Lattice models for large-scale simulations of coherent wave scattering

Shumin Wang and Fernando L. Teixeira

Magnetic Resonance Center, General Electric Company, 3200 North Grandview Boulevard, W-832, Waukesha, Wisconsin 53188, USA

ElectroScience Laboratory and Department of Electrical Engineering, The Ohio State University, 1320 Kinnear Road, Columbus, Ohio 43212, USA

(Received 6 June 2003; published 27 January 2004)

Lattice approximations for partial differential equations describing physical phenomena are commonly used for the numerical simulation of many problems otherwise intractable by pure analytical approaches. The discretization inevitably leads to many of the original symmetries to be broken or modified. In the case of Maxwell’s equations for example, invariance and isotropy of the speed of light in vacuum is invariably lost because of the so-called grid dispersion. Since it is a cumulative effect, grid dispersion is particularly harmful for the accuracy of results of large-scale simulations of scattering problems. Grid dispersion is usually combated by either increasing the lattice resolution or by employing higher-order schemes with larger stencils for the space and time derivatives. Both alternatives lead to increased computational cost to simulate a problem of a given physical size. Here, we introduce a general approach to develop lattice approximations with reduced grid dispersion error for a given stencil (and hence at no additional computational cost). The present approach is based on first obtaining stencil coefficients in the Fourier domain that minimize the maximum grid dispersion error for wave propagation at all directions (minimax sense). The resulting coefficients are then expanded into a Taylor series in terms of the frequency variable and incorporated into time-domain (update) equations after an inverse Fourier transformation. Maximally flat (Butterworth) or Chebyshev filters are subsequently used to minimize the wave speed variations for a given frequency range of interest. The use of such filters also allows for the adjustment of the grid dispersion characteristics so as to minimize not only the local dispersion error but also the accumulated phase error in a frequency range of interest.

DOI: 10.1103/PhysRevE.69.016701

PACS number(s): 02.60.Cb, 03.50.De, 41.20.–q

I. INTRODUCTION

Grid (numerical) dispersion is a major source of error in the simulation of wave phenomena using discretized equations on a lattice [1–5]. Grid dispersion manifests itself as a change on the phase velocity of the wave according to frequency and propagation angle and, because it is a cumulative effect, it poses serious limitations particularly for large-scale (time-domain) simulations of coherent wave scattering.

Grid dispersion is usually combated by either increasing the resolution of the lattice (i.e., approaching the “continuum limit”) or by employing higher-order schemes which utilize larger stencils to approximate space and time derivatives at the cost of loss of sparsity (“locality”) of the discrete model (spectral methods [4] can be thought of as the extreme examples in this direction). Both alternatives lead to an increased computational cost to simulate a problem of given physical size. For electrodynamics, several techniques have been developed over the years to reduce grid dispersion error [4–13]. In particular, space and time fourth-order accurate (FDTD) schemes have been shown to provide an attractive trade-off between increased computational cost and reduced dispersion error [14]. Grid dispersion on finite element solutions have been extensively discussed in, e.g., Refs. [15,16].

For most time-domain simulations of wave phenomena, the minimization of the local and accumulated dispersion error over a finite range of frequencies is often more relevant than the theoretical order of accuracy of the lattice approximation itself, where “order of accuracy” refers to the behavior of the truncation error of the scheme as the lattice spacing goes to zero (low frequency or long wavelength limit). This is because of two interconnected reasons. First, the limit is never taken in practice and, second, practical simulations involve computational domains of fixed physical size, and, as a result, long wavelengths imply electrically smaller domain where the accumulated phase error (which grows linearly with the electric size of the domain) is much less of a problem.

One relevant question is then how to construct, given a particular stencil (and hence, computational cost), optimal lattice approximations for wave problems in the time domain with optimal dispersion-relation-preserving (DRP) properties over a given (possibly wide) frequency band. Moreover, as alluded to above, since the high frequency spectrum is subject to larger cumulative phase error effect than the low frequency spectrum, it is also of interest to investigate the possibility of constructing lattice approximations where the (local) incurred dispersion error is actually lower at high frequencies than at low frequencies (traditionally, the opposite is true).

With the above observations in mind, we shall describe here a general methodology to develop lattice approximations of continuum equations with DRP properties. The methodology is based on first obtaining frequency dependent stencil coefficients in the Fourier domain to minimize the
maximum grid dispersion error for wave propagation at all directions (in a minimax sense). The resulting coefficients are then expanded into a Fourier series in terms of the frequency variable and incorporated into the time-domain (update) equations after an inverse Fourier transformation. Maximally flat (Butterworth) and Chebyshev filters are further used to minimize the wave speed error for a frequency range of interest. Moreover, the use of such filters allow for the adjustment of the grid dispersion characteristics so as to minimize both the local dispersion error and the accumulated phase error in a frequency range of interest. For concreteness, we employ Maxwell’s equations in 2+1 dimensions as example, but the same methodology is applicable for other dimensions and other linear wave phenomena as well.

II. METHODOLOGY

Because the dispersion error in FDTD simulations is in general a function of both frequency, propagation angle, and the particular Courant-Friedrich-Lewy (CFL) number, several definitions of minimum dispersion error [4] are possible. We define it here in the minimax (minimum maximum) sense, i.e., for a given CFL number, the maximum dispersion error for all angles is minimized up to a certain maximum frequency. More specifically, by denoting the dispersion error as \( \delta(f, k, \chi) \), where \( f \) denotes frequency, \( k = k|k| \) is the unit propagation vector, and \( \chi \) is the CFL number, the DRP-FDTD seeks to minimize the objective function \( F = \max\{\delta(f, k, \chi) : 0 \le f = f_m, \chi = \chi_0, \forall k\} \). To coincide with the case of most practical interest, the particular CFL number \( \chi_0 \) will most often be considered the maximum allowed from the CFL condition.

The DRP procedure here consists of three main steps: (1) For a scheme with given “order of accuracy in space” (stencil), the dispersion error is expanded in a Fourier series in terms of the propagation angle, the leading terms of which are made equal to zero, and analytical expressions for the DRP coefficients are subsequently derived as a function of frequency. (2) These analytical expressions are then cast into a form implementable in the fully discrete problem (FDTD update) by using polynomial expansions in terms of the frequency variable followed by an inverse Fourier transformation. (3) Filtering schemes (maximally flat or Chebyshev) are used to fine-tune the DRP two-dimensional (2D) FDTD coefficients for a (possibly broad) range of frequencies of interest.

The first step consists in considering a scheme with a given order of accuracy in space. Higher order of accuracy in time is introduced in the second step as we expand the analytical solution in series. The term order of accuracy in space is borrowed here from traditional higher-order schemes employing Taylor expansions only [6,7]. Since this work treats the problem from a different standpoint, this term does not retain its original meaning. The term order of accuracy should rather be considered here as referring to a class of spatial stencil sizes (and not necessarily to the order of the truncation error as the discretization cell approaches zero). As we will see, the major difference between the DRP higher-order schemes derived here and traditional higher-order schemes resides in the coefficients and not on the particular stencil.

A. Nonfiltered, DRP (4,2) schemes

A traditional leap-frog scheme with second order of accuracy in space can be written in general as

\[
E_{x,m+1/2,n}^{l+1} = E_{x,m+1/2,n}^l + \frac{\Delta t}{\epsilon \Delta x} \Gamma_x (H_{x,m+1/2,n+1/2}^{l+1/2} - H_{x,m+1/2,n-1/2}^{l+1/2}), \tag{1}
\]

\[
E_{y,m,n+1/2}^{l+1} = E_{y,m,n+1/2}^l + \frac{\Delta t}{\mu \Delta y} \Gamma_y (E_{y,m+1/2,n+1}^{l+1/2} - E_{y,m+1/2,n-1}^{l+1/2}), \tag{2}
\]

\[
H_{z,m+1/2,n+1/2}^{l+1} = H_{z,m+1/2,n+1/2}^l + \frac{\Delta t}{\epsilon \Delta x} \Gamma_z (E_{z,m+1/2,n+1}^{l+1/2} - E_{z,m+1/2,n+1}^{l+1/2}), \tag{3}
\]

where the subscripts denote the spatial location and the superscripts denote the time step. We treat \( \Gamma_x \) and \( \Gamma_y \) as unknown coefficients for the moment. In the above, reciprocity is explicitly enforced [17,18] in order to obtain a conditionally stable scheme. The \( E \) and \( H \) fields can be expanded into a discrete set of Fourier modes. For each mode

\[
E_{x,m+1/2}^l = E_x e^{i\omega t k_x} e^{-i j [k_x (m+1/2) \Delta \times + k_y (n+1/2) \Delta y]} \sin \left( \frac{k \sin (\theta) \Delta y}{2} \right), \tag{4}
\]

\[
E_{y,m+1/2}^l = E_y e^{i\omega t k_y} e^{-i j [k_y (m+1/2) \Delta \times + k_x (n+1/2) \Delta y]} \sin \left( \frac{k \cos (\theta) \Delta x}{2} \right), \tag{5}
\]

\[
H_{z,m+1/2}^l = H_z e^{i\omega t k_z} e^{-i j [k_z (m+1/2) \Delta \times + k_y (n+1/2) \Delta y]} \sin \left( \frac{k \cos (\theta) \Delta x}{2} \right) \cos (\theta), \tag{6}
\]

Substituting Eqs. (4)–(6) into Eqs. (1)–(3) and noticing that \( E_x = -E \sin (\theta) \), \( E_y = E \cos (\theta) \), and \( H_z = H \), we have

\[
\sin \left( \frac{\omega \Delta t}{2} \right) E \sin (\theta) = \sin \left( \frac{\Delta t}{\epsilon \Delta y} \Gamma_x \sin \left[ \frac{k \sin (\theta) \Delta y}{2} \right] \right), \tag{7}
\]

\[
\sin \left( \frac{\omega \Delta t}{2} \right) E \cos (\theta) = \sin \left( \frac{\Delta t}{\mu \Delta x} \Gamma_y \sin \left[ \frac{k \cos (\theta) \Delta x}{2} \right] \right), \tag{8}
\]

\[
\sin \left( \frac{\omega \Delta t}{2} \right) H = \sin \left( \frac{\Delta t}{\epsilon \Delta y} \Gamma_x \sin \left[ \frac{k \sin (\theta) \Delta y}{2} \right] \sin (\theta) \right) + \sin \left( \frac{\Delta t}{\mu \Delta x} \Gamma_y \sin \left[ \frac{k \cos (\theta) \Delta x}{2} \right] \cos (\theta) \right), \tag{9}
\]

where \( k_x = k \cos (\theta) \), \( k_y = k \sin (\theta) \). The numerical dispersion relationship can be derived from Eqs. (7)–(9) as
where \( v_p = \sqrt{\mu \epsilon} \). For a given set of coefficients \( \Gamma_x \) and \( \Gamma_y \), (in the classical Yee’s scheme [4] these are equal to unity), the above equation is traditionally used to analyze the discrete dispersion in the FDTD grid. The amount by which the discrete dispersion relation deviates from the continuum limit gives the local dispersion error. In this work, we shall adopt the reverse standpoint. That is, we shall enforce the exact relation between frequency and wave number, viz. \( \omega = v_p k \), for Eq. (10), and then solve for \( \Gamma_x \) and \( \Gamma_y \) as the unknowns (in an approximate sense to be clear later on). Ideally, the exact solutions for \( \Gamma_x \) and \( \Gamma_y \) should depend on both the frequency and propagation angle. We shall first expand the dispersion error in terms of a Fourier series in terms of the angular variable \( \theta \) and enforce coefficients on the series to be zero (the number of coefficients made equal to zero give the “order” of the method). In this manner, the maximum dispersion error for all angles is minimized simultaneously, and \( \Gamma_x \) and \( \Gamma_y \) become a function of frequency only. This latter property allows for incorporation of the DRP coefficients in FDTD algorithms after a polynomial expansion.

A simpler but equivalent way of solving for \( \Gamma_x \) and \( \Gamma_y \) in Eq. (10) is to enforce \( \omega = v_p k \) into Eqs. (7) and (8) and solve for \( \Gamma_x \) and \( \Gamma_y \), respectively. [Note that (from the symmetry of the problem) by letting \( \Delta x = \Delta y \) and \( \theta = \theta + \pi/2 \), Eq. (8) reduces to Eq. (7). In particular if \( \Delta x = \Delta y \), we should of course expect \( \Gamma_x = \Gamma_y \) as these coefficients are reduced to functions of frequency only.] We start by defining an error functional proportional to the difference between left-hand side (lhs) and right-hand side (rhs) of Eq. (7):

\[
\delta_2(\Gamma_x, \theta) = \frac{\sqrt{2}}{\lambda y} \sin \left( \frac{\pi q_y \lambda y}{\sqrt{2}} \right) \sin(\theta) - \Gamma_x \sin[\pi q_y \sin(\theta)],
\]

(11)

where \( \lambda y = \sqrt{2} v_p \Delta t / \Delta y \), \( q_y = \Delta y / \lambda \). Solving Eq. (7) with \( \omega = v_p k \) is therefore equivalent to letting \( \delta_2(\Gamma_x, \theta) = 0 \) and enforcing \( \delta \mathcal{E} / \delta \eta = \mathcal{H} \).

We expand \( \delta_2(\Gamma_x, \theta) \) in a Fourier series in terms of \( \theta \) and use the following identities:

\[
\int_0^{2\pi} \sin[\pi \sin(\theta)] \cos(n \theta) d \theta = 0
\]

for even \( n \) and

\[
\int_0^{2\pi} \sin[\pi \sin(\theta)] \sin(n \theta) d \theta = 2 \pi J_n(x)
\]

for odd \( n \) and

\[
\int_0^{2\pi} \sin[x \sin(\theta)] \cos(n \theta) d \theta = 0
\]

for even \( n \), where \( J_n(x) \) is the \( n \)th order first-kind Bessel function, to force the first nonzero (dominant) term of the series to be zero. This leads to

\[
\Gamma_x = \gamma_x^{(\infty)} = \frac{\sqrt{2}}{\lambda y} \sin \left( \frac{\pi q_y \lambda y}{\sqrt{2}} \right) / [2 J_1(\pi q_y)].
\]

(12)

Substituting Eq. (12) in Eq. (11), the residual error in Eq. (11) is given by

\[
\delta_2(\gamma_x^{(\infty)}, \theta) = \gamma_x^{(\infty)} J_3(\pi q_y) \sin(3 \theta) + \gamma_x^{(\infty)} J_5(\pi q_y) \sin(5 \theta) + \cdots.
\]

The above is an asymptotic series [19]. The magnitude of each coefficient in the series represents the maximum dispersion error from the corresponding angular mode (either at \( \sin(n \theta) = 1 \) or at \( \cos(n \theta) = 1 \)).

It is clear that the solution \( \gamma_x^{(\infty)} \) given in Eq. (12) cannot be implemented in a time domain method. Because of this, we expand \( \gamma_x^{(\infty)} \) in a Taylor series around \( q_y = 0 \) and retain the lowest order terms, i.e., \( \gamma_x^{(\infty)} = \gamma_x^{(2)} + O(q_y^3) \) with

\[
\gamma_x^{(2)} = 1 + \frac{\pi^2}{4} 4 \pi^2 q_y^2 (3 - 2 \lambda y^2).
\]

(13)

The above can be easily transformed back to time domain through \( \partial^2 / \partial t^2 \). However, if straightforward time discretization schemes are employed directly on the resulting equations (with third-order time derivatives), the update becomes unconditionally unstable [7]. Alternatively, the second-order time derivative can be further cast as a combination of spatial derivatives as \( v_p^{-2} \) (Helmholtz equation) and discretized as such [20]. This latter transformation is valid for staggered grids as long as \( \epsilon \) is uniform in the local stencil. In this manner, Eq. (13) becomes

\[
\gamma_x^{(2)} = 1 - \frac{(3 - 2 \lambda y^2) \Delta x^2 V^2}{96}.
\]

(14)

Following a similar procedure for \( \Gamma_y \), we find

\[
\gamma_y^{(2)} = 1 - \frac{(3 - 2 \lambda y^2) \Delta x^2 V^2}{96},
\]

(15)

where \( \lambda y = \sqrt{2} v_p \Delta t / \Delta x \). If only the first-order terms in Eqs. (14) and (15) are taken, the Yee’s scheme is recovered. Note that the second terms in Eqs. (14) and (15) are analogous to third-order time derivative terms in traditional schemes with fourth order of accuracy in time [7]. The difference resides in multiplicative factors and, in this sense, we call it a (nonfiltered, minimax) DRP (4,2) scheme.

To investigate the resulting grid dispersion, we define a dispersion error functional \( \delta_E(\Gamma_x, \Gamma_y, \theta) \) proportional to the difference between the lhs and the rhs of Eq. (10), i.e.,
FIG. 1. Comparison of the maximum value of the dispersion error $\delta_2(\Gamma_x, \Gamma_y, \theta)$ for all angles when using analytical solution, $(\Gamma_x, \Gamma_y) = (\gamma_x^{(2)}, \gamma_y^{(2)})$, second-order approximation $(\Gamma_x, \Gamma_y) = (\gamma_x^{(2)}, \gamma_y^{(2)})$ and Yee’s scheme $(\Gamma_x, \Gamma_y) = (1, 1)$.

$$
\delta_2(\Gamma_x, \Gamma_y, \theta) = \frac{\sqrt{2} \sin \left( \frac{\pi q \chi}{\sqrt{2}} \right)^2}{\chi} \frac{\Gamma_y^2 \sin^2 \left( \frac{k \sin(\theta) \Delta y}{2} \right)}{\Delta y^2} + \frac{\Gamma_y^2 \sin^2 \left( \frac{k \cos(\theta) \Delta x}{2} \right)}{\Delta x^2},
$$

where $\chi = \sqrt{2} v \Delta t/h$, $q = h/\lambda$, and $h = \min(\Delta x, \Delta y)$. By substituting $\Gamma_x = \gamma_x^{(2)}$ and $\Gamma_y = \gamma_y^{(2)}$ into Eq. (16), we obtain a limit value for the error $\delta^{(\infty)}_2(\theta) = \delta_2(\gamma_x^{(2)}, \gamma_y^{(2)}, \theta)$. In practical time-domain simulations employing Eqs. (14) and (15), we have an error $\delta^{(\infty)}_2(\theta) = \delta_2(\gamma_x^{(2)}, \gamma_y^{(2)}, \theta)$ instead, which inevitably introduces additional errors at high frequencies. The function $\delta^{(\infty)}_2(\theta)$ therefore serves as an inferior theoretical limit when an infinite Taylor expansion is considered. However, as we shall see in the following section, by expanding $\Gamma_x$ and $\Gamma_y$ in a different basis (e.g., Chebyshev polynomials) this limit can indeed be overcome over some finite, preassigned frequency range.

Figure 1 shows the maximum value for all angles of $\delta^{(\infty)}_2(\theta)$ (analytical), $\delta^{(2)}_2(\theta)$ (second order), and $\delta_2(1, 1, \theta)$ (Yee’s scheme), as a function of the number of wavelengths per cell (or, equivalently, the frequency). In these plots, $\Delta x = \Delta y$ and the CFL number $\chi = 3/4$ ($\chi = \chi_y = \chi_x$). From this figure, we observe that, for all the frequency range such that $q_y = 0.1$, the second-order approximation already gives results almost as accurate as using the analytical expression (12), and starts to deviate only slightly at high frequencies close to $q_y = 0.1$.

B. Nonfiltered, DRP (4,4) schemes

Traditional FDTD schemes with fourth order of accuracy in space employ larger stencils for the $E_x$ and $E_y$ update equations. In this case, the $E_x$ update can be rewritten as follows [compare with Eq. (1)]

$$
E_{x,m + 1/2,n} = E_{x,m + 1/2,n} + \frac{\Delta t}{\epsilon \Delta y} \Gamma_x (H_{c,m + 1/2,n} + 1/2)
- H_{c,m + 1/2,n - 1/2} + \frac{\Delta t}{\epsilon \Delta y} \Gamma_x (H_{c,m + 1/2,n} + 1/2
- H_{c,m + 1/2,n - 3/2}).
$$

(17)

We proceed as before and $\Gamma_{x1}$, $\Gamma_{x2}$ can be solved from

$$
\sin \left( \frac{\nu \Delta t}{2} \right) \mathcal{E} \sin(\theta) = \frac{\Delta t}{\epsilon \Delta y} \Gamma_x \mathcal{H} \left[ \frac{k \sin(\theta) \Delta y}{2} \right] + \frac{\Delta t}{\epsilon \Delta y} \Gamma_x \mathcal{H} \left[ \frac{3 k \sin(\theta) \Delta y}{2} \right].
$$

(18)

To obtain $\Gamma_{x1}$ and $\Gamma_{x2}$, we define an error functional in terms of the difference between the lhs and rhs of Eq. (18):

$$
\delta_4(\Gamma_{x1}, \Gamma_{x2}, \theta) = \sqrt{2} \sin \left( \frac{\pi q y X_x}{\sqrt{2}} \right) \sin(\theta)
- \frac{\Delta t}{\epsilon \Delta y} \Gamma_x \mathcal{H} \left[ \frac{k \sin(\theta) \Delta y}{2} \right],
$$

(19)

Similarly as before, we expand $\delta_4(\Gamma_{x1}, \Gamma_{x2}, \theta)$ in a Fourier series in terms of $\theta$. However, since there are two unknowns $\Gamma_{x1}$ and $\Gamma_{x2}$, we now may force the first two nonzero terms of the series to be zero [$\sin(\theta)$ and $\sin(3\theta)$ terms]. The solutions are given by

$$
\Gamma_{x1} = \gamma_{x1}^{(\infty)} = \frac{2 \sqrt{2} \sin \left( \frac{\pi q y X_x}{\sqrt{2}} \right) J_3(3 q y)}{2 X_x [J_1(3 q y) J_3(3 q y) - J_1(3 q y) J_3(3 q y)]},
$$

(20)

$$
\Gamma_{x2} = \gamma_{x2}^{(\infty)} = \frac{2 \sqrt{2} \sin \left( \frac{\pi q y X_x}{\sqrt{2}} \right) J_3(3 q y)}{2 X_x [J_1(3 q y) J_3(3 q y) - J_1(3 q y) J_3(3 q y)]},
$$

(21)

With these coefficients, the residual error in Eq. (19) becomes

$$
\delta_4(\gamma_{x1}^{(\infty)}, \gamma_{x2}^{(\infty)}, \theta) = \gamma_{x1}^{(\infty)} J_5(3 q y) \sin(5 \theta)
+ \gamma_{x2}^{(\infty)} J_5(3 q y) \sin(5 \theta) + \ldots.
$$

For incorporation into the FDTD update, $\gamma_{x1}^{(\infty)}$ and $\gamma_{x2}^{(\infty)}$ are approximated as

$$
\gamma_{x1}^{(\infty)} \approx \gamma_{x1}^{(2)} = \gamma_{x1}^{(2)} - \gamma_{x12}^{(2)} \Delta y^2 v^2,
$$

(22)

$$
\gamma_{x2}^{(\infty)} \approx \gamma_{x2}^{(2)} = \gamma_{x2}^{(2)} - \gamma_{x22}^{(2)} \Delta y^2 v^2.
$$

(23)
The CFL stability condition can be derived in a standard way with
\[\gamma_{x11}^{(2)} = 9g^{(2)} = -3(3-4\chi^2),\]
\[\gamma_{x21}^{(2)} = -\frac{1}{24}\gamma_{x22}^{(2)} = \frac{(27-4\chi^2)}{4608}.
\]
The lowest order terms in Eqs. (22) and (23) recover Fang’s (2,4) FDTD scheme [6].

Similarly, \(\gamma_{x1}^2\) and \(\gamma_{x2}^2\) are written as

\[
E_{x,m+1/2,n}^{t+1} = E_{x,m+1/2,n}^{t} + \frac{\Delta t}{\epsilon \Delta y} \left[ \gamma_{x11}^{(2)} + \gamma_{x12}^{(2)} \left( 3 + 2 \frac{\Delta y^2}{\Delta x^2} \right) - \gamma_{x22}^{(2)} \left( H_{z,m+1/2,n+1/2} - H_{z,m+1/2,n-1/2} \right) + \gamma_{x21}^{(2)} \left( 1 + \frac{\Delta y^2}{\Delta x^2} \right) \left( H_{z,m+1/2,n+3/2} - H_{z,m+1/2,n-3/2} \right) - \gamma_{x12}^{(2)} \left( H_{z,m+1/2,n+1/2} + H_{z,m+1/2,n-1/2} \right) - \gamma_{x11}^{(2)} \left( H_{z,m+1/2,n+1/2} + H_{z,m+1/2,n-1/2} \right) \right].
\]

The CFL stability condition can be derived in a standard way [4] and the result is

\[
\Delta t \leq \frac{1}{\sqrt{v_p} \sqrt{\frac{D_x^2}{\Delta y^2} + \frac{D_y^2}{\Delta x^2}}}, \quad (26)
\]

where

\[
D_x = 2 \frac{\Delta y^2}{\Delta x^2} \left[ \gamma_{x12}^{(2)} \sin \left( \frac{k_x \Delta y}{2} \right) + \gamma_{x22}^{(2)} \sin \left( \frac{3k_x \Delta y}{2} \right) \right] \cos(k_x \Delta x)
\]

\[
+ \gamma_{x21}^{(2)} \sin \left( \frac{5k_x \Delta y}{2} \right) - \gamma_{x11}^{(2)} \left[ 3 + \frac{\Delta y^2}{\Delta x^2} \right]
\]

\[
- \gamma_{x22}^{(2)} \sin \left( \frac{k_x \Delta y}{2} \right) - \gamma_{x21}^{(2)} \left[ 1 + \frac{\Delta y^2}{\Delta x^2} \right]
\]

and similarly for \(D_y\). For \(\Delta x = \Delta y\), Eq. (26) becomes

\[
\Delta t \leq \frac{h \chi}{v_p \sqrt{2}} = \frac{h}{v_p \sqrt{2} \sqrt{\frac{D_{x,y}^2}{\Delta y^2}}},
\]

with \(|D_{x,y}|_{\text{max}}\) denotes the maximum possible module of \(D_x\) and \(D_y\). The minimum maximum CFL number \(\chi\) with guaranteed stability is found to be 48/65. Note that this is a conservative bound. In reality, we have found \(\chi = 0.75\) to produce stable updates in our tests.

To estimate the approximation error, an error functional \(\tilde{\delta}_s(\Gamma_{x1}, \Gamma_{x2}, \Gamma_{y1}, \Gamma_{y2}, \theta)\) is constructed in an analogous way as the error functional in Eq. (16). Figure 2 shows the comp-

![Image](https://example.com/image.png)

**FIG. 2.** Comparison of the largest \(\tilde{\delta}_s\) among all angles by using different DRP schemes for the fourth-order stencil. See text for details.
parison of the largest $\delta_4$ (among all angles) by using different schemes when $\Delta x = \Delta y$ and $\chi = 48/65$. In this figure, the analytical result refers to the (ideal) choice 
$(\gamma_{x1}, \gamma_{x2}, \gamma_{y1}, \gamma_{y2}) = (\gamma_{x1}^{(m)}, \gamma_{x2}^{(m)}, \gamma_{y1}^{(m)}, \gamma_{y2}^{(m)})$, whereas the second-order result refers to the (implementable) approximation for the coefficients given by Eqs. (22)–(25), 
$(\gamma_{x1}, \gamma_{x2}, \gamma_{y1}, \gamma_{y2}) = (\gamma_{x1}^{(2)}, \gamma_{x2}^{(2)}, \gamma_{y1}^{(2)}, \gamma_{y2}^{(2)})$. Maximally flat and Chebyshev refer to (implementable) approximations using filtering schemes to be detailed in the following section.

C. Filtering

Instead of possibly using even higher-order terms in the Taylor series or other sophisticated (and more costly) time integration schemes, it is possible to improve the results of the preceding section (while maintaining the same spatial stencil sizes and computational cost) by using filters adjusted to some preassigned, finite frequency range.

The filters are designed to obtain both $\gamma_x$ and $\gamma_y$ (second-order stencil), or $\Gamma_{x1}$, $\Gamma_{x2}$, $\Gamma_{y1}$, and $\Gamma_{y2}$ (fourth-order stencil). The procedure is essentially the same for all coefficients and therefore, we will describe only the $\Gamma_{x1}$ case in detail.

1. Maximally flat (Butterworth) filters

Expanding Eq. (20) in a Taylor series, and approximating Eq. (20) in a finite series as in Eq. (22) noting that $\Delta y^2 \nabla^2 = -4\pi^2 q_y^2$, we have

$$\gamma_{x1}^{(m)} - \gamma_{x1}^{(m)} = \left[ \frac{9}{8} - \gamma_{x1}^{(m)} \right] - \frac{3(4\chi_y^2 - 3)\pi^2}{128} + 4\pi^2 \gamma_{x1}^{(m)} q_y^2 + \frac{3\pi^4 \chi_y^4 (2\chi_y^2 - 5)}{2560} - q_y^4 + O(q_y^6).$$

The superscript $(m)$ in $\gamma_{x1}^{(m)}$ and $\gamma_{x1}^{(m)}$ above refer to maximally flat coefficients. Previously, we have simply chosen $\gamma_{x1}^{(m)} = \gamma_{x1}^{(2)}$ and $\gamma_{x1}^{(m)} = \gamma_{x1}^{(2)}$ so as to make the first and second terms of the right-hand side of the above identically zero [and obtain a $O(q_y^4)$ truncation error]. These coefficients are treated as unknowns for the moment. The above equation can be rewritten as

$$\gamma_{x1}^{(m)} - \gamma_{x1}^{(m)} = (d_0, 0, d_2, 0, d_4) V^T + O(q_y^6),
\tag{29}$$

where

$$d_0 = \frac{9}{8} - \gamma_{x1}^{(m)},
\quad d_2 = -\frac{3(4\chi_y^2 - 3)\pi^2}{128} + 4\pi^2 \gamma_{x1}^{(m)}
\quad d_4 = \frac{3\pi^4 \chi_y^4 (2\chi_y^2 - 5)}{2560},
\quad V = (1, q_y, q_y^2, q_y^3, q_y^4).$$

By expanding the above in a new basis \{(q_y - q_y^*)^n\}, where $q_y^*$ refers to a center frequency of interest, the following relationship holds

$$Q^T = AV^T$$

with

$$Q = (1, (q_y - q_y^*)^2, (q_y - q_y^*)^3, (q_y - q_y^*)^4)$$

and, therefore, Eq. (29) can be rewritten as

$$\gamma_{x1}^{(m)} = \gamma_{x1}^{(m)} + (d_0, 0, d_2, 0, d_4) A^{-1} Q^T + O(q_y^6)$$

where

$$A = \begin{pmatrix}
*20c1 & 0 & 0 & 0 & 0 \\
-q_y^2 & 1 & 0 & 0 & 0 \\
-q_y^4 & -2q_y^2 & 1 & 0 & 0 \\
-q_y^6 & 3q_y^2 & -3q_y^4 & 1 & 0 \\
q_y^2 & -4q_y^4 & 6q_y^2 & -4q_y^4 & 1
\end{pmatrix}$$

To solve for $\gamma_{x1}^{(m)}$ and $\gamma_{x1}^{(m)}$, we force $d_0$ and $d_1$ to be zero and obtain

$$\gamma_{x1}^{(m)} = \frac{2880 + 3\pi^4 \chi_y^4 q_y^2 (2\chi_y^2 - 5)}{2560},$$

$$\gamma_{x1}^{(m)} = \frac{3 \{30 + \chi_y^2 \pi^2 q_y^2 (2\chi_y^2 - 5) - 40\} / 5120,$n}
\tag{33}
\gamma_{x1}^{(m)} = \frac{3 \{30 + \chi_y^2 \pi^2 q_y^2 (2\chi_y^2 - 5) - 40\} / 5120,$n}
\tag{34}

which are functions of $q_y^*$. Note that, if we let $q_y^* = 0$, we recover $\gamma_{x1}^{(m)} = \gamma_{x1}^{(2)}$ and $\gamma_{x1}^{(m)} = \gamma_{x1}^{(2)}$.

In this manner,

$$\gamma_{x1}^{(m)} - \gamma_{x1}^{(m)} = [d_2(q_y - q_y^*)^2 + d_3(q_y - q_y^*)^3 + d_4(q_y - q_y^*)^4]
\tag{29}$$

$+ O(q_y^6).$

At the center frequency, $q_y = q_y^*$ and $\delta = O(q_y^6)$. The remainder corresponds to a $O(q_y^6)$ error. Around the center frequency, the error is dominated by the $\Sigma_{n=2}^p d_n(q_y - q_y^*)^n$ term. In the above, we have illustrated the derivation of $\gamma_{x1}^{(m)}$ by using a $p = 4$ order polynomial for the Taylor expansion in Eq. (29). Theoretically, we can increase indefinitely the polynomial order $p$ at the (one-time) cost of inverting a
larger $A$ and hence make the response as close to an exact second-order maximally flat filter response as desired. This has the (one-time) cost of inverting an increasingly larger matrix $A$. In practice, we observe a fast convergence to the exact filter response since only negligible improvements are obtained for $p \geq 6$. We employ $p = 10$ throughout our numerical simulations. Also in practice, the design frequency $q_c$ may be chosen as the highest frequency of interest.

Figure 2 includes the $\tilde{\delta}_4$ error of the DRP scheme with maximally flat filters designed at $q^f = 0.1$ ($\Delta x = \Delta y$). From this figure, we observe that near the central frequency ($q^f = 0.1$), the $\tilde{\delta}_4$ using a second-order (with $p = 10$) maximally flat filter is almost indistinguishable from the one using the analytical solutions (20) and (21).

Following the same procedure as the maximally flat filter case, we can solve for the coefficients and obtain $\gamma_{x11}^{(c)}$ and $\gamma_{y12}^{(c)}$ of an approximate second-order Chebyshev filter. Here

$$\gamma_{x1}^{(c)} - \gamma_{x1}^{(c)} = \tilde{d}_2 T_2[(q_y - q_y^c)/\Delta q_y] + \tilde{d}_3 T_3[(q_y - q_y^c)/\Delta q_y] + \tilde{d}_4 T_4[(q_y - q_y^c)/\Delta q_y] + O(q_y^6),$$

where $\tilde{d}_2$, $\tilde{d}_3$, and $\tilde{d}_4$ denote the coefficients of the corresponding Chebyshev polynomials. Figure 2 (with $q_y = q_y^c = q$) clearly shows that, by using Chebyshev filters designed with $q^f = 0.1$ and $\Delta q = 0.02$, the error functional $\tilde{\delta}_4$ can be smaller than the one employing Eqs. (20) and (21) around the design frequency. ($\Delta q = 0.02$ is used here as an example. As shown later, $\Delta q$ can be fine-tuned for a better performance.)

Since the Chebyshev filter is of second order, larger ripples are expected to occur in the passband $q^f - \Delta q \leq q \leq q^f + \Delta q$ if a larger $\Delta q$ is used. In usual filter design, large ripples are undesired, but in our context, $\Delta q$ represents an extra degree of freedom that can be explored to reduce the accumulated phase error. This will be illustrated in the following section where we also set $p = 10$ for the Chebyshev filter approximation.

### III. Numerical Results

We compare the DRP (4,2) and (4,4) schemes against the Yee’s scheme, a traditional (2,4) scheme, and Deveze’s (4,4) scheme. Both nonfiltered, and the maximally flat and Chebyshev filtered versions of the DRP schemes are considered. The phase velocity in free space is solved from the transcendental dispersion relation by assuming a uniform 2D FDTD grid with $\Delta x = \Delta y$, or $\Gamma_x = \Gamma_y$, and, hence, the coordinate subscript in the coefficients is dropped in what follows.

The CFL number $\chi = 1$ is used for ordinary FDTD, while $\chi = 6/7$ for the (2,4) scheme, $\chi = 3/4$ for the (4,2) scheme and $\chi = 48/65$ for all (4,4) schemes. In fact, there is an optimal CFL number for each algorithm, which yields the minimum dispersion error among all possible CFL numbers. When comparing different algorithms, we employ their largest pos-

<table>
<thead>
<tr>
<th>TABLE I. Coefficients used in different (4,4) schemes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_{11}$ $\Gamma_{12}$</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Deveze</td>
</tr>
<tr>
<td>Nonfiltered DRP</td>
</tr>
<tr>
<td>Maximally Flat DRP</td>
</tr>
<tr>
<td>Chebyshev DRP</td>
</tr>
</tbody>
</table>
TABLE II. Coefficients used in different (4,4) schemes (cont’d).

<table>
<thead>
<tr>
<th>Scheme</th>
<th>(\Gamma_{21})</th>
<th>(\Gamma_{22})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deveze</td>
<td>(-4.166,667\times10^{-2})</td>
<td>(4.207,758\times10^{-4})</td>
</tr>
<tr>
<td>Nonfiltered DRP</td>
<td>(-4.166,667\times10^{-2})</td>
<td>(-5.386,002\times10^{-3})</td>
</tr>
<tr>
<td>Maximally flat DRP</td>
<td>(-4.159,725\times10^{-2})</td>
<td>(-5.733,253\times10^{-3})</td>
</tr>
<tr>
<td>Chebyshev DRP</td>
<td>(-4.159,842\times10^{-2})</td>
<td>(-5.744,642\times10^{-3})</td>
</tr>
</tbody>
</table>

Table I and Table II give some coefficients used in the DRP (4,4) schemes, where the maximally flat filter is designed with \(q^c=0.1\), and the Chebyshev filter is designed with \(q^c=0.1\) with \(\Delta q=0.02\). Table III and Table IV give the coefficients using Chebyshev filtering schemes for various pairs \((q_c, \Delta q)\).

Figures 3 and 4 show the maximum (for all angles) phase error per wavelength, defined as \(\left(\hat{k}/k-1\right)\times360\) with \(\hat{k}\) denotes the discrete wave number obtained by solving the transcendental dispersion relation and \(k\) refers to the exact (continuum) wave number. As we see from Fig. 3, the DRP (4,2) scheme is about as accurate as the (2,4) scheme (with opposite sign). Figure 4 shows that filtering schemes do reduce the dispersion error considerably around the specified frequency. We also observe that the Chebyshev filter performs better than the maximally flat filter in the full frequency range considered. Note that Fig. 4 does not correspond to Fig. 2 well after the crossover since Fig. 2 is just an analytical estimation for design purposes, while Fig. 4 depicts the actual performance of the algorithms.

We also note from Fig. 4 that, for DRP schemes with filtering, the dispersion error at high frequencies can indeed be made smaller than at low frequencies (contrary to nonfiltered DRP and traditional schemes). As mentioned in the Introduction, this is a desirable characteristic because, for a given computational domain size, high frequencies correspond to an electrically larger problem and hence are more impacted by the accumulated phase error. Nevertheless, unless the local dispersion error decreases faster than linearly with frequency, the largest accumulated phase error in the computational domain is still given by the highest frequency components. This issue will be further elaborated next.

We assume a FDTD simulation in a frequency range such that the lowest frequency corresponds to \(q=0.01\) and the highest frequency corresponds to \(q=0.1\) (which is typical). We calculate the phase error accumulated in a distance which equals the largest wavelength being considered. Figure 5 shows the largest (for all angles) phase error accumulated in such a distance by employing Chebyshev filtered DRP algorithms with different \(\Delta q\) (while \(q_c\) is kept fixed as 0.1). For \(\Delta q=0.02\), the usual notion that the largest accumulated phase error is dominated by the highest frequency is still valid. However, this is not true anymore if larger values for \(\Delta q\) are considered. In fact, when \(\Delta q=0.046\), the largest accumulated phase error is now determined by lower frequency components. Indeed, there is an optimal \(\Delta q\) around \(\Delta q=0.058\) for which the maximum accumulated magnitude of the phase error for all frequencies is a minimum.

In the case of narrow band simulations, we can obtain a pair \((q_c, \Delta q)\) which yields a minimum possible maximum phase error for all angles at some specific frequency. As illustrated in Fig. 6, if we specify \(q^c=0.1\) and \(\Delta q=0.046\), the maximum phase error is minimum at \(q=0.087\). This suggests that, in order to obtain a better performance at a specific frequency, we can simply design filters with a slightly larger \(q_c\) and a fine tuned \(\Delta q\) accordingly. Figure 6 shows another example with \(q=0.114\) and \(\Delta q=0.053\) which is intended to improve the performance at \(q=0.1\). As we see, a maximum phase error less than \(5\times10^{-3}\) is obtained at \(q=0.1\), while a good bandwidth is maintained.

Computational costs are compared for the same accuracy requirements. Since decreasing the cell size also decreases the time step, a fair comparison of computational costs

![FIG. 3. Comparison of the maximum (for all angles) phase error per wavelength using Yee’s, a traditional (2,4) scheme and the non-filtered DRP (4,2) scheme.](image-url)
should consider the same elapsed time duration, not the same number of time steps. Each field update of the DRP (4,4) schemes involves 18 floating point operations in contrast to three for the ordinary Yee’s scheme [4] and six for Fang’s (2,4) scheme. The present maximum CFL number is 48/65, in contrast to one for the Yee’s scheme, and 6/7 for Fang’s (2,4) scheme. For a wide band problem with frequency spectra in the range \( q = 0.01 - 0.1 \), the worst accuracy for the DRP scheme with Chebyshev filtering with \( (q^c = 0.1, \Delta q = 0.058) \) is about 0.073°/\( \lambda \). To achieve this accuracy, the cell size in Yee’s scheme needs to be reduced by about 25 times, and a problem of same physical size and elapsed time would require 625 times more memory and roughly 1925 times more central processing unit time than the DRP (4,4). A similar comparison can be made against Fang’s (2,4) scheme, and the results are summarized in Table V.

As a final example, we calculate the resonant frequencies of a \( TE_z \) single ridge resonator, as shown in Fig. 7, from DC to 15 GHz. The boundaries correspond to an perfect electric conductor. This 2D geometry is commonly used to find the cutoff frequencies of the corresponding 3D single ridge waveguide under \( TE_z \) excitation. The initial field is set to zero and a point source is located inside the resonator. To obtain adequate frequency resolution for the resonant frequencies, the time-domain simulations must be run over sufficiently long integration time. The resonant field at the cutoff frequencies corresponds to waves bouncing many times inside the resonator, where grid dispersion error will accumulate and affect the results. This is an example of a problem in a relatively small domain that can still be sensitive to grid dispersion error effects. Simulations are performed using Yee’s scheme, Fang’s (2,4) scheme, and DRP (4,4) scheme with Chebyshev filtering optimized with \( (q^c = 0.1, \Delta q = 0.058) \). The CFL numbers are the same as before. We run long enough time steps to guarantee 1 MHz frequency resolution in the FFT. As mentioned before, this implies running the same elapsed time for different simulations (and hence different number of time steps). All schemes use the same spatial resolution corresponding to 13 cells at short-
The dispersion error can be further reduced to a theoretical minimax limit (i.e., if an infinite Taylor series were used) by employing maximally flat (Butterworth) filters or even exceed that limit at some finite frequency band using Chebyshev filters. The behavior of the local dispersion error can also be adjusted by the filtering schemes in order to reduce the accumulated phase error in the computational domain on a preassigned given frequency band. This is done by making the local dispersion error at high frequencies smaller than at low frequencies.

IV. CONCLUSIONS

We have described a general approach to construct DRP schemes for large-scale 2D FDTD simulations of Maxwell’s equations. The maximum (for all angles) local dispersion error is minimized (minimax sense) by expanding the local dispersion error (for a given spatial stencil) in a Fourier series in terms of propagation angle and enforcing the leading terms of this series to be zero. DRP-FDTD coefficients are then obtained which are functions of frequency. Using a polynomial (Taylor) series of frequency, these coefficients can be incorporated into the FDTD update by using the first terms of this series. The same methodology can be used to derive optimized 3D schemes as well [21].

The dispersion error can be further reduced to a theoretical minimax limit (i.e., if an infinite Taylor series were used) by employing maximally flat (Butterworth) filters or even exceed that limit at some finite frequency band using Chebyshev filters. The behavior of the local dispersion error can also be adjusted by the filtering schemes in order to reduce the accumulated phase error in the computational domain on a preassigned given frequency band. This is done by making the local dispersion error at high frequencies smaller than at low frequencies.