}u = 2i\sqrt{k_l^2}u^{0,l}_{\text{inc}} \quad \text{at } z = 0. \tag{20}$$

This implies that the overall solution will satisfy Eq. (15), subject to boundary condition (18a) at $z = z_{\max}$ and boundary condition (20) at $z = 0$. Indeed, by the linear superposition principle, the overall solution can be written as the radiation solution with the incoming power added, $u = u_{\text{radiation}} + u^{0,l}_{\text{inc}} e^{ik_l z}$, where $u_{\text{radiation}}$ satisfies (15) and (18). Boundary conditions such as (20) are sometimes referred to as *inhomogeneous radiation boundary conditions* [17]. A similar derivation in the finite-difference framework is presented in Section 6.5.

4.3. *Nested Iterations*

In summary, our solution algorithm consists of two nested iteration loops. On the outer loop (11) we perform iterations with respect to the nonlinearity for $n = 0, 1, 2, \ldots, \mathcal{N}$. On the inner loop (12) we solve the linear equation with variable coefficients (which we obtain at each nonlinear iteration) for $m = 0, 1, 2, \ldots, \mathcal{M}(n)$. The numbers $\mathcal{M} = \mathcal{M}(n)$ and $\mathcal{N}$,

at which we terminate the inner and outer iteration loops, respectively, are determined experimentally in the course of iterations.

Our particular choice of solver for the linear variable-coefficient equation (11) is motivated by the following two reasons:

(1) The inner loop iterations (12) require inverting a linear *constant-coefficient* operator (which is the discrete analogue to $\boldsymbol{L}_0$) rather than a variable-coefficient one. As a result, the inversion can be performed by a direct method that involves separation of variables and LU decomposition. Moreover, the implementation of the radiation boundary conditions, including the two-way ABC at $z = 0$, is particularly convenient to do with the operator $\boldsymbol{L}_0$.

(2) If we used some other linear Helmholtz solver, on each outer loop iteration (11) we would have had to invert a different linear operator $\boldsymbol{L}_{F^n}$. However, using our particular linear solver involves *a repeated inversion of the same operator throughout both inner and outer loops*. This implies that the actual inversion can be performed only once in the very beginning and then the inverse operator, which is stored in memory, can be applied repeatedly to the changing right-hand side. From the standpoint of numerical efficacy this is beneficial because the inversion of the discretized $\boldsymbol{L}_0$ amounts to performing LU decomposition of a family of sparse matrices obtained after the separation of variables. The result of the LU decomposition is also sparse, hence its application to a given right-hand side has only linear complexity. Since the number of iterations required for convergence is large (see Section 7), this yields substantial savings of computer resources.

5. DISCRETIZATION

We integrate the linear constant-coefficient equation (13) on a Cartesian grid of variables (r, z) in the finite rectangular computational domain $[0, r_{\max}] \times [0, z_{\max}]$. Since the original physical domain stretches all the way to $z = +\infty$, at the artificial boundary $z = z_{\max}$ we set a radiation boundary condition that guarantees the reflectionless propagation of right-going waves (see Section 6). On the physical boundary $z = 0$ we set a two-way radiation boundary condition that similarly guarantees the reflectionless propagation of left-going backscattered waves and also correctly prescribes the right-going incoming signal (Section 6). For the transverse direction r, we assume that the solution vanishes at $r = r_{\max}$:

$$E(r_{\max}, z) = 0, \quad z \geq 0. \tag{21}$$

Physically, this condition amounts to having a conducting surface at $r = r_{\max}$, which acts as a perfect reflector. Therefore, we make $r_{\max}$ sufficiently large so that reflections from this boundary do not contaminate the solution in the primary region of interest near $r = 0$. We also assume that E is symmetric with respect to $r = 0$, i.e.,

$$E(r, z) = E(-r, z), \quad z \geq 0. \tag{22}$$

This assumption is physically plausible and allows us to consider only half of the domain $[0, r_{\max}]$ in the r direction rather than the full domain $[-r_{\max}, r_{\max}]$.

We use a uniform Cartesian grid with size h_r and a total of M cells in the r direction ($h_r = r_{\max}/M$), and size h_z and a total of N cells in the z direction ($h_z = z_{\max}/N$). Accordingly, the grid nodes are

$$\{(r_m, z_n) \mid r_m = m \cdot h_r, z_n = n \cdot h_z, m = 0, 1, \ldots, M, n = 0, 1, \ldots, N\}. \tag{23}$$

We discretize Eq. (13) using a fourth-order accurate central-difference scheme,

$$\begin{aligned} &\boldsymbol{L}^h_{rr} E_{\cdot,n}\big|_{m,n} + \boldsymbol{L}^h_{zz} E_{m,\cdot}\big|_{m,n} + k_0^2 E_{m,n} = \Phi_{m,n}, \\ &m = 0, 1, \ldots, M-1, \quad n = 2, 3, \ldots, N-2, \end{aligned} \tag{24}$$

where

$$\boldsymbol{L}^h_{rr} E_{\cdot,n}\Big|_{m,n} = \frac{-E_{m-2,n} + 16E_{m-1,n} - 30E_{m,n} + 16E_{m+1,n} - E_{m+2,n}}{12h_r^2}, \tag{25a}$$

$$\boldsymbol{L}^h_{zz} E_{m,\cdot}\Big|_{m,n} = \frac{-E_{m,n-2} + 16E_{m,n-1} - 30E_{m,n} + 16E_{m,n+1} - E_{m,n+2}}{12h_z^2}. \tag{25b}$$

The index n that corresponds to the coordinate z runs from 2 to $N-2$ in Eq. (24) because the stencil, which is five nodes wide in each direction, obviously cannot be applied to any of the boundary nodes $n = 0$, 1, $N-1$, and N located near $z = 0$ and $z = z_{\max}$. The treatment of these near-boundary grid nodes is discussed in Section 6 in the framework of the discrete radiation boundary conditions.

Similarly, the direct application of the transverse part $\boldsymbol{L}^h_{rr}$ of the discrete operator in (24) may also require a special treatment of the near-boundary nodes $m = 0$, 1, and $M-1$. This treatment should take into account the transverse boundary conditions at $r = 0$ (22) and at $r = r_{\max}$ (21). We can avoid this, however, by expanding the solution $E_{m,n}$, for each n, in a finite series with respect to eigenfunctions of the transverse discrete operator $\boldsymbol{L}^h_{rr}$, which also satisfy the two boundary conditions (21) and (22) [this is a discrete analog to the continuous Fourier expansion (14)]. This discrete eigenfunction expansion allows us to treat the operator $\boldsymbol{L}^h_{rr}$ in the transformed space from the very beginning and never implement it directly on the grid. In addition, the radiation boundary conditions in the z direction are most natural to implement in the transformed space separately for each longitudinal (i.e., z-aligned) mode, as we have seen in the continuous formulation in Section 4.2.1.

We shall now derive the discrete eigenfunction expansion for $E_{m,n}$. Let us introduce the space of all grid functions that are equal to zero at $m = M$, i.e.,

$$V = \{\psi_m \mid m = 0, 1, \ldots, M, \psi_M = 0\}.$$

Clearly, for each n, the function $E_{\cdot,n} \in V$. We can define a weighted inner product on V:

$$\langle \psi, \phi \rangle = \frac{1}{2M}\psi_0\phi_0 + \frac{1}{M}\sum_{m=1}^{M-1} \psi_m\phi_m. \tag{26}$$

PROPOSITION 5.1. *Let us consider a family of M one-dimensional grid functions of the argument m*:

$$\psi_m^{(k)} = \cos((2k-1)m\Delta\theta), \quad \Delta\theta \equiv \frac{\pi h_r}{2r_{\max}} = \frac{\pi}{2M}, \quad k = 1, 2, \ldots, M. \tag{27}$$

Then,

(I) $\{\psi^{(k)}\}_{k=1}^M \subset V$.

(II) *The functions $\psi^{(k)}$ are orthogonal with respect to the inner product* (26), *i.e.*,

$$\left\langle \psi^{(k)}, \psi^{(l)} \right\rangle = 0 \quad \textit{for } k \neq l. \tag{28}$$

(III) *The set* $\{\psi^{(k)}\}_{k=1}^{M}$ *forms a basis in* V.

(IV) $\psi_m^{(k)}$ *are even functions of the argument* m, *i.e., symmetric with respect to* $m = 0$:

$$\psi_m^{(k)} = \psi_{-m}^{(k)}.$$

(V) $\psi^{(k)}$ *are eigenfunctions of the transverse component of the finite-difference operator of* (24) *with eigenvalues* λ_k,[4] *i.e.,*

$$\begin{aligned} \boldsymbol{L}_{rr}^{h}\psi^{(k)} &= -\lambda_k \psi^{(k)}, \\ \lambda_k &= \frac{1}{3h_r^2}\left[16\sin^2\left(\frac{(2k-1)\Delta\theta}{2}\right) - \sin^2((2k-1)\Delta\theta)\right]. \end{aligned} \tag{29}$$

Proof. The inclusion (I) follows from the definition of the space V and the explicit form of the functions $\psi_m^{(k)}$ (27). To show the orthogonality (II), we calculate

$$\begin{aligned} M\cdot\left\langle \psi^{(k)}, \psi^{(l)}\right\rangle &= \sum_{m=0}^{M-1} \psi_m^{(k)}\psi_m^{(l)} - \frac{1}{2} \\ &= \sum_{m=0}^{M-1} \cos((2k-1)m\Delta\theta)\cos((2l-1)m\Delta\theta) - \frac{1}{2} \\ &= \frac{1}{2}\sum_{m=0}^{M-1}[\cos((2k+2l-2)m\Delta\theta) + \cos((2k-2l)m\Delta\theta)] - \frac{1}{2} \\ &= \frac{1}{2}\sum_{m=0}^{M-1}[\cos(2qm\Delta\theta) + \cos(2sm\Delta\theta)] - \frac{1}{2} \\ &= \frac{1}{4}\sum_{m=0}^{M-1} e^{i2qm\Delta\theta} + e^{-i2qm\Delta\theta} + e^{i2sm\Delta\theta} + e^{-i2sm\Delta\theta} - \frac{1}{2} \\ &= \frac{1}{4}\left[\frac{1-e^{i2qM\Delta\theta}}{1-e^{i2q\Delta\theta}} + \frac{1-e^{-i2qM\Delta\theta}}{1-e^{-i2q\Delta\theta}}\right] + \frac{1}{4}\left[\frac{1-e^{i2sM\Delta\theta}}{1-e^{i2s\Delta\theta}} + \frac{1-e^{-i2sM\Delta\theta}}{1-e^{-i2s\Delta\theta}}\right] - \frac{1}{2} = 0. \end{aligned}$$

We indeed obtain zero, because out of the two integer numbers $q = k + l - 1$ and $s = k - l$, one is always odd and the other is even, and thus one of the expressions in rectangular brackets on the last line in the previous chain of equalities is always equal to zero and the other one is equal to two. Property (III) follows easily from the orthogonality (II) because the orthogonality implies that the M functions $\psi^{(k)}$, $k = 1, \ldots, M$, are linearly independent, and the space V is obviously M-dimensional. Property (IV) is trivial and immediately follows from the definition (27). Finally, property (V), including the explicit expression for the eigenvalue λ_k given in (29), is obtained by directly applying the operator $\boldsymbol{L}_{rr}^{h}$ of (25a) to each $\psi^{(k)}$, $k = 1, \ldots, M$. The application of $\boldsymbol{L}_{rr}^{h}$ to a $\psi^{(k)}$ in the near-boundary nodes requires using the symmetry property (IV) and also noticing that all $\psi^{(k)}$, $k = 1, \ldots, M$, are in fact antisymmetric with respect to $m = M$, which again immediately follows from the definition (27). ■

[4] Note that for small wavenumbers the discrete eigenvalues and eigenfunctions are similar to those in the continuous formulation (cf. (14) and (15)) as $\lambda_k \approx (k-1/2)^2(\pi/r_{\max})^2$ and $\psi|_{r=mh_r} = \psi_m^{(k)} = \cos((k-1/2)\pi r/r_{\max})$.

Proposition 5.1 shows that the system $\{\psi^{(k)}\}_{k=1}^{M}$ forms an orthogonal basis of the space V, composed of the eigenfunctions of the operator $\boldsymbol{L}_{rr}^{h}$, which are symmetric with respect to $m = 0$ and vanish at $m = M$. For all n we can construct the expansion with respect to these eigenfunctions according to

$$u_{k,n} = \left\langle E_{\cdot,n}, \psi^{(k)} \right\rangle = \frac{1}{2M} E_{0,n} + \frac{1}{M} \sum_{m=1}^{M-1} E_{m,n} \cos((2k-1)m\Delta\theta), \qquad k = 1, 2, \ldots, M, \tag{30a}$$

so that

$$E_{m,n} = 2 \sum_{k=1}^{M} u_{k,n} \cos\left((2k-1)m\Delta\theta\right) = 2 \sum_{k=1}^{M} u_{k,n} \psi_m^{(k)}, \qquad m = 0, 1, \ldots, M. \tag{30b}$$

Representation (30b) can be easily verified by directly substituting $u_{k,n}$ of (30a) and performing the transformations similar to those performed when proving Proposition 5.1. Obviously, formulae (30a) and (30b) are particular versions of the direct and inverse discrete Fourier transforms, respectively.

The above eigenfunction expansion can be used to implement the transverse discrete differentiation along with the boundary conditions at $r = 0$ and $r = r_{\max}$. Indeed, if we expand $E_{m,n}$ and the RHS $\Phi_{m,n}$ in the form (30b) with the coefficients $u_{k,n}$ and $f_{k,n}$, respectively, obtained using (30a), then because of the orthogonality of the eigenfunctions $\psi^{(k)}$ (28), we arrive at the following family of one-dimensional discrete equations:[5]

$$\begin{aligned} &\boldsymbol{L}_{zz}^{h} u_{k,\cdot}\big|_{k,n} - \lambda_k u_{k,n} + k_0^2 u_{k,n} \\ &\quad \equiv \frac{-u_{k,n-2} + 16u_{k,n-1} - 30u_{k,n} + 16u_{k,n+1} - u_{k,n+2}}{12h_z^2} + k_c^2 u_{k,n} = f_{k,n}, \\ &\qquad k_c^2 = k_0^2 - \lambda_k, \quad k = 1, 2, \ldots, M, \quad n = 2, 3, \ldots, N-2, \end{aligned} \tag{31}$$

where the eigenvalues $\{\lambda_k\}$ are defined in (29). Each of the M equations of (31) is independent of the others and will be solved separately using the methodology of Section 6. Having obtained the modal solutions $u_{k,n}$ for all $k = 1, 2, \ldots, M$, we then recover the overall solution $E_{m,n}$ by means of the inverse transformation (30b).

5.1. *Implementation of Transformations (30) Using FFT*

It is convenient to implement the direct and inverse transformations (30a) and (30b) using the standard discrete Fourier transform, for which library subroutines optimized for performance are available (fast Fourier transforms). To do that, we note again (see end of the proof of Proposition 5.1) that representation (30b) allows us to extend $E_{m,n}$ for any n beyond $m = 0$ and $m = M$ using the explicit form of the basis functions $\psi^{(k)}$; see (27). The extension for negative m's is symmetric with respect to $m = 0$, and the extension beyond $m = M$ is

[5] Note the analogy to (15).

antisymmetric with respect to $m = M$. For a given function $E_{m,n}$, $m = 0, 1, \ldots, M$, it is convenient to first extend it antisymmetrically with respect to $m = M$ (so that the function is defined for $m = 0, 1, \ldots, 2M$), and then also extend it symmetrically with respect to $m = 0$ (so that it finally is defined for $m = -2M, \ldots, \mathbf{0}, \ldots, 2M$). In doing so, we arrive at a periodic grid function with the period $4M$. It is easy to see that for a function extended in this particular way the standard discrete Fourier transform

$$u_{l,n} = \frac{1}{4M} \sum_{m=-2M}^{2M-1} E_{m,n} e^{-ilm\Delta\theta}, \quad l = -2M, \ldots, 2M-1, \tag{32a}$$

reduces to (30a). Indeed, as $E_{m,n}$ is real we will always have $u_{l,n} = \bar{u}_{-l,n}$, and in this particular case the symmetry with respect to $m = 0$ implies that all $u_{l,n}$ are also real and thus $u_{l,n} = u_{-l,n}$. Consequently, we can consider only $2M + 1$ independent real coefficients $u_{l,n}$ for $l = 0, 1, \ldots, 2M$. Then, the antisymmetry with respect to $m = M$ will yield that $u_{l,n} = 0$ for all even $l = 0, 2, 4, \ldots, 2M$ and we are thus left with only the coefficients $u_{l,n}$ for odd $l = 1, 3, 5, \ldots, 2M - 1$. In other words, we can rewrite (32a) as

$$u_{l,n} = \frac{1}{2M} E_{0,n} + \frac{1}{M} \sum_{m=1}^{M-1} E_{m,n} \cos(lm\Delta\theta), \quad l = 1, 3, \ldots, 2M-1,$$

and conclude that it indeed coincides with (30a) if we change notations from $l = 1, 3, 5, \ldots, 2M - 1$ to $k = (l + 1)/2$, $k = 1, 2, \ldots, M$. Similarly, it is easy to see that because of the aforementioned properties of $u_{l,n}$ ($u_{l,n} = u_{-l,n}$, $u_{l,n}$ real, and $u_{l,n} = 0$ for $l = 0, 2, 4, \ldots, 2M$), the standard inverse discrete Fourier transform

$$E_{m,n} = \frac{1}{4M} \sum_{l=-2M}^{2M-1} u_{l,n} e^{ilm\Delta\theta}, \quad m = -2M, \ldots, 2M, \tag{32b}$$

reduces to (30b).

6. THE ONE-DIMENSIONAL DISCRETE HELMHOLTZ EQUATION

In this section we analyze the discrete one-dimensional linear nonhomogeneous Helmholtz equation (31), paying special attention to the treatment of the boundary conditions for $z = 0$ and $z = z_{\max}$. We recall that the boundary conditions at $z = z_{\max}$ should guarantee that this boundary be transparent for all waves traveling to the right (i.e., a standard radiation ABC). The boundary conditions at $z = 0$ should guarantee that this boundary be transparent for all backscattered waves traveling to the left, and at the same time impose the given incoming wave field (*two-way ABC*). We emphasize that we have not discussed a particular discrete form of these boundary conditions until now, since typically the ABCs are most convenient to set in the transformed space rather than in the original space [34].

To simplify the notations, we drop the subscript k, so that Eq. (31) takes the form

$$\frac{-u_{n-2} + 16u_{n-1} - 30u_n + 16u_{n+1} - u_{n+2}}{12h_z^2} + k_c^2 u_n = f_n, \quad n = 2, 3, \ldots, N-2. \tag{33}$$

Equation (33) is a fourth-order difference equation. It is obtained, however, as a fourth-order accurate difference approximation to the second-order differential equation. Therefore, compared to its original continuous counterpart, the difference equation (33) requires additional boundary conditions. A total of four boundary conditions are needed to guarantee the solvability and uniqueness for Eq. (33). Two extra boundary conditions that are not present in the continuous case are a pure numerical artifact. They are accounted for by the presence of two extra evanescent waves among the solutions of the homogeneous version of Eq. (33) in addition to the two standard traveling or evanescent waves (see Section 6.1). Altogether, these four boundary conditions should ensure the desired behavior of the solution near $z = 0$ and near $z = z_{\max}$. We also reiterate that the finite-difference equation itself obviously cannot be written in the form (33) for the grid nodes $n = 0, 1, N-1$, and N. A special form of the discrete equation for these four grid nodes is therefore required; this special form will actually constitute the boundary conditions and make the total number of equations in the linear system equal to the number of unknowns.

6.1. *The Discrete Homogeneous Problem*

We start by analyzing the homogeneous counterpart to the finite-difference equation (33) over an infinite grid domain, i.e.,

$$\frac{-u_{n-2} + 16u_{n-1} - 30u_n + 16u_{n+1} - u_{n+2}}{12h_z^2} + k_c^2 u_n = 0, \qquad n = 0, \pm 1, \pm 2, \ldots. \tag{34}$$

PROPOSITION 6.1. *Let $\alpha = (h_z k_c)^2$ be such that either $0 < \alpha < 16/3$ or $-3 \le \alpha < 0$. Then, the general solution to Eq.* (34) *has the form*

$$u_n = c_1 q_1^n + c_2 q_2^n + c_{-1} q_1^{-n} + c_{-2} q_2^{-n}, \tag{35}$$

where c_1, c_2, c_{-1}, and c_{-2} are arbitrary constants, and q_1 and q_2 are roots of the characteristic equation that corresponds to (34).

In addition,

(I) *when $0 < \alpha < 16/3$, q_1^n and q_1^{-n} are waves propagating to the right and to the left, respectively. In particular, when $0 < \alpha \ll 1$, then*

$$q_1 = e^{ik_c h_z} + \mathcal{O}((k_c \cdot h_z)^5), \tag{36a}$$

$$q_2 = e^{-ik_c h_z} + \mathcal{O}((k_c \cdot h_z)^5), \tag{36b}$$

and as such, q_1^n and q_1^{-n} are the discrete analogues of the right and left traveling waves $e^{ik_c z}$ and $e^{-ik_c z}$, respectively, with fourth-order accuracy.

(II) *When $-3 \le \alpha < 0$, q_1^n and q_1^{-n} are evanescent waves decaying to the right and to the left, respectively.*

(III) *In both cases, i.e., for $0 < \alpha < 16/3$ and for $-3 \le \alpha < 0$, q_2^n and q_2^{-n} are evanescent waves decaying to the right and to the left, respectively.*

Proof. Let us introduce the characteristic algebraic equation

$$-1 + 16q + (12\alpha - 30)q^2 + 16q^3 - q^4 = 0 \tag{37}$$

for the homogeneous finite-difference equation (34). It is generally known (see, e.g., [15]) that if all the roots q_j of a given characteristic algebraic equation are distinct, then the general solution to the corresponding homogeneous finite-difference equation is obtained as a linear span of the grid functions q_j^n, where the power n is determined by the grid location. In the specific case that we are studying, Eq. (37) is a quartic algebraic equation and thus provided that its four roots $\{q_j\}_{j=1}^4$ are distinct, the general solution to the homogeneous equation (34) has the form

$$u_n = c_1 q_1^n + c_2 q_2^n + c_3 q_3^n + c_4 q_4^n, \tag{38}$$

where $\{c_j\}_{j=1}^4$ are arbitrary constants.

Hereafter, we restrict ourselves only to the case when the roots $\{q_j\}_{j=1}^4$ of Eq. (37) are distinct. By explicitly calculating $\{q_j\}_{j=1}^4$ (see below), we will show that multiple roots are only possible for the two cases $\alpha = 0$ and $\alpha = 16/3$, which are easy to avoid in practical computations.

To simplify the actual calculation of the roots of quartic Eq. (37), we first note that by dividing (37) by q^4 we arrive at exactly the same equation for $1/q$. Therefore, if q is a root, then q^{-1} is also a root (this follows, of course, from the fact that the discretization (33) is symmetric). Accordingly, we rename the four roots of Eq. (37) as q_1, q_2, q_1^{-1}, and q_2^{-1}, and write

$$\begin{aligned} &-1 + 16q + (12\alpha - 30)q^2 + 16q^3 - q^4 \\ &\quad = -(q - q_1)\big(q - q_1^{-1}\big)(q - q_2)\big(q - q_2^{-1}\big) \\ &\quad = -(q^2 - d_1 q + 1)(q^2 - d_2 q + 1) \\ &\quad = -1 + (d_1 + d_2)q - (2 + d_1 d_2)q^2 + (d_1 + d_2)q^3 - q^4, \end{aligned} \tag{39}$$

where

$$d_1 = q_1 + q_1^{-1}, \quad d_2 = q_2 + q_2^{-1}. \tag{40}$$

By comparing the beginning and the end in the chain of equalities (39) we obtain system of equations for d_1 and d_2

$$d_1 + d_2 = 16, \quad -2 - d_1 d_2 = 12\alpha - 30,$$

from which we find that

$$d_1 = 8 - 6\sqrt{1 + \alpha/3}, \quad d_2 = 8 + 6\sqrt{1 + \alpha/3}. \tag{41}$$

From formulae (41) we conclude that both d_1 and d_2 are real provided that $\alpha \geq -3$. If, for example, $h_r \approx h_z$ (the cell aspect ratio of the discretization is close to one), then the definition of k_c (see (31)), where λ_k is given by (29), along with the definition of $\alpha = (h_z k_z)^2$, suggests that even for negative α's their absolute values are sufficiently small, and thus we can always assume that $\alpha \geq -3$ and consequently consider d_1 and d_2 real. However, allowing for the complex values of d_1 and d_2 may make the analysis more cumbersome, but it does not change any of the results hereafter. This, in particular, is corroborated by the computations of Section 7.1, which were conducted on the grids with cell aspect ratios $20/1$ and $20/3$.

From (40) we know that q_j and q_j^{-1} are the roots of the quadratic equation

$$q^2 - d_j q + 1 = 0, \quad j = 1, 2. \tag{42}$$

Let us analyze the case $j = 1$ first. For $0 < \alpha < 16/3$, Eq. (42) has two complex conjugate roots

$$q_1 = \frac{d_1 + i\sqrt{4 - d_1^2}}{2}, \quad q_1^{-1} = \frac{d_1 - i\sqrt{4 - d_1^2}}{2}. \tag{43}$$

From (43) it follows that $|q_1| = |q_1^{-1}| = 1$ and, in addition, that when $0 < \alpha \ll 1$ then (36) holds.

When $-3 \leq \alpha < 0$, we have

$$q_1 = \frac{d_1 - \sqrt{d_1^2 - 4}}{2}, \quad q_1^{-1} = \frac{d_1 + \sqrt{d_1^2 - 4}}{2}. \tag{44}$$

Therefore, both roots are real and satisfy $|q_1| < 1$ and $|q_1^{-1}| > 1$, showing that q_1^n and q_1^{-n} of (44) are discrete analogues of two evanescent waves. We note that as α changes from positive to negative in formulae (43), the right-propagating wave q_1^n changes into an exponential decreasing to the right and the left-propagating q_1^{-n} wave changes into an exponential decreasing to the left, a fact that simplifies the identification of the right and left traveling and decaying waves in the actual implementation of the boundary conditions at $z = 0$ and $z = z_{\max}$.

The case $\alpha \geq 16/3$ remains to be considered. For the positive values of k_c^2, we can introduce the wavelength $\lambda_c = 2\pi/k_c$ and for this range of α obtain $\lambda_c/h_z \leq \sqrt{3}\pi/2$. Thus, we see that $\alpha \geq 16/3$ implies a poor "points per wavelength" resolution even for the long waves $\lambda_c > \lambda_0 = 2\pi/k_0$. This makes the choice $\alpha \geq 16/3$ inappropriate for practical computations. Finally, regarding the last case that has not yet been considered, $\alpha = 0$, we note that for this value of α Eq. (42) will have a double root $q_1 = q_{-1} = 1$. However, formulae (29) and (31) show that the case $\alpha = 0 \Leftrightarrow k_c^2 = 0$ can be easily avoided by slightly changing the parameters of the discretization.

For $j = 2$, we find from Eq. (42) that

$$q_2 = \frac{d_2 - \sqrt{d_2^2 - 4}}{2}, \quad q_2^{-1} = \frac{d_2 + \sqrt{d_2^2 - 4}}{2}. \tag{45}$$

Clearly, $|q_2| < 1$, $|q_2^{-1}| > 1$ for all relevant values of α ($\alpha \geq -3$), i.e., the two components q_2^n and q_2^{-n} of (45) always correspond to evanescent waves. ■

6.2. *Discrete One-Way Helmholtz Equations*

In analogy with the continuous description in Section 4.2.1, we now construct the discrete one-way Helmholtz equations based on the solution (35) of the homogeneous finite-difference scheme (34). From the very beginning, we think of these discrete one-way Helmholtz equations as the relations to be used as boundary conditions for Eq. (33).

According to Proposition 6.1, the discrete homogeneous equation (34) has four linearly independent eigenfunctions, two of which are either traveling or evanescent waves and

two of which are always evanescent waves; the presence of the latter (in contrast to the continuous case) is due to the fact that (34) is a fourth–order finite-difference equation that approximates the original second-order differential equation. When constructing the discrete one-way Helmholtz equations, we of course first need to make sure that they handle the first pair of discrete waves, q_1^n and q_1^{-n}, in the same way that Eqs. (17) handle the corresponding continuous waves. In addition, we need to decide how the discrete one-way Helmholtz equations will handle the second pair of discrete waves, q_2^n and q_2^{-n}, which are purely numerical (i.e., due to the use of a fourth-order difference scheme). It is natural to require that the one-way-to-the-right discrete Helmholtz equation admit the right traveling/evanescent wave q_1^n and the right evanescent wave q_2^n and that the other two waves from representation (35), q_1^{-n} (left traveling/evanescent) and q_2^{-n} (left evanescent), be suppressed by this equation. Indeed, q_1^{-n} may either be traveling "the wrong way" or grow without bound as $n \to +\infty$, and q_2^{-n} will always grow without bound as $n \to +\infty$.[6] Clearly, if we use the one-way-to-the-right equation that possesses such properties as a boundary condition for (34) near $n = N$, it will guarantee that the corresponding far-field solution ($n > N$) is always bounded at infinity and also that this solution is only composed of outgoing (right propagating and/or right decaying) waves. In other words, the one-way-to-the-right discrete Helmholtz equation implies that in the far field $n > N$ one can represent the solution u_n in the "restricted" form

$$u_n = c_1 q_1^n + c_2 q_2^n, \tag{46}$$

as opposed to the general form (35). Formula (46) is equivalent to requiring that the vector $[u_{N-3}, u_{N-2}, u_{N-1}, u_N]$ be a linear combination of the two vectors $[1, q_1, q_1^2, q_1^3]$ and $[1, q_2, q_2^2, q_2^3]$, which is the same as requiring that

$$\text{Rank} \begin{bmatrix} u_{N-3} & u_{N-2} & u_{N-1} & u_N \\ 1 & q_1 & q_1^2 & q_1^3 \\ 1 & q_2 & q_2^2 & q_2^3 \end{bmatrix} = 2. \tag{47}$$

Relation (47) immediately yields the two linearly independent conditions

$$\det \begin{bmatrix} u_{N-3} & u_{N-2} & u_{N-1} \\ 1 & q_1 & q_1^2 \\ 1 & q_2 & q_2^2 \end{bmatrix} = 0, \quad \det \begin{bmatrix} u_{N-2} & u_{N-1} & u_N \\ q_1 & q_1^2 & q_1^3 \\ q_2 & q_2^2 & q_2^3 \end{bmatrix} = 0,$$

which reduce to

$$q_1 q_2 u_{N-3} - (q_1 + q_2) u_{N-2} + u_{N-1} = 0 \tag{48a}$$

and

$$q_1 q_2 u_{N-2} - (q_1 + q_2) u_{N-1} + u_N = 0. \tag{48b}$$

The two scalar Eqs. (48a) and (48b) constitute the one-way-to-the-right discrete Helmholtz equation.

[6] Besides being "natural," this choice is also motivated by the well-posedness considerations, as the analysis of [14, 23] suggests.

The one-way-to-the-left discrete Helmholtz equation is constructed similarly. Symmetrically to the previous case, we require that it admit the left traveling/evanescent wave q_1^{-n} and the left evanescent wave q_2^{-n} and that the other two waves from representation (35), q_1^n (right traveling/evanescent) and q_2^n (right evanescent), be prohibited by this equation. (From the standpoint of physics the two waves, q_1^{-n} and q_2^{-n}, account for the phenomenon of backscattering.) The waves q_1^n and q_2^n are to be suppressed in this case because q_1^n may either be traveling "the wrong way," i.e., to the right, or grow without bound as $n \to -\infty$, and q_2^n will always grow without bound as $n \to -\infty$. If the one-way-to-the-left discrete Helmholtz equation is used as the boundary condition for (34) near $n = 0$, it will guarantee that the corresponding far-field solution ($n < 0$) is always bounded as $z \to -\infty$, and also that this solution is only composed of outgoing (left propagating and/or left decaying) waves. In other words, the one-way-to-the-left discrete Helmholtz equation implies that in the far field $n < 0$ one can represent the solution u_n in the "restricted" form

$$u_n = c_{-1}q_1^{-n}n + c_{-2}q_2^{-n}, \tag{49}$$

as opposed to the general form (35). To make sure that representation (49) holds, we require that the vector $[u_0, u_1, u_2, u_3]$ be a linear combination of $[1, q_1^{-1}, q_1^{-2}, q_1^{-3}]$ and $[1, q_2^{-1}, q_2^{-2}, q_2^{-3}]$:

$$\text{Rank}\begin{bmatrix} u_0 & u_1 & u_2 & u_3 \\ 1 & q_1^{-1} & q_1^{-2} & q_1^{-3} \\ 1 & q_2^{-1} & q_2^{-2} & q_2^{-3} \end{bmatrix} = 2. \tag{50}$$

Relation (50) is equivalent to the two linearly independent homogeneous conditions

$$u_0 - (q_1 + q_2)u_1 + q_1q_2u_2 = 0 \tag{51a}$$

and

$$u_1 - (q_1 + q_2)u_2 + q_1q_2u_3 = 0, \tag{51b}$$

which constitute the one-way-to-the-left discrete Helmholtz equation.

We note that splitting the general solution (35) into right- and left-going waves (Eqs. (46) and (49), respectively), and allowing for only one direction while prohibiting the other at the corresponding edges of the interval, constitutes the *radiation principle* in the finite-difference discrete framework.

Having constructed the one-way discrete Helmholtz equations (48) and (51), we now implement them as boundary conditions for the discrete homogeneous equation (34). If we consider the finite grid $n = 0, 1, \ldots, N$ on the interval $[0, z_{\max}]$, the five-node difference stencil cannot be centered at the near-edge nodes $n = 0$, 1, $N - 1$, and N. As a consequence, the number of equations in the linear system is four less than the number of unknowns. To make the number of equations and the number of unknowns equal, we supplement Eqs. (34) on the grid $n = 2, 3, \ldots, N - 2$ by Eqs. (51a) and (51b) for $n = 0$ and $n = 1$, respectively, and by Eqs. (48a) and (48b) for $n = N - 1$ and $n = N$, respectively. In doing so, we arrive at a linear homogeneous algebraic system with $N + 1$ equations and $N + 1$ unknowns,

$$\boldsymbol{Au} = \boldsymbol{0}, \tag{52}$$

where

$$\boldsymbol{A} = \frac{1}{12h_z^2}\begin{bmatrix} 1 & -(q_1+q_2) & q_1q_2 & 0 & 0 & \dots & 0 \\ 0 & 1 & -(q_1+q_2) & q_1q_2 & 0 & \dots & 0 \\ -1 & 16 & (12\alpha-30) & 16 & -1 & \dots & 0 \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \dots & -1 & 16 & (12\alpha-30) & 16 & -1 \\ 0 & \dots & 0 & q_1q_2 & -(q_1+q_2) & 1 & 0 \\ 0 & \dots & 0 & 0 & q_1q_2 & -(q_1+q_2) & 1 \end{bmatrix} \tag{53}$$

and, obviously, $\boldsymbol{u} = [u_0, u_1, \dots, u_N]^T$.

The following Proposition (6.2) establishes the solvability and uniqueness of the solution for the nonhomogeneous counterpart of system (52).

PROPOSITION 6.2. *The linear nonhomogeneous system of equations* $\boldsymbol{Au} = \boldsymbol{f}$ *with the matrix* $\boldsymbol{A}$ *given by* (53) *is uniquely solvable for any right-hand side* $\boldsymbol{f} = [f_0, f_1, \dots, f_N]^T$.

Proof. We show that the corresponding homogeneous system (52) has only a trivial solution. Indeed, the only solution to any of the equations of $\boldsymbol{Au} = \boldsymbol{0}$ except the first two and the last two is a linear combination of the type (35). However, each of the components of (35) is explicitly eliminated by one of the boundary conditions (48a), (48b), (51a), or (51b), i.e., by one of the one-way discrete Helmholtz equations (the first two and the last two equations of $\boldsymbol{Au} = \boldsymbol{0}$). Therefore, the only solution to the homogeneous system is the trivial one.[7] ■

Although we have just shown that one can find the solution to $\boldsymbol{Au} = \boldsymbol{f}$ for any given $\boldsymbol{f} = [f_0, f_1, \dots, f_N]$, this solution will not, in fact, correctly approximate the corresponding solution of the nonhomogeneous differential equation, or in other words, it will not, generally speaking, be the discrete radiation solution from the sources $\boldsymbol{f} = [f_0, f_1, \dots, f_N]$. The reason for this discrepancy is that the one-way Helmholtz equations which are used in the first two rows and the last two rows of the matrix $\boldsymbol{A}$ have been constructed for the homogeneous case. As a result, these four equations will not handle the near-boundary source terms correctly, which may, generally speaking, be present. The "cure" to this problem, in the form of a local modification to $\boldsymbol{f}$, is derived in Section 6.4.

In our simulations (see Section 7), we solve the finite-difference Helmholtz equation by inverting the matrix $\boldsymbol{A}$ of (53). However, for the purpose of deriving the two-way ABCs that would correctly handle the near-boundary inhomogeneities, we now show how to construct the solution $\boldsymbol{u}$ by using the Green's function of the finite-difference operator of (34). As we shall see, this approach is rather expensive numerically and thus not useful for actual computing. However, it provides the most conceptually straightforward way to build the radiation solution. Moreover, the analysis that employs the Green's function reveals the mechanism of the aforementioned discrepancy between the radiation from the sources $\boldsymbol{f} = [f_0, f_1, \dots, f_N]$ and the solution to $\boldsymbol{Au} = \boldsymbol{f}$.

[7] This solvability result is obviously similar to the one in the continuous case; see Section 4.2.1.

6.3. *Radiation Solution by Means of the Green's Function*

In this section, we introduce a problem very similar to (33), except that the solution $\boldsymbol{u}$ is now defined on the infinite grid $n = 0, \pm 1, \pm 2, \ldots$, and the right-hand side f_n is compactly supported:

$$\begin{gathered}\frac{-u_{n-2} + 16u_{n-1} - 30u_n + 16u_{n+1} - u_{n+2}}{12h_z^2} + k_c^2 u_n = f_n, \\ n = 0, \pm 1, \pm 2, \ldots, \\ f_n = 0 \quad \text{for } n < 0 \quad \text{and} \quad n > N.\end{gathered} \tag{54}$$

We also require that the solution u_n of (54) satisfy the radiation principle in the areas of homogeneity $n < 0$ and $n > N$. In other words, we require that for $n \le 0$ one can represent u_n in the form (49) and for $n \ge N$ in the form (46). This is the most general formulation of the problem of finding the solution that corresponds to the radiation of waves by the sources $\boldsymbol{f} = [f_0, f_1, \ldots, f_N]^T$ in the finite-difference framework.

To solve this problem, we introduce the fundamental solution G^n (free-space Green's function) for the one-dimensional discrete Helmholtz operator, which is defined on the entire infinite grid $n = 0, \pm 1, \pm 2, \ldots$ and is the solution of the equation

$$\begin{gathered}\frac{-G^{n-2} + 16G^{n-1} - 30G^n + 16G^{n+1} - G^{n+2}}{12h_z^2} + k_c^2 G^n = \delta_n, \\ n = 0, \pm 1, \pm 2, \ldots,\end{gathered} \tag{55}$$

where

$$\delta_n = \begin{cases} 1, & n = 0 \\ 0, & n \ne 0. \end{cases}$$

We also require that the Green's function G^n satisfy the radiation principle as $n \to \pm\infty$, or in other words, that it can be represented in the following form:

$$G^n = \begin{cases} a_1 q_1^n + a_2 q_2^n, & n \ge 0 \\ b_1 q_1^{-n} + b_2 q_2^{-n}, & n \le 0. \end{cases} \tag{56}$$

PROPOSITION 6.3. *The values of the constants* a_1, a_2, b_1, b_2 *in* (56) *are given by*

$$a_1 = \frac{12h_z^2 q_1}{\left(q_2^{-1} - q_1\right)\left(q_1^{-1} - q_1\right)(q_2 - q_1)}, \tag{57a}$$

$$a_2 = \frac{-12h_z^2 q_2}{\left(q_2^{-1} - q_2\right)\left(q_1^{-1} - q_2\right)(q_2 - q_1)}, \tag{57b}$$

$$b_1 = \frac{-12h_z^2 q_1^{-1}}{\left(q_2^{-1} - q_1^{-1}\right)\left(q_1^{-1} - q_2\right)\left(q_1^{-1} - q_1\right)}, \tag{57c}$$

$$b_2 = \frac{12h_z^2 q_2^{-1}}{\left(q_2^{-1} - q_1^{-1}\right)\left(q_2^{-1} - q_2\right)\left(q_2^{-1} - q_1\right)}. \tag{57d}$$

Proof. To find these four constants, we need four equations. By matching the two branches (56) of the Green's function G^n at $n = 0$ we immediately obtain one equation:

$$a_1 + a_2 = b_1 + b_2. \tag{58a}$$

The other three equations for the coefficients of (56) are obtained from the original Eq. (55) written for the nodes $n = 0$, 1, and -1. For $n = 0$ we have

$$-G^{-2} + 16G^{-1} + (12\alpha - 30)G^0 + 16G^1 - G^2 = 12h_z^2,$$

or

$$\begin{aligned} &-\left(b_1q_1^2 + b_2q_2^2\right) + 16\left(b_1q_1 + b_2q_2\right) + (12\alpha - 30)(a_1 + a_2) + 16(a_1q_1 + a_2q_2) \\ &\quad -\left(a_1q_1^2 + a_2q_2^2\right) = 12h_z^2. \end{aligned}$$

The previous equation can be simplified by subtracting from it the relation

$$\begin{aligned} &-\left(a_1q_1^{-2} + a_2q_2^{-2}\right) + 16\left(a_1q_1^{-1} + a_2q_2^{-1}\right) + (12\alpha - 30)(a_1 + a_2) + 16(a_1q_1 + a_2q_2) \\ &\quad -\left(a_1q_1^2 + a_2q_2^2\right) = 0, \end{aligned}$$

which comes from the fact that each branch of the Green's function (the right branch $a_1q_1^n + a_2q_2^n$ in this particular instance) satisfies the homogeneous finite-difference equation (34). The subtraction yields

$$-\left(b_1q_1^2 + b_2q_2^2\right) + 16(b_1q_1 + b_2q_2) - 16\left(a_1q_1^{-1} + a_2q_2^{-1}\right) + \left(a_1q_1^{-2} + a_2q_2^{-2}\right) = 12h_z^2. \tag{58b}$$

For $n = 1$, Eq. (55) takes the form

$$-G^{-1} + 16G^0 + (12\alpha - 30)G^1 + 16G^2 - G^3 = 0,$$

and again, using the homogeneous equation for the right branch of the Green's function, we obtain

$$-(b_1q_1 + b_2q_2) + \left(a_1q_1^{-1} + a_2q_2^{-1}\right) = 0. \tag{58c}$$

Finally, for $n = -1$ we have

$$-G^{-3} + 16G^{-2} + (12\alpha - 30)G^{-1} + 16G^0 - G^1 = 0.$$

Combining this relation with the homogeneous difference equation for the left branch of the Green's function, we arrive at

$$\left(b_1q_1^{-1} + b_2q_2^{-1}\right) - (a_1q_1 + a_2q_2) = 0. \tag{58d}$$

Now we need to solve Eqs. (58) for a_1, a_2, b_1, and b_2. First, we multiply (58c) by 16 and substitute it into (58b), and then we rewrite all four equations as follows:

$$\begin{bmatrix} q_1^{-2} & q_2^{-2} & -q_1^2 & -q_2^2 \\ q_1^{-1} & q_2^{-1} & -q_1 & -q_2 \\ 1 & 1 & -1 & -1 \\ q_1 & q_2 & -q_1^{-1} & -q_2^{-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 12h_z^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{59}$$

The determinant of system (59) is easily reduced to a Vandermonde determinant, which eventually leads to expressions (57). ■

From the definition of G^n, we have the following.

PROPOSITION 6.4. *For any given right-hand side f_n compactly supported on $[0, 1, \ldots, N]$, the solution to* (54), *subject to the radiation principle, is given by the convolution*

$$u_n = \sum_{m=0}^{m=N} f_m G^{n-m}, \quad n = 0, \pm 1, \pm 2, \ldots. \tag{60}$$

6.4. *Radiation Solution by Means of Inverting the Matrix* $\boldsymbol{A}$

The cost of calculating the convolutions in (60) for $n = 0, 1, \ldots, N$ is $\mathcal{O}(N^2)$. We now show that the portion of the solution (60) that we are interested in, namely u_n for $n = 0, 1, \ldots, N$, can be recovered by means of inverting the matrix $\boldsymbol{A}$ of (53). The cost of this inversion will be only $\mathcal{O}(N)$ operations because the matrix $\boldsymbol{A}$ is pentadiagonal; see Section 6.7 for additional details.

PROPOSITION 6.5. *Let $\boldsymbol{A}$ be defined by* (53) *and $\boldsymbol{u} = [u_0, u_1, \ldots, u_N]^T$ be defined by* (60) *for $n = 0, 1, \ldots, N$. Denote $\boldsymbol{f} = [f_0, f_1, \ldots, f_{N-1}, f_N]^T$. Then, $\boldsymbol{Au} = \tilde{\boldsymbol{f}}$, where*

$$\tilde{\boldsymbol{f}} \stackrel{\text{def}}{=} \begin{bmatrix} 0 \\ 0 \\ f_2 \\ \vdots \\ f_{N-2} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \tilde{f}_0 \\ \tilde{f}_1 \\ 0 \\ \vdots \\ 0 \\ \tilde{f}_{N-1} \\ \tilde{f}_N \end{bmatrix}, \tag{61}$$

$$\tilde{f}_0 \stackrel{\text{def}}{=} \frac{1}{12h_z^2}[(f_0G^0 + f_1G^{-1}) - (q_1 + q_2)(f_0G^1 + f_1G^0) + q_1q_2(f_0G^2 + f_1G^1)], \tag{62a}$$

$$\tilde{f}_1 \stackrel{\text{def}}{=} \frac{1}{12h_z^2}[(f_0G^1 + f_1G^0 + f_2G^{-1}) - (q_1 + q_2)(f_0G^2 + f_1G^1 + f_2G^0) + q_1q_2(f_0G^3 + f_1G^2 + f_2G^1)], \tag{62b}$$

$$\tilde{f}_{N-1} \stackrel{\text{def}}{=} \frac{1}{12h_z^2}[q_1q_2(f_{N-2}G^{-1} + f_{N-1}G^{-2} + f_NG^{-3}) - (q_1 + q_2)(f_{N-2}G^0 + f_{N-1}G^{-1} + f_NG^{-2}) + (f_{N-2}G^1 + f_{N-1}G^0 + f_NG^{-1})], \tag{62c}$$

and

$$\tilde{f}_N \stackrel{\text{def}}{=} \frac{1}{12h_z^2}[q_1q_2(f_{N-1}G^{-1} + f_NG^{-2}) - (q_1 + q_2)(f_{N-1}G^0 + f_NG^{-1}) + (f_{N-1}G^1 + f_NG^0)]. \tag{62d}$$

Proof. By definition of the Green's function G^n (see Section 6.3), $(\boldsymbol{Au})_n = f_n$ for $2 \le n \le N-2$. Indeed, for $2 \le n \le N-2$ we have

$$
\begin{aligned}
(\boldsymbol{Au})_n &= \frac{1}{12h_z^2}[-u_{n-2} + 16u_{n-1} + (12\alpha - 30)u_n + 16u_{n+1} - u_{n+2}] \\
&= \frac{1}{12h_z^2}\left[-\sum_{m=0}^{N} f_m G^{n-2-m} + 16\sum_{m=0}^{N} f_m G^{n-1-m} + (12\alpha-30)\sum_{m=0}^{N} f_m G^{n-m}\right. \\
&\quad \left. + 16\sum_{m=0}^{N} f_m G^{n+1-m} - \sum_{m=0}^{N} f_m G^{n+2-m}\right] \\
&= \frac{1}{12h_z^2}\sum_{m=0}^{N} f_m(-G^{n-2-m} + 16G^{n-1-m} + (12\alpha-30)G^{n-m} + 16G^{n+1-m} - G^{n+2-m}) \\
&= \sum_{m=0}^{N} \delta_{n-m} f_m = f_n.
\end{aligned}
$$

As for $(\boldsymbol{Au})_0$, $(\boldsymbol{Au})_1$, $(\boldsymbol{Au})_{N-1}$, and $(\boldsymbol{Au})_N$, these four components need to be calculated separately. They will, generally speaking, differ from f_0, f_1, f_{N-1}, and f_N, respectively, because of the special structure of the first two and the last two rows of the matrix $\boldsymbol{A}$, which admit waves going in only one direction; see Section 6.2.

We start the analysis from the left edge of the interval. Clearly, any f_m for $m \ge 2$ is not going to contribute to $(\boldsymbol{Au})_0$ because when substituting $\boldsymbol{u}$ of (60) into (51a), we in fact, substitute only the left branch of the Green's function G^{n-m}; see (56). Indeed, in formula (51a) we only need the values of u_n for $n = 0, 1, 2$, and if $m \ge 2$ this implies $n - m \le 0$. The left branch of the Green's function (56) by definition turns (51a) into an identity; therefore $(\boldsymbol{Au})_0$ is not affected by f_m for $m \ge 2$. Consequently,

$$(\boldsymbol{Au})_0 = (\boldsymbol{A}[\, f_0 G^n + f_1 G^{n-1}])_0,$$

which proves (62a). Similarly, substitution of the left branch of the Green's function into (51b) suggests that any f_m for $m \ge 3$ is not going to contribute to $(\boldsymbol{Au})_1$. Therefore,

$$(\boldsymbol{Au})_1 = (\boldsymbol{A}[\, f_0 G^n + f_1 G^{n-1} + f_2 G^{n-2}])_1,$$

which proves (62b).

Similar analysis is conducted for the right edge of the interval. Only f_N and f_{N-1} affect $(\boldsymbol{Au})_N = \tilde{f}_N$ because for all other components of the RHS $\boldsymbol{f}$ the contribution to the solution $\boldsymbol{u}$ at $n = N-2, N-1, N$ is given by the right branch of the Green's function only; then the explicit form of the solution (60) and the definition of $\boldsymbol{A}$ (53) easily yield expression (62d). Analogously, only three components of the right-hand side, f_N, f_{N-1}, and f_{N-2}, contribute to $(\boldsymbol{Au})_{N-1} = \tilde{f}_{N-1}$, which together with (60) and (53) implies (62c). ■

From the standpoint of the original physical model the situation near $z = z_{\max}$ differs substantially from the situation near $z = 0$, because we can always make the effect of nonlinearity and/or variation of coefficients near $z = z_{\max}$ negligible, by taking $z_{\max}$ sufficiently large. *Therefore, from here on we will always assume that* $f_N = f_{N-1} = f_{N-2} = 0$.

Obviously, if we use the RHS $\boldsymbol{f} = [\,f_0,\ f_1, \ldots,\ f_{N-3}, 0, 0, 0]^T$ of this particular kind as source terms in (54), then for the corresponding solution $\boldsymbol{u} = [u_0, u_1, \ldots, u_N]$ we will have $(\boldsymbol{Au})_{N-1} = \tilde{f}_{N-1} = 0$ (see 62c) and $(\boldsymbol{Au})_N = \tilde{f}_N = 0$ (see 62d). In other words, the modified right-hand side $\tilde{\boldsymbol{f}}$ of (60) in this case becomes $\tilde{\boldsymbol{f}} = [\,\tilde{f}_0,\ \tilde{f}_1,\ f_2, \ldots,\ f_{N-3}, 0, 0, 0]^T$.

Let us emphasize that $\tilde{f}_0 = (\boldsymbol{Au})_0$ (see 62a) depends on f_0 and f_1, and $\tilde{f}_1 = (\boldsymbol{Au})_1$, (see 62b) depends on f_0, f_1, and f_2. Likewise, in order to obtain $\tilde{f}_{N-1} = (\boldsymbol{Au})_{N-1} = 0$ (see 62c) and $\tilde{f}_N = (\boldsymbol{Au})_N = 0$ (see 62d), in addition to the obvious requirement that $f_N = f_{N-1} = 0$, we also need to impose $f_{N-2} = 0$.

Propositions 6.2 and 6.5 guarantee that the *only* solution of the linear system $\boldsymbol{Au} = \tilde{\boldsymbol{f}}$, where $\tilde{\boldsymbol{f}} = [\,\tilde{f}_0,\ \tilde{f}_1,\ f_2, \ldots,\ f_{N-3}, 0, 0, 0]^T$, is the solution $\boldsymbol{u}$ of (54) with the RHS $\boldsymbol{f} = [\,f_0,\ f_1,\ f_2, \ldots,\ f_{N-3}, 0, 0, 0]$ subject to the radiation principle. Thus, we have addressed the concern raised at the end of Section 6.2, namely which modifications to the right-hand side $\boldsymbol{f}$ are needed to ensure that the solution obtained by inverting the matrix $\boldsymbol{A}$ will coincide with the pure radiation solution from these particular sources $\boldsymbol{f}$. Provided that near the right edge of the interval the RHS is zero, that is, $f_N = f_{N-1} = f_{N-2} = 0$, it turns out that these modifications are local and require only the replacement of the two old quantities f_0 and f_1 near the left edge of the interval by the new quantities $\tilde{f}_0$ and $\tilde{f}_1$, respectively. It is also important to mention that formulae (62a), (62b) are by themselves local as well, and therefore the overall modification $\boldsymbol{f} \mapsto \tilde{\boldsymbol{f}}$ amounts to only local, and thus numerically inexpensive, operations on the grid near $n = 0$. Finally, we note that besides the aforementioned unique solvability, the well-posedness of the discrete formulation needs to be established; for the problem studied in this paper it can be done by applying the theory of [14, 23].

6.5. *Adding the Incoming Power*

The boundary conditions at $z = 0$ should guarantee the complete transparency of this boundary for all backscattered waves and at the same time be capable of accurately prescribing the incoming signal; the combination of these two properties has been referred to as *the two-way ABCs*. Similar to the continuous case analyzed in Section 4.2.2, the incoming signal u^0_{inc} results in a forward propagating wave, given by $u^0_{\text{inc}} q_1^n$. The grid function $v_n \equiv u^0_{\text{inc}} q_1^n$ solves all equations of the homogeneous system $\boldsymbol{Av} = \boldsymbol{0}$ except for the first two, which are the one-way-to-the-left discrete Helmholtz equation (51). Therefore, by applying the matrix $\boldsymbol{A}$ of (53) to the vector $\boldsymbol{v}$ we create a right-hand side that we denote by $\boldsymbol{g}$. It is easy to see that

$$\boldsymbol{g} = \frac{u^0_{\text{inc}}}{12h_z^2} \begin{bmatrix} 1 - (q_1 + q_2)q_1 + q_1^3 q_2 \\ q_1(1 - (q_1 + q_2)q_1 + q_1^3 q_2) \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{63}$$

Proposition 6.2 guarantees that the only solution of the system of equations $\boldsymbol{Av} = \boldsymbol{g}$, where $\boldsymbol{g}$ is given by formula (63), is $\boldsymbol{v} = u^0_{\text{inc}} q_1^n$. Note that the inhomogeneity $\boldsymbol{g}$ of (63) is a discrete counterpart of the right-hand side of relation (20) (and 19) obtained when introducing the incoming signal in the continuous framework; see Sections 4.2.2 and 4.2.3.

6.6. *Obtaining the Overall Solution*

We can, finally, put together the foregoing stages of the derivation. Assume that there is a given RHS of the original equation (33) $\boldsymbol{f} = [f_0,\ f_1,\ f_2, \ldots,\ f_{N-3}, 0, 0, 0]^T$. To obtain the solution with the incoming power $u^0_{\text{inc}} q^n_1$ added, we first construct the new RHS $\tilde{\boldsymbol{f}}$ on the basis of $\boldsymbol{f}$ according to formulae (61) and (62a), (62b). Then, we construct the additional source terms $\boldsymbol{g}$ according to formula (63). Due to the linear superposition principle and according to Proposition 6.2 that guarantees solvability and uniqueness, we immediately see that the grid function $\boldsymbol{u}$ that we recover by solving the overall system

$$\boldsymbol{A}\boldsymbol{u} = \tilde{\boldsymbol{f}} + \boldsymbol{g} \tag{64}$$

is the solution we seek. Indeed, including $\tilde{\boldsymbol{f}}$ on the right-hand side of (64) guarantees the radiation from the original sources $\boldsymbol{f}$ both to the left and to the right, and including $\boldsymbol{g}$ on the right-hand side of (64) guarantees that the correct incoming signal $u^0_{\text{inc}} q^n_1$ will be added. The system (64) is, of course, solved by inverting the matrix $\boldsymbol{A}$ only once and not by solving separately with the RHSs $\tilde{\boldsymbol{f}}$ and $\boldsymbol{g}$.

Thus, setting the desired boundary conditions at $z = 0$ and $z = z_{\max}$ is reduced to building and inverting the special matrix $\boldsymbol{A}$ of (53) and also modifying the right-hand side of the equation $\boldsymbol{f} \mapsto \tilde{\boldsymbol{f}} + \boldsymbol{g}$. We again emphasize that the latter modification is not computationally expensive as both $\tilde{\boldsymbol{f}}$ and $\boldsymbol{g}$ are obtained by only local operations on the grid near $n = 0$. These operations will come at virtually no cost when implementing the algorithm numerically.

6.7. *Solution of* $\boldsymbol{A}\boldsymbol{u} = \tilde{\boldsymbol{f}} + \boldsymbol{g}$

We solve the system $\boldsymbol{A}\boldsymbol{u} = \tilde{\boldsymbol{f}} + \boldsymbol{g}$ using standard LU decomposition; for a pentadiagonal matrix $\boldsymbol{A}$ the components of this decomposition will obviously be banded as well. As the equation $\boldsymbol{A}\boldsymbol{u} = \tilde{\boldsymbol{f}} + \boldsymbol{g}$ needs to be solved many times with a changing source term but with the same $\boldsymbol{A}$, at the beginning of a simulation we calculate the LU decomposition of $\boldsymbol{A}$ once and use it throughout the iterations. Therefore, the costs per iteration in terms of solving this equation are only due to the backward substitution, which is $\mathcal{O}(N)$ arithmetic operations.

7. NUMERICAL EXPERIMENTS

To assess the numerical performance of our algorithm, we first solve a linear problem with variable coefficients in several different settings.

7.1. *Linear Problem with Variable Coefficients and Backscattering*

On a slender rectangular domain in the (r, z) coordinates $[0, r_{\max}] \times [0, z_{\max}]$, where $r_{\max} = \pi/2$ is fixed and $z_{\max}$ will vary as an essential part of testing the methodology, we recover the solution

$$E = E_{\text{right}} + C \cdot E_{\text{left}}, \tag{65}$$

where C is a constant, and the right and left propagating components E_{right} and E_{left} are given by

$$E_{\text{right}} = e^{i\sqrt{k_0^2 - \nu^2}\, z} \cos(\nu r)[1 + \epsilon z^4 e^{-z}], \tag{66a}$$

$$E_{\text{left}} = e^{-i\sqrt{k_0^2 - \nu^2}\, z} \cos(\nu r) e^{-(z/\beta)^2}. \tag{66b}$$

In the framework of our study, the left propagating component E_{left} of (66b) is interpreted as backscattering. Several parameters that control the actual shape of the solution (65) are as follows: k_0 is the wavenumber that corresponds to the homogeneous medium (see Sections 1 and 2); ν is the transversal frequency; ϵ in (66a) determines the extent of deviation from the constant-coefficient case for the right propagating mode (see below); and β in (66b) determines the spatial (longitudinal) extent to which the backscattered waves are present in the solution. In the linear case we of course introduce the backscattered waves artificially, but we are trying to follow the physically interesting situation when these waves are generated inside the domain and propagate toward and through the left boundary $z = 0$. The constant C is introduced in (65) so as to control the magnitude of the backscattered signal relative to the forward propagating signal and in particular to be able to fully eliminate backscattering ($C = 0$) if desired.

Substituting E_{right} of (66a) into Eq. (13a), we obtain

$$\begin{aligned}\Delta E_{\text{right}} + k_0^2 E_{\text{right}} &= \epsilon e^{i\sqrt{k_0^2-\nu^2}\,z}\cos(\nu r)e^{-z}z^2\left[2i\sqrt{k_0^2-\nu^2}(4z-z^2)+12-8z+z^2\right]\\ &= \epsilon\frac{e^{-z}z^2\left[2i\sqrt{k_0^2-\nu^2}(4z-z^2)+12-8z+z^2\right]}{1+\epsilon\cdot z^4e^{-z}}E_{\text{right}}\\ &= -\epsilon k_0^2 F_{\text{right}}\cdot E_{\text{right}}. \qquad (67)\end{aligned}$$

We therefore conclude that E_{right} of (66a) satisfies the variable-coefficient equation

$$\Delta E_{\text{right}} + k_{\text{right}}^2(z)E_{\text{right}} = 0,$$

where $k_{\text{right}}^2(z) = k_0^2(1+\epsilon F_{\text{right}}(z))$ and $F_{\text{right}}(z)$ is defined by equalities (67). We indeed see that ϵ controls the extent of spatial variation of the wavenumber k_{right}. The solution E_{right} is driven by the incoming wave

$$E_{\text{inc}} = e^{i\sqrt{k_0^2-\nu^2}\,z}\cos(\nu r), \quad z \le 0. \qquad (68)$$

Similarly, the backscattered solution E_{left} of (66b) satisfies the variable-coefficient equation

$$\Delta E_{\text{left}} + k_{\text{left}}^2(z)E_{\text{left}} = 0,$$

where $k_{\text{left}}^2(z) = k_0^2(1+F_{\text{left}}(z))$ and

$$F_{\text{left}}(z) = -\frac{1}{k_0^2}\left[4i\sqrt{k_0^2-\nu^2}\,\frac{z}{\beta^2} - \frac{2}{\beta^2} + \frac{4z^2}{\beta^2}\right]. \qquad (69)$$

For the overall solution E of (65) we obviously have

$$\Delta E + k^2(z)E = 0, \qquad (70)$$

where

$$k^2(z) = k_{\text{right}}^2\frac{E_{\text{right}}}{E} + k_{\text{left}}^2\frac{C\cdot E_{\text{left}}}{E}.$$

The driving incoming signal for Eq. (70) is E_{inc} of (68), evaluated at $z = 0$. The variable-coefficient linear equation (70) for E will be solved on the domain $[0, r_{\max}] \times [0, z_{\max}]$ with

the homogeneous radiation boundary condition (18a) at $z = z_{\max}$ and the non homogeneous (two-way) radiation boundary condition (20) at $z = 0$. The boundary conditions at $r = 0$ and $r = r_{\max}$ are symmetry and zero Dirichlet, respectively, which correspond to the general construction of Section 5, as well as to the particular explicit form of the solution (65), (66) that we use here. The solution will be obtained by iterations described in Section 4.2; the corresponding discrete solution methodology is delineated in Sections 5 and 6.

Our primary goal when solving numerically the foregoing linear problem is to demonstrate that the algorithm that we have constructed indeed possesses the design properties, i.e., (1) it converges with the fourth order of accuracy when the grid is refined, and (2) it properly handles the radiation of waves (including backscattering) or in other words, introduces no reflection from the boundaries $z = 0$ and $z = z_{\max}$ back into the domain. A secondary goal is deriving the guidelines for subsequent nonlinear simulations, for example, how geometric parameters such as domain size may affect the solution.

The forthcoming series of computational experiments corroborates our expectations in terms of grid convergence and handling the backscattered waves and also provides for a comparison between the following two algorithms: the one constructed in this paper with the two-way ABC at the boundary $z = 0$, and a more traditional one with the Dirichlet boundary condition at $z = 0$ (at the far-field boundary $z = z_{\max}$ we set the same radiation ABC in both cases).

7.1.1. Traditional approach—Dirichlet boundary condition. The algorithm that we have just referred to as a more traditional one is formulated with the Dirichlet boundary condition for E at $z = 0$. In fact, from the standpoint of physics one can already anticipate that this algorithm is not going to perform well when backscattering is present. Indeed, the physical setup of the model implies that all the information available at $z = 0$ *pertains only to the incoming wave*. Thus, we basically cannot say anything about the backscattered signal ahead of time because it is generated inside the domain (in the current example we of course know everything because we simply construct a sample solution including the backscattering, then produce the corresponding sources/inhomogeneities, and finally recover the same solution by the numerical method, but this is done only for demonstration purposes.) When constructing the two-way ABCs, we do not make and do not need any assumptions regarding the backscattered wave; we simply make the boundary transparent for all such waves. In contrast, in the Dirichlet case we can only specify the incoming wave as the boundary data because no explicit information about other waves is available. Mathematically, this amounts to making the assumption/approximation that

$$E(r, 0) = E^0_{\text{inc}}(r), \tag{71}$$

which, as opposed to (5), prescribes the entire field at $z = 0$, rather than its incoming component only. Consequently, the Dirichlet boundary condition will essentially reflect all backscattered waves reaching $z = 0$ back into the medium, in contrast with the two-way ABC, which will let them go through. We therefore expect that the algorithm with the Dirichlet boundary condition (71) at $z = 0$ may produce reasonable results only if no backscattered waves are present in the solution. Otherwise, the error should be roughly of the magnitude of the backscattered signal. The numerical results below corroborate these expectations.

Note that to enforce the Dirichlet boundary condition at $z = 0$ for the discretization we obviously assign a prescribed value to the solution at the leftmost grid node $n = 0$.

Besides, in the framework of the fourth-order scheme we are using, we need an additional relation to be specified right next to the boundary at $n = 1$. This is similar to obtaining the discrete one-way Helmholtz equations in the form of *two scalar relations*; see Section 6.2. The additional relation for the Dirichlet boundary conditions should be merely an approximation of the underlying differential equation at $n = 1$, but this cannot be the same approximation that we are using for the interior nodes ($n \geq 2$) because the latter employs a five-node-wide symmetric stencil. Thus, either a one-sided difference or a compact Padé-type approximation needs to be used at $n = 1$. We chose the fourth-order Padé [6] on a three-node-wide stencil in the particular form proposed in [30] because as opposed to the "long" nonsymmetric differences, it preserves the pentadiagonal structure of the matrix.

7.1.2. Results. For the simulations in the linear case we have chosen the following particular values of parameters (see formulae (65), (66)): $k_0 = 20$, $\epsilon = 0.2$, $\nu = 3$ or $\nu = 1$, $z_{\max} = 30$ or $z_{\max} = 10$, $\beta = 3$, $C = 1/2$ for the case with backscattering, and $C = 0$ for the case with no backscattering. The wavelengths in the r and z directions are $\lambda_r = 2\pi/\nu$ and $\lambda_z = 2\pi/k_0$, respectively. We choose the grid sizes h_r and h_z accordingly as fractions of the corresponding wavelengths: For the grid convergence study we refine the grid synchronously in both the r and z directions. We note that having the same resolution (nodes per wavelength) in both directions yields the cell aspect ratio of $h_r/h_z = \lambda_r/\lambda_z = k_0/\nu$, which in our simulations is equal to either $20/1$ or $20/3$.

We have looked at the values of the relative error (the difference between the computed and exact solution normalized by the maximum of the exact solution over the domain) in the maximum norm:

$$\text{Error} = \frac{\max_{(r,z)} |E_{\text{computed}} - E_{\text{exact}}|}{\max_{(r,z)} |E_{\text{exact}}|}. \tag{72}$$

The results are summarized in Tables I and II for $\nu = 1$ and $\nu = 3$, respectively. In both tables all values, except those in the last column, correspond to $z_{\max} = 30$.

From Tables I and II we first conclude that, as expected, the Dirichlet boundary condition (71) provides no convergence when the backscattering is present (third column). In all other columns we observe *a fourth-order grid convergence, because every time the grid is refined*

TABLE I
Maximum Relative Error (72) of the Calculated Solution in the Linear Case for $\nu = 1$

	Backscattering				
	Off ($C = 0$)		On ($C = 1/2$)		
	Boundary condition at $z = 0$				
	Dirichlet	Two-way	Dirichlet	Two-way	Two-way
Grid sizes	$z_{\max} = 30$				$z_{\max} = 10$
$h_r = \lambda_r/10$, $h_z = \lambda_z/10$	0.256	0.257	0.33	0.24	0.16
$h_r = \lambda_r/20$, $h_z = \lambda_z/20$	0.0165	0.0165	0.33	0.016	0.01
$h_r = \lambda_r/40$, $h_z = \lambda_z/40$	0.001	0.001	0.33	0.001	0.0012
$h_r = \lambda_r/80$, $h_z = \lambda_z/80$	$6.5 \cdot 10^{-5}$	$6.5 \cdot 10^{-5}$	0.33	$6.5 \cdot 10^{-5}$	0.00075

TABLE II
Maximum Relative Error (72) of the Calculated Solution in the Linear Case for $\nu = 3$

	Backscattering				
	Off ($C = 0$)		On ($C = 1/2$)		
	Boundary condition at $z = 0$				
	Dirichlet	Two-way	Dirichlet	Two-way	Two-way
Grid sizes	$z_{\max} = 30$				$z_{\max} = 10$
$h_r = \lambda_r/10$, $h_z = \lambda_z/10$	0.25	0.25	0.33	0.24	0.089
$h_r = \lambda_r/20$, $h_z = \lambda_z/20$	0.016	0.016	0.33	0.015	0.0064
$h_r = \lambda_r/40$, $h_z = \lambda_z/40$	0.001	0.001	0.33	0.001	0.0012
$h_r = \lambda_r/80$, $h_z = \lambda_z/80$	$6.3 \cdot 10^{-5}$	$6.3 \cdot 10^{-5}$	0.33	$6.3 \cdot 10^{-5}$	0.00075

by a factor of two in each direction, the value of the error drops by approximately a factor of sixteen (except for the last column of each table, as discussed below). Thus, the algorithm we have constructed indeed possesses the design convergence properties. Besides, we clearly see that the left propagating waves in the solution present no problem from the standpoint of numerics for the algorithm with the two-way ABC at $z = 0$.

Let us now return to the data appearing in the rightmost column of both Table I and Table II. These data clearly do not demonstrate the fourth-order grid convergence. The only difference between these data and all other data in the tables is that the rightmost columns correspond to a smaller computational domain in the z direction, $z_{\max} = 10$, as opposed to $z_{\max} = 30$. The breakdown of the grid convergence that we observe on the small domain has the following explanation.

The boundary condition that we specify at $z = z_{\max}$ is the homogeneous radiation boundary condition (18a), which is approximated by the one-way-to-the-right discrete Helmholtz Eq. (48). Both the continuous (18a) and discrete (18) radiation boundary conditions at $z = z_{\max}$ were obtained under the key assumption that the governing equation near $z = z_{\max}$ reduces to the constant-coefficient Helmholtz equation $\Delta E + k_0^2 E = 0$. In other words, this means that the mode E_{right} of (66a) has to reduce to the "pure" propagating mode $e^{i\sqrt{k_0^2 - \nu^2}z}\cos(\nu r)$, and that the mode E_{left} of (66b) has to effectively vanish at $z = z_{\max}$. From the specific form of the modes that we analyze (see 66), we conclude that the larger we make the domain $[0, z_{\max}]$, the better the quality of the agreement with the desired properties near $z = z_{\max}$. In other words, for the smaller domain $z_{\max} = 10$ we are essentially trying to apply a homogeneous radiation boundary condition to the equation, which is not "sufficiently homogeneous" itself, and therefore the error is dominated by this discrepancy, rather than by the actual truncation error associated with the discretization of the differential operator. As a consequence, we do not observe the fourth-order grid convergence for the smaller domain. This demonstrates the importance of choosing a $z_{\max}$ that is sufficiently large, so that the homogeneous radiation boundary conditions can be applied successfully.

Another interesting phenomenon we would like to discuss in the framework of the linear case is the behavior of the error as a function of the coordinate z. A typical example in Fig. 1, which corresponds to the case of no backscattering, shows a linear growth of the error with z except in the area of a small "bump" near the boundary $z = 0$. The actual

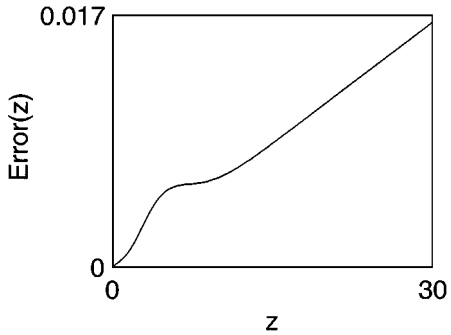

FIG. 1. Behavior of the error (73) for $\nu = 1$, two-way ABC at $z = 0$, $h_r = \lambda_r/20$, $h_z = \lambda_z/20$, $\beta = 3$, and $z_{\max} = 30$; no backscattering, $C = 0$.

quantity represented in Fig. 1 is

$$\text{Error}(z) = \frac{\max_r |E_{\text{computed}} - E_{\text{exact}}|}{\max_{(r,z)} |E_{\text{exact}}|}. \tag{73}$$

A similar error pattern is obtained for the case with backscattering, as shown in Fig. 2. The curve in Fig. 2 can be described as an oscillatory region next to the boundary $z = 0$ associated with backscattering (the magnitude of the error is still small there) followed again by a stretch of linear growth.

It is, in fact, easy to see where this linear growth comes from. Proposition 6.1 implies that the discrete right propagating mode q_1^n approximates the continuous right propagating mode $e^{ik_c z} \equiv e^{ik_c \cdot h_z n}$ (in the notations of this section, $k_c^2 = \sqrt{k_0^2 - \nu^2}$). Indeed, assuming that $k_c \cdot h_z$ is small, we have obtained $q_1 = e^{ik_c \cdot h_z} + \mathcal{O}((k_c \cdot h_z)^5)$; see formula (36a). Consequently, under the same assumption we have $q_1^n = e^{ik_c \cdot h_z n} + \mathcal{O}(n(k_c \cdot h_z)^5) = e^{ik_c z} + \mathcal{O}(z h_z^4)$ because $z = h_z n$. As $0 \le z \le z_{\max}$, we see that the error grows linearly in z and that the maximal error is $\mathcal{O}(z_{\max} \cdot h_z^4)$. The aforementioned linear growth of the error explains, in particular, why *on coarser grids* we obtain a smaller maximal error for $z_{\max} = 10$ (fifth column) than for $z_{\max} = 30$ (fourth column); see Tables I and II.

It is, in fact, instructive to see how the error curve similar to those displayed in Figs. 1 and 2 would look for a solution computed on the small domain $z_{\max} = 10$. In Fig. 3 we show such a curve for exactly the same set of parameters used for computations that led to Fig. 1,

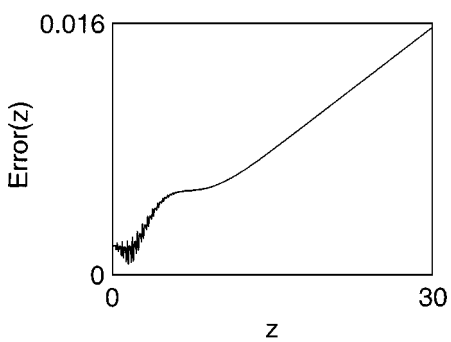

FIG. 2. Same as Fig. 1 but with backscattering, $C = 1/2$.

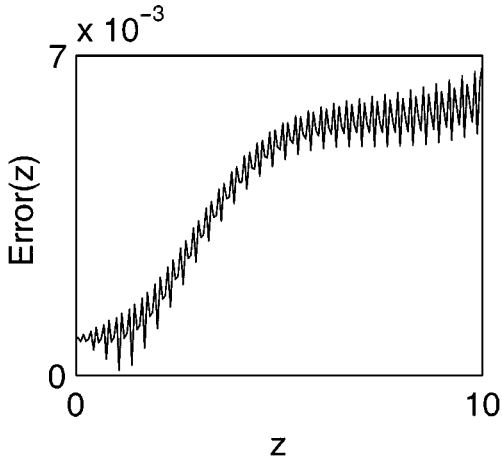

FIG. 3. Same as Fig. 1 with $z_{\max} = 10$.

except that $z_{\max}$ is equal to 10 instead of 30. Although the magnitude of the error is small, we observe oscillations throughout the entire domain. As we have no backscattering in this case ($C = 0$), the oscillations may come only from the right (far-field) boundary $z = z_{\max}$. In fact, these oscillations are an early manifestation of the phenomenon we discussed above. On small domains, the application of the homogeneous far-field radiation boundary conditions (18a) and (48) is not fully "legitimate" because the governing equation itself is not yet sufficiently close to the constant-coefficient version $\Delta E + k_0^2 E = 0$. The inconsistency gives rise to the oscillations shown in Fig. 3. For finer grids this inconsistency, as we have seen, prevents the methodology from converging on small domains with the theoretically prescribed rate of $\mathcal{O}(h^4)$.

7.2. *Nonlinear Problem*

Having corroborated the design properties of the numerical algorithm in the linear regime in Section 7.1, we now address its performance for the nonlinear case. In all cases that we analyze hereafter we take the value of $k_0 = 8$ and as before denote $\lambda_z = 2\pi/k_0$. In addition, in all simulations the solution is driven by the incoming signal

$$E^0_{\text{inc}}(r) = e^{-r^2}. \tag{74}$$

The key quantity in the NLS model, as far as nonlinear self-focusing and singularity formation are concerned, is the ratio of the power of E^0_{inc} and the critical power N_{c} (see Section 3.1). Therefore, we now briefly review the calculation of the critical power for the NLS (8).

7.2.1. Critical power. Weinstein [36] had proved that the critical power for singularity formation in the critical NLS, N_{c}, is equal to the power of the so-called waveguide solution. In the case of the $(1 + 1)$D critical NLS (8), the waveguide solutions are of the form

$$\psi(z, r) = \exp(iaz)\, Q(r; a).$$

Substitution of this solution in (8) shows that the waveguide profile Q satisfies

$$Q_{rr} - aQ + Q^5 = 0, \quad Q'(0) = 0, \quad Q(\infty) = 0.$$

Integration of this equation yields

$$Q(r; a) = (3a)^{1/4}\mathrm{sech}^{1/2}(2\sqrt{a}r).$$

Therefore, a necessary condition for singularity formation in (8) is that

$$\int_0^\infty |\psi_0(\tilde{r})|_2^2 \, d\tilde{r} \geq N_c,$$

where

$$N_c = \int_0^\infty Q^2(r)\, dr = \frac{\sqrt{3}\pi}{4}.$$

In dimensional variables, this condition is

$$\int_0^\infty \left|E_{\mathrm{inc}}^0(r)\right|_2^2 dr \geq \frac{N_c}{k_0\sqrt{\epsilon}}.$$

Therefore, the fractional critical power of E_{inc}^0 of (74) is

$$p = \frac{\int_0^\infty \left|E_{\mathrm{inc}}^0\right|^2 dr}{N_c/k_0\sqrt{\epsilon}} = \sqrt{\frac{2}{3\pi}}k_0\sqrt{\epsilon}. \tag{75}$$

7.2.2. Results. We start with a moderate nonlinearity in Eq. (7), $\epsilon = 0.04$, which according to (75), corresponds to 74% of the critical power when $k_0 = 8$. Our goal is to first demonstrate the grid convergence of the algorithm. We also compare the two-way ABC against the standard Dirichlet boundary condition at $z = 0$, as we did in the linear case, both from the standpoint of accuracy of the solution and the rate of convergence of our iterative scheme.

For the grid convergence study we first choose the following parameters: $z_{\max} = 20$, $r_{\max}/z_{\max} = 1, h_z = \lambda_z/10, h_r = \lambda_z/2$. In our computations we have observed that changing the discretization parameters in the r direction may exert a more noticeable influence on the solution than changing the discretization in the z direction. Therefore, we initially refine the grid in the r direction only and in Fig. 4 present three solution curves: absolute value $|E_{\mathrm{computed}}(0, z)|$ on the axis of symmetry $r = 0$ as a function of z for $h_r = \lambda_z/2$, $\lambda_z/4$, and $\lambda_z/8$. We see that the last two curves that correspond to $h_r = \lambda_z/4$ and $\lambda_z/8$ are virtually indistinguishable from one another and both differ noticeably from the first one obtained on a coarser grid $h_r = \lambda_z/2$. We therefore conclude that *as the grid is refined the numerical solution does converge*; even so in this nonlinear case we do not know what the exact solution is and consequently cannot explicitly find the error.

We note that we plot the values of the computed solution on the axis of symmetry $r = 0$ because this is the most interesting location in the domain where the genuinely nonlinear phenomena take place. A clear manifestation of these nonlinear phenomena is the "bump," or peak, on the solution curve in Fig. 4, whose value is *higher* than that of the incoming

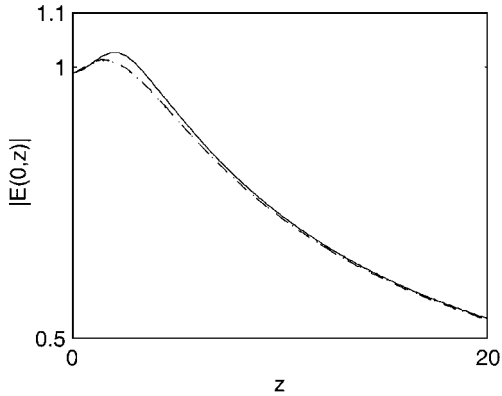

FIG. 4. Grid convergence for $\epsilon = 0.04$, $z_{\max} = 20$, $r_{\max}/z_{\max} = 1$, $h_z = \lambda_z/10$, for $h_r = \lambda_z/2$ (solid line), $h_r = \lambda_z/4$ (dotted line), and $h_r = \lambda_z/8$ (dashed line).

wave $E^0_{\text{inc}}(0) = 1$. Clearly, in the absence of nonlinear effects (i.e., $\epsilon = 0$), an unfocused input beam, such as (74), would simply diffract while propagating to the right, i.e., toward large z's, with its maximum amplitude monotonically decreasing. The amplification of the incoming signal due to the nonlinear response of the medium is called *self-focusing*, and is wellknown within the NLS framework.

Another interesting phenomenon, which is actually the one that our methodology has been specifically designed to capture, is *backscattering*. In the previous linear studies in Section 7.1, the extent of backscattering was predetermined by the value of C. To estimate the extent of backscattering in the current nonlinear case, we plot the quantity $|E_{\text{computed}}(r, 0) - E^0_{\text{inc}}(r)|$ as a function of r. In Fig. 5 we show the corresponding graph for $\epsilon = 0.04$, $z_{\max} = 20$, $r_{\max}/z_{\max} = 1$, $h_z = \lambda_z/10$, and $h_r = \lambda_z/4$. From Fig. 5 we conclude that most backscattering occurs around the axis of symmetry $r = 0$, and that the magnitude of backscattering there is about 1.2% of the incoming power. Backscattering obviously accounts for the deviation of the solution curve at $z = 0$ in Fig. 4 from the incoming signal value there, which is equal to 1.

A comprehensive grid refinement study should, of course, include refinement in the z direction along with the refinement in the r direction. In addition to the cases reported

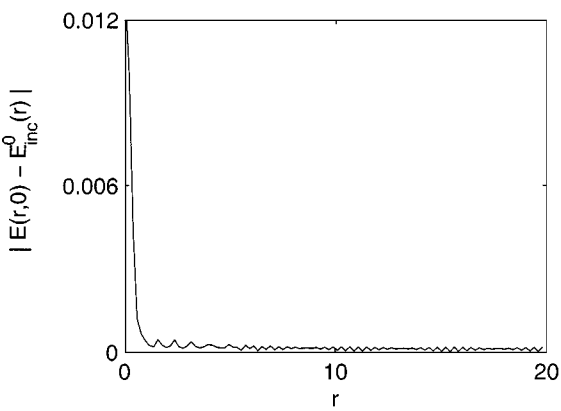

FIG. 5. Backscattering for $\epsilon = 0.04$, $z_{\max} = 20$, and $r_{\max}/z_{\max} = 1$. Gridsizes: $h_z = \lambda_z/10$ and $h_r = \lambda_z/4$.

TABLE III
Grid Refinement and Domain Enlargement Study for $\epsilon = 0.04$

$z_{\max}$	$r_{\max}/z_{\max}$	h_z	h_r	Maximum self-focusing	Maximum backscattering
20	1	$\lambda_z/10$	$\lambda_z/4$	1.0136	0.013
20	1	$\lambda_z/10$	$\lambda_z/8$	1.0129	0.0128
20	2	$\lambda_z/10$	$\lambda_z/4$	1.0135	0.0128
40	1	$\lambda_z/10$	$\lambda_z/4$	1.0132	0.0127
20	1	$\lambda_z/20$	$\lambda_z/4$	1.0124	0.0112
20	1	$\lambda_z/20$	$\lambda_z/8$	1.0119	0.0111

previously, we have run several others, refining the grid either separately in each direction or synchronously in both directions, and also changing the size of the computational domain. Note that determining the correct, i.e., sufficiently large, size of the computational domain is important, because choosing one too small in the z direction may cause reflections from the boundary $z = z_{\max}$ (Section 7.1), and choosing a domain that is too small in the r direction is dangerous because the boundary $r = r_{\max}$ is reflecting and the reflections may, in fact, completely destroy the solution (we have actually observed the latter phenomenon in our computations).

Basically, the solutions that we have obtained on all grids finer than $h_r = \lambda_z/2$, $h_z = \lambda_z/10$ (i.e., finer than the coarsest of the previous grids), and all domains larger than or equal to $z_{\max} = r_{\max} = 20$, are almost identical. We do not plot these solutions as they are very close to one another but rather summarize the results of our computations in Table III, in which the two key quantities for each case are presented: the maximum value of self-focusing, defined as $\max_z |E(0, z)|$ (i.e., the peak on the curve similar to those shown in Fig. 4), and the maximum backscattering at $z = 0$, defined as $\max_r |E(r, 0) - E^0_{\text{inc}}(r)|$ (i.e., the peak on the curve similar to those shown in Figs. 5 and 6).

From Table III we see that all values of maximum self-focusing we have computed on different grids and different domains differ from one another by at most 0.17%. This indicates that for the ranges of parameters (grid sizes and domain sizes) we have used, the numerical solution is already "well converged." The level of backscattering in all our simulations is between 1.1 and 1.3% of the incoming power, which again constitutes an error of only about 0.2% (relative to the maximum of the solution). One should probably regard the computational variant presented in the last row of Table III as the most accurate one because it was computed on the finest grid. The corresponding backscattering profile (for $h_z = \lambda_z/20$, $h_r = \lambda_z/8$) is shown in Fig. 6. We again see that this profile is practically the same as the one from Fig. 5, which corresponds to the grid twice as coarse in each direction.

We now look at the convergence histories for our numerical solutions. Let us recall that the iteration scheme we employ is nested. On the inner loop we solve a variable-coefficient linear equation, whereas on the outer loop we iterate with respect to the nonlinearity. Currently, we update the coefficient $k^2 = k_0^2 \left(1 + \epsilon |E|^4\right)$, i.e., make one nonlinear iteration, every ten linear iterations [i.e., in the notations of Section 4, $\mathcal{M}(n) = 10$ in (12)]. In Figs. 7 and 8 we show the convergence histories for the two cases we have discussed already—those that correspond to the first and last rows of Table III (Figs. 7 and 8, respectively).

The actual quantity shown in Figs. 7 and 8 is the maximum absolute difference between the two consecutive iterations. The sawtooth character of both curves is accounted for by

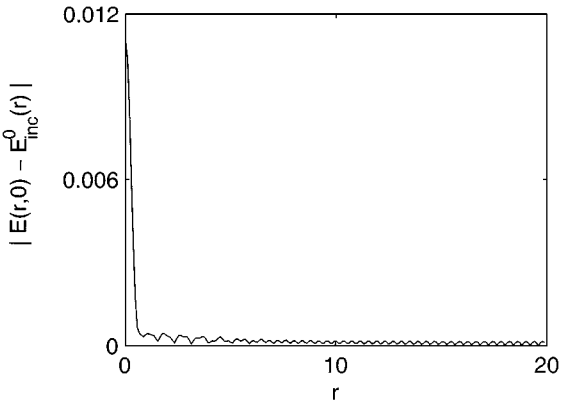

FIG. 6. Same as Fig. 5, but with the gridsizes $h_z = \lambda_z/20$ and $h_r = \lambda_z/8$.

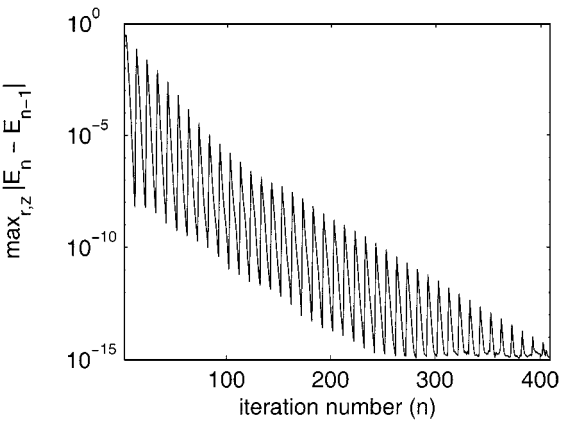

FIG. 7. Convergence of iterations for $\epsilon = 0.04$, $z_{\max} = 20$, $r_{\max}/z_{\max} = 1$. Gridsizes: $h_z = \lambda_z/10$ and $h_r = \lambda_z/4$.

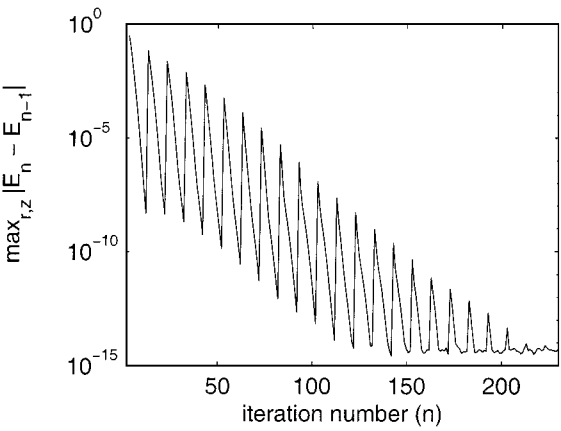

FIG. 8. Same as Fig. 7 but with the gridsizes $h_z = \lambda_z/20$ and $h_r = \lambda_z/8$.

the nested structure of the iterative procedure. The fast-scale decay followed by a jump back up is the convergence of linear iterations on the inner loop with subsequent update of k^2. The slow-scale decay all the way up to machine zero corresponds to the convergence of nonlinear iterations on the outer loop.

Figures 7 and 8 demonstrate the convergence of iterations. Besides, we notice that on a finer grid (see Fig. 8), this convergence is faster (about twice as fast) than on the coarser one (see Fig. 7). In fact, we have observed in various simulations that the geometry in the r direction influences the rate of convergence most noticeably. The larger the domain size $r_{\max}$ and/or the finer the grid size h_r, the faster the iterations converge. As of yet, we do not have a rigorous explanation of this computational phenomenon. We can only assume that both refining the grid in the r direction and putting the boundary $r = r_{\max}$ further away somehow reduce the adverse influence of this reflecting boundary on the solution.

As stated at the beginning of this section, a major goal of the nonlinear simulations is to compare the performance of the new two-way ABC against that of the traditional Dirichlet boundary condition at $z = 0$ (71). In Fig. 9 we compare the actual computed solutions with the two boundary conditions for the case that we have analyzed before: $\epsilon = 0.04$, $z_{\max} = 20$, $r_{\max}/z_{\max} = 1$, $h_z = \lambda_z/10$, $h_r = \lambda_z/4$. We see a noticeable discrepancy between the two curves. The dotted line that corresponds to the Dirichlet boundary conditions is above the solid one, which corresponds to the two-way ABC. The extent of the aforementioned discrepancy is roughly equal to the level of backscattering that we have recovered previously, which is clearly a natural result.

We also compare the rates of convergence of the iterative algorithm for the two types of boundary conditions that we set at $z = 0$. The convergence history for the two-way ABC is shown in Fig. 7; the convergence history for the Dirichlet boundary conditions is shown in Fig. 10. We see that the convergence with the two-way ABCs is about 1.5 times faster than that with the Dirichlet boundary conditions, which presents another advantage of using the new methodology. Let us mention that the phenomenon of convergence speedup for iterative solvers caused by the application of highly accurate nonlocal ABCs (similar to those developed in this paper) has been noticed previously by several authors, although in completely different settings primarily associated with fluid flow computations; see [34].

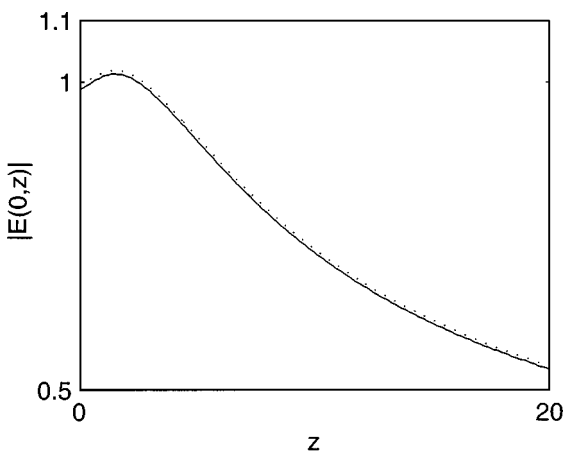

FIG. 9. $|E_{\text{computed}}(0, z)|$ for the two-way ABC (solid) and for the Dirichlet boundary conditions (dots).

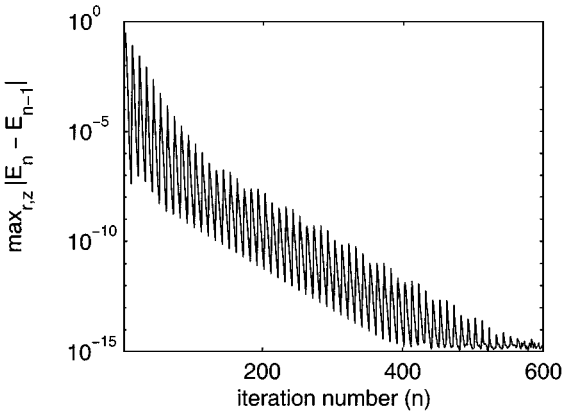

FIG. 10. Same as Fig. 7, with the Dirichlet boundary condition at $z = 0$.

We now consider the case $\epsilon = 0.06$, for which the input beam power is 90% of the critical power. Basically, the results have the same qualitative features as for the case $\epsilon = 0.04$. In particular, the convergence of iterations is faster for finer grids and larger computational domains, as well as for the two-way ABC compared with the traditional Dirichlet boundary condition at $z = 0$. Moreover, we note that for $\epsilon = 0.06$ some cases with the Dirichlet boundary condition did not converge at all.

In Fig. 11a, we plot the on-axis amplitude raised to the power 4 for the domain of the same size as corresponds to Fig. 4 (but with a finer grid). We plot this particular quantity

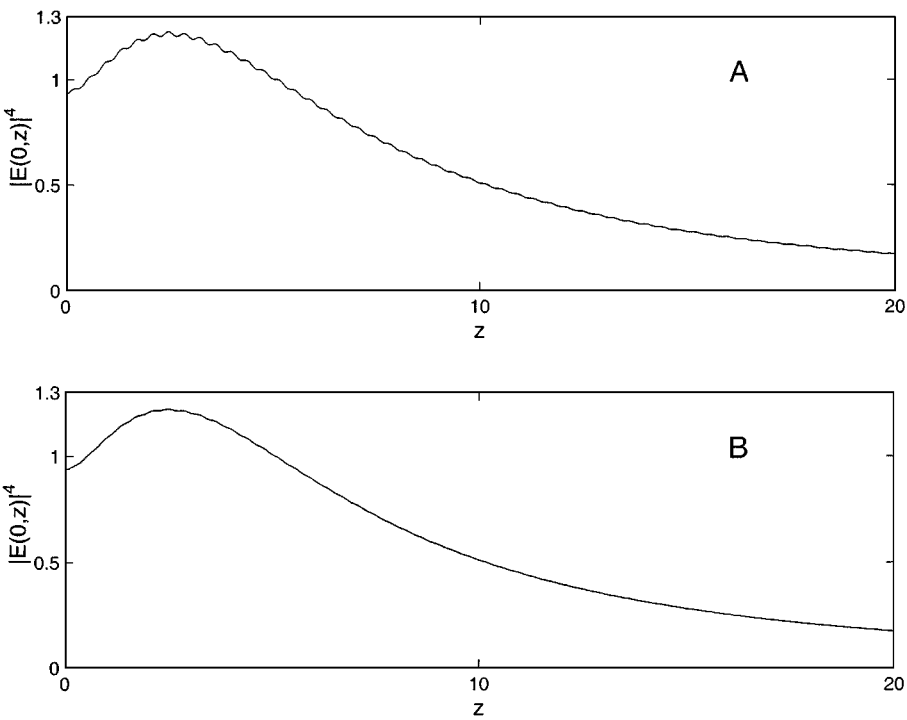

FIG. 11. $|E_{\text{computed}}(0, z)|^4$ for $\epsilon = 0.06$, $h_z = \lambda_z/20$, $h_r = \lambda_z/8$, $r_{\max}/z_{\max} = 1$. (a) $z_{\max} = 20$; (b) $z_{\max} = 40$.

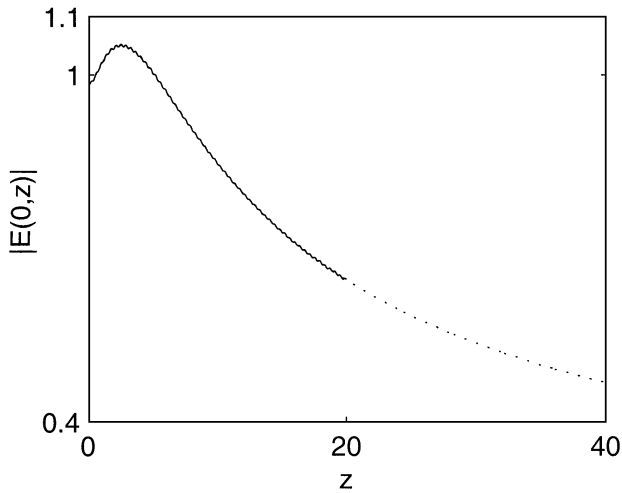

FIG. 12. $|E_{\text{computed}}(0, z)|$ for $\epsilon = 0.06$, $h_z = \lambda_z/20$, $h_r = \lambda_z/8$, $r_{\max}/z_{\max} = 1$. Solid line—$z_{\max} = 20$; dotted line—$z_{\max} = 40$.

because it is the one that controls the relative magnitude of nonlinearity, which is crucial for our study, and it also allows us to see most clearly that the solution for $z_{\max} = 20$ has small oscillations throughout the domain, which are reminiscent of those seen in Fig. 3. In order to verify that these oscillations are indeed due to the right boundary $z = z_{\max}$ being placed too close, we reran the same simulation but with the right boundary located at twice the previous distance, i.e., $z_{\max} = 40$. The corresponding profile of $|E_{\text{computed}}(0, z)|^4$ is shown in Fig. 11b, but only for the half of the new range from $0 \le z \le 20$, to make the scale the same as that in Fig. 11a. From Fig. 11b we see that in the case $z_{\max} = 40$ the little wiggles have almost disappeared, suggesting that this is indeed a numerical artifact, rather than a true physical phenomenon. Apart from the little wiggles, the two solutions seem to be identical, as Fig. 12 indicates.

The explanation for the appearance of the small wiggles throughout the domain when the right boundary is too close is the same as in the linear case, namely, that in order for the ABC at $z_{\max}$ to perform well, $\epsilon|E|^4$ should be sufficiently small there so that $k^2 \approx k_0^2$ with sufficient accuracy. Therefore, at higher ϵ, one needs more decay in $|E|^4$ for this approximation to hold. On top of that, at higher powers self-focusing is stronger, implying that $|E|^4$ would decay slower in z. This, in turn, means that we may need to use larger and larger domains at higher powers, otherwise, the quality of the computed solution will deteriorate. Besides, the convergence rate of our iterations may also be affected by the location of the boundary $z = z_{\max}$. For higher powers on those domains we have considered it becomes prohibitively slow (if there is convergence at all). This is why, at present, we could not go above $\epsilon = 0.06$. We should note, however, that besides enlarging the domain, changing the iterative algorithm itself to a more efficient one may alleviate the aforementioned problem. This issue will be studied in the future.

The results of the grid convergence study for $\epsilon = 0.06$ are summarized in Table IV. Comparison of Table III with Table IV shows that as the input power increases (relative to the critical power), more energy gets backscattered and the self-focusing peak becomes higher, which is expected from physical considerations.

TABLE IV
Grid Refinement and Domain Enlargement Study for $\epsilon = 0.06$

z_{max}	r_{max}/z_{max}	h_z	h_r	Maximum self-focusing	Maximum backscattering
20	1	$\lambda_z/10$	$\lambda_z/4$	1.0567	0.0188
20	1	$\lambda_z/20$	$\lambda_z/8$	1.0528	0.0188
20	1	$\lambda_z/20$	$\lambda_z/16$	1.0526	0.0188
20	2	$\lambda_z/20$	$\lambda_z/8$	1.0527	0.0188
20	1	$\lambda_z/40$	$\lambda_z/8$	1.0518	0.0179
40	1	$\lambda_z/20$	$\lambda_z/8$	1.0512	0.0173

8. DISCUSSION

In this section we briefly describe the approaches that have been used previously in the literature for solving similar problems. We then discuss the motivation behind making some particular choices when constructing our algorithm, present the conclusions, and outline directions for future research.

8.1. *Previous Approaches for Solving the NLH*

Feit and Fleck [7] solved the NLH by splitting the wave into its forward and backward components and solving only for the forward propagating part. Under this approach it was assumed that the "transverse variation in [k] is sufficiently small." As for backscattering, their algorithm "removes power that cannot propagate in the forward direction without accounting explicitly for where it goes" [7]. Akhmediev and collaborators [1, 2] solved an initial-value problem which can be viewed as a "modified" NLH. However, they neglected the ψ_{zz} term, as well as backscattering.

In contrast to the aforementioned approaches, in this paper we solve the Helmholtz equation as a true "unabridged" boundary-value problem. By doing that, we can account correctly for the backscattering, without introducing any adhoc assumptions, the validity of which is unclear.

8.2. *Discontinuity at the Interface $z = 0$*

In the current study we consider the simplest possible model for the interface $z = 0$, where we assume that this interface is nonreflecting, i.e., the wavenumber k is continuous across $z = 0$ (Section 2.1). From the standpoint of physics this is, of course, not necessarily true. For example, an incoming laser beam traveling through air which impinges on a water interface would be partially reflected, due to the difference in the (linear) index of refraction between air and water. The easiest way to incorporate the discontinuity in k at $z = 0$ into the model would be to do that for the linear constant-coefficient Eq. (12) in the framework of the iteration scheme, as we do all other boundary conditions. After the transverse Fourier transform, we obtain a collection of one-dimensional Helmholtz equations. For each of the latter, the application of the standard elliptic interface conditions, which for the second-order equations are the continuity of the solution and its flux across the interface, yields the standard expressions for the reflection and transmission coefficients, once the incoming

wave is given. If we want to use the transmitted wave (i.e., already past the interface) as the primary data for the problem, the same expressions will yield the amount of reflection and the original incoming signal. Moreover, they will also apply to treating the possible reflection of the backscattered waves by the interface $z = 0$.

8.3. *Nonlinear Iterations*

The primary motivation behind our choice of the nonlinear iteration scheme (see Section 4) was its simplicity. We note that Eqs. (9), (10) have been obtained by simply freezing the nonlinear term rather than differentiating it in the sense of Frechet. For complex-valued solutions E (which is the case in our study), the nonlinearity in Eq. (7) is obviously nondifferentiable and consequently, the direct implementation of Newton's method is not possible. However, as mentioned by Bayliss [3], Newton-type iterations may still apply to Eq. (7) if it is solved separately for the real and imaginary components of E. We did not try to implement this idea in the current study. We acknowledge, however, that among the different parts of our algorithm the nonlinear iteration scheme is apparently the primary candidate for improvements in order to achieve convergence with higher input power, i.e., for larger ϵ.

8.4. *Linear Solver*

The solver that we employ for the variable-coefficient linear Helmholtz equation is also iterative and fits as the inner loop of the overall nonlinear solver. This choice is, of course, by no means unique. In general, one can solve the linear Helmholtz equation with variable coefficients using a variety of other methods, such as the Ricatti method [18]. A recent review by Turkel of different approaches for solving the linear Helmholtz equation can be found in [35]. We note, however, that combining a Helmholtz solver with global ABCs, and in particular, a two-way ABC of the type constructed in this paper, presents a rather difficult task, since the speed of propagation of plane waves in the z direction depends on their transverse wavenumber. Indeed, most of the solvers available in the literature deal with simpler boundary conditions, such as those of the Dirichlet type. The solver that we have constructed involves a direct inversion of the constant-coefficient operator on every iteration using the separation of variables. This approach, as mentioned, is most natural for incorporating global ABCs into the model.

8.5. *Fourth-Order Scheme*

In this study we chose a fourth-order method, rather than a conventional second-order one, for our simulations. The motivation behind this choice is, in fact, standard, and relies primarily on the possibility of having less points per wavelength and accordingly reducing the required overall grid dimension for a given level of accuracy. Besides, our numerical simulations corroborate that the extent of backscattering in the model we study is indeed small. In such cases, i.e., when the interesting phenomenon is small in magnitude compared to the background, it is generally acknowledged that higher order methods perform better than lower order ones.

We note in this connection that the construction of one-way discrete Helmholtz equations and radiation ABCs for a second-order scheme would be conceptually the same as the construction described in Section 6 but substantially less cumbersome in both derivation

and implementation, as it would not require taking care of an extra pair of evanescent waves. However, having a higher order method justifies, in our opinion, the additional work invested in obtaining the more sophisticated ABCs.

8.6. *Discrete Implementation*

The implementation in this study of both the conventional radiation boundary condition at $z = z_{\max}$ and the two-way ABC at $z = 0$, is done after the continuous problem has been replaced with a discretized finite-difference formulation. The "fully discrete" approach that we have adopted is quite different from the more common technique of deriving continuous ABCs and subsequently discretizing them; see [34]. The advantage of working completely at the discrete level is that discretizing (20) may be nontrivial, especially in the framework of a higher order method. Another advantage of building the ABCs at the discrete level is the "automatic" well-posedness of the resulting formulation. As mentioned, the corresponding solvability and well-posedness analysis for general one-dimensional systems of finite-difference equations can be found in [14, 23]. To the best of our knowledge, the proposed full-fledged discrete construction of the two-way ABCs for a higher order finite-difference scheme is unique.

8.7. *Time-Dependence*

In the present study we have focused on the nonlinear Helmholtz equation (NLH), which models the propagation of stationary (cw) laser beams in a Kerr medium. It is important to note that from the standpoint of physics, the natural formulation of the Kerr effect is a cubic nonlinearity in the time-harmonic framework described by the NLH (4) and, accordingly, a nonlinear convolution for time-dependent problems (of propagation of laser *pulses*) (see [4]) rather than the other way around. As mentioned, the NLS in nonlinear optics is derived from the NLH under the paraxial approximation; see Section 3. Therefore, our results can be used to assess, for example, the effect of nonparaxiality and that of backscattering on soliton propagation in the one-dimensional cubic NLS. We reemphasize that the "dynamic" NLS (6a) in nonlinear optics describes stationary propagation, since the axial coordinate variable z plays the role of "time" (recall that the initial condition (6b) is prescribed at $z = 0$ for all $0 \le r < \infty$).

In nonstationary models that are used in nonlinear optics for propagation of laser pulses, one can formally consider time as an additional spatial variable, because the field intensity ("initial condition") is prescribed at $z = 0$ for all $0 \le r < \infty$ and all $-\infty < t < \infty$. In that case, the extension of our methodology is straightforward. Solving "genuine" time-dependent problems in other nonlinear wave models can, of course, present a significant independent interest. Without delving into details, we mention that construction of accurate ABCs for time-dependent problems is often more demanding (both theoretically and computationally) than for similar steady-state or time-harmonic problems. An approach to constructing accurate global fully discrete ABCs for time-dependent wave propagation problems is presented in [27, 28].

8.8. *Conclusions*

In the current paper we have developed and implemented a fourth-order finite-difference method for solving the nonlinear scalar Helmholtz equation that accounts for the phenomena

of self-focusing and backscattering. The method is supplemented by the highly accurate global ABCs that make the external artificial boundaries fully transparent for all outgoing waves (including the backscattered waves) and at the same time are capable of correctly prescribing the incoming signal at the outer boundary of the computational domain. To the best of our knowledge this is the first attempt ever to construct global ABCs that possess the foregoing two-way capability.

The fourth-order grid convergence of the method has been directly verified by solving model linear problems. In the presence of backscattering, the new method clearly outperforms a traditional technique based on the Dirichlet boundary condition. We have also conducted a comprehensive experimental study of the nonlinear case in the regime where the input power is below the critical one for blowup. As with the linear case, this study corroborates the convergence of the method and its superiority over the traditional approach.

The new method allows for a systematic quantitative study of backscattering in nonlinear self-focusing. To the best of our knowledge, this is the first study that allows, for example, a calculation of the actual extent of backscattering, its dependence on the input power, etc. As mentioned, the new extended capabilities are accounted for by the fact that, unlike previous studies, we solve the NLH as a true nonlinear boundary-value problem, without introducing any simplifying assumptions on the continuous level prior to the discretization. Therefore, the only error that we are actually left with is the truncation error associated with the discrete approximation of derivatives.

8.9. *Future Work*

In this paper we have developed a new numerical methodology for solving the true boundary-value problem for the NLH. We believe that our approach can be extended to address various other issues that are not covered by the present study. For example, it is interesting to conduct a systematic comparison of NLH simulations with the corresponding NLS simulations. Such a comparison would enhance our understanding of the role of nonparaxiality and backscattering. It is also interesting to compare our NLH simulations with earlier approaches for solving the NLH, which did not treat the NLH as a true boundary-value problem. In addition, future studies should attempt to go above the critical power for blowup. If successful, this would provide strong support for the current belief that there is no blowup in the presence of nonparaxiality.

In this study we have primarily focused on the NLH which corresponds to the critical NLS. However, our numerical approach can be applied for both subcritical NLS (e.g., calculating the amount of backscattering for solitons), as well as the supercritical case.

We finally note that the nonlocal homogeneous radiation ABC at $z = z_{\max}$, as well as the nonlocal nonhomogeneous two-way ABC at $z = 0$, can be cast into the general framework of pseudo-differential boundary equations and projection operators of Calderon's type (the Calderon equation in the case of the two-way ABC will be nonhomogeneous as well) and the difference potentials method by Ryaben'kii; see [5, 20, 24–26, 29]. This, in particular, may allow considering curvilinear outer boundaries if necessary, as opposed to only linear boundaries considered in the current study. Besides, such a reformulation will be generally useful from the standpoint of understanding the fundamental connections between global ABCs of various types that appear in the scientific computing literature.

ACKNOWLEDGMENTS

We acknowledge useful discussions with A. Bayliss, G. Eskin, J. B. Keller, G. Papanicolaou, P. Petropoulos, E. Turkel, and X. P. Wang. We also thank the reviewers of the paper for their helpful comments.

REFERENCES

1. N. Akhmediev, A. Ankiewicz, and J. Soto-Crespo, Does the nonlinear Schrödinger equation correctly describe beam propagation?, *Opt. Lett.* **18**, 411 (1993).
2. N. Akhmediev and J. Soto-Crespo, Generation of a train of three-dimensional optical solitons in a self-focusing medium, *Phys. Rev. A* **47**, 1358 (1993).
3. A. Bayliss, personal communication, 2000.
4. R. Boyd, *Nonlinear Optics* (Academic Press, Boston, 1992).
5. A. P. Calderon, Boundary-value problems for elliptic equations, in *Proceedings of the Soviet–American Conference on Partial Differential Equations at Novosibirsk* (Fizmatgiz, Moscow, 1963), pp. 303–304.
6. L. Collatz, *The Numerical Treatment of Differential Equations*, 3rd ed. (Springer-Verlag, Berlin, 1960).
7. M. Feit and J. Fleck, Beam nonparaxiality, filament formation, and beam breakup in the self-focusing of optical beams, *J. Opt. Soc. Am. B* **5**, 633 (1988).
8. G. Fibich, *Self-Focusing in the Nonlinear Schrödinger Equation for Ultrashort Laser–Tissue Interactions*, Ph.D. thesis (Courant Institute, New York University, 1994).
9. G. Fibich, Small beam nonparaxiality arrests self-focusing of optical beams, *Phys. Rev. Lett.* **76**, 4356 (1996).
10. G. Fibich and G. Papanicolaou, Self-focusing in the perturbed and unperturbed nonlinear Schrödinger equation in critical dimension, *SIAM J. Appl. Math.* **60**, 183 (1999).
11. G. Fibich and G. Papanicolaou, A modulation method for self-focusing in the perturbed critical nonlinear Schrödinger equation, *Phys. Lett. A* **239**, 167 (1998).
12. D. Givoli and I. Patlashenko, Finite element schemes for nonlinear problems in infinite domains, *Int. J. Numer. Meth. Eng.* **42**, 341 (1998).
13. D. Givoli and I. Patlashenko, Finite-element solution of nonlinear time-dependent exterior wave problems, *J. Comput. Phys.* **143**, 241 (1998).
14. S. K. Godunov and V. S. Ryaben'kii, Canonical forms of systems of ordinary linear difference equations with constant coefficients, *U.S.S.R. Comput. Math. and Math. Phys.* **3**, 281 (1963).
15. S. K. Godunov and V. S. Ryaben'kii, *Difference Schemes* (North-Holland, Amsterdam, 1987).
16. P. Kelley, Self-focusing of optical beams, *Phys. Rev. Lett.* **15**, 1005 (1965).
17. J. B. Keller, personal communication, 2000.
18. Y. Y. Lu and J. R. McLaughlin, The Riccati method for the Helmholtz equation, *J. Acoust. Soc. Am.* **100**, 1432 (1996).
19. V. Malkin, On the analytical theory for stationary self-focusing of radiation, *Physica D* **64**, 251 (1993).
20. S. G. Mikhlin, N. F. Morozov, and M. V. Paukshto, *The Integral Equations of the Theory of Elasticity* (Teubner, Stuttgart, 1995).
21. A. Newell and J. Moloney, *Nonlinear Optics* (Addison-Wesley, Redwood City, CA, 1992).
22. I. Patlashenko and D. Givoli, A numerical method for problems in infinite strips with irregularities extending to infinity, *Numer. Meth. Partial Differential Eq.* **14**, 233 (1998).
23. V. S. Ryaben'kii, Necessary and sufficient conditions for good definition of boundary value problems for systems of ordinary difference equations, *U.S.S.R. Comput. Math. and Math. Phys.* **4**, 43 (1964).
24. V. S. Ryaben'kii, Boundary equations with projections, *Russ. Math. Surv.* **40**, 147 (1985).
25. V. S. Ryaben'kii, *Difference Potentials Method for Some Problems of Continuous Media Mechanics* (Nauka, Moscow, 1987). In Russian.
26. V. S. Ryaben'kii, Difference potentials method and its applications, *Math. Nachr.* **177**, 251 (1996).
27. V. S. Ryaben'kii, S. V. Tsynkov, and V. I. Turchaninov, Long-time numerical computation of wave-type solutions driven by moving sources, *Appl. Numer. Math.*, to appear.

28. V. S. Ryaben'kii, S. V. Tsynkov, and V. I. Turchaninov, *Global Discrete Artificial Boundary Conditions for Time-Dependent Wave Propagation*, ICASE Report, NASA Langley Res. Center, Hampton, VA, *J. Comput. Phys.*, submitted to appear.
29. R. T. Seeley, Singular integrals and boundary value problems, *Am. J. Math.* **88**, 781 (1966).
30. I. Singer and E. Turkel, High-order finite-difference methods for the Helmholtz equation, *Comput. Meth. Appl. Mech. Eng.* **163**, 343 (1998).
31. W. Strauss, *Nonlinear Wave Equations* (Am. Math. Soc. Providence, 1989).
32. C. Sulem and P. Sulem, *The Nonlinear Schrödinger Equation* (Springer-Verlag, New York, 1999).
33. L. N. Trefethen and L. Halpern, Well-posedness of one-way wave equations and absorbing boundary conditions, *Math. Comput.* **47**, 421 (1986).
34. S. V. Tsynkov, Numerical solution of problems on unbounded domains. A review, *Appl. Numer. Math.* **27**, 465 (1998).
35. E. Turkel, Numerical difficulties solving time harmonic equations, in *Multiscale Computational Methods in Chemistry*, edited by A. Brandt, J. Bernholc, and K. Binder, NATO Sci. series, Comput. Syst. Sci. subseries IOS Press, Amsterdam, 2000.
36. M. Weinstein, Nonlinear Schrödinger equations and sharp interpolation estimates, *Commun. Math. Phys.* **87**, 567 (1983).