Journal of Computational Physics 394 (2019) 100–135

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Dispersion analysis of finite difference and discontinuous Galerkin schemes for Maxwell's equations in linear Lorentz media

Yan Jiang [a], Puttha Sakkaplangkul [b,*], Vrushali A. Bokil [c,1], Yingda Cheng [d,2], Fengyan Li [e,3]

[a] School of Mathematical Sciences, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China
[b] Department of Mathematics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok 10520, Thailand
[c] Department of Mathematics, Oregon State University, Corvallis, OR 97331, USA
[d] Department of Mathematics, Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA
[e] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

ARTICLE INFO

Article history:
Received 27 September 2018
Received in revised form 15 May 2019
Accepted 19 May 2019
Available online 24 May 2019

Keywords:
Maxwell's equations
Lorentz model
Numerical dispersion
Finite differences
Discontinuous Galerkin finite elements

ABSTRACT

In this paper, we consider Maxwell's equations in linear dispersive media described by a single-pole Lorentz model for electronic polarization. We study two classes of commonly used spatial discretizations: finite difference methods (FD) with arbitrary even order accuracy in space and high spatial order discontinuous Galerkin (DG) finite element methods. Both types of spatial discretizations are coupled with second order semi-implicit leap-frog and implicit trapezoidal temporal schemes. By performing detailed dispersion analysis for the semi-discrete and fully discrete schemes, we obtain rigorous quantification of the dispersion error for Lorentz dispersive dielectrics. In particular, comparisons of dispersion error can be made taking into account the model parameters, and mesh sizes in the design of the two types of schemes. This work is a continuation of our previous research on energy-stable numerical schemes for nonlinear dispersive optical media [6,7]. The results for the numerical dispersion analysis of the reduced linear model, considered in the present paper, can guide us in the optimal choice of discretization parameters for the more complicated and nonlinear models. The numerical dispersion analysis of the fully discrete FD and DG schemes, for the dispersive Maxwell model considered in this paper, clearly indicate the dependence of the numerical dispersion errors on spatial and temporal discretizations, their order of accuracy, mesh discretization parameters and model parameters. The results obtained here cannot be arrived at by considering discretizations of Maxwell's equations in free space. In particular, our results contrast the advantages and disadvantages of using high order FD or DG schemes and leap-frog or trapezoidal time integrators over different frequency ranges using a variety of measures

* Corresponding author.
E-mail addresses: jiangy@ustc.edu.cn (Y. Jiang), puttha.sa@kmitl.ac.th (P. Sakkaplangkul), bokilv@math.oregonstate.edu (V.A. Bokil), ycheng@msu.edu (Y. Cheng), lif@rpi.edu (F. Li).

[1] Research is supported by NSF grant DMS-1720116.
[2] Research is supported by NSF grants DMS-1453661, DMS-1720023 and the Simons Foundation under award number 558704.
[3] Research is supported by NSF grant DMS-1719942.

https://doi.org/10.1016/j.jcp.2019.05.022

of numerical dispersion errors. Finally, we highlight the limitations of the second order accurate temporal discretizations considered.

1. Introduction

The electromagnetic (EM) field inside a material is governed by the macroscopic Maxwell's equations along with constitutive laws that account for the response of the material to the electromagnetic field. In this work, we consider a linear dispersive material in which the delayed response to the EM field is modeled as a damped vibrating system for the polarization accounting for the average dipole moment per unit volume over the atomic structure of the material. The corresponding mathematical equations are called the Lorentz model for electronic polarization. Such dielectric materials have actual physical dispersion. The complex-valued electric permittivity of such a dispersive material is frequency dependent and includes physical dissipation, or attenuation. It is well known that numerical discretizations of (systems of) partial differential equations (PDEs) will have numerical errors. These errors include dissipation, the dampening of some frequency modes, and dispersion, the frequency dependence of the phase velocity of numerical wave modes [45]. To preserve the correct physics, it is important that the dispersion and dissipation effects are accurately captured by numerical schemes, particularly for long time simulations. Thus, an understanding of how numerical discretizations affect the dispersion relations of PDEs is important in constructing good numerical schemes that correctly predict wave propagation over long distances. When the PDEs have physical dispersion modeling a retarded response of the material to the imposed electromagnetic field, the corresponding numerical discretizations will support numerical dispersion errors that have a complicated dependence on mesh step sizes, spatial and temporal accuracy and model parameters.

In this paper, we perform dispersion analysis of high spatial order discontinuous Galerkin (DG) and a class of high order finite difference (FD) schemes, both coupled with second order implicit trapezoidal or semi-implicit leap-frog temporal discretizations for Maxwell's equations in linear Lorentz media. The fully discrete time domain (TD) methods are the leap-frog DGTD or FDTD methods and the trapezoidal DGTD or FDTD methods. This paper is a continuation of our recent efforts on energy stable numerical schemes for nonlinear dispersive optical media. In [6,7], we developed fully discrete energy stable DGTD and FDTD methods, respectively, for Maxwell's equations with linear Lorentz and nonlinear Kerr and Raman responses via the *auxiliary differential equation* (ADE) approach. These schemes include second order modified leap-frog or trapezoidal temporal schemes combined with high order DG or FD methods for the spatial discretization. In the ADE approach, ordinary differential equations (ODEs) for the evolution of the electric polarization are appended to Maxwell's equations. The two spatial discretizations that were used, the DG method and the FD method are very popular methods for electromagnetic simulations in the literature. The DG methods, which are a class of finite element methods using discontinuous polynomial spaces, have grown to be broadly adopted for EM simulations in the past two decades. They have been developed and analyzed for time dependent linear models, including Maxwell's equations in free space (e.g., [11,12,20]), and in dispersive media (e.g., [16,24,31,33]). The Yee scheme [47] is a leap-frog FDTD method that was initially developed for Maxwell's equations in linear dielectrics, and is one of the gold standards for numerical simulation of EM wave propagation in the time domain. The Yee scheme has been extended to linear dispersive media [25,30,29] (see the books [43,44] and references therein), and then to nonlinear dispersive media [50,18,25,21,42]. Additional references for Yee and other FDTD methods for EM wave propagation in linear and nonlinear Lorentz dispersion can be found in [27,26,9,19,39] for the 1D case, and in [15,28,50] for 2D and 3D cases.

In our recent work [6,7], we proved energy stability of fully discrete new FDTD and DGTD schemes for Maxwell's equations with Lorentz, Kerr and Raman effects. Both types of schemes employ second order time integrators, while utilizing high order discretizations in space. The schemes are benchmarked on several one-dimensional test examples and their performance in stability and accuracy are validated. The objective of the present work is to conduct numerical dispersion analysis of the aforementioned DGTD and FDTD schemes for Maxwell's equations in linear Lorentz media, and this can guide us in the optimal choice of numerical discretization parameters for more general dispersive and nonlinear models.

There has been abundant study on the dispersion analysis of DG methods. Most work was carried out for semi-discrete schemes, e.g., for scalar linear conservation laws [1,23,22,41,3], and for the second-order wave equation [4]. Dispersive behavior of fully discrete DGTD schemes is studied for the one-way wave equation [46,2] and two-way wave equations [10]. Particularly, in [40] the accuracy order of the dispersion and dissipation errors of nodal DG methods with Runge-Kutta time discretization for Maxwell's equations in free space are analyzed numerically. The stability and dispersion properties of a variety of FDTD schemes applied to Maxwell's equations in free space are also well known (see [43]). Additionally, various time domain finite element methods have been devised for the numerical approximation of Maxwell's equations in free space (see [34,32] and the references therein). There has been relatively less work on phase error analysis for dispersive dielectrics; see [43,37,38,8] for finite difference methods and [5] for finite element methods.

To the best of our knowledge, the present work is the first in the literature to conduct dispersion analysis of fully discrete DGTD methods for Maxwell's equations in Lorentz dispersive media and providing comparisons of the numerical dispersion errors with those of fully discrete FDTD methods. By rigorous quantification of the numerical dispersion error for such dispersive Maxwell systems, we make comparisons of the DGTD and FDTD methods taking into account the model

parameters, spatial and temporal accuracy and mesh sizes in the design of the schemes. Given the popularity of both DGTD and FDTD methods in science and engineering, such a comparison of errors between the two schemes will provide practitioners of these methods with guidelines on their proper implementation.

Our dispersion analysis indicates that there is a complicated dependence of dispersion errors on the model parameters, orders of spatial and temporal discretizations, CFL conditions as well as mesh discretization parameters. We compute and plot a variety of different measures of numerical dispersion errors as functions of the quantity $\frac{\omega}{\omega_1}$, where ω_1 is the resonance frequency of the Lorentz material, and ω is an angular frequency. These measures include normalized phase and group velocities, attenuation constants and an energy velocity [17]. The parameter range of the quantity $\frac{\omega}{\omega_1}$ separates the response of the material into distinct bands. We find that some counterintuitive results can occur for high-loss materials where a low order scheme can have smaller numerical dispersion error than a higher order scheme. Since this situation does not occur in non-dispersive dielectrics, our results demonstrates the need to analyze and study the numerical dispersion relation for the Lorentz media beyond those for the case of free space that commonly appear in the literature. We have made quantitative comparisons of the high order FD and DG schemes based on the metrics discussed above. We also identify the differences in numerical dispersion due to the temporal integrator used. In particular, our results clearly identify the limitation of the second order temporal accuracy of our time discretizations, by identifying distinct bands in the frequency parameter ranges where the high order spatial accuracy of either the DG or FD schemes is unable to alleviate the error in numerical dispersion due to time discretization.

The rest of the paper is organized as follows. In Section 2 we introduce Maxwell's equations in a one spatial dimensional Lorentz dispersive material. In Sections 3 and 4, we present and analyze the dispersion relations and the relative phase errors for the PDE model, and two semi-discrete in time finite difference numerical schemes, respectively. In Sections 5 and 6 numerical dispersion errors in semi-discrete in space staggered FD methods, and fully space-time discrete FDTD methods, respectively, are considered, while numerical dispersion errors in semi-discrete in space DG methods and fully discrete DGTD methods are studied in Sections 7, and 8, respectively. In Section 9, we define four quantities that provide different measures of numerical dispersion error and compare these for the FDTD and DGTD methods. Interpretations and conclusions of our results are made in Section 10.

2. Maxwell's equations in a linear Lorentz dielectric

We begin by introducing Maxwell's equations in a non-magnetic, non-conductive medium $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, from time 0 to T, containing no free charges, that govern the dynamic evolution of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{H}$ in the form

$$\partial_t \mathbf{B} + \nabla \times \mathbf{E} = 0, \text{ in } (0, T] \times \Omega, \tag{2.1a}$$

$$\partial_t \mathbf{D} - \nabla \times \mathbf{H} = 0, \text{ in } (0, T] \times \Omega, \tag{2.1b}$$

$$\nabla \cdot \mathbf{B} = 0, \ \nabla \cdot \mathbf{D} = 0, \text{ in } (0, T] \times \Omega, \tag{2.1c}$$

along with initial data that satisfies the Gauss laws (2.1c), and appropriate boundary data. System (2.1) has to be completed by constitutive laws on $[0, T] \times \Omega$. The electric flux density $\mathbf{D}$, and the magnetic induction $\mathbf{B}$, are related to the electric field and magnetic field, respectively, via the constitutive laws

$$\mathbf{D} = \epsilon_0(\epsilon_\infty \mathbf{E} + \mathbf{P}), \quad \mathbf{B} = \mu_0 \mathbf{H}. \tag{2.2}$$

The parameter ϵ_0 is the electric permittivity of free space, while μ_0 is the magnetic permeability of free space. The term $\epsilon_0\epsilon_\infty \mathbf{E}$ captures the linear instantaneous response of the medium to the EM fields, with ϵ_∞ defined as the relative electric permittivity in the limit of infinite frequencies. The macroscopic *(electric) retarded polarization* $\mathbf{P}$ is modeled as a single pole resonance Lorentz dispersion mechanism, in which the time dependent evolution of the polarization follows the second order ODE [17,43]

$$\frac{\partial^2 \mathbf{P}}{\partial t^2} + 2\gamma \frac{\partial \mathbf{P}}{\partial t} + \omega_1^2 \mathbf{P} = \omega_p^2 \mathbf{E}. \tag{2.3}$$

In the ODE (2.3), ω_1 and ω_p are the resonance and plasma frequencies of the medium, respectively, and γ is a damping constant. The plasma frequency is related to the resonance frequency via the relation $\omega_p^2 = (\epsilon_s - \epsilon_\infty)\omega_1^2 := \epsilon_d \omega_1^2$. Here ϵ_s is defined as the relative permittivity at zero frequency, and ϵ_d measures the strength of the electric field coupling to the linear Lorentz dispersion model. We note that the limit $\epsilon_d \to 0$, or $\epsilon_s \to \epsilon_\infty$ corresponds to a linear dispersionless dielectric.

In this paper, we focus on a one dimensional Maxwell model on $\Omega = \mathbb{R}$ that is obtained from (2.1), (2.2) and (2.3) by assuming an isotropic and homogeneous material in which electromagnetic plane waves are linearly polarized and propagate in the x direction. Thus, the electric field is represented by one scalar component $E := E_z$, while the magnetic field is represented by the one component $H := H_y$. All the other variables are similarly represented by single scalar components.

We convert the second order ODE (2.3) for the linear retarded polarization P to first order form by introducing the linear polarization current density J,

$$\frac{\partial P}{\partial t} = J, \quad \frac{\partial J}{\partial t} = -2\gamma J - \omega_1^2 P + \omega_p^2 E. \tag{2.4}$$

We consider a rescaled formulation of the resulting one spatial dimensional Maxwell-Lorentz system with the following scaling: let the reference time scale be t_0, and reference space scale be x_0 with $x_0 = ct_0$ and $c = 1/\sqrt{\mu_0\epsilon_0}$. Henceforth, the rescaled fields and constants are defined based on a reference electric field E_0 as follows,

$$(H/E_0)\sqrt{\mu_0/\epsilon_0} \to H, \quad D/(\epsilon_0 E_0) \to D, \quad P/E_0 \to P, \quad (J/E_0)t_0 \to J, \quad E/E_0 \to E,$$
$$\omega_1 t_0 \to \omega_1, \quad \omega_p t_0 \to \omega_p, \quad \gamma t_0 \to \gamma,$$

where for simplicity, we have used the same notation to denote the scaled and original variables. In summary, we arrive at the following dimensionless Maxwell's equations with linear Lorentz dispersion in one dimension:

$$\frac{\partial H}{\partial t} = \frac{\partial E}{\partial x}, \tag{2.5a}$$
$$\frac{\partial D}{\partial t} = \frac{\partial H}{\partial x}, \tag{2.5b}$$
$$\frac{\partial P}{\partial t} = J, \tag{2.5c}$$
$$\frac{\partial J}{\partial t} = -2\gamma J - \omega_1^2 P + \omega_p^2 E, \tag{2.5d}$$
$$D = \epsilon_\infty E + P. \tag{2.5e}$$

3. Dispersion relations

The Maxwell-Lorentz system (2.5) is a linear dispersive system, i.e. it admits plane wave solutions of the form $e^{i(kx-\omega t)}$ for all its unknown field variables, with the property that the speed of propagation of these waves is not independent of the wave number k or the angular frequency ω [45]. In this section, we derive the dispersion relation of (2.5) and highlight its main properties. We assume the space-time harmonic variation

$$X(x,t) \equiv X_0 e^{i(kx-\omega t)}, \tag{3.1}$$

of all field components $X \in \{H, E, P, J\}$. Substituting (3.1) in (2.5) yields the system

$$\omega H_0 + kE_0 = 0, \tag{3.2a}$$
$$kH_0 + \epsilon_\infty \omega E_0 + \omega P_0 = 0, \tag{3.2b}$$
$$i\omega P_0 + J_0 = 0, \tag{3.2c}$$
$$\omega_p^2 E_0 - \omega_1^2 P_0 + (i\omega - 2\gamma)\, J_0 = 0. \tag{3.2d}$$

Define the vector $\mathbf{U} = [H_0, E_0, P_0, J_0]^T$ containing all amplitudes of the field solution, then (3.2) can be rewritten as a linear system, given by

$$\mathcal{A}\mathbf{U} = \mathbf{0}, \quad \text{with} \quad \mathcal{A} = \begin{pmatrix} \omega & k & 0 & 0 \\ k & \epsilon_\infty \omega & \omega & 0 \\ 0 & 0 & i\omega & 1 \\ 0 & \omega_p^2 & -\omega_1^2 & i\omega - 2\gamma \end{pmatrix}. \tag{3.3}$$

By solving $\det(\mathcal{A}) = 0$, we obtain the exact dispersion relation for (2.5) as

$$k = \pm k^{\text{ex}}, \qquad \text{with} \quad k^{\text{ex}} = \omega\sqrt{\epsilon(\widehat{\omega};\mathbf{p})}, \quad \text{and} \quad \epsilon(\widehat{\omega};\mathbf{p}) = \epsilon_\infty\left(1 - \frac{\epsilon_d/\epsilon_\infty}{\widehat{\omega}^2 + 2i\widehat{\gamma}\widehat{\omega} - 1}\right). \tag{3.4}$$

Here, $\epsilon(\widehat{\omega};\mathbf{p})$ is the permittivity of the medium dependent on the "relative" frequency $\widehat{\omega} = \omega/\omega_1$ and the parameter set $\mathbf{p} = [\epsilon_s, \epsilon_\infty, \widehat{\gamma}]$, with $\widehat{\gamma} = \gamma/\omega_1$. The permittivity is clearly frequency dependent and displays the dispersive nature of the system. A major goal in the design and construction of numerical methods for linear dispersive PDEs is to devise methods that accurately capture the medium's complex permittivity [43]. We will assume that $\epsilon_s > 0, \epsilon_\infty > 0$ and $\epsilon_d = \epsilon_s - \epsilon_\infty > 0$. These assumptions are based on physical considerations [43]. In the dispersion analysis, we assume ω is a real number,

Table 3.1
Notations used and the place of their first appearance. The symbol $*$ in the superscript can be LF (for the leap-frog temporal scheme) or TP (for the trapezoidal temporal scheme).

	$\widehat{\omega}$	$\widehat{\gamma}$	W	W_1	K	$\epsilon(\widehat{\omega};\mathbf{p})$	$\delta(\widehat{\omega};\mathbf{p})$	$\Psi^*(\widehat{\omega})$
Def	ω/ω_1	γ/ω_1	$\omega\Delta t$	$\omega_1\Delta t$	$k^{\text{ex}}h$	$\epsilon_\infty - \dfrac{\epsilon_d}{\widehat{\omega}^2+2i\widehat{\gamma}\,\widehat{\omega}-1}$	$\dfrac{\epsilon_d\,\widehat{\omega}\,(\widehat{\omega}+i\widehat{\gamma})}{(\widehat{\omega}^2+2i\widehat{\gamma}\,\widehat{\omega}-1)^2}$	$\left\lvert\dfrac{k^{\text{ex}}(\widehat{\omega})-k^*(\widehat{\omega})}{k^{\text{ex}}(\widehat{\omega})}\right\rvert$
Eqn	(3.4)	(3.4)	(4.5)	(4.5)	(5.10)	(3.4)	(4.9)	(4.10), (4.15)

	$\Psi_{\text{FD},2M}(\widehat{\omega})$	$\Psi^*_{\text{FD},2M}(\widehat{\omega})$	$\Psi_{\text{DG},p}(\widehat{\omega})$	$\Psi^*_{\text{DG},p}(\widehat{\omega})$
Def	$\left\lvert\dfrac{k^{\text{ex}}(\widehat{\omega})-k_{\text{FD},2M}(\widehat{\omega})}{k^{\text{ex}}(\widehat{\omega})}\right\rvert$	$\left\lvert\dfrac{k^{\text{ex}}(\widehat{\omega})-k^*_{\text{FD},2M}(\widehat{\omega})}{k^{\text{ex}}(\widehat{\omega})}\right\rvert$	$\left\lvert\dfrac{k^{\text{ex}}(\widehat{\omega})-k_{\text{DG},p}(\widehat{\omega})}{k^{\text{ex}}(\widehat{\omega})}\right\rvert$	$\left\lvert\dfrac{k^{\text{ex}}(\widehat{\omega})-k^*_{\text{DG},p}(\widehat{\omega})}{k^{\text{ex}}(\widehat{\omega})}\right\rvert$
Eqn	(5.22)	Fig. 6.3	Fig. 7.1	Fig. 8.1

and restrict $\omega \geq 0$ in this work. Note that $\epsilon(\widehat{\omega};\mathbf{p})$ and k can be complex, depending on the values that certain parameters assume.

For lossless materials (i.e. $\widehat{\gamma}=0$), the *medium absorption band* is defined by $\widehat{\omega}\in[1,\sqrt{\epsilon_s/\epsilon_\infty}]$, in which $\epsilon(\widehat{\omega};\mathbf{p})\leq 0$ and k^{ex} is an imaginary number or zero. Outside the medium absorption band, i.e. for other $\widehat{\omega}$ values, we have $\epsilon(\widehat{\omega};\mathbf{p})>0$ and k^{ex} is a real number. Moreover, it is easy to check $|k^{\text{ex}}|\to\infty$ as $\widehat{\omega}$ approaches 1 (the *resonance frequency*, which is also the lower bound of the medium absorption band) and $k^{\text{ex}}=0$ at the upper bound $\widehat{\omega}=\sqrt{\epsilon_s/\epsilon_\infty}$. In this paper, we are mainly interested in low-loss materials, i.e. $\widehat{\gamma}>0$ with $\widehat{\gamma}\ll 1$. In this case, the dispersion relation retains similar properties, which means $|k^{\text{ex}}|$ is a large number around $\widehat{\omega}=1$ and a small number near $\widehat{\omega}=\sqrt{\epsilon_s/\epsilon_\infty}$. This behavior of the exact dispersion relation has implications for the numerical dispersion errors, as illustrated in later sections.

Remark 3.1. In the literature, dispersion relations can be presented in two ways; 1) representing the continuous or discrete angular frequency $\omega\in\mathbb{C}$ as a function of the exact and continuous wave number $k\in\mathbb{R}$ (and also of the model parameters and possible mesh parameters); 2) representing the continuous or discrete wave number $k\in\mathbb{C}$ as a function of the exact and continuous angular frequency $\omega\in\mathbb{R}$ [43]. In the first approach, we will obtain a fourth order polynomial for ω as a function of k and other parameters. We provide some insight into approach 1 in Appendix A, for the semi-discrete in space FDTD discretizations in which the effect of high order FDTD spatial approximations on the dispersion relation is clearly evident in terms of the *symbol* of the spatial discretization operators. In this paper, we mainly use the second approach since in this approach we are able to explicitly identify the effects of discretization on the permittivity of the Maxwell-Lorentz model (2.5).

Before we proceed, for convenience of the readers, we gather some notations frequently used in the paper, together with the place of their first appearances in Table 3.1.

4. Second order accurate temporal discretizations

This section concerns the dispersion analysis of the semi-discrete in time schemes. Continuing from our previous work [6, 7], we consider two types of commonly used second-order time schemes for the linear system (2.5), both implicit in the ODE parts. Let $\Delta t>0$ be a temporal mesh step. Suppose $u^n(x)$ is the solution at time $t^n=n\Delta t, n\in\mathbb{N}$, with $u=H,E,D,P,J$. Then, we compute $u^{n+1}(x)$ at time $t^{n+1}=t^n+\Delta t$ by the following methods. The first scheme uses a staggered leap-frog discretization in time for the PDE part, with the magnetic field H staggered in time from the rest of the field components. The scheme is given by:

$$\frac{H^{n+1/2}-H^n}{\Delta t/2}=\frac{\partial E^n}{\partial x}, \tag{4.1a}$$

$$\frac{D^{n+1}-D^n}{\Delta t}=\frac{\partial H^{n+1/2}}{\partial x}, \tag{4.1b}$$

$$\frac{P^{n+1}-P^n}{\Delta t}=\frac{1}{2}\left(J^n+J^{n+1}\right), \tag{4.1c}$$

$$\frac{J^{n+1}-J^n}{\Delta t}=-\gamma\left(J^n+J^{n+1}\right)-\frac{\omega_1^2}{2}\left(P^n+P^{n+1}\right)+\frac{\omega_p^2}{2}\left(E^n+E^{n+1}\right), \tag{4.1d}$$

$$D^{n+1}=\epsilon_\infty E^{n+1}+P^{n+1}, \tag{4.1e}$$

$$\frac{H^{n+1}-H^{n+1/2}}{\Delta t/2}=\frac{\partial E^{n+1}}{\partial x}. \tag{4.1f}$$

The second scheme, which is a fully implicit scheme based on the trapezoidal rule, is given as follows:

$$\frac{H^{n+1}-H^n}{\Delta t}=\frac{1}{2}\left(\frac{\partial E^{n+1}}{\partial x}+\frac{\partial E^n}{\partial x}\right), \tag{4.2a}$$

$$\frac{D^{n+1}-D^n}{\Delta t}=\frac{1}{2}\left(\frac{\partial H^{n+1}}{\partial x}+\frac{\partial H^n}{\partial x}\right), \tag{4.2b}$$

$$\frac{P^{n+1}-P^n}{\Delta t}=\frac{1}{2}\left(J^n+J^{n+1}\right), \tag{4.2c}$$

$$\frac{J^{n+1}-J^n}{\Delta t}=-\gamma\left(J^n+J^{n+1}\right)-\frac{\omega_1^2}{2}\left(P^n+P^{n+1}\right)+\frac{\omega_p^2}{2}\left(E^n+E^{n+1}\right), \tag{4.2d}$$

$$D^{n+1}=\epsilon_\infty E^{n+1}+P^{n+1}. \tag{4.2e}$$

Similar to the continuous case, we can perform dispersion analysis on the semi-discrete schemes (4.1) and (4.2) by assuming the time discrete plane wave solution as

$$X^n(x)\equiv X_0 e^{i(k^* x-\omega t_n)}, \tag{4.3}$$

where $*$ can be LF (with respect to the leap-frog scheme (4.1)) or TP (with respect to the trapezoidal scheme (4.2)). Define $\mathbf{U}=[H_0, E_0, P_0, J_0]^T$ as the vector containing all amplitudes of the field solutions. Substituting (4.3) in the schemes (4.1) or (4.2), we obtain linear systems for each case in the form

$$\mathcal{A}^*\mathbf{U}=\mathbf{0}. \tag{4.4}$$

The semi-discrete numerical dispersion relation can be then obtained from $\det(\mathcal{A}^*)=0$.

For the leap-frog scheme (4.1), we have

$$\mathcal{A}^{\mathrm{LF}}=\begin{pmatrix}\sin\left(\frac{W}{2}\right) & \frac{\Delta t}{2}k^{\mathrm{LF}} & 0 & 0\\ \frac{\Delta t}{2}k^{\mathrm{LF}} & \epsilon_\infty\sin\left(\frac{W}{2}\right) & \sin\left(\frac{W}{2}\right) & 0\\ 0 & 0 & i\sin\left(\frac{W}{2}\right) & \frac{\Delta t}{2}\cos\left(\frac{W}{2}\right)\\ 0 & \frac{\Delta t}{2}\omega_p^2\cos\left(\frac{W}{2}\right) & -\frac{\Delta t}{2}\omega_1^2\cos\left(\frac{W}{2}\right) & i\sin\left(\frac{W}{2}\right)-\gamma\Delta t\cos\left(\frac{W}{2}\right)\end{pmatrix}, \tag{4.5}$$

where $W:=\omega\Delta t=\widehat{\omega}W_1$, with $W_1:=\omega_1\Delta t$. This yields the dispersion relation

$$k^{\mathrm{LF}}=\pm\omega\sqrt{\epsilon(\widehat{\omega}^{\mathrm{LF}};\mathbf{p}^{\mathrm{LF}})},\quad\text{with}\quad\epsilon(\widehat{\omega}^{\mathrm{LF}};\mathbf{p}^{\mathrm{LF}})=\epsilon_\infty^{\mathrm{LF}}\left(1-\frac{\epsilon_d^{\mathrm{LF}}/\epsilon_\infty^{\mathrm{LF}}}{\left(\widehat{\omega}^{\mathrm{LF}}\right)^2+2i\widehat{\gamma}^{\mathrm{LF}}\widehat{\omega}^{\mathrm{LF}}-1}\right), \tag{4.6}$$

where $s_\omega:=\frac{\sin(\frac{W}{2})}{\frac{W}{2}}$ and $r_\omega:=\frac{\tan(\frac{W}{2})}{\frac{W}{2}}$ as in [37], and $\widehat{\omega}^{\mathrm{LF}}=\widehat{\omega}r_\omega$, $\mathbf{p}^{\mathrm{LF}}=[\epsilon_s^{\mathrm{LF}},\epsilon_\infty^{\mathrm{LF}},\widehat{\gamma}^{\mathrm{LF}}]$, with components given by the identities

$$\epsilon_s^{\mathrm{LF}}=\epsilon_s s_\omega^2,\quad\epsilon_\infty^{\mathrm{LF}}=\epsilon_\infty s_\omega^2,\quad\widehat{\gamma}^{\mathrm{LF}}=\widehat{\gamma}. \tag{4.7}$$

In this form, we can clearly identify how the leap-frog time discretization misrepresents the permittivity by misrepresenting the parameters of the model. These misrepresentations are solely due to the discretizations of the ODEs by the leap-frog time integrator. The misrepresentations depend on the value of the (exact) angular frequency that is chosen, and in particular as $\frac{W}{2}$ approaches zero, the discrete parameters approach the continuous ones. Thus, a guideline for practitioners using this time integrator to control these misrepresentations, is to choose Δt so that $\cos\left(\frac{W}{2}\right)\approx 1$ across the range of frequencies present in the short pulse that propagates in the medium [37].

To further analyze the dispersion error, we consider the regime when $W\ll 1$, and obtain the Taylor expansion of (4.6) with respect to W as

$$k^{\mathrm{LF}}=\pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}(W^4)\right), \tag{4.8}$$

where

$$\delta(\widehat{\omega};\mathbf{p})=\frac{\epsilon_d\,\widehat{\omega}\,(\widehat{\omega}+i\,\widehat{\gamma})}{\left(\widehat{\omega}^2+2i\,\widehat{\gamma}\,\widehat{\omega}-1\right)^2}. \tag{4.9}$$

We define the *relative phase error* for the LF scheme to be the ratio

$$\Psi^{\mathrm{LF}}(\widehat{\omega}) := \left| \frac{k^{\mathrm{LF}}(\widehat{\omega}) - k^{\mathrm{ex}}(\widehat{\omega})}{k^{\mathrm{ex}}(\widehat{\omega})} \right| = \left| \frac{\sqrt{\epsilon(\widehat{\omega}^{\mathrm{LF}}; \mathbf{p}^{\mathrm{LF}})} - \sqrt{\epsilon(\widehat{\omega}; \mathbf{p})}}{\sqrt{\epsilon(\widehat{\omega}; \mathbf{p})}} \right|. \tag{4.10}$$

Here, we consider k^{LF} in (4.6) with plus sign in front. A similar definition will be used for all semi-discrete and fully discrete schemes that appear in this paper, and provides quantitative measurement of the numerical dispersion error. Equation (4.8) verifies a second order dispersion error in time of the leap-frog scheme in the small time step limit.

Similarly, for the trapezoidal method (4.2), we can obtain

$$\mathcal{A}^{\mathrm{TP}} = \begin{pmatrix} \sin\left(\frac{W}{2}\right) & \frac{\Delta t}{2} k^{\mathrm{TP}} \cos\left(\frac{W}{2}\right) & 0 & 0 \\ \frac{\Delta t}{2} k^{\mathrm{TP}} \cos\left(\frac{W}{2}\right) & \epsilon_\infty \sin\left(\frac{W}{2}\right) & \sin\left(\frac{W}{2}\right) & 0 \\ 0 & 0 & i \sin\left(\frac{W}{2}\right) & \frac{\Delta t}{2} \cos\left(\frac{W}{2}\right) \\ 0 & \frac{\Delta t}{2} \omega_p^2 \cos\left(\frac{W}{2}\right) & -\frac{\Delta t}{2} \omega_1^2 \cos\left(\frac{W}{2}\right) & i \sin\left(\frac{W}{2}\right) - \gamma \Delta t \cos\left(\frac{W}{2}\right) \end{pmatrix}. \tag{4.11}$$

This leads to the dispersion relation

$$k^{\mathrm{TP}} = \pm\omega \sqrt{\epsilon(\widehat{\omega}^{\mathrm{TP}}; \mathbf{p}^{\mathrm{TP}})} = \frac{s_\omega}{r_\omega} k^{\mathrm{LF}}, \quad \text{with} \quad \epsilon(\widehat{\omega}^{\mathrm{TP}}; \mathbf{p}^{\mathrm{TP}}) = \epsilon_\infty^{\mathrm{TP}} \left(1 - \frac{\epsilon_d^{\mathrm{TP}} / \epsilon_\infty^{\mathrm{TP}}}{\left(\widehat{\omega}^{\mathrm{TP}}\right)^2 + 2i\,\widehat{\gamma}^{\mathrm{TP}}\,\widehat{\omega}^{\mathrm{TP}} - 1} \right), \tag{4.12}$$

where $\widehat{\omega}^{\mathrm{TP}} = \widehat{\omega} r_\omega$, $\mathbf{p}^{\mathrm{TP}} = [\epsilon_s^{\mathrm{TP}}, \epsilon_\infty^{\mathrm{TP}}, \widehat{\gamma}^{\mathrm{TP}}]$, with components given as

$$\epsilon_s^{\mathrm{TP}} = \epsilon_s r_\omega^2, \quad \epsilon_\infty^{\mathrm{TP}} = \epsilon_\infty r_\omega^2, \quad \widehat{\gamma}^{\mathrm{TP}} = \widehat{\gamma}. \tag{4.13}$$

Again, we can clearly identify how the trapezoidal time discretization misrepresents the permittivity. In particular, this method misrepresents the dissipation and medium resonance in the same manner as the leap-frog method. However, the relative permittivities ϵ_∞ and ϵ_s are misrepresented in a different manner. Thus, the speeds of propagation of discrete plane waves are different in these two discretizations. In particular, the slow and fast speeds in the medium, corresponding to relative permittivities ϵ_s and ϵ_∞, respectively, are different.

In the small time step limit, for $W \ll 1$, we have

$$k^{\mathrm{TP}} = \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12} \left(\frac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} + 1 \right) W^2 + \mathcal{O}(W^4) \right), \tag{4.14}$$

which indicates second order accuracy in time for the *relative phase error* for the trapezoidal scheme defined, in a similar manner to the leap-frog scheme, as

$$\Psi^{\mathrm{TP}}(\widehat{\omega}) := \left| \frac{k^{\mathrm{TP}}(\widehat{\omega}) - k^{\mathrm{ex}}(\widehat{\omega})}{k^{\mathrm{ex}}(\widehat{\omega})} \right| = \left| \frac{\sqrt{\epsilon(\widehat{\omega}^{\mathrm{TP}}; \mathbf{p}^{\mathrm{TP}})} - \sqrt{\epsilon(\widehat{\omega}; \mathbf{p})}}{\sqrt{\epsilon(\widehat{\omega}; \mathbf{p})}} \right|. \tag{4.15}$$

Finally, we make qualitative comparisons of the leap-frog and trapezoidal temporal discretizations. For low-loss materials, the conclusions can be implied from considering the case of $\widehat{\gamma} = 0$. For this case, for a given set of parameters $\mathbf{p}$, $\epsilon(\widehat{\omega}; \mathbf{p})$ and $\delta(\widehat{\omega}; \mathbf{p}) \geq 0$ are real numbers. When $\widehat{\omega} \to \sqrt{\epsilon_s/\epsilon_\infty}$, we have $\epsilon(\widehat{\omega}; \mathbf{p}) \to 0$. This means $\dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} \to \infty$, and thus the leading error term of both temporal schemes would be approaching ∞. On the other hand, when $\widehat{\omega} \to 1$, it is easy to check that $\dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} \to \infty$ as well. Hence, both time schemes will give large dispersion error at $\widehat{\omega} = 1$ and $\widehat{\omega} = \sqrt{\epsilon_s/\epsilon_\infty}$, which are the two endpoints of the medium absorption band. In addition, when $W \ll 1$, if $\widehat{\omega} \in (1, \sqrt{\epsilon_s/\epsilon_\infty})$, i.e. for values in the interior of the medium absorption band, we can prove that $\dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} < -1$, which leads to the relation $\left| \dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} - \dfrac{1}{2} \right| \geq \left| \dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} + 1 \right|$. This means the leap-frog scheme has a larger relative phase error than the trapezoidal scheme in the interior of the medium absorption band. For other values of $\widehat{\omega}$ outside the medium absorption band we obtain $\epsilon(\widehat{\omega}; \mathbf{p}) > 0$ and $\left| \dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} - \dfrac{1}{2} \right| \leq \left| \dfrac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} + 1 \right|$. Hence, the leap-frog scheme would give a small relative phase error outside the absorption band. For low-loss Lorentz medium, i.e., when $\widehat{\gamma} \ll 1$, we believe that these conclusions are still valid with a slight change in two peak positions (see Fig. 4.1).

Now we choose the following set of parameters, which are the same as in [17], representing a low-loss Lorentz medium:

$$\epsilon_s = 5.25, \quad \epsilon_\infty = 2.25, \quad \widehat{\gamma} = 0.01. \tag{4.16}$$

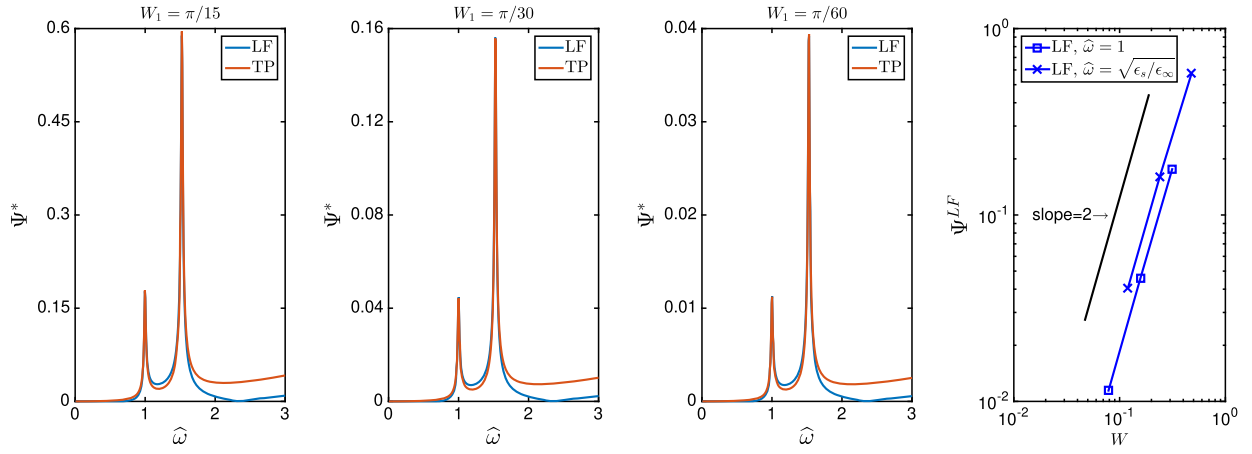

Fig. 4.1. The relative phase error of leap-frog (LF) and trapezoidal (TP) time discretizations. In the first three plots we fix $W_1 \in \{\pi/15, \pi/30, \pi/60\}$, respectively, while we vary $\widehat{\omega} \in [0, 3]$. In the fourth plot, we fix $\widehat{\omega} = 1$ or $\widehat{\omega} = \sqrt{\epsilon_s/\epsilon_\infty}$ and consider three different values of W corresponding to $W_1 \in \{\pi/15, \pi/30, \pi/60\}$, with leap-frog time discretization.

Taking W_1 as $\{\pi/15, \pi/30, \pi/60\}$, the relative phase errors are plotted against $\widehat{\omega} \in [0, 3]$ in Fig. 4.1. We can observe that the phase errors always have two peaks around $\widehat{\omega} = 1$ and $\widehat{\omega} = \sqrt{\epsilon_s/\epsilon_\infty} \approx 1.527$. The relative phase errors in the two temporal schemes are basically the same for small $\widehat{\omega}$. As $\widehat{\omega}$ is increased towards 1, we see that the error in the leap-frog scheme is slightly smaller than that in the trapezoidal scheme. When $\widehat{\omega}$ is between 1 and 1.527, the leap-frog scheme presents larger error than the trapezoidal method. Beyond 1.527, the trapezoidal scheme generates larger error than the leap-frog scheme. There is no obvious difference between the dispersion errors of two time discretizations at the peaks. Therefore, in the last graph of Fig. 4.1, we only plot the errors of the leap-frog time schemes at the peaks. We verify the second order accuracy of the method when the mesh size is varied. These observations are consistent with our analysis.

5. Spatial discretization: high order staggered finite difference methods

In this section, we consider semi-discrete in space staggered finite difference schemes for (2.5). The spatial discretizations that we consider here for system (2.5) combined with a nonlinear instantaneous Kerr response and a Raman retarded nonlinear response have been recently developed in [7]. The electric and magnetic fields are staggered in space and the discrete spatial operators have arbitrary even order, $2M$, $M \in \mathbb{N}$, accuracy in space. Below, we describe the semi-discrete spatial schemes, denoted as the FD2M scheme, and then we obtain and discuss dispersion relations of these schemes.

As in [7], we define two staggered grids on $\mathbb{R}$ with spatial step size h, the primal grid G_p, and the dual grid G_d, defined respectively, as

$$G_p = \{jh \mid j \in \mathbb{Z}\}, \quad \text{and} \quad G_d = \{(j + \frac{1}{2})h \mid j \in \mathbb{Z}\}. \tag{5.1}$$

The discrete magnetic field will be approximated at spatial nodes on the dual grid. These approximations are denoted by $H_{j+1/2}$, termed as *degrees of freedom* (DoF) of H. All the other discrete fields will have their DoF at spatial nodes on the primal grid. For a continuous field variable V, V_h denotes its corresponding *grid function*, defined as the set of all DoF on its respective grid. The semi-discrete scheme is given as follows:

$$\frac{\partial H_{j+1/2}}{\partial t} = \left(\mathcal{D}_h^{(2M)} E_h\right)_{j+\frac{1}{2}}, \tag{5.2a}$$

$$\frac{\partial D_j}{\partial t} = \left(\widetilde{\mathcal{D}}_h^{(2M)} H_h\right)_j, \tag{5.2b}$$

$$\frac{\partial P_j}{\partial t} = J_j, \tag{5.2c}$$

$$\frac{\partial J_j}{\partial t} = -2\gamma J_j - \omega_1^2 P_j + \omega_p^2 E_j, \tag{5.2d}$$

$$D_j = \epsilon_\infty E_j + P_j, \tag{5.2e}$$

where $\mathcal{D}_h^{(2M)}$ and $\widetilde{\mathcal{D}}_h^{(2M)}$ are the $2M$-th order finite difference approximations (with $M \in \mathbb{N}$) of the spatial differential operator ∂_x, on the primal and dual grids, respectively. These approximations are defined as

$$\left(\mathcal{D}_h^{(2M)} E_h\right)_{j+1/2} = \frac{1}{h}\sum_{p=1}^{M} \frac{\lambda_{2p-1}^{2M}}{(2p-1)}\left(E_{j+p} - E_{j-p+1}\right), \tag{5.3a}$$

$$\left(\widetilde{\mathcal{D}}_h^{(2M)} H_h\right)_{j} = \frac{1}{h}\sum_{p=1}^{M} \frac{\lambda_{2p-1}^{2M}}{(2p-1)}\left(H_{j+p-\frac{1}{2}} - H_{j-p+\frac{1}{2}}\right), \tag{5.3b}$$

and λ_{2p-1}^{2M}, is given as [7]

$$\lambda_{2p-1}^{2M} = \frac{2(-1)^{p-1}[(2M-1)!!]^2}{(2M+2p-2)!!(2M-2p)!!(2p-1)}, \tag{5.4}$$

with the double factorial $n!!$ defined as

$$n!! = \begin{cases} n\cdot(n-2)\cdot(n-4)\ldots 5\cdot 3\cdot 1, & n>0,\ \text{odd} \\ n\cdot(n-2)\cdot(n-4)\ldots 6\cdot 4\cdot 2, & n>0,\ \text{even} \\ 1, & n=-1,0. \end{cases} \tag{5.5}$$

5.1. Semi-discrete in space dispersion analysis

In this section we analyze the spatial semi-discrete system (5.2), i.e., the FD2M scheme. We assume that the semi-discrete system (5.2) has plane wave solutions of the form

$$X_j(t) \equiv X_0 e^{i\left(k_{\mathrm{FD},2M}\, jh - \omega t\right)}, \tag{5.6}$$

where $k_{\mathrm{FD},2M}$ represents the numerical wave number of the semi-discrete FD2M scheme. By substituting (5.6) in (5.2) we obtain the linear system

$$\mathcal{A}_{\mathrm{FD},2M}\mathbf{U}_{\mathrm{FD}} = \mathbf{0}, \tag{5.7}$$

where the vector $\mathbf{U}_{\mathrm{FD}} = [H_0, E_0, P_0, J_0]^T$, and the matrix $\mathcal{A}_{\mathrm{FD},2M}$ is given by

$$\mathcal{A}_{\mathrm{FD},2M} = \begin{pmatrix} \omega & \Lambda_{2M} & 0 & 0 \\ \Lambda_{2M} & \epsilon_\infty\omega & \omega & 0 \\ 0 & 0 & i\omega & 1 \\ 0 & \omega_p^2 & -\omega_1^2 & i\omega - 2\gamma \end{pmatrix} \quad \text{with} \quad \Lambda_{2M} = \frac{2}{h}\sum_{p=1}^{M} \frac{\lambda_{2p-1}^{2M}}{(2p-1)} \sin\left[\left(p-\frac{1}{2}\right) k_{\mathrm{FD},2M} h\right]. \tag{5.8}$$

The numerical dispersion relation of the FD2M method is obtained by solving the characteristic equation of matrix $\mathcal{A}_{\mathrm{FD},2M}$ and is given as

$$\Lambda_{2M} = \pm\omega\sqrt{\epsilon(\widehat{\omega};\mathbf{p})} = \pm\omega\sqrt{\epsilon_\infty\left(1 - \frac{\epsilon_d/\epsilon_\infty}{\widehat{\omega}^2 + 2i\,\widehat{\gamma}\,\widehat{\omega} - 1}\right)} = \pm k^{\mathrm{ex}}. \tag{5.9}$$

Using results from [7], we can rewrite this numerical dispersion relation as

$$\frac{1}{2}h\,\Lambda_{2M} = \sum_{p=1}^{M} \frac{[(2p-3)!!]^2}{(2p-1)!}\sin^{2p-1}\left(\frac{k_{\mathrm{FD},2M}h}{2}\right) = \pm\frac{1}{2}K, \tag{5.10}$$

with $K := k^{\mathrm{ex}}h$. In general, for any $M \geq 1$, we will have $(4M-2)$ discrete wave numbers $k_{\mathrm{FD},2M}$ that satisfy (5.10). In particular, when $M=1$, i.e. for the FD2 scheme, the numerical dispersion relation (5.10) is

$$\sin\left(\frac{k_{\mathrm{FD},2}h}{2}\right) = \pm\frac{1}{2}K. \tag{5.11}$$

Thus, considering $K \ll 1$, and performing a Taylor expansion of (5.11) we obtain

$$k_{\mathrm{FD},2} = \pm k^{\mathrm{ex}}\left(1 + \frac{1}{24}K^2 + \frac{3}{640}K^4 + \mathcal{O}(K^6)\right), \tag{5.12}$$

which indicates that the numerical dispersion error of the FD2 scheme is second order accurate in space.

For the case $M=2$, i.e. for the FD4 scheme, the numerical dispersion relation (5.10) becomes

$$\frac{1}{6}\sin^3\left(\frac{k_{\mathrm{FD},4}h}{2}\right) + \sin\left(\frac{k_{\mathrm{FD},4}h}{2}\right) = \pm\frac{1}{2}K. \tag{5.13}$$

The Taylor expansions of all roots in equation (5.13) are given by

$$k_{\mathrm{FD\,phys},4} = \pm k^{\mathrm{ex}}\left(1 + \frac{3}{640}K^4 - \frac{1}{3584}K^6 + \mathcal{O}(K^8)\right), \tag{5.14a}$$

$$k_{\mathrm{FD\,spur1},4} = \pm k^{\mathrm{ex}}\left(i\frac{\operatorname{arcsinh}(2\sqrt{42})}{K} - \frac{1}{2\sqrt{7}} + i\frac{9}{1568}\sqrt{42}K + \mathcal{O}(K^2)\right), \tag{5.14b}$$

$$k_{\mathrm{FD\,spur2},4} = \pm k^{\mathrm{ex}}\left(-i\frac{\operatorname{arcsinh}(2\sqrt{42})}{K} - \frac{1}{2\sqrt{7}} - i\frac{9}{1568}\sqrt{42}K + \mathcal{O}(K^2)\right), \tag{5.14c}$$

where $k_{\mathrm{FD\,phys},4}$ and $k_{\mathrm{FD\,spur1},4}, k_{\mathrm{FD\,spur2},4}$ are wave numbers corresponding to the physical modes and spurious modes, respectively, of the FD4 scheme. The physical modes indicate a fourth order accurate numerical dispersion error, while the leading terms in the spurious modes of k are proportional to $\mathcal{O}(1/h)$, indicating an exponential increase or damping corresponding to the opposite sign in front.

The existence of spurious, parasitic or non-physical modes for a variety of problems and their discretizations has been extensively discussed in the literature, see for example e.g., [4,13,14,35,49]. In [35], the author analyzes spurious modes in finite element discretizations of the wave equation and shows that the spurious modes have a contribution to the numerical error that behaves in a reasonable manner, so that higher-order elements can be more accurate than lower-order elements. In [49], the author considers spurious modes in high order finite difference methods that can occur due to spectral or non-spectral pollution. Here the author shows the dependence of the spurious modes on boundary approximations and closures. In [13], the author analyzes dispersion error, in particular, errors in the discrete group velocities in terms of a numerical angular frequency, for a variety of high order finite difference schemes for the second order wave equation. Here the author identifies spurious modes (parasitic waves) in high order fully discrete approximations of the wave equation. For a symmetric fully fourth order method the author notes that the parasitic wave has a velocity that tends to infinity as the time step goes to zero, and remarks that such waves have an amplitude decreasing with the time step. We would like to note, that to the best of the authors' knowledge, the existence of spurious modes for high order FD discretizations for the Maxwell Lorentz model has not been analytically identified in the literature. Equations (5.14b) and (5.14c) provide explicit formulas for the spurious modes for discretizations of the Maxwell-Lorentz system that we have not found in the literature.

Below, we focus on the physical modes, and prove that for the FD scheme of order $2M$ (FD2M), the dispersion error is of $2M$-th order. Thus, the dispersion error is of the same order as the local truncation error for the finite difference schemes.

Theorem 5.1. *The physical modes of the dispersion relation (5.10), for the spatial semi-discrete finite difference method FD2M, result in the dispersion error identity*

$$k_{FD\,phys,2M} = \pm k^{ex}\left(1 + \varsigma_{2M}\right), \tag{5.15}$$

for any $M \geq 1$, where

$$\varsigma_{2M} = \frac{[(2M-1)!!]^2}{2^{2M}(2M+1)!}K^{2M} + \mathcal{O}(K^{2M+2}). \tag{5.16}$$

In other words, the dispersion error of the FD2M scheme (5.2) is of order $2M$.

Proof. Here, we only consider $k_{\mathrm{FD\,phys},2M}$ with plus sign in front. Define $\varsigma_{2M} := \dfrac{k_{\mathrm{FD\,phys},2M} - k^{\mathrm{ex}}}{k^{\mathrm{ex}}}$. Then, substituting from (5.9) for k^{ex}, rearranging, and (using results from [7]) we obtain the identity

$$k^{\mathrm{ex}}\varsigma_{2M} = k_{\mathrm{FD\,phys},2M} - \frac{2}{h}\sum_{p=1}^{M}\frac{\lambda_{2p-1}^{2M}}{(2p-1)}\sin\left[\left(p-\frac{1}{2}\right)k_{\mathrm{FD\,phys},2M}h\right],$$

i.e., $k_{\mathrm{FD\,phys},2M}$ satisfies (5.15) for ς_{2M} as defined in (5.16). Next, we prove the identity (5.16). Because we only consider the physical modes here, it is reasonable to assume that $k_{\mathrm{FD\,phys},2M} = k^{\mathrm{ex}}\left(1 + O(K^{\tau})\right)$ for some $\tau > 0$. Hence, when $k_{\mathrm{FD\,phys},2M}h = K\left(1 + \mathcal{O}(K^{\tau})\right)$ is small enough, performing a Taylor expansion with respect to $k_{\mathrm{FD\,phys},2M}h$ we get

$$\begin{aligned} k^{\mathrm{ex}}\varsigma_{2M} &= k_{\mathrm{FD\,phys},2M} - \frac{2}{h}\sum_{p=1}^{M}\frac{\lambda_{2p-1}^{2M}}{(2p-1)}\sum_{\ell=0}^{\infty}\frac{(-1)^{\ell}}{(2\ell+1)!}\left[\frac{1}{2}(2p-1)\,k_{\mathrm{FD\,phys},2M}h\right]^{2\ell+1} \\ &= k_{\mathrm{FD\,phys},2M} - \frac{2}{h}\sum_{\ell=0}^{\infty}\sum_{p=1}^{M}\frac{\lambda_{2p-1}^{2M}}{(2p-1)}\frac{(-1)^{\ell}}{(2\ell+1)!}\left[\frac{1}{2}(2p-1)\,k_{\mathrm{FD\,phys},2M}h\right]^{2\ell+1} \end{aligned}$$

$$= k_{\text{FD phys},2M} - k_{\text{FD phys},2M} \sum_{\ell=0}^{\infty} \left[\sum_{p=1}^{M} \lambda_{2p-1}^{2M} (2p-1)^{2\ell} \right] \frac{(-1)^{\ell}}{2^{2\ell}(2\ell+1)!} \left(k_{\text{FD phys},2M} h \right)^{2\ell}.$$

Based on the derivation of λ_{2p-1}^{2M} as discussed in [7], we have the following identities

$$\sum_{p=1}^{M} \lambda_{2p-1}^{2M} = 1,$$

$$\sum_{p=1}^{M} \lambda_{2p-1}^{2M} (2p-1)^{2\ell} = 0, \quad \text{for } \ell = 1, 2, .., M-1,$$

$$\sum_{p=1}^{M} \lambda_{2p-1}^{2M} (2p-1)^{2M} = (-1)^{M+1} \left[(2M-1)!! \right]^2.$$

Therefore,

$$\begin{aligned} k^{\text{ex}} \varsigma_{2M} &= k_{\text{FD phys},2M} - k_{\text{FD phys},2M} \left[1 - \frac{1}{2^{2M}} \frac{\left[(2M-1)!! \right]^2}{(2M+1)!} \left(k_{\text{FD phys},2M} h \right)^{2M} + \mathcal{O}\left(\left(k_{\text{FD phys},2M} h \right)^{2M+2} \right) \right] \\ &= -k_{\text{FD phys},2M} \left[-\frac{1}{2^{2M}} \frac{\left[(2M-1)!! \right]^2}{(2M+1)!} (K)^{2M} + \mathcal{O}\left(K^{2M+\tau} + K^{2M+2} \right) \right] \\ &= -k^{\text{ex}} \left[-\frac{1}{2^{2M}} \frac{\left[(2M-1)!! \right]^2}{(2M+1)!} (K)^{2M} + \mathcal{O}\left(K^{2M+\tau} + K^{2M+2} \right) \right], \end{aligned} \tag{5.17}$$

which proves (5.16). Hence, with the assumption $k_{\text{FD phys},2M} = k^{\text{ex}} \left(1 + \mathcal{O}(K^{\tau}) \right)$ and $\tau > 0$, we can deduce that $k_{\text{FD phys},2M} = k^{\text{ex}} \left(1 + \mathcal{O}(K^{2M}) \right)$. □

From equation (5.10) and using results in [8], we can show that K is bounded by

$$K \leq 2 \sum_{p=1}^{M} \frac{\left[(2p-3)!! \right]^2}{(2p-1)!} \leq 2 \cdot \frac{\pi}{2} = \pi. \tag{5.18}$$

We define

$$\widetilde{K}_{FD,2M} = \frac{k_{\text{FD},2M} h}{2} \quad \text{and} \quad \widetilde{K} = \frac{1}{2} K,$$

so equation (5.10) in terms of $\widetilde{K}_{FD,2M}$ and $\widetilde{K}$ becomes

$$\sum_{p=1}^{M} \frac{\left[(2p-3)!! \right]^2}{(2p-1)!} \sin^{2p-1}\left(\widetilde{K}_{FD,2M} \right) = \widetilde{K}. \tag{5.19}$$

To understand the behavior of high order schemes (large M) as the wave number (parameter K) increases, we consider the following two cases.

Case 1: We first analyze the relative error between the exact and numerical wave numbers as a function of the order M of the scheme, for different values of the exact wave number. Using the numerical dispersion relation (5.19), we define the relative error in the numerical wave number with respect to the exact value of $\widetilde{K}$ as a function of M,

$$\text{Relative Error}_{\widetilde{K}}(M) := \left| \frac{\widetilde{K}_{FD,2M}(M, \widetilde{K})}{\widetilde{K}} - 1 \right|. \tag{5.20}$$

In Fig. 5.1, we numerically plot the relative error (5.20) for several fixed values of $\widetilde{K}$. One can observe that when $\widetilde{K}$ is within the bound given in (5.18), the relative error decreases as M increases. However, the decrease in the error becomes less significant as $\widetilde{K}$ increases. Thus, the greatest benefit of high order FD2M is seen in small values of $\widetilde{K}$.

Case 2: Next, we analyze the relative error between the exact and numerical wave numbers as a function of the exact wave number, for different orders M of the FD2M scheme.
We now define the relative error in the numerical wave number, $\widetilde{K}_{FD,2M}$ for fixed values of M in a similar way as

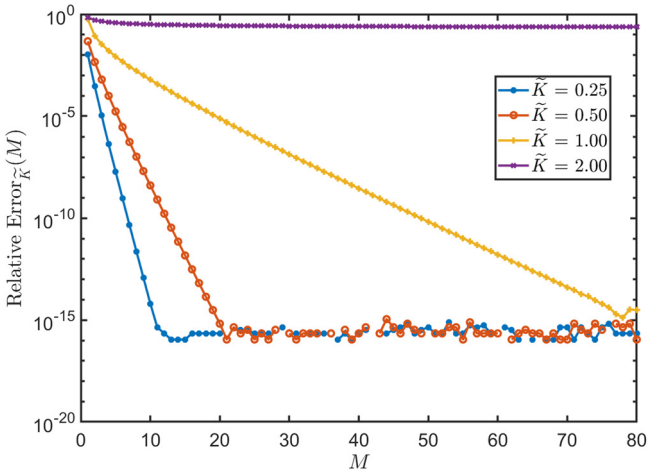

Fig. 5.1. The relative error, Relative Error$_{\widetilde{K}}(M)$, with respect to the order of scheme, M.

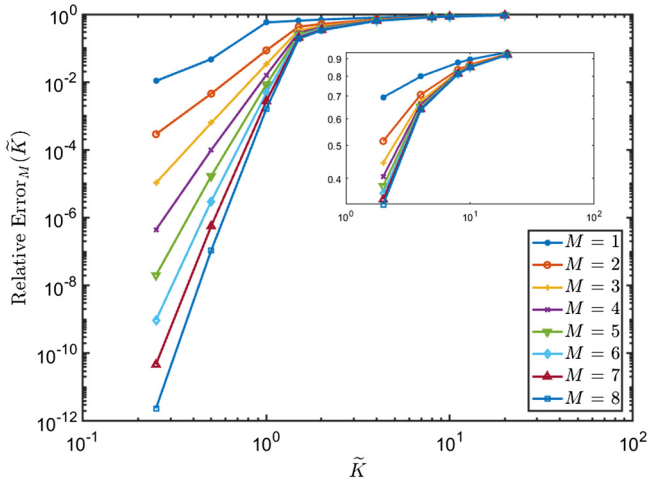

Fig. 5.2. The relative error, Relative Error$_M(\widetilde{K})$, with respect to the parameter $\widetilde{K}$.

$$\text{Relative Error}_M(\widetilde{K}) := \left| \frac{\widetilde{K}_{FD,2M}(M,\widetilde{K})}{\widetilde{K}} - 1 \right|. \tag{5.21}$$

In Fig. 5.2, we plot the error (5.21) as a function of the parameter $\widetilde{K}$. In this plot, we can see that the relative error decreases as M increases. However, again the decrease in the error is most significant for small $\widetilde{K}$. The figure also shows that the error is reasonable when $\widetilde{K} \le \pi/2$, yet it increases and approaches 1 when $\widetilde{K}$ further increases.

Next, we illustrate the relative phase errors of the FD2M scheme for (5.2), $M = 1, \ldots, 5$, with the parameter set $\mathbf{p}$ fixed at values given in (4.16). The numerical wave number $k_{\text{FD},2M}$ is obtained by solving (5.10) exactly or with the help of a Newton solver (we set the tolerance at 10^{-18}). Since

$$k^{\text{ex}}h = \widehat{\omega}\sqrt{\epsilon(\widehat{\omega};\mathbf{p})}\,\omega_1 h,$$

then $k_{\text{FD},2M}h$ depends on $\widehat{\omega}$, $\mathbf{p}$, $\omega_1 h$, and M, and so does the relative phase error

$$\Psi_{\text{FD},2M}(\widehat{\omega}) := \left| \frac{k_{\text{FD},2M}(\widehat{\omega}) - k^{\text{ex}}(\widehat{\omega})}{k^{\text{ex}}(\widehat{\omega})} \right| = \left| \frac{k_{\text{FD},2M}(\widehat{\omega})h - k^{\text{ex}}(\widehat{\omega})h}{k^{\text{ex}}(\widehat{\omega})h} \right|. \tag{5.22}$$

First, we fix $\omega_1 h = \pi/30$, and present the relative phase errors as functions of $\widehat{\omega} \in [0,3]$ in Fig. 5.3. Because the leading error term in the numerical wave number for the FD2M scheme is proportional to K^{2M}, we expect $2M$ order accuracy of the relative phase error with respect to K at a fixed angular frequency. We observe that all schemes have significantly larger error around $\widehat{\omega} = 1$, while the error fades out near $\widehat{\omega} = \sqrt{\epsilon_s/\epsilon_\infty}$, where K is close to zero. As expected from analysis, higher order spatial accuracy does result in reduced relative phase errors. We present the relative phase errors at $\widehat{\omega} = 1$

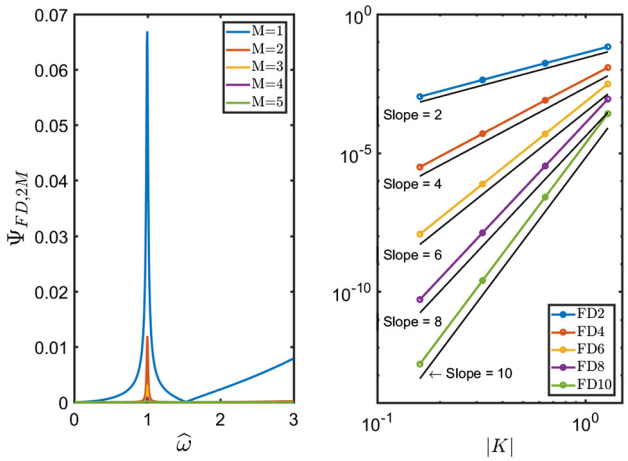

Fig. 5.3. The relative phase error of physical modes for the spatial discretization FD2M. Left: fix $\omega_1 h = \pi/30$ with $\widehat{\omega} \in [0, 3]$; right: fix $\widehat{\omega} = 1$ with different $\omega_1 h \in \{\pi/30, \pi/60, \pi/120, \pi/240\}$.

with $\omega_1 h = \pi/30$ in the left plot in Fig. 5.3, while in the right plot we depict the $2M$ order convergence of relative phase errors with respect to K for fixed $\widehat{\omega} = 1$. The slopes of phase errors in this plot are shown to be the same as those of reference lines with slope $2M$, indicating the $2M$th order of accuracy for each FD2M scheme, which agrees with the results in Theorem 5.1. We note the presence of just one peak in these plots as compared to the presence of two peaks in analogous plots of phase errors for temporal discretizations presented in Section 4.

6. Fully discrete FDTD methods

In this section, we consider the high order staggered spatial discretizations (5.2) combined with either the leap-frog scheme in time (4.1) or the trapezoidal scheme in time (4.2) presented in Section 4. These fully discrete methods are second order accurate in time and $2M$-th order accurate in space, thus we denote them as $(2, 2M)$ leap-frog FDTD schemes or $(2, 2M)$ trapezoidal FDTD methods. In particular, the $(2, 2)$ leap-frog method is the extension of the standard Yee FDTD method to Lorentz dispersive media. Finally, comparisons will be made among all finite difference schemes under considerations.

We first compute the dispersion relation for the fully discrete $(2, 2M)$ schemes. To do so, we assume the plane wave solutions

$$X_j^n \equiv X_0 e^{i\left(k^*_{\mathrm{FD},2M} jh - \omega n \Delta t\right)}, \tag{6.1}$$

where $*$ is either LF or TP. Substituting (6.1) into the appropriate $(2, 2M)$ FDTD method (which we have not explicitly written out here for brevity), we obtain the linear system

$$\mathcal{A}^*_{\mathrm{FD},2M} \mathbf{U}^*_{\mathrm{FD}} = \mathbf{0}, \tag{6.2}$$

where the coefficient matrix for the two schemes will be discussed in the next two sections.

6.1. Fully discrete dispersion analysis: $(2, 2M)$ leap-frog-FDTD schemes

We first consider the $(2, 2M)$ leap-frog FDTD scheme. For the leap-frog temporal discretization the coefficient matrix in the linear system (6.2) is given as

$$\mathcal{A}^{\mathrm{LF}}_{\mathrm{FD},2M} = \begin{pmatrix} \sin\left(\frac{W}{2}\right) & \Lambda^{\mathrm{LF}}_{2M} & 0 & 0 \\ \Lambda^{\mathrm{LF}}_{2M} & \epsilon_\infty \sin\left(\frac{W}{2}\right) & \sin\left(\frac{W}{2}\right) & 0 \\ 0 & 0 & i \sin\left(\frac{W}{2}\right) & \frac{\Delta t}{2} \cos\left(\frac{W}{2}\right) \\ 0 & \frac{\Delta t}{2}\omega_p^2 \cos\left(\frac{W}{2}\right) & -\frac{\Delta t}{2}\omega_1^2 \cos\left(\frac{W}{2}\right) & i \sin\left(\frac{W}{2}\right) - \gamma \Delta t \cos\left(\frac{W}{2}\right) \end{pmatrix}, \tag{6.3}$$

with

$$\Lambda^{\mathrm{LF}}_{2M} = \frac{\Delta t}{h} \sum_{p=1}^{M} \frac{\lambda^{2M}_{2p-1}}{(2p-1)} \sin\left[\left(p - \frac{1}{2}\right) k^{\mathrm{LF}}_{\mathrm{FD},2M} h\right].$$

Based on previous discussions, we can derive the identity

$$\sum_{p=1}^{M} \frac{[(2p-3)!!]^2}{(2p-1)!} \sin^{2p-1}\left(\frac{k^{\mathrm{LF}}_{\mathrm{FD},2M} h}{2}\right) = \frac{1}{2} k^{\mathrm{LF}} h. \tag{6.4}$$

For both the fully discrete methods, we focus our discussions on the physical modes. For the case of $W \ll 1$ and $K \ll 1$, we analyze the Taylor expansion of the physical modes for the fully discrete leap-frog FDTD schemes and observe the following pattern:

$$k^{\mathrm{LF}}_{\mathrm{FD\,phys},2M} = \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} - \frac{1}{2}\right) W^2 + \frac{[(2M-1)!!]^2}{2^{2M}(2M+1)!} K^{2M} + \mathcal{O}(K^{2M+2} + K^{2M} W^2 + W^4)\right), \quad M \geq 1. \tag{6.5}$$

Furthermore, with the relation $K = \dfrac{\sqrt{\epsilon(\widehat{\omega};\mathbf{p})}}{\nu\sqrt{\epsilon_\infty}} W$, we can treat $k^{\mathrm{LF}}_{\mathrm{FD\,phys},2M}$ as a function of W and $\nu = \dfrac{\Delta t}{h\sqrt{\epsilon_\infty}}$, the CFL (Courant-Friedrich-Lewy) number subject to the stability constraint for the $(2, 2M)$ leap-frog FDTD scheme. Assuming $\nu = O(1)$, we have

$$k^{\mathrm{LF}}_{\mathrm{FD\,phys},2M} = \begin{cases} \pm k^{\mathrm{ex}} \left(1 + \dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} - \dfrac{1}{2} + \dfrac{\epsilon(\widehat{\omega};\mathbf{p})}{2\epsilon_\infty \nu^2}\right) W^2 + \mathcal{O}(W^4)\right), & M = 1, \\ \pm k^{\mathrm{ex}} \left(1 + \dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} - \dfrac{1}{2}\right) W^2 + \mathcal{O}(W^4)\right), & M \geq 2. \end{cases} \tag{6.6}$$

We can see that due to the second order time discretizations employed, the fully discrete scheme always results in a second order dispersion error. Particularly for all $M \geq 2$, the leading term in the dispersion error is identical, and independent of ν which comes solely from the temporal discretization. To compare the performance of the scheme for $M = 1$ and $M \geq 2$, we will focus on comparison of the coefficients of leading error terms in (6.6).

We first consider $\widehat{\gamma} = 0$. We can make the following conclusions.

- For $\widehat{\omega}$ in the medium absorption band, i.e. $\widehat{\omega} \in (1, \sqrt{\epsilon_s/\epsilon_\infty})$, it is easy to check that

$$\left|\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} - \frac{1}{2} + \frac{\epsilon(\widehat{\omega};\mathbf{p})}{2\epsilon_\infty \nu^2}\right| \geq \left|\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} - \frac{1}{2}\right|$$

  based on the inequalities $\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} \leq -1$ and $\dfrac{\epsilon(\widehat{\omega};\mathbf{p})}{2\epsilon_\infty \nu^2} \leq 0$, which implies that the high order schemes reduce the dispersion error as one would expect. This is true independent of other parameter choices.
- For other $\widehat{\omega}$ values, the outcome will depend on the parameters. We can show that the general condition for the higher order scheme ($M \geq 2$) to be more accurate in its dispersion error is equivalent to the inequality

$$\left(\frac{\epsilon_d}{\epsilon_\infty}\right)^2 + (1-2\nu^2)(\widehat{\omega}^2 - 1)^2 - 2\left(\frac{\epsilon_d}{\epsilon_\infty}\right)\left(\widehat{\omega}^2 - 1 + \nu^2(1 - 3\widehat{\omega}^2)\right) \geq 0. \tag{6.7}$$

  This is a quadratic inequality in $\widehat{\omega}^2$. We can conclude that with the CFL condition $\nu \leq 1$, which is a necessary condition to ensure the fully discrete $(2, 2M)$ leap-frog-FDTD scheme is stable for any $M \geq 1$ [8,7], we have
  - if $0 < \nu \leq \frac{1}{\sqrt{2}}$, the condition (6.7) always holds for all $\widehat{\omega} \geq 0$.
  - if $\frac{1}{\sqrt{2}} < \nu < 1$ and
    - if $0 < \epsilon_d/\epsilon_\infty \leq 2\nu^2 - 1$, then the condition (6.7) holds on

$$\widehat{\omega}_L \leq \widehat{\omega} \leq \widehat{\omega}_R,$$

    - if $\epsilon_d/\epsilon_\infty \geq 2\nu^2 - 1$, then the condition (6.7) holds on

$$0 \leq \widehat{\omega} \leq \widehat{\omega}_R,$$

  where

$$\widehat{\omega}_L = \sqrt{\frac{-1 - \epsilon_d/\epsilon_\infty + 2\nu^2 + 3\epsilon_d/\epsilon_\infty \nu^2 - \nu\sqrt{-4\epsilon_d/\epsilon_\infty - 4(\epsilon_d/\epsilon_\infty)^2 + 8\epsilon_d/\epsilon_\infty \nu^2 + 9(\epsilon_d \nu/\epsilon_\infty)^2}}{2\nu^2 - 1}},$$

$$\widehat{\omega}_R = \sqrt{\frac{-1 - \epsilon_d/\epsilon_\infty + 2\nu^2 + 3\epsilon_d/\epsilon_\infty \nu^2 + \nu\sqrt{-4\epsilon_d/\epsilon_\infty - 4(\epsilon_d/\epsilon_\infty)^2 + 8\epsilon_d/\epsilon_\infty \nu^2 + 9(\epsilon_d \nu/\epsilon_\infty)^2}}{2\nu^2 - 1}}.$$

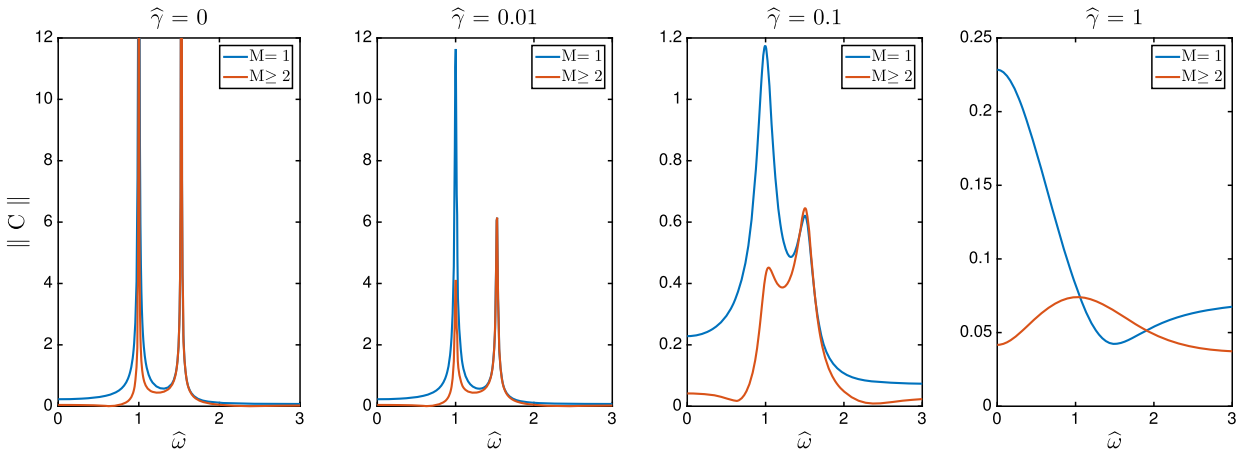

Fig. 6.1. Absolute value of coefficients of leading error terms in (6.6) (denoted by C) for the $(2, 2M)$ leap-frog FDTD scheme.

The case of $\widehat{\gamma} > 0$ is even more complicated. For low loss materials, in general we expect similar conclusions as in the lossless case. We now perform a numerical study, and compare the leading error terms in (6.6) with $\nu = 0.6$ (which is small enough to guarantee that the scheme is stable for arbitrary M (see next section and [7]). The absolute values of coefficients of leading error terms are plotted in Fig. 6.1, with $\epsilon_\infty = 2.25$, $\epsilon_s = 5.25$ and various $\widehat{\gamma}$ values. It is observed that we can not determine which method performs better for the general case. From Fig. 6.1, it's clear that higher order schemes have smaller dispersion error for $\widehat{\gamma} = 0, 0.01$ in the range $\widehat{\omega} \in [0, 3]$. This is no longer true for $\widehat{\gamma} = 0.1, 1$. The discussion here reveals an interesting fact. For some parameter values, we can have counterintuitive results that the lower order scheme performs better than higher order scheme when numerical dispersion is present.

6.2. *Fully discrete dispersion analysis:* $(2, 2M)$ *trapezoidal-FDTD schemes*

We repeat the analysis done in the previous section for the fully discrete $(2, 2M)$ trapezoidal FDTD schemes. We obtain the numerical dispersion relation for these schemes by setting the determinant of the matrix

$$\mathcal{A}^{\mathrm{TP}}_{\mathrm{FD},2M} = \begin{pmatrix} \sin\left(\frac{W}{2}\right) & \Lambda^{\mathrm{TP}}_{2M} & 0 & 0 \\ \Lambda^{\mathrm{TP}}_{2M} & \epsilon_\infty \sin\left(\frac{W}{2}\right) & \sin\left(\frac{W}{2}\right) & 0 \\ 0 & 0 & i \sin\left(\frac{W}{2}\right) & \frac{\Delta t}{2}\cos\left(\frac{W}{2}\right) \\ 0 & \frac{\Delta t}{2}\omega_p^2 \cos\left(\frac{W}{2}\right) & -\frac{\Delta t}{2}\omega_1^2 \cos\left(\frac{W}{2}\right) & i \sin\left(\frac{W}{2}\right) - \gamma\, \Delta t \cos\left(\frac{W}{2}\right) \end{pmatrix} \tag{6.8}$$

to zero. In the above, we have

$$\Lambda^{\mathrm{TP}}_{2M} = \frac{\Delta t}{h} \sum_{p=1}^{M} \frac{\lambda^{2M}_{2p-1}}{(2p-1)} \sin\left[\left(p - \frac{1}{2}\right) k^{\mathrm{TP}}_{\mathrm{FD},2M} h\right] \cos\left(\frac{W}{2}\right).$$

The numerical dispersion is given by

$$\sum_{p=1}^{M} \frac{[(2p-3)!!]^2}{(2p-1)!} \sin^{2p-1}\left(\frac{k^{\mathrm{TP}}_{\mathrm{FD},2M} h}{2}\right) = \frac{1}{2} k^{\mathrm{TP}} h. \tag{6.9}$$

By requiring $W \ll 1$ and $K \ll 1$, we can obtain the physical modes in the form

$$k^{\mathrm{TP}}_{\mathrm{FD\,phys},2M} = \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} + 1\right) W^2 + \frac{[(2M-1)!!]^2}{2^{2M}(2M+1)!} K^{2M} + \mathcal{O}(K^{2M+2} + K^{2M} W^2 + W^4)\right), \quad M \geq 1. \tag{6.10}$$

For $W \ll 1$ with $\nu = \mathcal{O}(1)$, Taylor expansion gives us

$$k^{\mathrm{TP}}_{\mathrm{FD\,phys},2M} = \begin{cases} \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} + 1 + \frac{\epsilon(\widehat{\omega}; \mathbf{p})}{2\epsilon_\infty \nu^2}\right) W^2 + \mathcal{O}(W^4)\right), & M = 1, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega}; \mathbf{p})}{\epsilon(\widehat{\omega}; \mathbf{p})} + 1\right) W^2 + \mathcal{O}(W^4)\right), & M \geq 2. \end{cases} \tag{6.11}$$

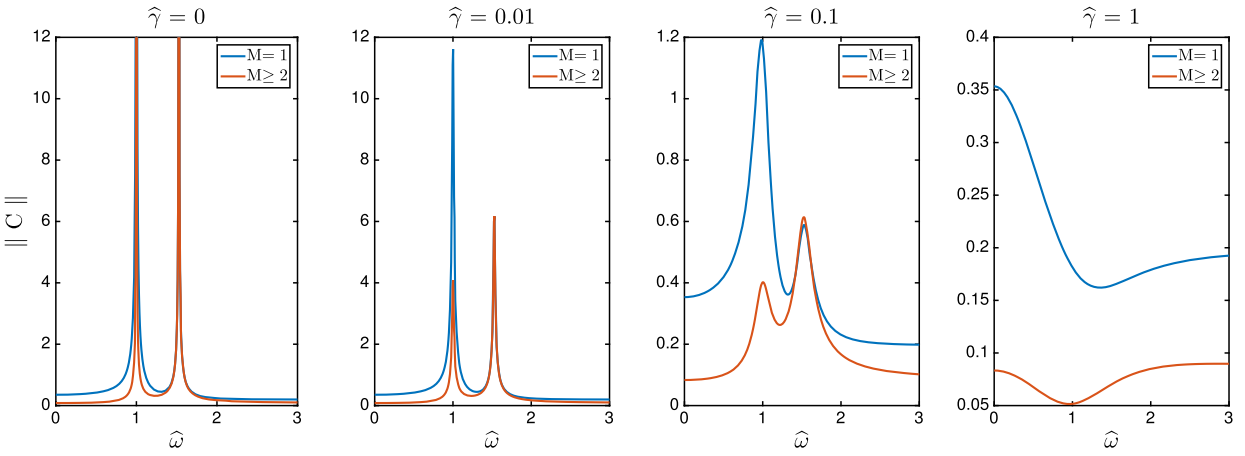

Fig. 6.2. Absolute value of coefficients of leading error terms in (6.11) (denoted by C) for the $(2, 2M)$ trapezoidal FDTD scheme.

This shows second order dispersion error in all cases. When $\widehat{\gamma} = 0$, it is easy to check that $\left|\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1 + \frac{\epsilon(\widehat{\omega};\mathbf{p})}{2\epsilon_\infty \nu^2}\right| \geq \left|\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right|$ for any $\widehat{\omega} \geq 0$ and $\nu > 0$. Hence, the high order FDTD schemes with $M \geq 2$ always have smaller dispersion error than the $(2, 2)$ FDTD scheme. On the other hand, numerical tests comparing the coefficients of leading order error terms in the trapezoidal FDTD schemes are provided in Fig. 6.2, with various $\widehat{\gamma}$ values and the same parameters as used in Fig. 6.1. The plots indicate that it is again difficult to determine which coefficient (for $M = 1$ or $M \geq 2$) is larger when $\widehat{\gamma}$ is large.

6.3. Comparison among fully discrete FDTD schemes

Here, we will present comparisons of the relative phase error for both the leap-frog and trapezoidal FDTD schemes using the parameters values fixed as in (4.16). For the fully discrete schemes, $\omega_1 \Delta t$ and $\omega_1 h$ are needed to determine $k^*_{\mathrm{FD},2M}$. As shown in [6,7], the schemes based on the trapezoidal rule are unconditionally stable, while the leap-frog schemes are conditionally stable, with the stability condition as $\nu \leq \nu^{2M}_{max}$, with ν^{2M}_{max} defined as the largest CFL number of the $(2, 2M)$ leap-frog-FD scheme, given by the formula [7,8]

$$\nu^{2M}_{max} = \frac{1}{\sum_{p=1}^{M} \frac{[(2p-3)!!]^2}{(2p-1)!}}. \tag{6.12}$$

We note that as M increases, ν^{2M}_{max} decreases but is bounded from below by $\nu^{\infty}_{max} = 2/\pi$, i.e. in the limiting case ($M \to \infty$), ν^{2M}_{max} approaches $2/\pi$ [8].

First, we will consider the schemes with a normalized CFL number $\nu/\nu^{2M}_{max} = 0.7$ for both types of temporal discretizations. Relative phase errors are plotted in the range $\widehat{\omega} \in [0, 3]$. In Fig. 6.3, we show errors of LF$(2, 2M)$ and TP$(2, 2M)$ with $W_1 = \pi/30$. The fully discrete schemes do give two peaks near $\widehat{\omega} = 1$ and $\widehat{\omega} = \sqrt{\epsilon_s/\epsilon_\infty}$. As seen in Section 4, the phase errors for schemes based on semi-discretizations in time have two peaks in this range, while only one peak is observed for the semi-discrete spatial schemes as seen in Section 5.1. Thus, it is reasonable to believe that the second peak results from time discretization, while the first one is associated with both space and time discretization. Comparing FDTD schemes with the same time discretization, phase error for the scheme with $M = 2$ is smaller than that of the second order scheme for $M = 1$. However, there is no significant difference among the phase errors with $M \geq 2$ indicating that dispersion errors are dominated by time discretizations when $M \geq 2$. These observations are consistent with our analysis. On the other hand, difference in phase error plots between LF$(2, 2M)$ and TP$(2, 2M)$ is similar to the results obtained for the semi-discrete in time schemes as seen in the second plot of Fig. 4.1.

In the second experiment, we will consider the fully discrete trapezoidal FDTD scheme with various CFL numbers. We give the contour plots of the dispersion error at $\widehat{\omega} = 1$ in Fig. 6.4, with $W_1 \in [0.05, 0.3]$ and $\omega_1 h \in [0.01, 0.1]$. Here, the vertical coordinate is $W = \widehat{\omega} W_1$ and the horizontal coordinate is $|K| = |\widehat{\omega}\sqrt{\epsilon(\widehat{\omega};\mathbf{p})}\,\omega_1 h|$. In this coordinate system, for the range of values considered, the dispersion error of TP(2,2) can be improved by both taking smaller time steps and/or refining the spatial grid. With fixed time step and spatial grid, we can also reduce the phase error by increasing the scheme to fourth order. The contour lines in Fig. 6.4 of higher order ($M = 2, 3$) schemes are horizontal, and the contours for TP(2,4) and TP(2,6) have no visible difference. Neither decreasing space mesh size nor increasing spatial order can reduce the phase error, which also illustrates the dominant role of temporal errors.

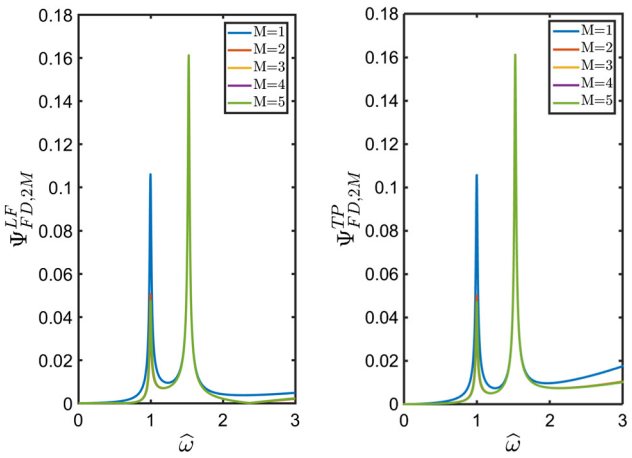

Fig. 6.3. The relative phase error of fully discrete FDTD schemes for the physical modes with $v/v^{2M}_{max} = 0.7$ and $W_1 = \pi/30$. Left: the leap-frog scheme; right: the trapezoidal scheme.

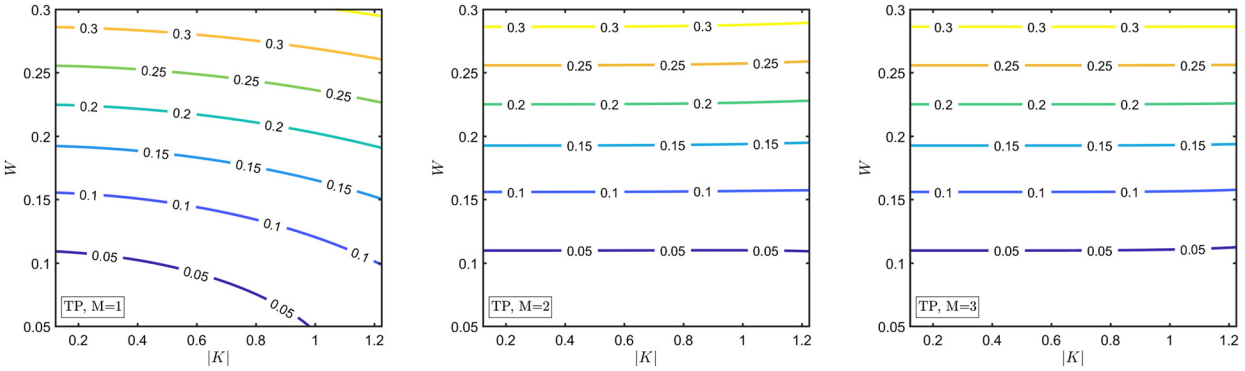

Fig. 6.4. The contour plots of relative phase error of fully discrete FD schemes for the physical modes with trapezoidal scheme. $\widehat{\omega} = 1$.

Both our analysis and figures demonstrate that FDTD schemes with $M \geq 3$ do not improve the phase error of fully discrete schemes significantly beyond that achieved for $M = 2$. Hence, LF$(2,4)$ and TP$(2,4)$ seem to be the "best" schemes to work with from this perspective for most parameter choices (except for materials with large loss, or low-loss materials with certain range of frequencies as shown in Figs. 6.1 and 6.2).

7. Spatial discretization: discontinuous Galerkin schemes

In this section and next one, similar to Section 5 and Section 6, we perform semi-discrete and fully discrete analysis when the spatial variable is discretized by DG schemes. Here, we define the grid as $x_{j+1/2} = (j+1/2)h,\ j \in \mathbb{Z}$, with uniform mesh size h. Let $I_j = [x_{j-1/2}, x_{j+1/2}]$ be a mesh element, with $x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})$ as its center. We now define a finite dimensional discrete space,

$$V_h^p = \{v : v|_{I_j} \in P^p(I_j),\ j \in \mathbb{Z}\}, \tag{7.1}$$

which consists of piecewise polynomials of degree up to p with respect to the mesh. For any $v \in V_h^p$, let $v^+_{j+\frac{1}{2}}$ (resp. $v^-_{j+\frac{1}{2}}$) denote the limit value of v at $x_{j+\frac{1}{2}}$ from the element I_{j+1} (resp. I_j), $[v]_{j+\frac{1}{2}} = v^+_{j+\frac{1}{2}} - v^-_{j+\frac{1}{2}}$ denote its jump, and $\{v\}_{j+\frac{1}{2}} = \frac{1}{2}(v^+_{j+\frac{1}{2}} + v^-_{j+\frac{1}{2}})$ be its average, again at $x_{j+\frac{1}{2}}$.

The semi-discrete DG method for the system (2.5) is formulated as follows: find $H_h(t,\cdot)$, $D_h(t,\cdot)$, $E_h(t,\cdot)$, $P_h(t,\cdot)$, $J_h(t,\cdot) \in V_h^p$, such that $\forall j$,

$$\int_{I_j} \partial_t H_h \phi dx + \int_{I_j} E_h \partial_x \phi dx - (\widehat{E_h}\phi^-)_{j+1/2} + (\widehat{E_h}\phi^+)_{j-1/2} = 0, \quad \forall \phi \in V_h^p, \tag{7.2a}$$

$$\int_{I_j} \partial_t D_h \phi dx + \int_{I_j} H_h \partial_x \phi dx - (\widetilde{H_h}\phi^-)_{j+1/2} + (\widetilde{H_h}\phi^+)_{j-1/2} = 0, \quad \forall \phi \in V_h^p, \tag{7.2b}$$

$$\partial_t P_h = J_h, \tag{7.2c}$$

$$\partial_t J_h = -2\gamma J_h - \omega_1^2 P_h + \omega_p^2 E_h, \tag{7.2d}$$

$$D_h = \epsilon_\infty E_h + P_h. \tag{7.2e}$$

Both the terms $\widehat{E_h}$ and $\widetilde{H_h}$ are numerical fluxes, and they are single-valued functions defined on the cell interfaces and should be designed to ensure numerical stability and accuracy. In the present work, we consider the following general form of numerical fluxes similar to the ones introduced in [10],

$$\widehat{E_h} = \{E_h\} + \alpha[E_h] + \beta_1[H_h], \tag{7.3a}$$

$$\widetilde{H_h} = \{H_h\} - \alpha[H_h] + \beta_2[E_h]. \tag{7.3b}$$

Here, α, β_1 and β_2 are constants that are taken to be $\mathcal{O}(1)$, with β_1 and β_2 being non-negative for stability. For example, if we take $\alpha = \beta_1 = \beta_2 = 0$, we have the central flux

$$\widehat{E_h} = \{E_h\}, \quad \widetilde{H_h} = \{H_h\}; \tag{7.4}$$

if $\alpha = \pm 1/2$ and $\beta_1 = \beta_2 = 0$, we have the alternating flux

$$\widehat{E_h} = E_h^-, \quad \widetilde{H_h} = H_h^+; \quad \text{or} \quad \widehat{E_h} = E_h^+, \quad \widetilde{H_h} = H_h^-; \tag{7.5}$$

and if $\alpha = 0$, $\beta_1 = 1/(2\sqrt{\epsilon_\infty})$, and $\beta_2 = \sqrt{\epsilon_\infty}/2$, we have the "upwind" flux for the Maxwell's equations neglecting Lorentz dispersion

$$\widehat{E_h} = \{E_h\} + \frac{1}{2\sqrt{\epsilon_\infty}}[H_h], \quad \widetilde{H_h} = \{H_h\} + \frac{\sqrt{\epsilon_\infty}}{2}[E_h]. \tag{7.6}$$

In particular, when using the alternating flux with $p = 0$, it is easy to check that the DG scheme is equivalent to FD2 discretization.

7.1. Semi-discrete in space dispersion analysis

In order to carry out the dispersion analysis for piecewise P^p polynomials, we assume that the semi-discrete system has plane wave solutions of the form

$$X_h(x,t)|_{I_j} = e^{i(k_{\mathrm{DG},p} jh - \omega t)} X_0\left(2(x - x_j)/h\right), \tag{7.7}$$

where $X_0(\cdot) \in P^p[-1,1]$, and $k_{\mathrm{DG},p}$ representing the numerical wave number of the semi-discrete DG scheme. To find $k_{\mathrm{DG},p}$ for any $p \geq 0$, we will follow the idea given in [1,4,3], in which the system will be simplified by using Jacobi polynomials as the basis functions of P^p.

In our analysis, we will transform the interval I_j to the reference interval $[-1,1]$ with $s = 2(x - x_j)/h$. An inner product is defined on $[-1,1]$ as

$$< f, g >= \int_{-1}^{1} f(s)g(s)ds.$$

Then, (7.2) leads to the following system for any $\phi \in P^p([-1,1])$:

$$\begin{aligned}
&-\frac{1}{2}i\omega h < H_0, \phi > - < \partial_\xi E_0, \phi > \\
&\quad + \left(\left(\alpha - \frac{1}{2}\right)\left(e^{ik_{\mathrm{DG},p}h}E_0(-1) - E_0(1)\right) + \beta_1\left(e^{ik_{\mathrm{DG},p}h}H_0(-1) - H_0(1)\right)\right)\phi(1) \\
&\quad - \left(\left(\alpha + \frac{1}{2}\right)\left(E_0(-1) - e^{-ik_{\mathrm{DG},p}h}E_0(1)\right) + \beta_1\left(H_0(-1) - e^{-ik_{\mathrm{DG},p}h}H_0(1)\right)\right)\phi(-1) = 0,
\end{aligned} \tag{7.8a}$$

$$\begin{aligned}
&-\frac{1}{2}i\omega h < D_0, \phi > - < \partial_\xi H_0, \phi > \\
&\quad + \left(\left(\alpha - \frac{1}{2}\right)\left(e^{ik_{\mathrm{DG},p}h}H_0(-1) + H_0(1)\right) + \beta_2\left(e^{ik_{\mathrm{DG},p}h}E_0(-1) - E_0(1)\right)\right)\phi(1)
\end{aligned}$$

$$-\left(\left(\alpha+\frac{1}{2}\right)\left(H_0(-1)+e^{-ik_{\mathrm{DG},p}h}H_0(1)\right)+\beta_2\left(E_0(-1)-e^{-ik_{\mathrm{DG},p}h}E_0(1)\right)\right)\phi(-1)=0, \tag{7.8b}$$

$$-i\omega h<P_0,\phi>=<J_0,\phi>, \tag{7.8c}$$

$$-i\omega h<J_0,\phi>=-2\gamma<J_0,\phi>-\omega_1^2<P_0,\phi>+\omega_p^2<E_0,\phi>, \tag{7.8d}$$

$$<D_0,\phi>=\epsilon_\infty<E_0,\phi>+<P_0,\phi>. \tag{7.8e}$$

Note that (7.8c)-(7.8e) give us that

$$<D_0,\phi>=\left(\epsilon_\infty-\frac{\omega_p^2}{\omega^2-2i\gamma\omega+\omega_1^2}\right)<E_0,\phi>=\epsilon<E_0,\phi>, \tag{7.9}$$

with $\epsilon=\epsilon(\widehat{\omega};\mathbf{p})$ given in (3.4). Hence, we can have a system only in the variables E_0, H_0 and ϕ by plugging (7.9) into (7.8a) and (7.8b). Furthermore, define two polynomials in $P^p([-1,1])$

$$u_1=\sqrt{\epsilon}\,E_0+H_0,\quad\text{and}\quad u_2=\sqrt{\epsilon}\,E_0-H_0.$$

Then, u_1 and u_2 satisfy the following system,

$$<\mathcal{L}_\epsilon^- u_1,\phi>+\mathcal{R}^-(u_1,u_2,\phi;\alpha,\beta_1,\beta_2,\epsilon,e^{ik_{\mathrm{DG},p}h})=0, \tag{7.10a}$$

$$<\mathcal{L}_\epsilon^+ u_2,\phi>+\mathcal{R}^+(u_1,u_2,\phi;\alpha,\beta_1,\beta_2,\epsilon,e^{ik_{\mathrm{DG},p}h})=0, \tag{7.10b}$$

with the differential operators depending on ϵ

$$\mathcal{L}_\epsilon^\pm v=\mp(\frac{1}{2}i\sqrt{\epsilon}\,\omega h)\,v+\partial_\xi v=\mp\frac{1}{2}iKv+\partial_\xi v,\quad K=k^{ex}h. \tag{7.11}$$

And the second terms $\mathcal{R}^\pm$ are given as

$$\begin{aligned}&\mathcal{R}^-(u_1,u_2,\phi;\alpha,\beta_1,\beta_2,\epsilon,e^{ik_{\mathrm{DG},p}h})\\&=\frac{1}{2}\phi(1)\left\{\left(1+\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(e^{ik_{\mathrm{DG},p}h}u_1(-1)-u_1(1)\right)+\left(2\alpha-\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(e^{ik_{\mathrm{DG},p}h}u_2(-1)-u_2(1)\right)\right\}\\&-\frac{1}{2}\phi(-1)\left\{\left(-1+\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(u_1(-1)-e^{-ik_{\mathrm{DG},p}h}u_1(1)\right)+\left(2\alpha-\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(u_2(-1)-e^{-ik_{\mathrm{DG},p}h}u_2(1)\right)\right\},\end{aligned} \tag{7.12a}$$

$$\begin{aligned}&\mathcal{R}^+(u_1,u_2,\phi;\alpha,\beta_1,\beta_2,\epsilon,e^{ik_{\mathrm{DG},p}h})\\&=\frac{1}{2}\phi(1)\left\{\left(2\alpha+\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(e^{ik_{\mathrm{DG},p}h}u_1(-1)-u_1(1)\right)+\left(1-\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(e^{ik_{\mathrm{DG},p}h}u_2(-1)-u_2(1)\right)\right\}\\&-\frac{1}{2}\phi(-1)\left\{\left(2\alpha+\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(u_1(-1)-e^{-ik_{\mathrm{DG},p}h}u_1(1)\right)+\left(-1-\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(u_2(-1)-e^{-ik_{\mathrm{DG},p}h}u_2(1)\right)\right\}.\end{aligned} \tag{7.12b}$$

In particular, if the test function vanishes at the two endpoints, i.e., $\phi(-1)=\phi(1)=0$, the second terms $\mathcal{R}^\pm=0$. Next, we will look at the system (7.10) instead of (7.8).

When $p=0$, u_1 and u_2 are constant. Taking the test function $\phi=1$, (7.10) gives us

$$\begin{aligned}0=&iKu_1+\frac{1}{2}\left(\lambda-\lambda^{-1}+(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})(\lambda-2+\lambda^{-1})\right)u_1+\frac{1}{2}\left(2\alpha-\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)(\lambda-2+\lambda^{-1})u_2,\\0=&-iKu_2+\frac{1}{2}\left(2\alpha+\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}\right)(\lambda-2+\lambda^{-1})u_1+\frac{1}{2}\left(\lambda-\lambda^{-1}-(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})(\lambda-2+\lambda^{-1})\right)u_2,\end{aligned}$$

with $\lambda=e^{ik_{\mathrm{DG},p}h}$. This system has non-trivial solutions u_1 and u_2 if and only if

$$Det\begin{pmatrix}iK+\frac{1}{2}\left(\lambda-\lambda^{-1}+(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})(\lambda-2+\lambda^{-1})\right) & \frac{1}{2}\left(2\alpha-\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)(\lambda-2+\lambda^{-1})\\ \frac{1}{2}\left(2\alpha+\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}\right)(\lambda-2+\lambda^{-1}) & -iK+\frac{1}{2}\left(\lambda-\lambda^{-1}-(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})(\lambda-2+\lambda^{-1})\right)\end{pmatrix}=0.$$

For general $p\geq 1$, use the fact that

$$<\mathcal{L}_\epsilon^- u_1,\phi>=<\mathcal{L}_\epsilon^+ u_2,\phi>=0,\quad\forall\phi=(1-s)(1+s)w(s),\ \text{with}\ w(s)\in P^{p-2}[-1,1],$$

we obtain that

$$\mathcal{L}_\epsilon^- u_1 = \widetilde{a}^- P_p^{(0,1)} + \widetilde{b}^- P_p^{(1,0)}, \quad \mathcal{L}_\epsilon^+ u_1 = \widetilde{a}^+ P_p^{(0,1)} + \widetilde{b}^+ P_p^{(1,0)},$$

where $P_p^{(m,n)}$ denotes the Jacobi polynomial of type (m,n) and degree p. Moreover, recall the following polynomials of degree p from [1],

$$\Phi_p^{1,\pm}(s) = \sum_{m=0}^{p} (\pm iK)^m \frac{(2p+1-m)!}{(2p+1)!} P_m^{(p-m,p-m+1)}(s), \tag{7.13a}$$

$$\Phi_p^{2,\pm}(s) = \sum_{m=0}^{p} (\pm iK)^m \frac{(2p+1-m)!}{(2p+1)!} P_m^{(p-m+1,p-m)}(s). \tag{7.13b}$$

And they have the property that

$$\begin{aligned}
\mathcal{L}_\epsilon^- \Phi_p^{1,-}(s) &= -\frac{(-iK)^{p+1}}{2} \frac{(p+1)!}{(2p+1)!} P_p^{(0,1)}(s),\\
\mathcal{L}_\epsilon^- \Phi_p^{2,-}(s) &= -\frac{(-iK)^{p+1}}{2} \frac{(p+1)!}{(2p+1)!} P_p^{(1,0)}(s),\\
\mathcal{L}_\epsilon^+ \Phi_p^{1,+}(s) &= -\frac{(iK)^{p+1}}{2} \frac{(p+1)!}{(2p+1)!} P_p^{(0,1)}(s),\\
\mathcal{L}_\epsilon^+ \Phi_p^{2,+}(s) &= -\frac{(iK)^{p+1}}{2} \frac{(p+1)!}{(2p+1)!} P_p^{(0,1)}(s).
\end{aligned}$$

Therefore,

$$u_1 = a^- \Phi_p^{1,-} + b^- \Phi_p^{2,-}, \quad u_2 = a^+ \Phi_p^{1,+} + b^+ \Phi_p^{2,+}.$$

Then, the coefficients $a^\pm$ and $b^\pm$ satisfy four algebraic equations corresponding to choosing test function in (7.10) with test function $\phi = 1 \pm s$. This leads a 4×4 system

$$M \cdot (a^+, a^-, b^+, b^-)^T = 0, \tag{7.14}$$

with the matrix M given as the following

$$\begin{pmatrix}
(1+\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})\lambda & (-1)^p(-1+\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}) & -(-1)^p(2\alpha-\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})\lambda & -(2\alpha-\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})\\
(2\alpha+\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}})\lambda & (-1)^p(2\alpha+\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}}) & -(-1)^p(1-\beta_1\sqrt{\epsilon}-\frac{\beta_2}{\sqrt{\epsilon}})\lambda & (1+\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}})\\
\lambda F_{p+1}^+ - F_p^- & \lambda F_p^+ - F_{p+1}^- & 0 & 0\\
0 & 0 & \lambda F_{p+1}^- - F_p^+ & \lambda F_p^- - F_{p+1}^+
\end{pmatrix}$$

and

$$F_p^\pm = \sum_{m=0}^{\infty} \frac{(-p)_m}{(-2p-1)_m} \frac{(\pm iK)^m}{m!}, \qquad F_{p+1}^\pm = \sum_{m=0}^{\infty} \frac{(-p-1)_m}{(-2p-1)_m} \frac{(\pm iK)^m}{m!},$$

where $(a)_0 = 1$ and $(a)_m = a(a+1)\cdots(a+m-1)$. Then, the numerical dispersion relation can be obtained by solving $Det(M) = 0$.

In summary, we have the following theorem which characterizes the dispersion relation satisfied by $k_{\mathrm{DG},p}$.

Theorem 7.1. *Consider the DG scheme (7.2) with V_h^p as the discrete space. Then, $k_{DG,p}$ satisfies the following equation for any $p \geq 0$,*

$$a_p \left(e^{ik_{DG,p}h} + e^{-ik_{DG,p}h}\right)^2 + b_p \left(e^{ik_{DG,p}h} + e^{-ik_{DG,p}h}\right) + c_p = 0. \tag{7.15}$$

Here, for $p = 0$,

$$\begin{aligned}
a_0 &= 1 - 4(\alpha^2 + \beta_1\beta_2),\\
b_0 &= 16(\alpha^2 + \beta_1\beta_2) - 4iK\left(\beta_1\sqrt{\epsilon} + \frac{\beta_2}{\sqrt{\epsilon}}\right),\\
c_0 &= -4\left(1 + 4(\alpha^2 + \beta_1\beta_2)\right) + 8iK\left(\beta_1\sqrt{\epsilon} + \frac{\beta_2}{\sqrt{\epsilon}}\right) + 4K^2.
\end{aligned}$$

While for $p \geq 1$,

$$
\begin{aligned}
a_p =&(-1)^p \left(1-4(\alpha^2+\beta_1\beta_2)\right) F_p^+ F_p^-,\\
b_p =&-(-1)^p \left(1-4(\alpha^2+\beta_1\beta_2)\right)\left(F_p^+ F_{p+1}^+ + F_p^- F_{p+1}^-\right)\\
&+\left(1+4(\alpha^2+\beta_1\beta_2)\right)\left(F_p^+ F_{p+1}^- + F_p^- F_{p+1}^+\right)\\
&+2\left(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)\left(F_p^+ F_{p+1}^- - F_p^- F_{p+1}^+\right),\\
c_p =&2(-1)^p \left(1-4(\alpha^2+\beta_1\beta_2)\right)\left(F_{p+1}^+ F_{p+1}^- - F_p^+ F_p^-\right)\\
&-\left(1-4(\alpha^2+\beta_1\beta_2)\right)\left((F_p^+)^2+(F_p^-)^2+(F_{p+1}^+)^2+(F_{p+1}^-)^2\right)\\
&-2\left(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right)\left((F_p^+)^2-(F_p^-)^2-(F_{p+1}^+)^2+(F_{p+1}^-)^2\right).
\end{aligned}
$$

In particular, $k_{DG,p}$ are the roots of a quartic polynomial equation in terms of $\lambda = e^{ik_{DG,p}h}$ if $\alpha^2+\beta_1\beta_2 \neq 1/4$, and $k_{DG,p}$ are the roots of a quadratic polynomial equation in terms of $\lambda = e^{ik_{DG,p}h}$ when $\alpha^2+\beta_1\beta_2 = 1/4$.

By Theorem 7.1, we can see that the dispersion relation is more complicated than that of the FD scheme, caused by the dependence on the flux parameters α, β_1, β_2, and the coupling of the local degrees of freedom. For the DG scheme (7.2) employing the central flux (7.4) ($\alpha=\beta_1=\beta_2=0$), there are four discrete wave numbers $k_{\mathrm{DG},p}$, corresponding to two physical modes and two spurious modes. While for the alternating fluxes (7.5) and the upwind flux (7.6) ($\alpha^2+\beta_1\beta_2=1/4$), there are only two discrete wave numbers $k_{\mathrm{DG},p}$, corresponding to the physical modes. This conclusion holds for arbitrary p. Unlike the FD scheme, when we increase the order of the accuracy of the scheme, the number of modes won't change when the dispersion relation is expressed by representing the discrete wavenumber as a function of the angular frequency.

Next, with the help of Theorem 7.1, we can obtain the analytical dispersion relation formula for general $p \geq 0$ based on the small wave number limit $K \to 0$. In the following, we write

$$
b = \omega\left(\beta_1\,\epsilon(\widehat{\omega};\mathbf{p})+\beta_2\right), \quad \text{and} \quad B = b\,h = K\left(\beta_1\sqrt{\epsilon}+\frac{\beta_2}{\sqrt{\epsilon}}\right). \tag{7.16}
$$

Note that, $b(\omega)=0$ if and only if $\beta_1=\beta_2=0$. And we assume $B/K = b/k^{ex} = \mathcal{O}(1)$, which means $B \ll 1$ as well. The results are given as follows.

Theorem 7.2. *For the spatial semi-discrete DG method with V^p as the discrete space.*

- *When $\alpha=\beta_1=\beta_2=0$, there are four discrete wave numbers. Two of them correspond to the physical*

$$
k_{DG^{phys},p} = \begin{cases}
\pm k^{ex}\left(1+\dfrac{p+1}{2(2p+3)}\left(\dfrac{p!}{(2p+1)!}\right)^2 K^{2p+2}+\mathcal{O}\left(K^{2p+4}\right)\right), & p\geq 0 \text{ and even},\\
\pm k^{ex}\left(1-\dfrac{2p+1}{2(p+1)}\left(\dfrac{p!}{(2p+1)!}\right)^2 K^{2p}+\mathcal{O}\left(K^{2p+2}\right)\right), & p\geq 0 \text{ and odd}.
\end{cases} \tag{7.17}
$$

- *When $\alpha^2+\beta_1\beta_2=1/4$, we have two physical modes*

$$
k_{DG^{phys},p} = \begin{cases}
\pm k^{ex}\left(1+\dfrac{1}{2}iB+\dfrac{1}{24}\left(K^2-9B^2\right)+\mathcal{O}(iK^2B+iB^3+K^4)\right), & p=0,\\
\pm k^{ex}\left(1+\dfrac{1}{2}\left(\dfrac{p!}{(2p+1)!}\right)^2 iK^{2p}B+\dfrac{1}{2(2p+1)(2p+3)}\right.\\
\quad\left(\dfrac{p!}{(2p+1)!}\right)^2\left(K^{2p+2}-(2p+3)K^{2p}B^2\right)\\
\quad\left.+\mathcal{O}\left(iK^{2p+2}B+iK^{2p}B^3+K^{2p+4}\right)\right), & p\geq 1.
\end{cases} \tag{7.18}
$$

Proof. The results of $\alpha=\beta_1=\beta_2=0$ are special cases of the general results of Theorem in [1]. So we only consider the second results with $\alpha^2+\beta_1\beta_2=1/4$ here. In this case, (7.15) gives us

$$0 = b_p(e^{ik_{\mathrm{DG},p}h} + e^{-ik_{\mathrm{DG},p}h}) + c_p = 2b_p\cos(k_{\mathrm{DG},p}h) + c_p.$$

When $p = 0$, this reduces to

$$\cos(k_{\mathrm{DG},p}h) = \frac{8 - 8iB - 4K^2}{8 - 8iB} = 1 - \frac{1}{2}K^2 - \frac{1}{2}i\left(\frac{B}{K}\right)K^3 + \frac{1}{2}\left(\frac{B}{K}\right)^2 K^4 + \mathcal{O}(K^5).$$

Furthermore, take $\cos^{-1}$ on both sides and we can obtain

$$\begin{aligned} k_{\mathrm{DG},p}h &= \pm\left(K + \frac{1}{2}i\left(\frac{B}{K}\right)K^2 + \frac{1}{24}\left(1 - 9\left(\frac{B}{K}\right)^2\right)K^3 + \mathcal{O}\left(i\left(\frac{B}{K}\right)K^4 + i\left(\frac{B}{K}\right)^3 K^4 + K^5\right)\right) \\ &= \pm k^{ex}h\left(1 + \frac{1}{2}iB + \frac{1}{24}\left(K^2 - 9B^2\right) + \mathcal{O}(iBK^3 + iB^3 + K^4)\right). \end{aligned}$$

In the following, we will look at the general $p \geq 1$. Denote

$$\Pi_p^{\pm} = (F_p^{\mp})^2 e^{\pm iK}, \quad \text{and} \quad \Theta_p^{\pm} = \frac{e^{\pm iK} - [p+1/p]_{e^{\pm iK}}}{e^{\pm iK}},$$

with $[p+1/p]_{e^{\pm iK}} = \dfrac{F_{p+1}^{\pm}}{F_p^{\mp}}$ being the $[p+1/p]$-Padé approximation of $e^{\pm iK}$. Then

$$\begin{aligned} \cos(k_{\mathrm{DG},p}h) &= -\frac{-2\left(\Pi_p^+\left(1 - \frac{B}{K}\right)\left(e^{-iK} + e^{iK}(1-\Theta_p^+)^2\right) + \Pi_p^-\left(1 + \frac{B}{K}\right)\left(e^{iK} + e^{-iK}(1-\Theta_p^-)^2\right)\right)}{2\times 2\left(\Pi_p^+\left(1 - \frac{B}{K}\right)(1-\Theta_p^+) + \Pi_p^-\left(1 + \frac{B}{K}\right)(1-\Theta_p^-)\right)} \\ &= \cos(K) - i\frac{1}{2}\sin(K)\left[\left(\Theta_p^+ - \Theta_p^-\right) + \left(\Theta_p^+ + \Theta_p^-\right)\frac{-\Pi_p^+\left(1 - \frac{B}{K}\right)\Theta_p^+ + \Pi_p^-\left(1 + \frac{B}{K}\right)\Theta_p^-}{\Pi_p^+\left(1 - \frac{B}{K}\right)\Theta_p^+ + \Pi_p^-\left(1 + \frac{B}{K}\right)\Theta_p^-}\right] \\ &\quad + \mathcal{O}\left((\Theta_p^+)^2 + (\Theta_p^-)^2 + \Theta_p^+\Theta_p^-\right). \end{aligned}$$

Next, we will estimate the second term of the right hand side in the case where $K \ll 1$. Corollary 1 in [1] showed that

$$\Theta_p^{\pm} = -\frac{1}{2}K^{2p+2}\left(\frac{p!}{(2p+1)!}\right)^2\left(1 - \frac{\pm iK(2p+2)}{(2p+1)(2p+2)} + \mathcal{O}(K^2)\right).$$

Therefore,

$$\Theta_p^+ + \Theta_p^- = -K^{2p+2}\left(\frac{p!}{(2p+1)!}\right)^2 + \mathcal{O}(K^{2p+4}), \tag{7.19a}$$

$$\Theta_p^+ - \Theta_p^- = iK^{2p+3}\left(\frac{p!}{(2p+1)!}\right)^2\frac{(2p+2)}{(2p+1)(2p+2)} + \mathcal{O}(iK^{2p+5}). \tag{7.19b}$$

Note that $B/K = \mathcal{O}(1)$. Then, we can obtain the Taylor series expansions in K

$$\begin{aligned} \frac{-\Pi_p^+\left(1 - \frac{B}{K}\right)\Theta_p^+ + \Pi_p^-\left(1 + \frac{B}{K}\right)\Theta_p^-}{\Pi_p^+\left(1 - \frac{B}{K}\right)\Theta_p^+ + \Pi_p^-\left(1 + \frac{B}{K}\right)\Theta_p^-} &= -\left(\frac{B}{K}\right) + i\left(1 - \left(\frac{B}{K}\right)^2\right)\frac{K}{2p+1} - \left(\frac{B}{K}\right)\left(1 - \left(\frac{B}{K}\right)^2\right)\frac{K^2}{(2p+1)^2} \\ &\quad + i\left(1 - \left(\frac{B}{K}\right)^2\right)\left(\frac{p}{2p+1} - \left(\frac{B}{K}\right)^2\right)\frac{K^2}{(2p+1)^3} + \mathcal{O}(K^4). \end{aligned} \tag{7.20}$$

Combining series (7.19), (7.20) and the fact that $\sin K = K - \frac{1}{6}K^3 + O(K^5)$, we have

$$\cos(k_{\mathrm{DG},p}h) = \cos(K) - iC_1K^{2p+3} - C_2K^{2p+4} + \mathcal{O}\left(iBK^{2p+4} + iB^3K^{2p+2} + K^{2p+6}\right), \tag{7.21}$$

with

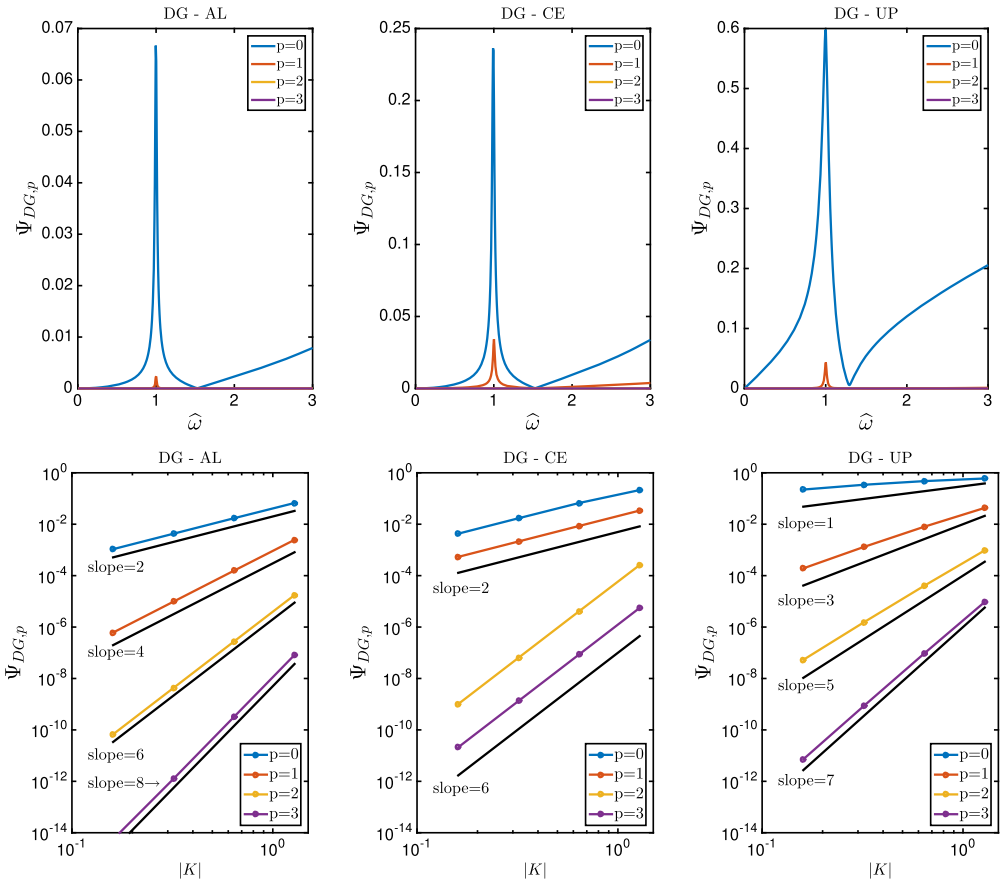

Fig. 7.1. The relative phase error of semi-discrete DG scheme for the physical modes. First row: fix $\omega_1 h = \pi/30$ with $\widehat{\omega}_1 \in [0, 3]$; second row: fix $\widehat{\omega}_1 = 1$ with different $\omega_1 h \in \{\frac{\pi}{30}, \frac{\pi}{60}, \frac{\pi}{120}, \frac{\pi}{240}\}$.

$$C_1 = i\frac{1}{2}\left(\frac{B}{K}\right)\left(\frac{p!}{(2p+1)!}\right)^2, \quad C_2 = \frac{1}{2(2p+1)(2p+3)}\left(\frac{p!}{(2p+1)!}\right)^2\left(1-(2p+3)\left(\frac{B}{K}\right)^2\right).$$

Finally, take $\cos^{-1}$ on both sides

$$k_{\mathrm{DG},p}h = \pm\left(K + C_1K^{2p+2} + C_2K^{2p+3} + \mathcal{O}\left(iBK^{2p+3} + iB^3K^{2p+1} + K^{2p+5}\right)\right),$$

and the result is proved. □

Remark 7.1. These formulas show that, when using the central flux ($\alpha = \beta_1 = \beta_2 = 0$), the physical modes have a dispersion error with order

$$\begin{cases} 2p+2, & \text{if } p \text{ is even,} \\ 2p, & \text{if } p \text{ is odd.} \end{cases} \tag{7.22}$$

When using the alternating fluxes ($\alpha^2 + \beta_1\beta_2 = 1/4$ and $B = 0$), the scheme has a dispersion error of order $(2p+2)$. In particular, the dispersion errors for $\alpha = 1/2$ and $\alpha = -1/2$ are the same. For the upwind flux ($\alpha^2 + \beta_1\beta_2 = 1/4$ and $B \neq 0$), we can observe a $(2p+1)$-th order dispersion error, which is related to K and B at the same time.

It is clear that the order of dispersion error for DG scheme is higher than that of the L^2 convergence. This is an advantage of DG schemes, and differs from FD schemes significantly.

To verify the results above, in Fig. 7.1, we study the relative phase error of the physical modes of the semi-discrete DG scheme (7.2) for $p = 0, 1, 2, 3$, with parameters in (4.16) using the alternating flux (DG-AL) (only the result of one version of the alternating fluxes is shown, because they are identical to each other), the central flux (DG-CE) and the upwind flux (DG-UP). The numerical wave number $k_{\mathrm{DG},p}$ is obtained by solving (7.15) exactly. First, we fix $\omega_1 h = \pi/30$, and plot the dependence of relative phase error as a function of $\widehat{\omega}$, see the first row of Fig. 7.1. It is clear that DG-AL always gives smallest error when the same discrete space is used. This can also be verified by comparing the orders and coefficients in

(7.17) and (7.18). All schemes have significant larger errors around $\widehat{\omega}=1$. For DG-AL and DG-CE, the phase errors approach zero near $\widehat{\omega}=\sqrt{\epsilon_s/\epsilon_\infty}$ where K is close to zero. For DG-UP with $p=0$, the error is dominated by B (see equation (7.18)). Therefore, the "zero" point would shift to the "zero" point of B, which is about $\sqrt{1+\epsilon_d/(2\epsilon_\infty)}\approx 1.291$. Comparing FD2 (Fig. 5.3) and DG-AL with P^0, they have the same performance. However, once we increase the order to $p=1$, DG-AL has significantly smaller error than FD4, resulting from the smaller coefficients in leading error terms, see (5.12), (5.14a) and (7.18). In the second row of Fig. 7.1, we present the errors at $\widehat{\omega}_1=1$ with mesh refinement. Slopes indicate the order of accuracy for each scheme, which agree with our analysis in (7.17) and (7.18).

8. Fully discrete discontinuous Galerkin methods

Here, we consider the DG scheme (7.2) coupled with the leap-frog time discretization (4.1) and the trapezoidal time discretization (4.2). To analyze dispersion relation for those fully discrete schemes, we assume numerical solutions in the form of

$$X_h^n(x)|_{I_j}=e^{i(k^*_{\mathrm{DG},p}h-\omega n\Delta t)}X_0(2(x-x_j)/h),$$

where, $*$ can be LF and TP.

8.1. Fully discrete dispersion analysis: leap-frog-DG schemes

Here, we define

$$u_1=\sqrt{\epsilon^{\mathrm{LF}}}E_0+\frac{2}{W}\sin\left(\frac{W}{2}\right)H_0,\quad u_2=\sqrt{\epsilon^{\mathrm{LF}}}E_0-\frac{2}{W}\sin\left(\frac{W}{2}\right)H_0,$$

with $\epsilon^{\mathrm{LF}}=\epsilon(\widehat{\omega}^{\mathrm{LF}};\mathbf{p}^{\mathrm{LF}})$ given in (4.6). Then, the fully discrete leap-frog-DG schemes reduce to the following problem:

$$<\mathcal{L}^-_{\epsilon^{\mathrm{LF}}}u_1,\phi>+\mathcal{R}^-(u_1,u_2,\phi;\alpha,\beta_1\frac{W/2}{\tan(W/2)},\beta_2\frac{\sin(W/2)\cos(W/2)}{W/2},\epsilon^{\mathrm{LF}},e^{ik^{\mathrm{LF}}_{\mathrm{DG},p}h})=0 \tag{8.1a}$$

$$<\mathcal{L}^+_{\epsilon^{\mathrm{LF}}}u_2,\phi>+\mathcal{R}^+(u_1,u_2,\phi;\alpha,\beta_1\frac{W/2}{\tan(W/2)},\beta_2\frac{\sin(W/2)\cos(W/2)}{W/2},\epsilon^{\mathrm{LF}},e^{ik^{\mathrm{LF}}_{\mathrm{DG},p}h})=0 \tag{8.1b}$$

Note that (8.1) and (7.10) are given in the same form with different parameters. Hence, (7.15) is still valid for the fully discrete leap-frog DG scheme, with parameters modified as the following:

$$\epsilon\to\epsilon^{\mathrm{LF}},\quad K\to\sqrt{\epsilon^{\mathrm{LF}}}\omega h,\quad \beta_1\to\beta_1\frac{W/2}{\tan(W/2)},\quad \beta_2\to\beta_2\frac{\sin(W/2)\cos(W/2)}{W/2}.$$

In this case

$$a_p=(-1)^p\left(1-4(\alpha^2+\beta_1\beta_2\cos^2(W/2))\right).$$

Therefore, if $\alpha=\pm\frac{1}{2}$ and $\beta_1=\beta_2=0$, then $k^{\mathrm{LF}}_{\mathrm{DG},p}$ are the roots of a quadratic polynomial equation in terms of $\lambda=e^{ik^{\mathrm{LF}}_{\mathrm{DG},p}h}$. Otherwise, $k^{\mathrm{LF}}_{\mathrm{DG},p}$ are the roots of a quartic polynomial equation in terms of $\lambda=e^{ik^{\mathrm{LF}}_{\mathrm{DG},p}h}$.

We can see that for the upwind flux, there are four discrete wave numbers $k^{\mathrm{LF}}_{\mathrm{DG},p}$ corresponding to two physical modes and two spurious modes, which differs from the semi-discrete results in Theorem 7.1.

Below, we analyze numerical dispersion property of the physical modes when $W\ll 1$. The formulas can be obtained by repeating the proof of Theorem 7.2, and details would not be given here. Note that $B=bh=\dfrac{b/\omega}{\sqrt{\epsilon_\infty}\,\nu}W\ll 1$ with a fixed CFL number ν, with B given in (7.16).

- When using the central flux, i.e. $\alpha=\beta_1=\beta_2=0$, we have four solutions, and two of them correspond to the physical modes,

$$k^{\mathrm{LF}}_{\mathrm{DG}\,\mathrm{phys},p}=\begin{cases}\pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\dfrac{1}{2}\right)W^2+\dfrac{1}{6}K^2+\mathcal{O}\left(W^4+W^2K^2+K^4\right)\right), & p=0,\\ \pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\dfrac{1}{2}\right)W^2-\dfrac{1}{48}K^2+\mathcal{O}\left(W^4+W^2K^2+K^4\right)\right), & p=1,\\ \pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\dfrac{1}{2}\right)W^2+\mathcal{O}\left(W^4+K^{2p+2}\right)\right), & p\geq 2, even,\\ \pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\dfrac{1}{2}\right)W^2+\mathcal{O}\left(W^4+K^{2p}\right)\right), & p\geq 2, odd,\end{cases} \tag{8.2}$$

in the case of $K \ll 1$ and $W \ll 1$. They can be further written as

$$k^{\mathrm{LF}}_{\mathrm{DG}\,\mathrm{phys},p} = \begin{cases} \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}+\frac{2\epsilon(\widehat{\omega};\mathbf{p})}{\epsilon_\infty \nu^2}\right)W^2+\mathcal{O}\left(W^4\right)\right), & p=0, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}-\frac{\epsilon(\widehat{\omega};\mathbf{p})}{4\epsilon_\infty \nu^2}\right)W^2+\mathcal{O}\left(W^4\right)\right), & p=1, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}\left(W^4\right)\right), & p\geq 2, \end{cases} \tag{8.3}$$

with $W \ll 1$ and a fixed CFL number ν.

- When using the alternating flux, i.e. $\alpha = \pm 1/2$ and $\beta_1 = \beta_2 = 0$, there are only two solutions, corresponding to the physical modes,

$$k^{\mathrm{LF}}_{\mathrm{DG}\,\mathrm{phys},p} = \begin{cases} \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\frac{1}{24}K^2+\mathcal{O}\left(W^4+W^2K^2+K^4\right)\right), & p=0, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}\left(W^4+K^{2p+2}\right)\right), & p\geq 1, \end{cases} \tag{8.4}$$

in the case of $K \ll 1$ and $W \ll 1$, or

$$k^{\mathrm{LF}}_{\mathrm{DG}\,\mathrm{phys},p} = \begin{cases} \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}+\frac{\epsilon(\widehat{\omega};\mathbf{p})}{2\epsilon_\infty \nu^2}\right)W^2+\mathcal{O}\left(W^4\right)\right), & p=0, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}\left(W^4\right)\right), & p\geq 1, \end{cases} \tag{8.5}$$

with $W \ll 1$ and a fixed CFL number ν.

- When using the upwind flux, i.e. $\alpha = 0$, $\beta_1 = \frac{1}{2\sqrt{\epsilon_\infty}}$ and $\beta_2 = \frac{\sqrt{\epsilon_\infty}}{2}$, there are four solutions. The two physical modes are

$$k^{\mathrm{LF}}_{\mathrm{DG}\,\mathrm{phys},p} = \begin{cases} \pm k^{\mathrm{ex}}\left(1+\frac{1}{2}iB+\mathcal{O}\left(W^2+K^2\right)\right), & p=0, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}\left(W^4+iK^{2p}B\right)\right), & p\geq 1, \end{cases} \tag{8.6}$$

which can also be written as

$$k^{\mathrm{LF}}_{\mathrm{DG}\,\mathrm{phys},p} = \begin{cases} \pm k^{\mathrm{ex}}\left(1+i\frac{b/\omega}{2\sqrt{\epsilon_\infty}\,\nu}W+\mathcal{O}\left(W^2\right)\right), & p=0, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}\left(iW^3\right)\right), & p=1, \\ \pm k^{\mathrm{ex}}\left(1+\frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}-\frac{1}{2}\right)W^2+\mathcal{O}\left(W^4\right)\right), & p\geq 2, \end{cases} \tag{8.7}$$

with $W \ll 1$ and a fixed CFL number ν.

The formulations above demonstrate that all fully discrete schemes are second order accurate in numerical dispersion, except for the upwind flux with P^0, for which the error is of first order. Comparing the leading error terms with the same flux but with different p values, we can see that the temporal error would be dominant when $p \geq 2$ by upwind flux or central flux and $p \geq 1$ by alternating fluxes. Moreover, it is observed that the leading error terms for high order schemes, when the temporal error dominates, are the same with the FD schemes (6.5) and (6.6), which come from the time discretization (4.8). In particular, the leading terms of DG scheme with alternating flux are the same as those of FD scheme. Hence, we can also have counterintuitive results that the lower order scheme performs better than higher order scheme when numerical dispersion is concerned for some given dispersive media and discretization parameters. Similar results are observed for other numerical fluxes as well as for the trapezoidal-DG schemes in next section.

8.2. Fully discrete dispersion analysis: trapezoidal-DG schemes

We define

$$u_1 = \sqrt{\epsilon^{\mathrm{TP}}}\,E_0 + \frac{2}{W}\tan\left(\frac{W}{2}\right)H_0, \quad u_2 = \sqrt{\epsilon^{\mathrm{TP}}}\,E_0 - \frac{2}{W}\tan\left(\frac{W}{2}\right)H_0,$$

with $\epsilon^{\mathrm{TP}} = \epsilon(\widehat{\omega}^{\mathrm{TP}}; \mathbf{p}^{\mathrm{TP}})$ given in (4.12). And we can turn the problem in the similar form of (7.10):

$$< \mathcal{L}^{-}_{\epsilon^{\mathrm{TP}}} u_1, \phi > + \mathcal{R}^{-}(u_1, u_2, \phi; \alpha, \beta_1 \frac{W/2}{\tan(W/2)}, \beta_2 \frac{\tan(W/2)}{W/2}, \epsilon^{\mathrm{TP}}, e^{ik^{\mathrm{TP}}_{\mathrm{DG},p} h}) = 0 \tag{8.8a}$$

$$< \mathcal{L}^{+}_{\epsilon^{\mathrm{TP}}} u_2, \phi > + \mathcal{R}^{+}(u_1, u_2, \phi; \alpha, \beta_1 \frac{W/2}{\tan(W/2)}, \beta_2 \frac{\tan(W/2)}{W/2}, \epsilon^{\mathrm{TP}}, e^{ik^{\mathrm{TP}}_{\mathrm{DG},p} h}) = 0 \tag{8.8b}$$

Again, (7.15) is still valid for the fully discrete trapezoidal-DG scheme, with parameters as the following:

$$\epsilon \to \epsilon^{\mathrm{TP}}, \quad K \to \sqrt{\epsilon^{\mathrm{TP}}}\omega h, \quad \beta_1 \to \beta_1 \frac{W/2}{\tan(W/2)}, \quad \beta_2 \to \beta_2 \frac{\tan(W/2)}{W/2}.$$

In this case

$$a_p = (-1)^p \left(1 - 4(\alpha^2 + \beta_1 \beta_2)\right)$$

which is the same as that for semi-discrete scheme. Therefore, $k^{\mathrm{TP}}_{\mathrm{DG},p}$ are the roots of a quartic polynomial equation in terms of $\lambda = e^{ik^{\mathrm{TP}}_{\mathrm{DG},p} h}$ if $\alpha^2 + \beta_1\beta_2 \neq 1/4$, and $k_{\mathrm{DG},p}$ are the roots of a quadratic polynomial equation in terms of $\lambda = e^{ik^{\mathrm{TP}}_{\mathrm{DG},p} h}$ when $\alpha^2 + \beta_1\beta_2 = 1/4$. This is the same as semi-discrete results.

Below, we list the physical modes $k^{\mathrm{TP}}_{\mathrm{DG},p}$ for $p \geq 0$, and perform an asymptotic analysis when $W \ll 1$ and $K \ll 1$.

- When using the central flux, i.e. $\alpha = \beta_1 = \beta_2 = 0$, we have four solutions, and two of them correspond to the physical modes. When $W \ll 1$ and $K \ll 1$, the physical solutions have the form as

$$k^{\mathrm{TP}}_{\mathrm{DG\,phys},p} = \begin{cases} \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \frac{1}{6}K^2 + \mathcal{O}\left(W^4 + W^2K^2 + K^4\right)\right), & p = 0, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 - \frac{1}{48}K^2 + \mathcal{O}\left(W^4 + W^2K^2 + K^4\right)\right), & p = 1, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \mathcal{O}\left(W^4 + K^{2p+2}\right)\right), & p \geq 2, even, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \mathcal{O}\left(W^4 + K^{2p}\right)\right), & p \geq 2, odd, \end{cases} \tag{8.9}$$

  and can be further rewritten into

$$k^{\mathrm{TP}}_{\mathrm{DG\,phys},p} = \begin{cases} \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1 + \frac{2\epsilon(\widehat{\omega};\mathbf{p})}{\epsilon_\infty \nu^2}\right) W^2 + \mathcal{O}\left(W^4\right)\right), & p = 0, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1 - \frac{\epsilon(\widehat{\omega};\mathbf{p})}{4\epsilon_\infty \nu^2}\right) W^2 + \mathcal{O}\left(W^4\right)\right), & p = 1, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \mathcal{O}\left(W^4\right)\right), & p \geq 2, \end{cases} \tag{8.10}$$

  in the case of $W \ll 1$ and with a fixed CFL number ν.
- When using the alternating flux, i.e. $\alpha = \pm 1/2$ and $\beta_1 = \beta_2 = 0$, there are only two solutions corresponding to the physical modes,

$$k^{\mathrm{TP}}_{\mathrm{DG\,phys},p} = \begin{cases} \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \frac{1}{24}K^2 + \mathcal{O}\left(W^4 + W^2K^2 + K^4\right)\right), & p = 0, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \mathcal{O}\left(W^4 + K^{2p+2}\right)\right), & p \geq 1, \end{cases} \tag{8.11}$$

  in the case of $W \ll 1$ and $K \ll 1$, or

$$k^{\mathrm{TP}}_{\mathrm{DG\,phys},p} = \begin{cases} \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1 + \frac{\epsilon(\widehat{\omega};\mathbf{p})}{2\epsilon_\infty \nu^2}\right) W^2 + \mathcal{O}\left(W^4\right)\right), & p = 0, \\ \pm k^{\mathrm{ex}} \left(1 + \frac{1}{12}\left(\frac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})} + 1\right) W^2 + \mathcal{O}\left(W^4\right)\right), & p \geq 1, \end{cases} \tag{8.12}$$

  in the case of $W \ll 1$ and with a fixed CFL number ν.

Table 8.1
ν^p_{max} for DG scheme and ν^{2M}_{max} for FD scheme.

	$p=0$	$p=1$	$p=2$	$p=3$		$p\to\infty$
DG-CE	1	0.211325	0.101287	0.0605268		0
DG-AL	1	0.192450	0.089115	0.0521629		0
DG-UP	1	0.211325	0.101287	0.0605268		0
	$M=1$	$M=2$	$M=3$	$M=4$	$M=5$	$M\to\infty$
FD	1	0.857143	0.805369	0.777418	0.759479	0.636620

- When using the upwind flux, i.e. $\alpha=0$, $\beta_1=\frac{1}{2\sqrt{\epsilon_\infty}}$ and $\beta_2=\frac{\sqrt{\epsilon_\infty}}{2}$, there are only two solutions corresponding to the physical modes,

$$k^{\mathrm{TP}}_{\mathrm{DG\,phys},p}=\begin{cases}\pm k^{\mathrm{ex}}\left(1+\dfrac{1}{2}iB+\mathcal{O}\left(W^2+K^2\right)\right), & p=0,\\ \pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}+1\right)W^2+\mathcal{O}\left(W^4+iK^{2p}B\right)\right), & p\geq 1,\end{cases} \tag{8.13}$$

which can be rewritten as

$$k^{\mathrm{TP}}_{\mathrm{DG\,phys},p}=\begin{cases}\pm k^{\mathrm{ex}}\left(1+i\dfrac{b/\omega}{2\sqrt{\epsilon_\infty}\nu}W+\mathcal{O}\left(W^2\right)\right), & p=0,\\ \pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}+1\right)W^2+\mathcal{O}\left(iW^3\right)\right), & p=1,\\ \pm k^{\mathrm{ex}}\left(1+\dfrac{1}{12}\left(\dfrac{\delta(\widehat{\omega};\mathbf{p})}{\epsilon(\widehat{\omega};\mathbf{p})}+1\right)W^2+\mathcal{O}\left(W^4\right)\right), & p\geq 2,\end{cases} \tag{8.14}$$

in the case of $W\ll 1$ and with a fixed CFL number ν.

We have similar conclusions as those for the fully discrete leap-frog DG schemes, except that the leading error terms for high order schemes come from the fully implicit time discretization (4.14).

8.3. Comparison among fully discrete DG schemes

In our previous work [6], we have proved that the fully discrete DG schemes based on the trapezoidal rule is unconditionally stable and the leap-frog schemes are conditionally stable. Following the proof in [6], we can find ν^p_{max} such that under the condition $\nu\leq\nu^p_{max}$, the leap-frog schemes using P^p space are stable. Those ν^p_{max} values for $p=0,\ldots,3,\infty$ are listed in Table 8.1. For comparison, we also list the CFL condition for FD scheme for various M values in the same table. We can see that the CFL number for DG scheme is much smaller than that for FD scheme, particularly for high order case.

In Fig. 8.1, we plot the relative phase errors of fully discrete DG schemes with leap-frog discretization for $W_1=\pi/30,\pi/300$ using material parameters (4.16). We can observe that the overall behavior of the plot with $W_1=\pi/30$ is quite different from the FD plots and the plots obtained with $W_1=\pi/300$, and the magnitude of the errors is very large. This phenomenon results from $\omega_1 h=\dfrac{1}{\sqrt{\epsilon_\infty}\nu}W_1$ and the tiny CFL numbers restricted by the stability condition which makes the mesh size h extra large. We conclude that the small CFL number is one disadvantage of high order DG schemes. When comparing the figures obtained with $W_1=\pi/300$ using the three numerical fluxes, it is clear that the alternating flux has the smallest error, while the error obtained by the upwind flux is the largest. The overall dependence of the error on $\widehat{\omega}$ is very similar to those from the FD schemes.

Next, we consider the unconditionally stable DG scheme with trapezoidal rule and varying CFL numbers. Fig. 8.2 shows the contour plots of the dispersion error at $\widehat{\omega}=1$ with $(W_1,\omega_1 h)\in[0.05,0.3]\times[0.01,0.1]$. It is observed that DG-AL with $p\geq 1$ have horizontal contour lines, indicating dispersion errors are dominated by temporal ones. In comparison, DG-UP and DG-CE have horizontal contour lines when $p\geq 2$. The values of numerical dispersion errors obtained by high order DG schemes are very similar, which also illustrates the dominant role of temporal errors. This observation is consistent with our theoretical analysis.

9. Benchmark on physical quantities

In this section, we will verify the performance of the finite difference and discontinuous Galerkin methods by plotting quantities that are important for wave propagation such as the normalized ratio between the numerical and exact phase velocity (also refractive index); normalized attenuation constant; normalized energy velocity; and normalized group velocity, to validate the performance of the numerical methods (see [17]). The model parameters in (4.16) are used in computing all the quantities and plots below.

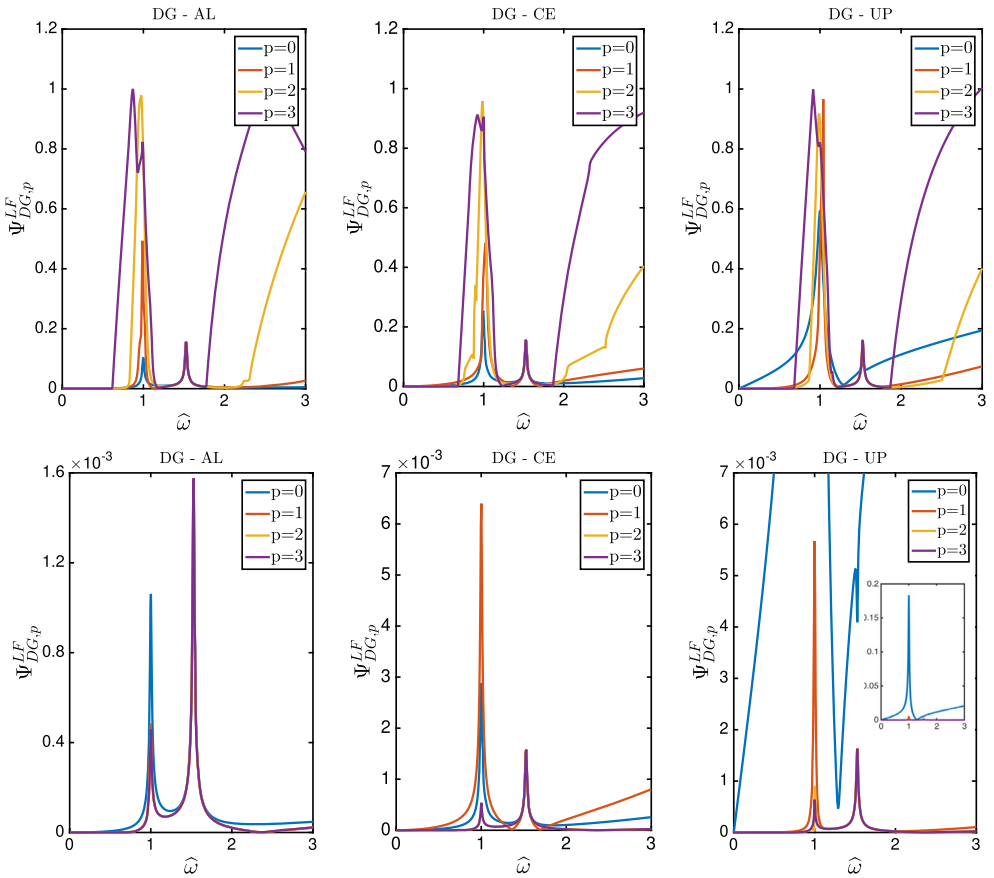

Fig. 8.1. The relative phase error in physical modes of the fully discrete DG schemes with leap-frog time discretization, using $\nu/\nu^p_{max} = 0.7$. First row: $W_1 = \pi/30$; second row: $W_1 = \pi/300$.

We first define ψ, given as

$$\psi = \frac{k}{\omega} = \sqrt{\epsilon(\widehat{\omega}; \mathbf{p})}. \tag{9.1}$$

We note that ψ is the complex index of refraction of the medium, whose real part is the real refractive index of the medium, whereas the imaginary part is related to the absorption or extinction coefficient [36]. We use $\Re$ and $\Im$ to denote the real and the imaginary parts of a complex number. Let the superscripts E and N denote the value of a quantity related to the exact solution of system (2.5) and a numerical approximation, respectively. We have the following definitions (see [17]):

- **Normalized Phase Velocity**: We consider the ratio between the real parts of the exact and numerical phase velocities, with the phase velocity, v_p, defined as $v_p = \omega/k = 1/\psi$. We define

$$\text{Normalized Phase Velocity} = \frac{\Re(v_p^N)}{\Re(v_p^E)}. \tag{9.2}$$

- **Normalized Attenuation Constant**: We consider the ratio between the imaginary parts of the exact and numerical ψ, which is also the ratio between the imaginary parts of the exact and numerical indices of refraction. We define

$$\text{Normalized Attenuation Constant} = \frac{\Im\left(\psi^N\right)}{\Im\left(\psi^E\right)}. \tag{9.3}$$

- **Normalized Energy Velocity**: The *velocity of energy transport* of a (monochromatic) plane-wave field is an important concept of wave propagation in a dispersive medium. In [36] this velocity is defined as a ratio of the time-average value of the Poynting vector to the total time-average electromagnetic energy density stored in both the field and the medium.

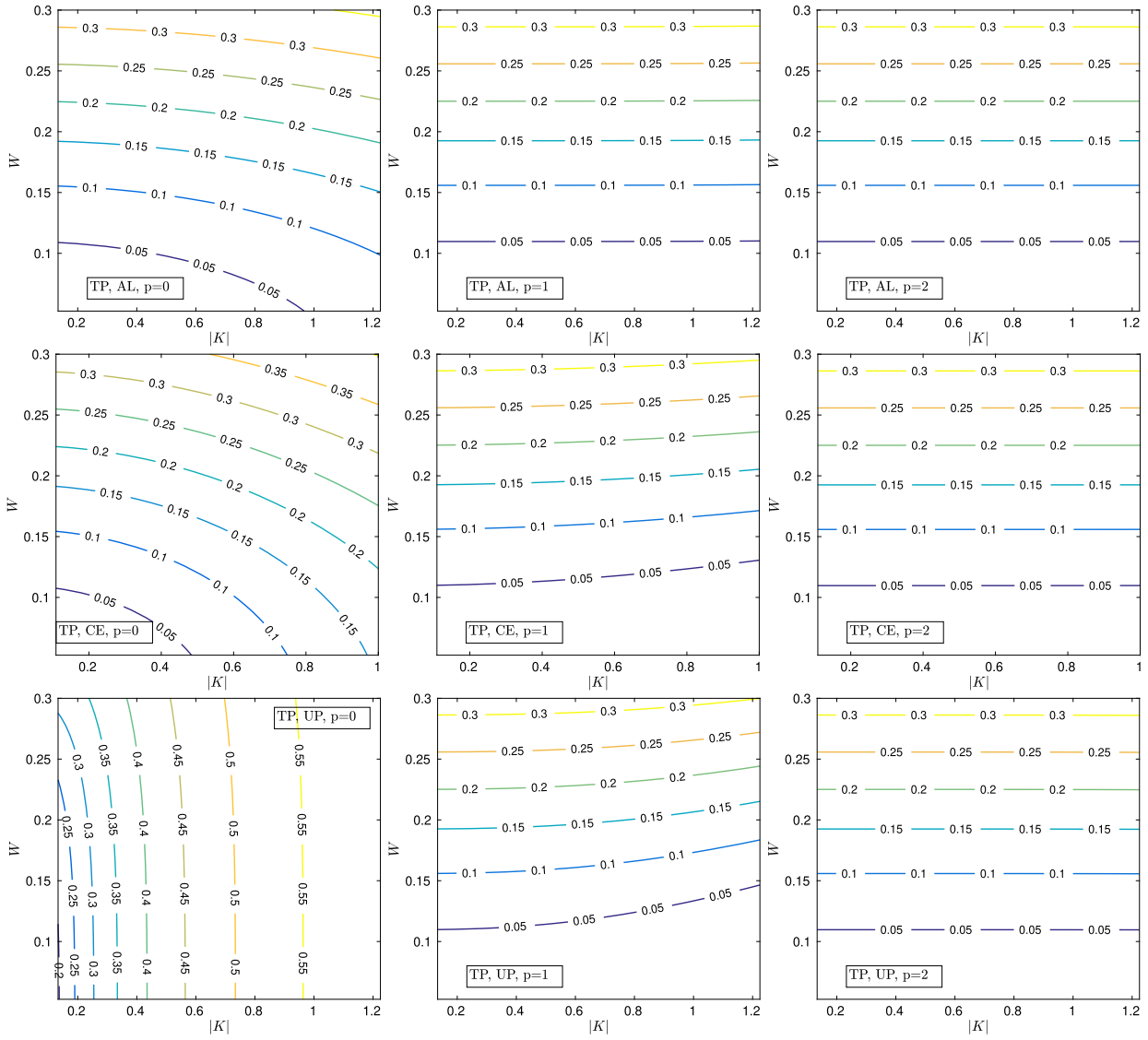

Fig. 8.2. The contour plot of relative phase error of fully discrete DG schemes for the physical modes with trapezoid rule. $\widehat{\omega} = 1$. First row: DG-AL; second row: DG-CE; third row: DG-UP.

The normalized energy velocity is a quantity that is defined (see [17,36]) as a function of the real and imaginary parts of the quantity ψ given as

$$\text{Energy Velocity} = \left[\Re(\psi) + \frac{\left(\Re\left(\psi^2\right) - \epsilon_s\right)\left(\Re\left(\psi^2\right) - \epsilon_\infty\right) + \left(\Im\left(\psi^2\right)\right)^2}{\left(\epsilon_s - \epsilon_\infty\right)\Re\left(\psi\right)} \right]^{-1}.$$

Based on the definition of the energy transport velocity, we define the ratio between the exact and numerical energy transport velocity to be the normalized energy velocity as

$$\text{Normalized Energy Velocity} = \frac{\text{Energy Velocity}^N}{\text{Energy Velocity}^E}. \tag{9.4}$$

- **Normalized Group Velocity**: We define the normalized group velocity to be the real part of the ratio of group velocities of the exact and numerical solutions. We have

$$\text{Normalized Group Velocity} = \Re\left(\frac{v_g^N}{v_g^E}\right), \tag{9.5}$$

where the group velocity is defined by $v_g = \frac{\partial \omega}{\partial k}$. Here, both v_g^E and v_g^N are obtained numerically by

$$(v_g)^{-1} = \frac{\partial k}{\partial \widehat{\omega}} \frac{\partial \widehat{\omega}}{\partial \omega} \approx \frac{k(\widehat{\omega} + 0.001) - k(\widehat{\omega})}{0.001} \frac{\partial \widehat{\omega}}{\partial \omega}.$$

In Figs. 9.1 and 9.2, we plot the four physical quantities defined in (9.2)-(9.5) for the leap-frog and trapezoidal FD schemes in various ranges of values for $\widehat{\omega}$: below resonance ($\widehat{\omega} < 1$), near resonance ($\widehat{\omega} \approx 1$), at the upper edge of the medium absorption band ($\widehat{\omega} \approx 1.527$), and far above resonance ($\widehat{\omega} > 3$). Fig. 9.1 offers excellent agreement with the plots in [17] for the (2,2) Yee scheme (leap-frog FDTD scheme with $M = 1$) and a (2,4) leap-frog FDTD scheme ($M = 2$). Both schemes have large errors at the resonance frequency and the upper edge of the medium absorption band $\widehat{\omega} = \sqrt{\epsilon_s/\epsilon_\infty}$. Higher order schemes have values for the physical quantities that are closer to 1, which indicates smaller dispersion error with increase in the spatial order of the scheme. We note that, while the increase in spatial order reduces the four physical quantities near resonance for both the leap-frog and trapezoidal FDTD methods, there is virtually no change with spatial order at the upper edge of the medium absorption band. This is also true for the DG schemes. A comparison between Figs. 9.1 and 9.2 suggests that the main differences between the two temporal discretizations can be observed for frequencies below and far beyond resonance. For $\widehat{\omega} < 1$, the plots obtained by the trapezoidal FDTD schemes are monotone. This is not the case for the leap-frog FDTD method as shown in Fig. 9.1. The results can be understood by comparing equations (4.8) with (4.14). The leading error coefficients in the two time schemes are different, with one being monotone on $\widehat{\omega}$ and the other not. For high frequencies, the leap-frog FDTD scheme can no longer resolve frequencies beyond 14.8, when the fields start to decay exponentially and have an increasing phase velocity. This number changes to around 10 for the trapezoidal scheme, which shows different resolution offered by the two temporal schemes.

In Fig. 9.3, we plot the four physical quantities defined in (9.2)-(9.5), obtained by DG-AL scheme using trapezoidal time discretization with a fixed CFL number $\nu = 0.7$. This choice is made based on previous observations that the DG-AL performs the best among all three fluxes. The overall behaviors of the physical quantities for FD and DG schemes are very similar when comparing Fig. 9.2 with Fig. 9.3. The main difference lies in the last column for high frequencies. The increasing resolution in the higher order DG scheme is evident, while increasing order does not impact this much for FD schemes. Thus, high order DG schemes in space can have a better performance in resolving high frequencies when compared with the FD scheme using the same mesh size. Similar conclusion holds with leap frog time discretization, and the plots are omitted for brevity.

10. Conclusions

In this paper, we studied the exact and numerical dispersion relations of a one-dimensional Maxwell's equations in a linear dispersive material characterized by a single pole Lorentz model for electronic polarization with low loss (i.e. when $\widehat{\gamma}$ is small). We consider two different high order spatial discretizations, the FD and DG methods, each coupled with two different second order temporal discretizations, leap-frog and trapezoidal integrators, to construct both semi-discrete and fully discrete schemes. In addition, for the DG schemes we have considered three different types of fluxes: central, upwind and alternating fluxes. Comparisons based on dispersion analysis are made of the FD and DG methods and the leap-frog and trapezoidal time discretizations.

It is well known that the FD and DG (which are a class of finite element methods) schemes, both being very popular discretizations, differ quite a lot in how they simulate wave phenomenon in their discrete grids. For example, DG schemes work well for multi-dimensional problems and can be constructed on unstructured meshes for complicated geometries. The FD schemes are simpler to code, and are mostly defined on structured meshes. The extension of the FD methods to non-uniform and unstructured meshes are cumbersome.

Both types of spatial discretizations can be designed with high spatial order accuracy. The FD scheme achieves this by extending the stencil of the discretization, while higher order polynomials are needed for the DG construction. When we express the dispersion relation for the discrete wavenumber as a function of the angular frequency ω, the number of spurious modes will increase with M (the accuracy order) of the FD scheme, while for DG schemes, the number of spurious modes is independent of p (the polynomial order). However, as shown in the Appendix of the FD scheme, using an alternative description of phase error, when the discrete angular frequency ω is expressed as a function of the wave number k the conclusions are reversed. Namely, there are no spurious modes for FD schemes, while more spurious modes will be present for higher order DG schemes, see [10] for relevant discussions in free space.

When comparing the order of numerical errors, the FD schemes manifest the same order of accuracy of the dispersion error and point-wise convergence error, while the DG schemes have higher order of accuracy in dispersion error than in the L^2 errors [13,14] (superconvergence in dispersion error). The CFL numbers for the two methods when coupled with an explicit time stepping are also different. It is known that the CFL number will approach a constant other than zero when $M \to \infty$ for the FD scheme, but the CFL number will go to zero when $p \to \infty$ for the DG scheme. Therefore, high order DG schemes require much smaller time steps than high order FD schemes.

Based on the numerical dispersion results in this paper, we observe that the physical dispersion of the material plays an important role in the numerical dispersion errors. For the low-loss materials considered, we can observe that the error is largest near the resonance frequency. This is no longer true for materials with high loss (i.e. when $\widehat{\gamma}$ is not small).

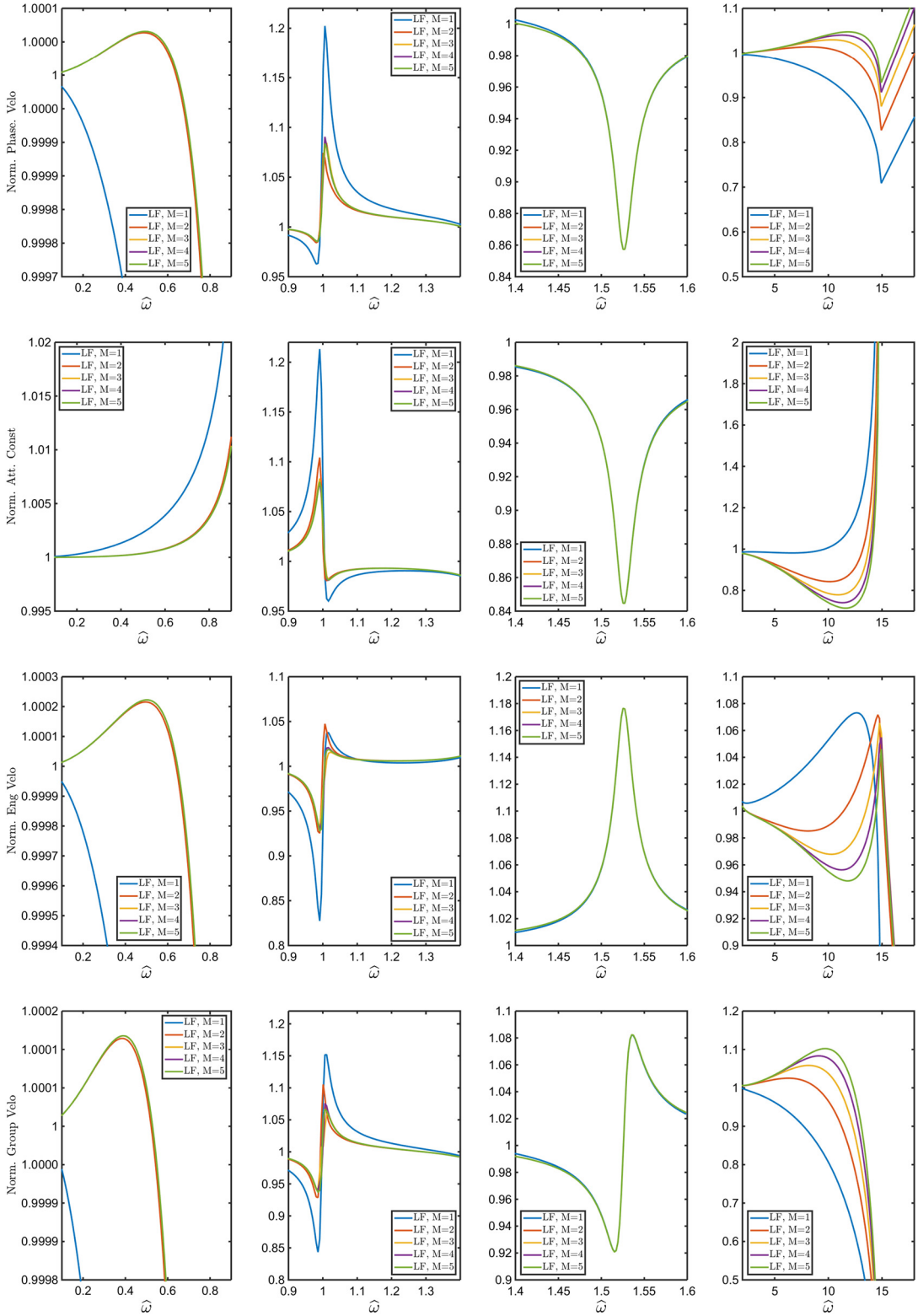

Fig. 9.1. Results for the leap-frog time discretization and FD2M with CFL number $\nu/\nu_{max}^{2M} = 0.7$. First row: normalized phase velocity; Second row: normalized attenuation constants; Third row: normalized energy velocity; Fourth row: normalized group velocity.

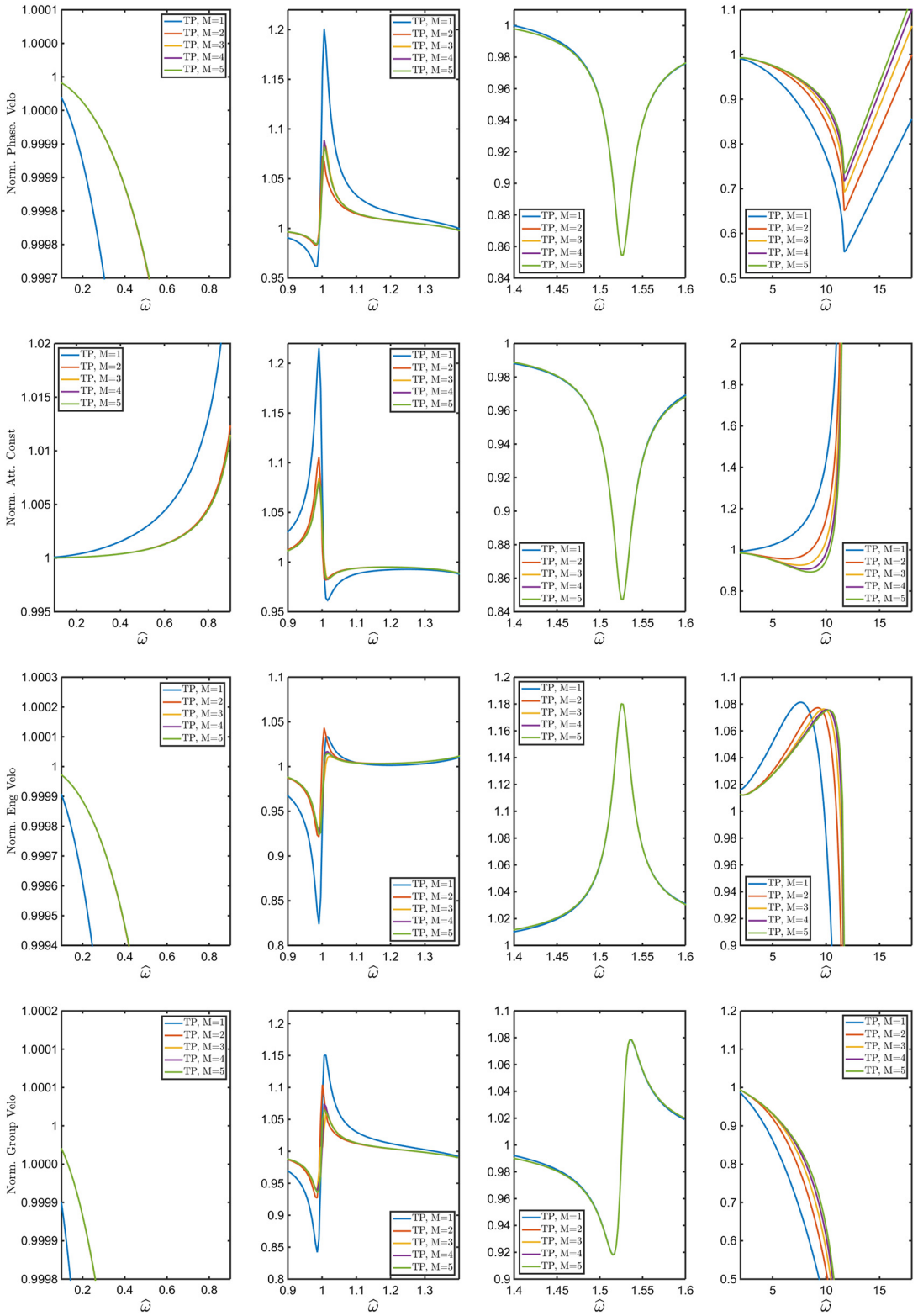

Fig. 9.2. Results for the trapezoidal time discretization and FD2M with CFL number $\nu/\nu_{max}^{2M} = 0.7$. First row: normalized phase velocity; Second row: normalized attenuation constants; Third row: normalized energy velocity; Fourth row: normalized group velocity.

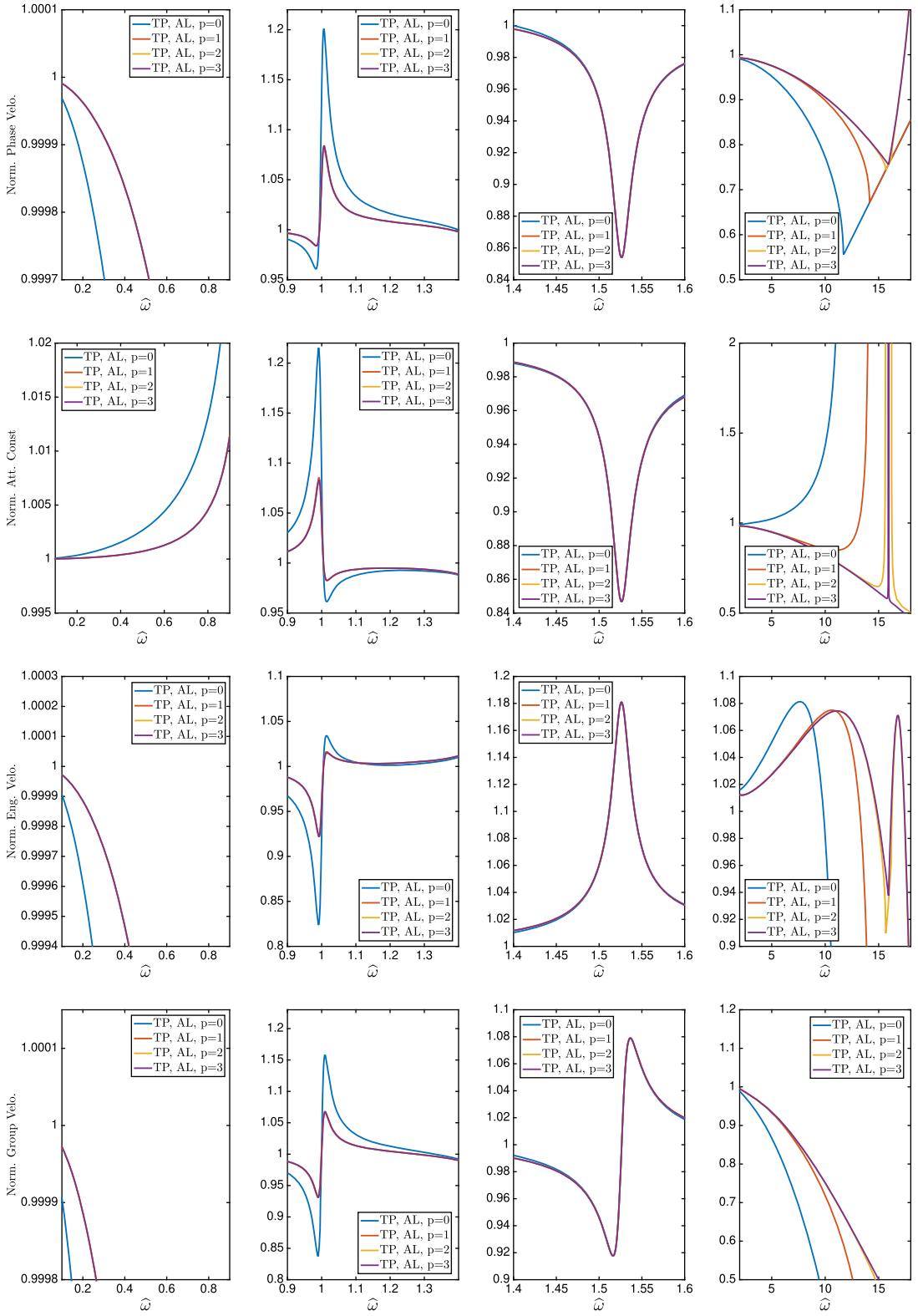

Fig. 9.3. Results for the trapezoidal time discretization and DG-AL with CFL number $\nu = 0.7$. First row: normalized phase velocity; Second row: normalized attenuation constants; Third row: normalized energy velocity; Fourth row: normalized group velocity.

An interesting finding is that for some materials and discretization parameters, we observe counterintuitive results that the dispersion error of a low order scheme can be potentially smaller than that of high order schemes (see for example Fig. 6.1). This demonstrates that the dispersion analysis conducted for free space may not be revealing for general dispersive media.

We find that the second order accuracy of the temporal discretizations limits the accuracy of the numerical dispersion errors, and is a good motivator for considering high order temporal discretizations, which are non-trivial to construct for the case of dispersive Maxwell models [48]. This limiting behavior in the medium absorption band is made clear by the difference in errors in the semi-discrete schemes versus the fully discrete schemes. In our future work we will investigate higher order temporal discretizations.

Appendix A. An alternative dispersion analysis for semi-discrete finite difference schemes

In this appendix, we provide an alternative method of analyzing the dispersion error of the semi-discrete in space high order FD schemes (FD2M). We express the discrete angular frequency ω as a function of the continuous wavenumber $k \in \mathbb{R}$, and measure the relative errors that result for different $M, M \in \mathbb{N}$, with $2M$ being the spatial accuracy of the schemes.

We introduce the following definitions

$$\widehat{k} := kh, \qquad F_{2M}(\widehat{k}) := 2\sum_{p=1}^{M} \frac{[(2p-3)!!]^2}{(2p-1)!} \sin^{2p-1}\left(\frac{\widehat{k}}{2}\right). \tag{A.1}$$

For the exact dispersion relation of Maxwell's equations in a one spatial dimensional Lorentz dielectric, by solving $\det(\mathcal{A}) = 0$ with $\mathcal{A}$ given by (3.3), we get the following quartic equation for the continuous angular frequency $\widehat{\omega}^{\text{ex}} = \omega^{\text{ex}}/\omega_1$,

$$(\widehat{\omega}^{\text{ex}})^4 + 2i\,\widehat{\gamma}\,(\widehat{\omega}^{\text{ex}})^3 - \frac{1}{\epsilon_\infty}\left(\epsilon_s + \frac{\widehat{k}^2}{(\omega_1 h)^2}\right)(\widehat{\omega}^{\text{ex}})^2 - \frac{2i}{\epsilon_\infty}\widehat{\gamma}\,\frac{\widehat{k}^2}{(\omega_1 h)^2}\,\widehat{\omega}^{\text{ex}} + \frac{1}{\epsilon_\infty}\frac{\widehat{k}^2}{(\omega_1 h)^2} = 0. \tag{A.2}$$

Similarly, considering the dispersion relation of semi-discrete FD2M scheme (5.10), we have

$$(\widehat{\omega}^{\text{FD},2M})^4 + 2i\,\widehat{\gamma}\,(\widehat{\omega}^{\text{FD},2M})^3 - \frac{1}{\epsilon_\infty}\left(\epsilon_s + \frac{F_{2M}(\widehat{k})^2}{(\omega_1 h)^2}\right)(\widehat{\omega}^{\text{FD},2M})^3 - \frac{2i}{\epsilon_\infty}\widehat{\gamma}\,\frac{F_{2M}(\widehat{k})^2}{(\omega_1 h)^2}\,\widehat{\omega}^{\text{FD},2M} + \frac{1}{\epsilon_\infty}\frac{F_{2M}(\widehat{k})^2}{(\omega_1 h)^2} = 0. \tag{A.3}$$

Clearly, both (A.2) and (A.3) have four (complex) roots each. Therefore, the FD scheme has no spurious modes for the discrete angular frequency. To better understand the errors, similar to previous sections, we first consider the lossless material ($\widehat{\gamma} = 0$) as an example. In this case, only even order terms appear in (A.2) and (A.3), and we can get

$$\widehat{\omega}^{\text{ex}}_{1,2}(\widehat{k}) = \pm\frac{1}{\sqrt{2}}\left[\frac{\epsilon_s}{\epsilon_\infty} + \frac{\widehat{k}^2}{\epsilon_\infty(\omega_1 h)^2} - \sqrt{\left(\frac{\epsilon_s}{\epsilon_\infty} + \frac{\widehat{k}^2}{\epsilon_\infty(\omega_1 h)^2}\right)^2 - \frac{4\widehat{k}^2}{\epsilon_\infty(\omega_1 h)^2}}\right]^{1/2}, \tag{A.4a}$$

$$\widehat{\omega}^{\text{ex}}_{3,4}(\widehat{k}) = \pm\frac{1}{\sqrt{2}}\left[\frac{\epsilon_s}{\epsilon_\infty} + \frac{\widehat{k}^2}{\epsilon_\infty(\omega_1 h)^2} + \sqrt{\left(\frac{\epsilon_s}{\epsilon_\infty} + \frac{\widehat{k}^2}{\epsilon_\infty(\omega_1 h)^2}\right)^2 - \frac{4\widehat{k}^2}{\epsilon_\infty(\omega_1 h)^2}}\right]^{1/2}, \tag{A.4b}$$

and $\widehat{\omega}^{\text{FD},2M}_{1,2}(\widehat{k}) = \widehat{\omega}^{\text{ex}}_{1,2}(F_{2M}(\widehat{k}))$, $\widehat{\omega}^{\text{FD},2M}_{3,4}(\widehat{k}) = \widehat{\omega}^{\text{ex}}_{3,4}(F_{2M}(\widehat{k}))$.

In Fig. A.1, we present the relative dispersion errors with $\widehat{k} \in [0, 2\pi]$ and the parameter values

$$\epsilon_s = 5.25, \quad \epsilon_\infty = 2.25, \quad \omega_1 h = \frac{\pi}{30}.$$

In this figure, we can observe the decrease of error when M (order of the scheme) increases. The numerical error in the first and second solutions of the discrete angular frequency, are smaller than that of the third and fourth solution, which can be understood if we consider the small wavenumber limit. In this case, we can derive expressions for the relative phase error as

$$\widehat{\Psi}_{\text{FD},2M}(\widehat{k}) := \left|\frac{\widehat{\omega}^{\text{ex}}_i(\widehat{k}) - \widehat{\omega}^{\text{FD},2M}_i(\widehat{k})}{\widehat{\omega}^{\text{ex}}_i(\widehat{k})}\right| = \begin{cases} \dfrac{[(2M-1)!!]^2}{2^{2M}(2M+1)!}\widehat{k}^{2M} + \mathcal{O}(\widehat{k}^{2M+2}), & i = 1,2, \\ \dfrac{[(2M-1)!!]^2}{2^{2M}(2M+1)!}\dfrac{\epsilon_d}{\epsilon_s^2}\dfrac{k^2}{\omega_1^2}\widehat{k}^{2M} + \mathcal{O}(\widehat{k}^{2M+2}), & i = 3,4, \end{cases} \tag{A.5}$$

which indicates a dispersion error of order $2M$ and is consistent to our previous conclusion (see Theorem 5.1). By comparing the coefficients, we verify that, for the parameters we consider, the leading error coefficient corresponding to $\widehat{\omega}^{\text{FD},2M}_{3,4}(\widehat{k})$ is indeed much larger than that for $\widehat{\omega}^{\text{FD},2M}_{1,2}(\widehat{k})$.

For low-loss material, e.g. $\widehat{\gamma} = 0.01$, the conclusions are very similar. The error plots show no visible difference from the no loss case, and are thus omitted.

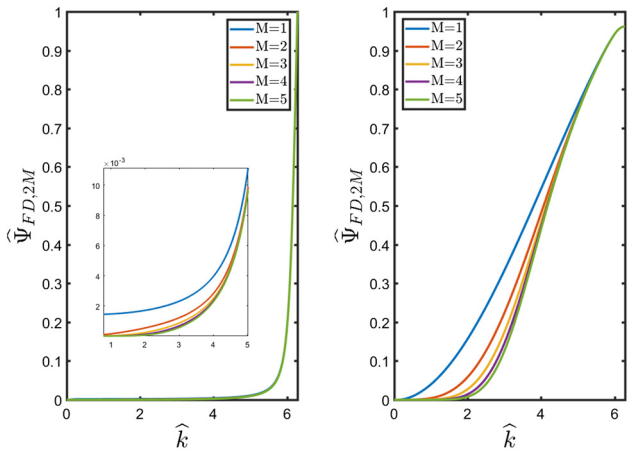

Fig. A.1. Relative phase error (A.5) for the spatial discretization FD2M with $\widehat{\gamma} = 0$. $\widehat{k} \in [0, 2\pi]$. Left: $i = 1, 2$; The inset in the left plot displays a zoomed-in region of the relative phase error for low values of $\widehat{k}$; Right: $i = 3, 4$.

References

[1] M. Ainsworth, Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods, J. Comput. Phys. 198 (2004) 106–130.
[2] M. Ainsworth, Dispersive behaviour of high order finite element schemes for the one-way wave equation, J. Comput. Phys. 259 (2014) 1–10.
[3] M. Ainsworth, G. Fu, Dispersive behavior of an energy-conserving discontinuous Galerkin method for the one-way wave equation, arXiv preprint, arXiv:1806.04306, 2018.
[4] M. Ainsworth, P. Monk, W. Muniz, Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation, J. Sci. Comput. 27 (2006) 5–40.
[5] H. Banks, V. Bokil, N. Gibson, Analysis of stability and dispersion in a finite element method for Debye and Lorentz dispersive media, Numer. Methods Partial Differ. Equ. 25 (2009) 885–917.
[6] V.A. Bokil, Y. Cheng, Y. Jiang, F. Li, Energy stable discontinuous Galerkin methods for Maxwell's equations in nonlinear optical media, J. Comput. Phys. 350 (2017) 420–452.
[7] V.A. Bokil, Y. Cheng, Y. Jiang, F. Li, P. Sakkaplangkul, High spatial order energy stable FDTD methods for Maxwell's equations in nonlinear optical media in one dimension, J. Sci. Comput. (2018) 1–42.
[8] V.A. Bokil, N. Gibson, Analysis of spatial high-order finite difference methods for Maxwell's equations in dispersive media, IMA J. Numer. Anal. 32 (2012) 926–956.
[9] A. Bourgeade, B. Nkonga, Numerical modeling of laser pulse behavior in nonlinear crystal and application to the second harmonic generation, Multiscale Model. Simul. 4 (2005) 1059–1090.
[10] Y. Cheng, C.-S. Chou, F. Li, Y. Xing, L^2 stable discontinuous Galerkin methods for one-dimensional two-way wave equations, Math. Comput. 86 (2017) 121–155.
[11] E.T. Chung, P. Ciarlet, T.F. Yu, Convergence and superconvergence of staggered discontinuous Galerkin methods for the three-dimensional Maxwell's equations on Cartesian grids, J. Comput. Phys. 235 (2013) 14–31.
[12] B. Cockburn, F. Li, C.-W. Shu, Locally divergence-free discontinuous Galerkin methods for the Maxwell equations, J. Comput. Phys. 194 (2004) 588–610.
[13] G. Cohen, Higher-Order Numerical Methods for Transient Wave Equations, Springer Science & Business Media, 2013.
[14] G. Cohen, S. Pernet, Finite Element and Discontinuous Galerkin Methods for Transient Wave Equations, Springer, 2016.
[15] M. Fujii, M. Tahara, I. Sakagami, W. Freude, P. Russer, High-order FDTD and auxiliary differential equation formulation of optical pulse propagation in 2-D Kerr and Raman nonlinear dispersive media, IEEE J. Quantum Electron. 40 (2004) 175–182.
[16] S.D. Gedney, J.C. Young, T.C. Kramer, J.A. Roden, A discontinuous Galerkin finite element time-domain method modeling of dispersive media, IEEE Trans. Antennas Propag. 60 (2012) 1969–1977.
[17] L. Gilles, S. Hagness, L. Vázquez, Comparison between staggered and unstaggered finite-difference time-domain grids for few-cycle temporal optical soliton propagation, J. Comput. Phys. 161 (2000) 379–400.
[18] P.M. Goorjian, A. Taflove, R.M. Joseph, S.C. Hagness, Computational modeling of femtosecond optical solitons from Maxwell's equations, IEEE J. Quantum Electron. 28 (1992) 2416–2422.
[19] J.H. Greene, A. Taflove, General vector auxiliary differential equation finite-difference time-domain method for nonlinear optics, Opt. Express 14 (2006) 8305–8310.
[20] J.S. Hesthaven, T. Warburton, Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations, J. Comput. Phys. 181 (2002) 186–221.
[21] C.V. Hile, W.L. Kath, Numerical solutions of Maxwell's equations for nonlinear-optical pulse propagation, J. Opt. Soc. Am. B 13 (1996) 1135–1145.
[22] F.Q. Hu, H.L. Atkins, Eigensolution analysis of the discontinuous Galerkin method with nonuniform grids: I. One space dimension, J. Comput. Phys. 182 (2002) 516–545.
[23] F.Q. Hu, M. Hussaini, P. Rasetarinera, An analysis of the discontinuous Galerkin method for wave propagation problems, J. Comput. Phys. 151 (1999) 921–946.
[24] Y. Huang, J. Li, W. Yang, Interior penalty DG methods for Maxwell's equations in dispersive media, J. Comput. Phys. 230 (2011) 4559–4570.
[25] R.M. Joseph, S.C. Hagness, A. Taflove, Direct time integration of Maxwell's equations in linear dispersive media with absorption for scattering and propagation of femtosecond electromagnetic pulses, Opt. Lett. 16 (1991) 1412–1414.
[26] R.M. Joseph, A. Taflove, Spatial soliton deflection mechanism indicated by FD-TD Maxwell's equations modeling, IEEE Photonics Technol. Lett. 6 (1994) 1251–1254.
[27] R.M. Joseph, A. Taflove, FDTD Maxwell's equations models for nonlinear electrodynamics and optics, IEEE Trans. Antennas Propag. 45 (1997) 364–374.
[28] R.M. Joseph, A. Taflove, P.M. Goorjian, Direct time integration of Maxwell's equations in two-dimensional dielectric waveguides for propagation and scattering of femtosecond electromagnetic solitons, Opt. Lett. 18 (1993) 491–493.

[29] T. Kashiwa, I. Fukai, A treatment by the FD-TD method of the dispersive characteristics associated with electronic polarization, Microw. Opt. Technol. Lett. 3 (1990) 203–205.
[30] T. Kashiwa, N. Yoshida, I. Fukai, A treatment by the finite-difference time domain method of the dispersive characteristics associated with orientational polarization, IEEE Trans. IEICE 73 (1990) 1326–1328.
[31] S. Lanteri, C. Scheid, Convergence of a discontinuous Galerkin scheme for the mixed time-domain Maxwell's equations in dispersive media, IMA J. Numer. Anal. 33 (2013).
[32] J.-F. Lee, R. Lee, A. Cangellaris, Time-domain finite-element methods, IEEE Trans. Antennas Propag. 45 (1997) 430–442.
[33] T. Lu, P. Zhang, W. Cai, Discontinuous Galerkin methods for dispersive and lossy Maxwell's equations and PML boundary conditions, J. Comput. Phys. 200 (2004) 549–580.
[34] P. Monk, A comparison of three mixed methods for the time-dependent Maxwell's equations, SIAM J. Sci. Stat. Comput. 13 (1992) 1097–1122.
[35] W. Mulder, Spurious modes in finite-element discretizations of the wave equation may not be all that bad, Appl. Numer. Math. 30 (1999) 425–445.
[36] K.E. Oughstun, S. Shen, Velocity of energy transport for a time-harmonic field in a multiple-resonance Lorentz medium, J. Opt. Soc. Am. B 5 (1988) 2395–2398.
[37] P.G. Petropoulos, Stability and phase error analysis of FD-TD in dispersive dielectrics, IEEE Trans. Antennas Propag. 42 (1994) 62–69.
[38] K. Prokopidis, E. Kosmidou, T. Tsiboukis, An FDTD algorithm for wave propagation in dispersive media using higher-order schemes, J. Electromagn. Waves Appl. 18 (2004) 1171–1194.
[39] O. Ramadan, Systematic wave-equation finite difference time domain formulations for modeling electromagnetic wave-propagation in general linear and nonlinear dispersive materials, Int. J. Mod. Phys. C 26 (2015) 1550046.
[40] D. Sármány, M.A. Botchev, J.J. van der Vegt, Dispersion and dissipation error in high-order Runge-Kutta discontinuous Galerkin discretisations of the Maxwell equations, J. Sci. Comput. 33 (2007) 47–74.
[41] S. Sherwin, Dispersion analysis of the continuous and discontinuous Galerkin formulations, in: Discontinuous Galerkin Methods, Springer, 2000, pp. 425–431.
[42] M.P. Sørensen, G.M. Webb, M. Brio, J.V. Moloney, Kink shape solutions of the Maxwell-Lorentz system, Phys. Rev. E 71 (2005) 036602.
[43] A. Taflove, S.C. Hagness, Computational Electrodynamics: The Finite-Difference Time-Domain Method, Artech House, 2005.
[44] A. Taflove, A. Oskooi, S.G. Johnson, Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology, Artech House, 2013.
[45] L.N. Trefethen, Group velocity in finite difference schemes, SIAM Rev. 24 (1982) 113–136.
[46] H. Yang, F. Li, J. Qiu, Dispersion and dissipation errors of two fully discrete discontinuous Galerkin methods, J. Sci. Comput. 55 (2013) 552–574.
[47] K. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE Trans. Antennas Propag. 14 (1966) 302–307.
[48] J. Young, A higher order FDTD method for EM propagation in a collisionless cold plasma, IEEE Trans. Antennas Propag. 44 (1996) 1283–1289.
[49] S. Zhao, On the spurious solutions in the high-order finite difference methods for eigenvalue problems, Comput. Methods Appl. Mech. Eng. 196 (2007) 5031–5046.
[50] R.W. Ziolkowski, J.B. Judkins, Nonlinear finite-difference time-domain modeling of linear and nonlinear corrugated waveguides, J. Opt. Soc. Am. B 11 (1994) 1565–1575.