Communications on Applied Mathematics and Computation
https://doi.org/10.1007/s42967-022-00212-2

ORIGINAL PAPER

Energy Stable Nodal DG Methods for Maxwell's Equations of Mixed-Order Form in Nonlinear Optical Media

Maohui Lyu[1] · Vrushali A. Bokil[2] · Yingda Cheng[3,4] · Fengyan Li[5]

Received: 26 March 2022 / Revised: 13 July 2022 / Accepted: 20 August 2022

Abstract
In this work, we develop energy stable numerical methods to simulate electromagnetic waves propagating in optical media where the media responses include the linear Lorentz dispersion, the instantaneous nonlinear cubic Kerr response, and the nonlinear delayed Raman molecular vibrational response. Unlike the first-order PDE-ODE governing equations considered previously in Bokil et al. (J Comput Phys 350: 420–452, 2017) and Lyu et al. (J Sci Comput 89: 1–42, 2021), a model of mixed-order form is adopted here that consists of the first-order PDE part for Maxwell's equations coupled with the second-order ODE part (i.e., the auxiliary differential equations) modeling the linear and nonlinear dispersion in the material. The main contribution is a new numerical strategy to treat the Kerr and Raman nonlinearities to achieve provable energy stability property within a second-order temporal discretization. A nodal discontinuous Galerkin (DG) method is further applied in space for efficiently handling nonlinear terms at the algebraic level, while preserving the energy stability and achieving high-order accuracy. Indeed with d_E as the number of the components of the electric field, only a $d_E \times d_E$ nonlinear algebraic system needs to be solved at each interpolation node, and more importantly, all these small nonlinear systems are completely decoupled over one time step, rendering very high parallel efficiency. We evaluate the proposed schemes by comparing them with the methods in Bokil et al. (2017) and Lyu et al. (2021) (implemented in nodal form) regarding the accuracy, computational efficiency, and energy stability, by a parallel scalability study, and also through the simulations of the soliton-like wave propagation in one dimension, as well as the spatial-soliton propagation and two-beam interactions modeled by the two-dimensional transverse electric (TE) mode of the equations.

Keywords Maxwell's equations · Kerr and Raman · Discontinuous Galerkin method · Energy stability

The research of Maohui Lyu is supported by China Postdoctoral Science Foundation grant 2020TQ0344, and the NSFC grants 11871139 and 12101597. The research of Vrushali A. Bokil is supported by the NSF grants DMS-1720116 and DMS-2012882. The research of Yingda Cheng is supported by the NSF grant DMS-2011838. The research of Fengyan Li is supported by the NSF grants DMS-1719942 and DMS-1913072.

✉ Fengyan Li
lif@rpi.edu

Extended author information available on the last page of the article

Mathematics Subject Classification 65M60 · 78M10

1 Introduction

In recent decades, there has been growing interest in the studies of the behavior of light when it interacts with optical media on the nanometer scale. These studies focus on not only the linear responses of the media, but also the nonlinear ones that contribute to phenomena such as high-order harmonic generations and frequency mixing, among others [6]. The nonlinear optical responses, though being a rather weak optical process, become more important at very high light intensities and arise in various practical applications. For example, in the laser industry, second harmonic generation is used to make green 532 nm laser from a 1 064 nm source [18]. It is well known that the dynamics of the light-media interactions is governed by the classical Maxwell's equations. The linear and nonlinear responses of the media can be described either at a microscopic level through the motion of charged particles in the media driven by electromagnetic fields, or at a macroscopic level through some phenomenological constitutive relations [1, 6, 15].

In this work, we consider Maxwell's equations along with a macroscopic description of the media responses, defined on a spatial domain $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ over a time interval $[0, T]$,

$$\mu_0 \partial_t \boldsymbol{H} + \nabla \times \boldsymbol{E} = \boldsymbol{0}, \tag{1a}$$

$$\partial_t \boldsymbol{D} - \nabla \times \boldsymbol{H} = \boldsymbol{0}, \tag{1b}$$

$$\boldsymbol{D} = \epsilon_0 \Big(\epsilon_\infty \boldsymbol{E} + \boldsymbol{P} + a(1-\theta)|\boldsymbol{E}|^2 \boldsymbol{E} + a\theta Q \boldsymbol{E} \Big), \tag{1c}$$

$$\partial_{tt} \boldsymbol{P} + \gamma \partial_t \boldsymbol{P} + \omega_0^2 \boldsymbol{P} = \omega_p^2 \boldsymbol{E}, \tag{1d}$$

$$\partial_{tt} Q + \gamma_v \partial_t Q + \omega_v^2 Q = \omega_v^2 |\boldsymbol{E}|^2 \tag{1e}$$

supplemented with boundary conditions and compatible initial conditions

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{x}, 0) =& \boldsymbol{H}_0(\boldsymbol{x}), \quad \boldsymbol{E}(\boldsymbol{x}, 0) = \boldsymbol{E}_0(\boldsymbol{x}), \quad \boldsymbol{P}(\boldsymbol{x}, 0) = \boldsymbol{P}_0(\boldsymbol{x}), \\ \partial_t \boldsymbol{P}(\boldsymbol{x}, 0) =& \boldsymbol{J}_0(\boldsymbol{x}), \quad Q(\boldsymbol{x}, 0) = Q_0(\boldsymbol{x}), \quad \partial_t Q(\boldsymbol{x}, 0) = \sigma_0(\boldsymbol{x}). \end{aligned}$$

Here, $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic fields, respectively. The dielectric parameters include the magnetic permeability μ_0 of free space, the relative electric permittivity ϵ_∞ in the limit of the infinite frequency, and the electric permittivity ϵ_0 of free space. The electric flux density $\boldsymbol{D}$ is related to the electric field through the constitutive law (1c), that models the linear and nonlinear optical responses of the media. Particularly, $\epsilon_0 \epsilon_\infty \boldsymbol{E}$ is for the instantaneous linear response and $\epsilon_0 a(1-\theta)|\boldsymbol{E}|^2 \boldsymbol{E}$ is for the instantaneous cubic nonlinear Kerr effect, while the delayed dispersive effects include the linear Lorentz response $\epsilon_0 \boldsymbol{P}$ and the nonlinear Raman molecular vibrational response $\epsilon_0 a\theta Q \boldsymbol{E}$, with each further modeled through the auxiliary differential equations (1d) and (1e), respectively. Moreover, in (1d), ω_0 and ω_p are the resonance and plasma frequencies of the medium, respectively. Additionally, γ is a damping constant. In (1e), ω_v is the resonance frequency of the

vibration, and γ_v is a damping constant. The constants a and θ measure the strength and relative strength of the Kerr and Raman nonlinearities. One can refer to [1, 4] for more discussions of the model.

The model (1) is said to be in the *mixed-order form* as it consists of the first-order PDE part of the Maxwell's equations (1a)–(1b) and the second-order ODE part of the auxiliary differential equations (1d)–(1e) for the linear and nonlinear dispersive responses. For model (1), one can derive the following energy relation under periodic boundary conditions:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(t) = -\frac{\epsilon_0\gamma}{\omega_p^2}\int_\Omega |\partial_t \boldsymbol{P}|^2 \mathrm{d}\Omega - \frac{\epsilon_0 a\theta\gamma_v}{2\omega_v^2}\int_\Omega (\partial_t Q)^2 \mathrm{d}\Omega \leqslant 0, \tag{2}$$

where the energy $\mathcal{E}(t)$ is defined as

$$\begin{aligned}\mathcal{E}(t) = \int_\Omega \Bigg(&\frac{\mu_0}{2}|\boldsymbol{H}|^2 + \frac{\epsilon_0\epsilon_\infty}{2}|\boldsymbol{E}|^2 + \frac{\epsilon_0}{2\omega_p^2}|\partial_t \boldsymbol{P}|^2 + \frac{\epsilon_0\omega_0^2}{2\omega_p^2}|\boldsymbol{P}|^2 \\ &+ \frac{\epsilon_0 a\theta}{4\omega_v^2}(\partial_t Q)^2 + \frac{\epsilon_0 a\theta}{2}Q|\boldsymbol{E}|^2 + \frac{3\epsilon_0 a(1-\theta)}{4}|\boldsymbol{E}|^4 + \frac{\epsilon_0 a\theta}{4}Q^2\Bigg)\mathrm{d}\Omega.\end{aligned} \tag{3}$$

If we further assume $\theta \in [0, \frac{3}{4}]$, then

$$\frac{\epsilon_0 a\theta}{2}Q|\boldsymbol{E}|^2 + \frac{3\epsilon_0 a(1-\theta)}{4}|\boldsymbol{E}|^4 + \frac{\epsilon_0 a\theta}{4}Q^2 = \frac{\epsilon_0 a\theta}{4}\left(|\boldsymbol{E}|^2 + Q\right)^2 + \frac{\epsilon_0 a(3-4\theta)}{4}|\boldsymbol{E}|^4 \geqslant 0,$$

and $\mathcal{E}(t) \geqslant 0$. The energy relation, therefore, gives an energy stability for the model. When the model is damping free with $\gamma = \gamma_v = 0$, the energy relation leads to energy conservation.

In this work, we propose and analyze a family of numerical schemes for the mixed-order model (1), that is accurate, efficient, and at the same time satisfies a provable discrete analogue of the energy relation and stability (2). The methods are second-order accurate in time and can be arbitrary order accurate in space. They can be implemented very efficiently especially in the parallel setting, as the nonlinear algebraic system to solve per time step is completely decoupled into small sub-systems, with each being of size $d_E \times d_E$ at one (interpolation) point in space. Here, d_E is the number of the components of the electric field that often but not always coincides with d. The main ingredients to design the methods are: (i) novel temporal treatments of the nonlinear terms (i.e., $|\boldsymbol{E}|^2\boldsymbol{E}$ and $Q\boldsymbol{E}$ in (1c), $|\boldsymbol{E}|^2$ in (1e)) along with the leap-frog time discretization for the PDE part and a central difference based two-step time discretization for the ODE part of the model, (ii) nodal discontinuous Galerkin (DG) methods in space with alternating numerical fluxes and with nonlinear terms handled in an interpolatory fashion [8, 14]. The methods satisfy a provable energy relation with respect to the discrete energy, and the result further gives the energy stability under a time step condition in addition to $\theta \in [0, \frac{3}{4}]$ as for the continuous model. It is worth mentioning that the time step condition is the same for the L^2 stability when the methods are applied to the model in the absence of the Kerr and Raman nonlinear effects as well as the linear Lorentz dispersion. DG methods are chosen due to their many attractive features, such as excellent numerical dispersive and dissipative properties in standard wave simulations, flexibility in adaptive implementation and high parallelization, and suitability for complicated geometry etc., [14]. Moreover, with the adoption of the nodal formulation of DG methods, the nonlinearities can be handled very efficiently (see Sect. 2.3 for more discussions).

There have been extensive research activities in devising finite difference methods, finite element methods, and DG methods to simulate Maxwell's equations in linear or nonlinear dispersive optical media (e.g., [9–13]). To our best knowledge, the methods proposed here are the first one to satisfy a provable energy relation/stability based on the *mixed-order form* of (1) that model the electromagnetic waves in optical media with the linear Lorentz dispersion, and nonlinear Kerr and Raman effects. The work is motivated by the previous developments of designing numerical methods with provable energy relation/stability when the underlying equations are entirely in the *first-order form*, that is, with (1d)–(1e) reformulated into

$$\partial_t \boldsymbol{P} = \boldsymbol{J}, \quad \partial_t \boldsymbol{J} + \gamma \boldsymbol{J} + \omega_0^2 \boldsymbol{P} = \omega_p^2 \boldsymbol{E}, \tag{4a}$$

$$\partial_t Q = \sigma, \quad \partial_t \sigma + \gamma_v \sigma + \omega_v^2 Q = \omega_v^2 |\boldsymbol{E}|^2. \tag{4b}$$

Examples of such efforts for the first-order form model (i.e., (1a)–(1c), (4)) include energy stable DG schemes in [4] and FDTD schemes in [5] in one dimension, and energy stable DG schemes in [17] in higher dimensions. The key for these methods to achieve a provable energy relation lies in the treatments of the nonlinear terms, with the one for the Kerr term most nontrivial. Particularly, the strategies in [4, 5, 17] all start with an auxiliary vector $\boldsymbol{Y} = |\boldsymbol{E}|^2 \boldsymbol{E}$ and the discretization is designed based on the differentiated form $\partial_t \boldsymbol{Y} = \partial_t (|\boldsymbol{E}|^2 \boldsymbol{E})$. When the media responses only contain the nonlinear Kerr effect (i.e., without the linear Lorentz response and $\theta = 0$), the idea above is also applied to an FDTD method for the transverse magnetic (TM) mode of Maxwell's equations in two dimensions and extended to a finite element method in three dimensions.

In relation to the work reviewed above, our proposed methods here are closest to those in [17], in their use of nodal DG spatial discretizations and the leap-frog method in time for the PDE part. The nonlinear Kerr term is handled similarly as in [4, 17] based on $\partial_t \boldsymbol{Y} = \partial_t (|\boldsymbol{E}|^2 \boldsymbol{E})$, yet with additional adaptation to suit the two-step temporal discretization setting. For the nonlinear Raman term $\boldsymbol{QE}$, unlike the algebraic strategy in [4, 5, 17] that works for the first-order form of the model, the model of the mixed-order form in this work together with the two-step temporal discretization strategy call for a second auxiliary vector $\boldsymbol{W} = \boldsymbol{QE}$, with the proposed numerical treatment based on the differentiated form $\partial_t \boldsymbol{W} = \partial_t (\boldsymbol{QE})$. Relatedly, a special discretization is employed to the nonlinear driving term $|\boldsymbol{E}|^2$ in (1e). The proposed discretizations for the nonlinear Raman term is one main contribution of this work. As one will see in Sect. 2.3, we will not directly solve and store $\boldsymbol{Y}$ and $\boldsymbol{W}$ in actual implementation.

The remainder of the paper is organized as follows. In Sect. 2, we formulate the proposed numerical schemes. Particularly, a second-order temporal discretization is presented in Sect. 2.1 with special attention given to the nonlinear terms; in Sect. 2.2, we further apply a family of nodal DG methods in space, illustrated using the two-dimensional (2D) model with the transverse electric (TE) mode. For both the semi-discrete in time and fully discrete schemes, the energy relation and stability are established. In Sect. 2.3, some discussions are made for implementation. In Sect. 3, the performance of the proposed schemes is demonstrated through a comparison with the methods in [4, 17] (implemented in nodal form) regarding the accuracy, computational efficiency, and energy stability, through a parallel scalability study, and also through simulations of the soliton-like wave propagation in one dimension, as well as the spatial-soliton propagation and two-beam interactions in two dimensions. Concluding remarks follow in Sect. 4.

2 Numerical Methods

In this section, we formulate and analyze the proposed numerical schemes for Maxwell's equations in nonlinear optical media in (1). Particularly, a second-order time discretization will be presented in Sect. 2.1, with the focus on handling the nonlinear terms to achieve a provable energy relation/stability. With nodal DG methods (of any formal accuracy) further applied to space, we will come to the fully discrete schemes in Sect. 2.2. Energy stability relations will be established for both the semi-discrete in time and fully discrete schemes. Some aspects of implementation will be discussed in Sect. 2.3.

2.1 Time Discretization

We start with some notation. Let $0 = t^0 < t^1 < \cdots < t^{N_t} = T$ be a uniform partition of $[0, T]$ with $t^n = n\Delta t$, $\Delta t = T/N_t$. For a grid function u with u^n as its value at t^n, we define

$$\overline{u^{n+1/2}} = \frac{u^{n+1} + u^n}{2}, \quad \overline{\overline{u^n}} = \frac{u^{n+1} + 2u^n + u^{n-1}}{4}.$$

Our proposed temporal discretization is second-order accurate, and it involves the leap-frog method for the PDE part of the first-order form in (1a)–(1b), as well as a central difference based two-step method for the ODE part of the second-order form in (1d)–(1e), with special attention paid to the nonlinear terms, namely, $|\boldsymbol{E}|^2\boldsymbol{E}$ and $Q\boldsymbol{E}$ in (1c), $|\boldsymbol{E}|^2$ in (1e), that are related to the nonlinear Kerr and Raman effects. Particularly, given u^{n-1}, u^n, we look for u^{n+1}, with $u = \boldsymbol{H}, \boldsymbol{E}, \boldsymbol{D}, \boldsymbol{P}$, and Q, satisfying

$$\mu_0 \frac{\boldsymbol{H}^{n+1/2} - \boldsymbol{H}^n}{\Delta t/2} = -\nabla\times\boldsymbol{E}^n, \tag{5a}$$

$$\frac{\boldsymbol{D}^{n+1} - \boldsymbol{D}^n}{\Delta t} = \nabla\times\boldsymbol{H}^{n+1/2}, \tag{5b}$$

$$\boldsymbol{D}^{n+1} = \epsilon_0\Big(\epsilon_\infty \boldsymbol{E}^{n+1} + \boldsymbol{P}^{n+1} + a(1-\theta)\boldsymbol{Y}^{n+1} + a\theta\boldsymbol{W}^{n+1}\Big), \tag{5c}$$

$$\begin{aligned}\boldsymbol{Y}^{n+1} - \boldsymbol{Y}^{n-1} = &\left(\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2 + \left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2 - \overline{\boldsymbol{E}^{n+1/2}}\cdot\overline{\boldsymbol{E}^{n-1/2}}\right)(\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n-1}) \\ &+ 2\left(\overline{\overline{\boldsymbol{E}^n}}\cdot(\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n-1})\right)\overline{\overline{\boldsymbol{E}^n}},\end{aligned} \tag{5d}$$

$$\boldsymbol{W}^{n+1} - \boldsymbol{W}^{n-1} = 2\Big(\overline{Q^{n+1/2}}\,\overline{\boldsymbol{E}^{n+1/2}} - \overline{Q^{n-1/2}}\,\overline{\boldsymbol{E}^{n-1/2}}\Big), \tag{5e}$$

$$\frac{\boldsymbol{P}^{n+1} - 2\boldsymbol{P}^n + \boldsymbol{P}^{n-1}}{(\Delta t)^2} + \gamma\frac{\boldsymbol{P}^{n+1} - \boldsymbol{P}^{n-1}}{2\Delta t} + \omega_0^2\overline{\overline{\boldsymbol{P}^n}} = \omega_p^2\overline{\overline{\boldsymbol{E}^n}}, \tag{5f}$$

$$\frac{Q^{n+1}-2Q^n+Q^{n-1}}{(\Delta t)^2}+\gamma_v\frac{Q^{n+1}-Q^{n-1}}{2\Delta t}+\omega_v^2\overline{\overline{Q^n}}=\omega_v^2\overline{\boldsymbol{E}^{n+1/2}}\cdot\overline{\boldsymbol{E}^{n-1/2}},\tag{5g}$$

$$\mu_0\frac{\boldsymbol{H}^{n+1}-\boldsymbol{H}^{n+1/2}}{\Delta t/2}=-\nabla\times\boldsymbol{E}^{n+1}.\tag{5h}$$

Here, the nonlinear Kerr term $|\boldsymbol{E}|^2\boldsymbol{E}$ is handled in a similar fashion as in [4, 17], with additional adaptation to the two-step temporal discretization setting. The key is to introduce an auxiliary vector $\boldsymbol{Y}=|\boldsymbol{E}|^2\boldsymbol{E}$ and to consider its differentiated form

$$\partial_t\boldsymbol{Y}=\big(3|\boldsymbol{E}|^2-2|\boldsymbol{E}|^2\big)\partial_t\boldsymbol{E}+2\big(\boldsymbol{E}\cdot\partial_t\boldsymbol{E}\big)\boldsymbol{E}.\tag{6}$$

Indeed, (5d) is a second-order discretization for (6), and this can be better seen through the following equivalent form of (5d):

$$\begin{aligned}\overline{\boldsymbol{Y}^{n+1/2}}-\overline{\boldsymbol{Y}^{n-1/2}}=&\left(3\frac{\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2+\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2}{2}-2\left|\frac{\overline{\boldsymbol{E}^{n+1/2}}+\overline{\boldsymbol{E}^{n-1/2}}}{2}\right|^2\right)\Big(\overline{\boldsymbol{E}^{n+1/2}}-\overline{\boldsymbol{E}^{n-1/2}}\Big)\\&+2\Big(\overline{\overline{\boldsymbol{E}^n}}\cdot\big(\overline{\boldsymbol{E}^{n+1/2}}-\overline{\boldsymbol{E}^{n-1/2}}\big)\Big)\overline{\overline{\boldsymbol{E}^n}}.\end{aligned}\tag{7}$$

For the nonlinear Raman term $Q\boldsymbol{E}$, unlike the algebraic strategy in [4, 17] that works for the first-order form of the model, the model of the mixed-order form in this work together with the two-step temporal discretization strategy call for a second auxiliary vector $\boldsymbol{W}=Q\boldsymbol{E}$ and its differentiated form

$$\partial_t\boldsymbol{W}=Q\partial_t\boldsymbol{E}+(\partial_tQ)\boldsymbol{E}.\tag{8}$$

Again, (5e) is a second-order discretization for (8), and this can be better seen through the following equivalent form of (5e):

$$\begin{aligned}\overline{\boldsymbol{W}^{n+1/2}}-\overline{\boldsymbol{W}^{n-1/2}}=&\frac{\overline{Q^{n+1/2}}+\overline{Q^{n-1/2}}}{2}\Big(\overline{\boldsymbol{E}^{n+1/2}}-\overline{\boldsymbol{E}^{n-1/2}}\Big)\\&+\frac{\overline{\boldsymbol{E}^{n+1/2}}+\overline{\boldsymbol{E}^{n-1/2}}}{2}\Big(\overline{Q^{n+1/2}}-\overline{Q^{n-1/2}}\Big).\end{aligned}\tag{9}$$

Relatedly, a special discretization is employed in (5g) to the nonlinear driving term $|\boldsymbol{E}|^2$.

The temporal discretizations are carefully designed for different terms of the model to achieve a provable energy relation/stability, a discrete analogue of (2), that is given in the next theorem.

Theorem 1 (Semi-discrete in time energy relation/stability) *Under the assumption of periodic boundary conditions, the semi-discrete in time scheme in* (5) *satisfies*

$$\begin{aligned}\mathcal{E}^{n+1/2}-\mathcal{E}^{n-1/2}=&-\frac{\epsilon_0\gamma\Delta t}{\omega_p^2}\int_\Omega\left|\frac{\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n-1}}{2\Delta t}\right|^2\mathrm{d}\Omega\\&-\frac{\epsilon_0 a\theta\gamma_v\Delta t}{2\omega_v^2}\int_\Omega\left(\frac{Q^{n+1}-Q^{n-1}}{2\Delta t}\right)^2\mathrm{d}\Omega\leqslant 0\end{aligned}\tag{10}$$

with the discrete energy defined as

$$\begin{aligned}\mathcal{E}^{n-1/2}=\int_\Omega\Bigg(&\frac{\mu_0}{2}\overline{\boldsymbol{H}^n}\cdot\overline{\boldsymbol{H}^{n-1}}+\frac{\epsilon_0\epsilon_\infty}{2}\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2+\frac{\epsilon_0}{2\omega_p^2}\left|\frac{\boldsymbol{P}^n-\boldsymbol{P}^{n-1}}{\Delta t}\right|^2+\frac{\epsilon_0\omega_0^2}{2\omega_p^2}\left|\overline{\boldsymbol{P}^{n-1/2}}\right|^2\\&+\frac{\epsilon_0 a\theta}{4\omega_v^2}\left(\frac{Q^n-Q^{n-1}}{\Delta t}\right)^2+\frac{\epsilon_0 a\theta}{2}\overline{Q^{n-1/2}}\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2+\frac{3\epsilon_0 a(1-\theta)}{4}\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^4\\&+\frac{\epsilon_0 a\theta}{4}\left(\overline{Q^{n-1/2}}\right)^2\Bigg)\mathrm{d}\Omega.\end{aligned}\tag{11}$$

Here, we write

$$\overline{\boldsymbol{H}^n}=\frac{\boldsymbol{H}^{n+1/2}+\boldsymbol{H}^{n-1/2}}{2}.\tag{12}$$

Proof The following two relations will be repeatedly used:

$$\frac{\boldsymbol{E}^{n+1}-\boldsymbol{E}^{n-1}}{2}=\overline{\boldsymbol{E}^{n+1/2}}-\overline{\boldsymbol{E}^{n-1/2}},\qquad\overline{\overline{\boldsymbol{E}^n}}=\frac{\overline{\boldsymbol{E}^{n+1/2}}+\overline{\boldsymbol{E}^{n-1/2}}}{2}.\tag{13}$$

Applying two time steps to (5a) and (5h), we have

$$\frac{\mu_0}{2}\frac{\boldsymbol{H}^{n+3/2}-\boldsymbol{H}^{n-1/2}}{\Delta t}=-\nabla\times\overline{\boldsymbol{E}^{n+1/2}},\tag{14}$$

and hence

$$\frac{\mu_0}{2}\left(\overline{\boldsymbol{H}^{n+1}}-\overline{\boldsymbol{H}^{n-1}}\right)=-\Delta t\nabla\times\overline{\overline{\boldsymbol{E}^n}}.\tag{15}$$

Based on (5b), we get

$$\overline{\boldsymbol{D}^{n+1/2}}-\overline{\boldsymbol{D}^{n-1/2}}=\Delta t\nabla\times\overline{\boldsymbol{H}^n}.\tag{16}$$

Take the dot product of (16) and $\overline{\overline{\boldsymbol{E}^n}}$, (15) and $\overline{\boldsymbol{H}^n}$, integrate the resulting equations over Ω and sum them up, and under the assumption of the periodic boundary conditions, we obtain

$$\frac{\mu_0}{2}\int_\Omega\left(\overline{\boldsymbol{H}^{n+1}}\cdot\overline{\boldsymbol{H}^n}-\overline{\boldsymbol{H}^n}\cdot\overline{\boldsymbol{H}^{n-1}}\right)\mathrm{d}\Omega+\int_\Omega\left(\overline{\boldsymbol{D}^{n+1/2}}-\overline{\boldsymbol{D}^{n-1/2}}\right)\cdot\overline{\overline{\boldsymbol{E}^n}}\mathrm{d}\Omega=0.\tag{17}$$

For the second term above, we proceed from (5c) and get

$$
\begin{aligned}
&\left(\overline{\boldsymbol{D}^{n+1/2}}-\overline{\boldsymbol{D}^{n-1/2}}\right)\cdot\overline{\overline{\boldsymbol{E}^{n}}}\\
&=\frac{\epsilon_0\epsilon_\infty}{2}\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2-\frac{\epsilon_0\epsilon_\infty}{2}\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2+\frac{\epsilon_0 a(1-\theta)}{2}(\boldsymbol{Y}^{n+1}-\boldsymbol{Y}^{n-1})\cdot\overline{\overline{\boldsymbol{E}^{n}}}\\
&\quad+\frac{\epsilon_0}{2}(\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n-1})\cdot\overline{\overline{\boldsymbol{E}^{n}}}+\frac{\epsilon_0 a\theta}{2}(\boldsymbol{W}^{n+1}-\boldsymbol{W}^{n-1})\cdot\overline{\overline{\boldsymbol{E}^{n}}}.
\end{aligned}
\tag{18}
$$

For the Kerr term in (18), using (5d), we have

$$
\begin{aligned}
&(\boldsymbol{Y}^{n+1}-\boldsymbol{Y}^{n-1})\cdot\overline{\overline{\boldsymbol{E}^{n}}}\\
&=\left(\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2+\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2-\overline{\boldsymbol{E}^{n+1/2}}\cdot\overline{\boldsymbol{E}^{n-1/2}}\right)\left(\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2-\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2\right)\\
&\quad+\frac{1}{2}\left(\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2-\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2\right)\left|\overline{\boldsymbol{E}^{n+1/2}}+\overline{\boldsymbol{E}^{n-1/2}}\right|^2\\
&=\frac{3}{2}\left(\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2+\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2\right)\left(\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2-\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2\right)\\
&=\frac{3}{2}\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^4-\frac{3}{2}\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^4.
\end{aligned}
\tag{19}
$$

For the Lorentz term in (18), using (5f), we have

$$
\begin{aligned}
&(\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n-1})\cdot\overline{\overline{\boldsymbol{E}^{n}}}\\
&=\frac{1}{\omega_p^2}(\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n-1})\cdot\left(\frac{\boldsymbol{P}^{n+1}-2\boldsymbol{P}^{n}+\boldsymbol{P}^{n-1}}{(\Delta t)^2}+\gamma\frac{\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n-1}}{2\Delta t}+\omega_0^2\overline{\overline{\boldsymbol{P}^{n}}}\right)\\
&=\frac{1}{\omega_p^2}\left|\frac{\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n}}{\Delta t}\right|^2-\frac{1}{\omega_p^2}\left|\frac{\boldsymbol{P}^{n}-\boldsymbol{P}^{n-1}}{\Delta t}\right|^2+\frac{2\gamma\Delta t}{\omega_p^2}\left|\frac{\boldsymbol{P}^{n+1}-\boldsymbol{P}^{n-1}}{2\Delta t}\right|^2\\
&\quad+\frac{\omega_0^2}{\omega_p^2}\left|\overline{\boldsymbol{P}^{n+1/2}}\right|^2-\frac{\omega_0^2}{\omega_p^2}\left|\overline{\boldsymbol{P}^{n-1/2}}\right|^2.
\end{aligned}
\tag{20}
$$

For the Raman term in (18), using (5e), we have

$$
\begin{aligned}
(\boldsymbol{W}^{n+1}-\boldsymbol{W}^{n-1})\cdot\overline{\overline{\boldsymbol{E}^{n}}}&=\overline{Q^{n+1/2}}\left|\overline{\boldsymbol{E}^{n+1/2}}\right|^2-\overline{Q^{n-1/2}}\left|\overline{\boldsymbol{E}^{n-1/2}}\right|^2\\
&\quad+\frac{1}{2}\left(Q^{n+1}-Q^{n-1}\right)\overline{\boldsymbol{E}^{n+1/2}}\cdot\overline{\boldsymbol{E}^{n-1/2}}
\end{aligned}
\tag{21}
$$

with the last term above further reformulated based on (5g),

$$
\begin{aligned}
&\left(Q^{n+1}-Q^{n-1}\right)\overline{\boldsymbol{E}^{n+1/2}}\cdot\overline{\boldsymbol{E}^{n-1/2}}\\
&=\frac{1}{\omega_v^2}\left(Q^{n+1}-Q^{n-1}\right)\left(\frac{Q^{n+1}-2Q^n+Q^{n-1}}{(\Delta t)^2}+\gamma_v\frac{Q^{n+1}-Q^{n-1}}{2\Delta t}+\omega_v^2\overline{\overline{Q^n}}\right)\\
&=\frac{1}{\omega_v^2}\left(\frac{Q^{n+1}-Q^n}{\Delta t}\right)^2-\frac{1}{\omega_v^2}\left(\frac{Q^n-Q^{n-1}}{\Delta t}\right)^2+\frac{2\gamma_v\Delta t}{\omega_v^2}\left(\frac{Q^{n+1}-Q^{n-1}}{2\Delta t}\right)^2\\
&\quad+\left(\overline{Q^{n+1/2}}\right)^2-\left(\overline{Q^{n-1/2}}\right)^2.
\end{aligned}
\tag{22}
$$

Combining (17)–(22), we will reach the semi-discrete in time energy relation (10) with the discrete energy defined in (11).

2.2 Fully Discrete Schemes

In this subsection, we will further apply a class of nodal DG methods in space to obtain fully discrete schemes. We will present the methods in two dimensions for simplicity, and the extension to three-dimensional (3D) and one-dimensional (1D) cases is straightforward. Since the Maxwell's system (1) in the TM mode has essentially 1D nonlinearity (i.e., $d_E = 1$), we will consider its TE mode with $d_E = 2$, namely,

$$\mu_0\partial_t H_z+\partial_x E_y-\partial_y E_x=0, \tag{23a}$$

$$\partial_t D_x-\partial_y H_z=0, \tag{23b}$$

$$\partial_t D_y+\partial_x H_z=0, \tag{23c}$$

$$\boldsymbol{D}=\epsilon_0\Big(\epsilon_\infty\boldsymbol{E}+\boldsymbol{P}+a(1-\theta)|\boldsymbol{E}|^2\boldsymbol{E}+a\theta Q\boldsymbol{E}\Big), \tag{23d}$$

$$\partial_{tt}\boldsymbol{P}+\gamma\partial_t\boldsymbol{P}+\omega_0^2\boldsymbol{P}=\omega_p^2\boldsymbol{E}, \tag{23e}$$

$$\partial_t Q+\gamma_v\partial_t Q+\omega_v^2 Q=\omega_v^2|\boldsymbol{E}|^2. \tag{23f}$$

Here, $\boldsymbol{U}=(U_x,U_y)$, with $\boldsymbol{U}=\boldsymbol{D},\boldsymbol{E},\boldsymbol{P}$, stand for vectors with two components U_x and U_y in the x-y plane.

DG methods of the nodal form, instead of the modal form, are adopted here, due to their efficiency to deal with the nonlinearity. To prepare, we will start with some notation. Assume the computational domain is $\Omega=[x_a,x_b]\times[y_a,y_b]$. Let

$$\mathcal{T}_h=\left\{K_{ij}=I_i\times J_j, I_i=[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}], J_j=[y_{j-\frac{1}{2}},y_{j+\frac{1}{2}}], 1\leqslant i\leqslant N_x, 1\leqslant j\leqslant N_y\right\}$$

be a mesh for Ω, that is based on a partition in the x direction: $x_a=x_{\frac{1}{2}}<x_{\frac{3}{2}}<\cdots<x_{N_x+\frac{1}{2}}=x_b$ and a partition in the y direction: $y_a=y_{\frac{1}{2}}<y_{\frac{3}{2}}<\cdots<y_{N_y+\frac{1}{2}}=y_b$. For a typical cell $K_{ij}=I_i\times J_j$, we denote its center as (x_i,y_j). We write $\Delta x_i=|I_i|$, $\Delta y_j=|J_j|$, and set

$h = \max_{i,j}(\Delta x_i, \Delta y_j)$. We further assume the mesh is quasi-uniform, namely, there exists a constant δ such that $\frac{h}{\min_{i,j}(\Delta x_i, \Delta y_j)} < \delta$ as the mesh is refined.

Associated with $\mathcal{T}_h$, we introduce a finite-dimensional discrete space

$$V_h^k = \left\{ v \in L^2(\Omega) : v|_{K_{ij}} \in Q^k(K_{ij}),\ 1 \leqslant i \leqslant N_x, 1 \leqslant j \leqslant N_y \right\} \tag{24}$$

with $Q^k(K_{ij})$ being the set of tensor-type Q^k polynomials with degree up to k in *each variable* on K_{ij}. Without confusion, we also use V_h^k to represent its vector version. Note that each function $v \in V_h^k$ is discontinuous at $x = x_{i+\frac{1}{2}}$ and at $y = y_{j+\frac{1}{2}}$. With $v(x^{\pm}_{i+\frac{1}{2}}, y) = \lim_{\nu \to 0\pm} v(x_{i+\frac{1}{2}} + \nu, y)$, $v(x, y^{\pm}_{j+\frac{1}{2}}) = \lim_{\nu \to 0\pm} v(x, y_{j+\frac{1}{2}} + \nu)$ being the one-sided traces, we write the jumps of $v \in V_h^k$ at $x = x_{i+\frac{1}{2}}$ and at $y = y_{j+\frac{1}{2}}$ as $[v]_{x_{i+\frac{1}{2}}} = v(x^+_{i+\frac{1}{2}}, y) - v(x^-_{i+\frac{1}{2}}, y)$, $[v]_{y_{j+\frac{1}{2}}} = v(x, y^+_{j+\frac{1}{2}}) - v(x, y^-_{j+\frac{1}{2}})$, respectively.

For the nodal DG formulation, we introduce the $(k+1)$-points Gauss-Legendre quadrature points as $\{\xi_m\}_{m=0}^k$ on $[-1, 1]$ with positive weights $\{\widehat{w}_m\}_{m=0}^k$. The corresponding quadrature formula is exact for polynomials of degrees up to $2k+1$ (i.e., for functions in $P^{2k+1}([-1, 1])$). Let $\{l_n(\xi)\}_{n=0}^k$ be the Lagrange basis of $P^k([-1, 1])$, satisfying $l_n(\xi_m) = \delta_{nm}$. Here, δ_{nm} is the Kronecker delta. Associated with each element K_{ij}, we define

$$\phi^{ij}_{mn}(x, y) = l_m\left(\frac{x - x_i}{\Delta x_i/2}\right) l_n\left(\frac{y - y_j}{\Delta y_j/2}\right),$$

and set $x_{im} = x_i + \frac{\Delta x_i}{2}\xi_m$, $y_{jn} = y_j + \frac{\Delta y_j}{2}\xi_n$. It is easy to see that $\left\{\phi^{ij}_{mn}(x, y)\right\}_{m,n=0}^k$ forms an orthogonal basis for $V_h^k|_{K_{ij}} = Q^k(K_{ij})$ with respect to the L^2 inner product on K_{ij}. Relatedly, we introduce a local interpolation operator $\mathcal{I}_h^{ij}: C(K_{ij}) \mapsto Q^k(K_{ij})$ that satisfies $\left(\mathcal{I}_h^{ij} f\right)(x_{im}, y_{jn}) = f(x_{im}, y_{jn})$ *for all* $m, n = 0, \cdots, k$, and equivalently,

$$\left(\mathcal{I}_h^{ij} f\right)(x, y) = \sum_{m,n=0}^{k} f(x_{im}, y_{jn}) \phi^{ij}_{mn}(x, y), \quad (x, y) \in K_{ij}. \tag{25}$$

We further define a global interpolation operator $\mathcal{I}_h: W_h(\Omega) \mapsto V_h^k$ with $\mathcal{I}_h|_{K_{ij}} = \mathcal{I}_h^{ij}$. Here $W_h(\Omega) = \left\{ f \in L^2(\Omega), f|_{K_{ij}} \in C(K_{ij}),\ 1 \leqslant i \leqslant N_x, 1 \leqslant j \leqslant N_y \right\}$ with $C(K_{ij})$ as the set of continuous functions on K_{ij}. The next lemma recalls some properties of the operator $\mathcal{I}_h$.

Lemma 1 [17] *For any* $f \in W_h(\Omega)$, *there hold*

(i) $\displaystyle\int_\Omega \mathcal{I}_h(f)\psi \mathrm{d}\Omega = \int_\Omega \mathcal{I}_h(f\psi)\mathrm{d}\Omega, \quad \forall \psi \in V_h^k;$

as a special case, $\int_\Omega \mathcal{I}_h(\phi)\psi \mathrm{d}\Omega = \int_\Omega \phi\psi \mathrm{d}\Omega$ *for all* $\phi, \psi \in V_h^k$;

(ii) *if, additionally, f is nonnegative, then* $\int_\Omega \mathcal{I}_h(f)\mathrm{d}\Omega \geqslant 0$.

We are now ready to present the fully discrete schemes, that combine the temporal discretization in the previous section, and nodal spatial DG methods with the discrete space

V_h^k $(k = 0, 1, 2, \cdots)$. Particularly, given $u_h^{n-1}, u_h^n \in V_h^k$ at t^{n-1} and t^n, we look for $u_h^{n+1} \in V_h^k$ at t^{n+1}, with $u = H_z$, $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{P}$, and Q, satisfying

$$\mu_0\left(\frac{H_{zh}^{n+1/2} - H_{zh}^n}{\Delta t/2}, \phi\right) + \mathcal{B}_h^{\boldsymbol{E}}(\boldsymbol{E}_h^n, \phi) = 0, \quad \forall \phi \in V_h^k, \tag{26a}$$

$$\left(\frac{D_{xh}^{n+1} - D_{xh}^n}{\Delta t}, \phi\right) + \mathcal{B}_{xh}^H(H_{zh}^{n+1/2}, \phi) = 0, \quad \forall \phi \in V_h^k, \tag{26b}$$

$$\left(\frac{D_{yh}^{n+1} - D_{yh}^n}{\Delta t}, \phi\right) + \mathcal{B}_{yh}^H(H_{zh}^{n+1/2}, \phi) = 0, \quad \forall \phi \in V_h^k, \tag{26c}$$

$$\boldsymbol{D}_h^{n+1} = \epsilon_0\left(\epsilon_\infty \boldsymbol{E}_h^{n+1} + \boldsymbol{P}_h^{n+1} + a(1-\theta)\boldsymbol{Y}_h^{n+1} + a\theta \boldsymbol{W}_h^{n+1}\right), \tag{26d}$$

$$\begin{aligned}\boldsymbol{Y}_h^{n+1} - \boldsymbol{Y}_h^{n-1} = \mathcal{I}_h\left(\left(\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^2 + \left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2 - \overline{\boldsymbol{E}_h^{n+1/2}} \cdot \overline{\boldsymbol{E}_h^{n-1/2}}\right)(\boldsymbol{E}_h^{n+1} - \boldsymbol{E}_h^{n-1})\right) \\ + 2\mathcal{I}_h\left(\left(\overline{\overline{\boldsymbol{E}_h^n}} \cdot (\boldsymbol{E}_h^{n+1} - \boldsymbol{E}_h^{n-1})\right)\overline{\overline{\boldsymbol{E}_h^n}}\right),\end{aligned} \tag{26e}$$

$$\boldsymbol{W}_h^{n+1} - \boldsymbol{W}_h^{n-1} = 2\mathcal{I}_h\left(\overline{Q_h^{n+1/2}}\,\overline{\boldsymbol{E}_h^{n+1/2}} - \overline{Q_h^{n-1/2}}\,\overline{\boldsymbol{E}_h^{n-1/2}}\right), \tag{26f}$$

$$\frac{\boldsymbol{P}_h^{n+1} - 2\boldsymbol{P}_h^n + \boldsymbol{P}_h^{n-1}}{(\Delta t)^2} + \gamma \frac{\boldsymbol{P}_h^{n+1} - \boldsymbol{P}_h^{n-1}}{2\Delta t} + \omega_0^2 \overline{\overline{\boldsymbol{P}_h^n}} = \omega_p^2 \overline{\overline{\boldsymbol{E}_h^n}}, \tag{26g}$$

$$\frac{Q_h^{n+1} - 2Q_h^n + Q_h^{n-1}}{(\Delta t)^2} + \gamma_v \frac{Q_h^{n+1} - Q_h^{n-1}}{2\Delta t} + \omega_v^2 \overline{\overline{Q_h^n}} = \omega_v^2 \mathcal{I}_h\left(\overline{\boldsymbol{E}_h^{n+1/2}} \cdot \overline{\boldsymbol{E}_h^{n-1/2}}\right), \tag{26h}$$

$$\mu_0\left(\frac{H_{zh}^{n+1} - H_{zh}^{n+1/2}}{\Delta t/2}, \phi\right) + \mathcal{B}_h^{\boldsymbol{E}}(\boldsymbol{E}_h^{n+1}, \phi) = 0, \quad \forall \phi \in V_h^k. \tag{26i}$$

Here, $(\cdot, \cdot)$ is the standard L^2 inner product for $L^2(\Omega)$, and the bilinear forms $\mathcal{B}_h^{\boldsymbol{E}}$, $\mathcal{B}_{xh}^H$, and $\mathcal{B}_{yh}^H$ are defined as

$$\begin{aligned}\mathcal{B}_h^{\boldsymbol{E}}(\boldsymbol{E}_h, \phi) = \sum_{j=1}^{N_y} \int_{x_a}^{x_b} \widehat{\widehat{E_{xh}}}(x, y_{j+\frac{1}{2}})[\phi]_{y_{j+\frac{1}{2}}}\,\mathrm{d}x + (E_{xh}, \partial_y \phi) \\ - \sum_{i=1}^{N_x} \int_{y_a}^{y_b} \widehat{E_{yh}}(x_{i+\frac{1}{2}}, y)[\phi]_{x_{i+\frac{1}{2}}}\,\mathrm{d}y - (E_{yh}, \partial_x \phi),\end{aligned} \tag{27a}$$

$$\mathcal{B}_{xh}^{H}(H_{zh},\phi)=\sum_{j=1}^{N_y}\int_{x_a}^{x_b}\widetilde{\widetilde{H_{zh}}}(x,y_{j+\frac{1}{2}})[\phi]_{y_{j+\frac{1}{2}}}\,\mathrm{d}x+(H_{zh},\partial_y\phi),\tag{27b}$$

$$\mathcal{B}_{yh}^{H}(H_{zh},\phi)=-\sum_{i=1}^{N_x}\int_{y_a}^{y_b}\widetilde{H_{zh}}(x_{i+\frac{1}{2}},y)[\phi]_{x_{i+\frac{1}{2}}}\,\mathrm{d}y-(H_{zh},\partial_x\phi),\tag{27c}$$

with $\widehat{E_{yh}}$, $\widehat{\widehat{E_{xh}}}$, $\widetilde{H_{zh}}$, and $\widetilde{\widetilde{H_{zh}}}$ as numerical fluxes, a key ingredient of DG methods for both numerical stability and accuracy. Specifically, we choose to use alternating fluxes as in [17]. Among the four possible choices of alternating numerical fluxes from [17], we consider the following two in this work. The important property the fluxes possess is that $\widehat{E_{yh}}$ and $\widetilde{H_{zh}}$ take traces from opposite sides of an interface, so do $\widehat{\widehat{E_{xh}}}$ and $\widetilde{\widetilde{H_{zh}}}$.

Alternating 1:

$$\widehat{E_{yh}}(x_{i+\frac{1}{2}},y)=E_{yh}(x^{+}_{i+\frac{1}{2}},y),\ \widehat{\widehat{E_{xh}}}(x,y_{j+\frac{1}{2}})=E_{xh}(x,y^{+}_{j+\frac{1}{2}}),\tag{28a}$$

$$\widetilde{H_{zh}}(x_{i+\frac{1}{2}},y)=H_{zh}(x^{-}_{i+\frac{1}{2}},y),\ \widetilde{\widetilde{H_{zh}}}(x,y_{j+\frac{1}{2}})=H_{zh}(x,y^{-}_{j+\frac{1}{2}}).\tag{28b}$$

Alternating 2:

$$\widehat{E_{yh}}(x_{i+\frac{1}{2}},y)=E_{yh}(x^{+}_{i+\frac{1}{2}},y),\ \widehat{\widehat{E_{xh}}}(x,y_{j+\frac{1}{2}})=E_{xh}(x,y^{-}_{j+\frac{1}{2}}),\tag{29a}$$

$$\widetilde{H_{zh}}(x_{i+\frac{1}{2}},y)=H_{zh}(x^{-}_{i+\frac{1}{2}},y),\ \widetilde{\widetilde{H_{zh}}}(x,y_{j+\frac{1}{2}})=H_{zh}(x,y^{+}_{j+\frac{1}{2}}).\tag{29b}$$

With the nodal form of DG spatial discretizations, nonlinear terms in (26e), (26f), and (26h) are handled in an interpolatory fashion [8, 14] through the interpolation operator $\mathcal{I}_h$. Indeed, (26f) in its strong form is equivalent to the following weak form, with any $\phi\in V_h^k$,

$$\begin{aligned}&\int_\Omega\left(\boldsymbol{W}_h^{n+1}-\boldsymbol{W}_h^{n-1}\right)\phi\mathrm{d}x\mathrm{d}y\\&=2\int_\Omega\mathcal{I}_h\left(\overline{Q_h^{n+1/2}}\,\overline{\boldsymbol{E}_h^{n+1/2}}-\overline{Q_h^{n-1/2}}\,\overline{\boldsymbol{E}_h^{n-1/2}}\right)\phi\mathrm{d}x\mathrm{d}y\\&=2\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\frac{\Delta x_i}{2}\frac{\Delta y_j}{2}\sum_{m,n=0}^{k}\hat{\omega}_m\hat{\omega}_n\left(\left(\overline{Q_h^{n+1/2}}\,\overline{\boldsymbol{E}_h^{n+1/2}}-\overline{Q_h^{n-1/2}}\,\overline{\boldsymbol{E}_h^{n-1/2}}\right)\phi\right)(x_{im},y_{jn}).\end{aligned}\tag{30}$$

The second equality is due to the definition of the interpolation operator and the exactness of the $(k+1)$-point Gauss-Legendre quadrature for polynomials of degrees up to $2k+1$. Equalities of this type also facilitate a more precise definition of a discrete energy, with respect to which the fully discrete methods have a provable energy stability, see Theorem 2. The nodal form of DG methods leads to great flexibility in practical simulation and boost the computational efficiency in solving the nonlinear system, partially illustrated through (26f) and its equivalent form

$$\left(\boldsymbol{W}_h^{n+1}-\boldsymbol{W}_h^{n-1}\right)(x_{im},y_{jn})=2\left(\overline{Q_h^{n+1/2}}\,\overline{\boldsymbol{E}_h^{n+1/2}}-\overline{Q_h^{n-1/2}}\,\overline{\boldsymbol{E}_h^{n-1/2}}\right)(x_{im},y_{jn}) \tag{31}$$

with $m,n=0,\cdots,k$, $i=1,\cdots,N_x$, $j=1,\cdots,N_y$ (also see Sect. 2.3), while preserving the energy relation (2)–(3) at the fully discrete level. Similar equivalent forms hold for (26e) and (26h).

Theorem 2 (Fully discrete energy relation and stability) *Under the assumption of periodic boundary conditions, the family of the fully discrete schemes in* (26)–(27), *with the discrete space* V_h^k $(k=0,1,2,\cdots)$ *and the alternating fluxes in either* (28) *or* (29), *satisfy*

$$\begin{aligned}\mathcal{E}_h^{n+1/2}-\mathcal{E}_h^{n-1/2}=&-\frac{\epsilon_0\gamma\Delta t}{\omega_p^2}\int_\Omega\left|\frac{\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1}}{2\Delta t}\right|^2\mathrm{d}\Omega\\&-\frac{\epsilon_0 a\theta\gamma_v\Delta t}{2\omega_v^2}\int_\Omega\left(\frac{Q_h^{n+1}-Q_h^{n-1}}{2\Delta t}\right)^2\mathrm{d}\Omega\leqslant 0\end{aligned} \tag{32}$$

with the discrete energy $\mathcal{E}_h^{n-1/2}$ *defined as*

$$\begin{aligned}\mathcal{E}_h^{n-1/2}=\int_\Omega\Bigg(&\frac{\mu_0}{2}\overline{H_{zh}^n}\,\overline{H_{zh}^{n-1}}+\frac{\epsilon_0\epsilon_\infty}{2}\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2+\frac{\epsilon_0}{2\omega_p^2}\left|\frac{\boldsymbol{P}_h^{n}-\boldsymbol{P}_h^{n-1}}{\Delta t}\right|^2\\&+\frac{\epsilon_0\omega_0^2}{2\omega_p^2}\left|\overline{\boldsymbol{P}_h^{n-1/2}}\right|^2+\frac{\epsilon_0 a\theta}{4\omega_v^2}\left(\frac{Q_h^{n}-Q_h^{n-1}}{\Delta t}\right)^2+\frac{\epsilon_0 a\theta}{2}\mathcal{I}_h\left(\overline{Q_h^{n-1/2}}\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2\right)\\&+\frac{3\epsilon_0 a(1-\theta)}{4}\mathcal{I}_h\left(\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^4\right)+\frac{\epsilon_0 a\theta}{4}\left(\overline{Q_h^{n-1/2}}\right)^2\Bigg)\mathrm{d}\Omega,\end{aligned} \tag{33}$$

and with $\overline{H_{zh}^n}=\frac{H_{zh}^{n+1/2}+H_{zh}^{n-1/2}}{2}$. *Moreover, when* $\theta\in[0,\frac{3}{4}]$ *and under a time step condition,*

$$\Delta t\leqslant\frac{\min(\mu_0,\epsilon_0\epsilon_\infty)}{C_\star}h,$$

the discrete energy $\mathcal{E}_h^{n-1/2}\geqslant 0$, *hence* (33) *gives an energy stability result. Here,* $C_\star$ *is a constant, dependent of the polynomial degree k and the mesh parameter δ.*

Proof Step 1: we will first establish the energy relation in (32). The proof will follow a similar flow as that in Theorem 1 for the semi-discrete in time scheme. Again, the relations in (13) will be repeatedly used. Applying two time steps to (26a) and (26i), we obtain

$$\frac{\mu_0}{2}\left(H_{zh}^{n+3/2}-H_{zh}^{n-1/2},\phi\right)+\Delta t\mathcal{B}_h^E\left(\overline{\boldsymbol{E}_h^{n+1/2}},\phi\right)=0,\quad\forall\phi\in V_h^k, \tag{34}$$

and hence,

$$\frac{\mu_0}{2}\left(\overline{H_{zh}^{n+1}}-\overline{H_{zh}^{n-1}},\phi\right)+\Delta t\mathcal{B}_h^{\boldsymbol{E}}\left(\overline{\overline{\boldsymbol{E}_h^n}},\phi\right)=0,\quad \forall\phi\in V_h^k. \tag{35}$$

Based on (26b)–(26c), we have

$$\left(\overline{\boldsymbol{D}_h^{n+1/2}}-\overline{\boldsymbol{D}_h^{n-1/2}},\phi\right)+\Delta t\mathcal{B}_{xh}^H(\overline{H_{zh}^n},\phi)+\Delta t\mathcal{B}_{yh}^H(\overline{H_{zh}^n},\phi)=0,\quad \forall\phi\in V_h^k. \tag{36}$$

Taking $\phi=\overline{H_{zh}^n}$ in (35), $\phi=\overline{\overline{\boldsymbol{E}_h^n}}$ in (36), summing them up, and using periodic boundary conditions, we reach

$$\frac{\mu_0}{2}\left(\overline{H_{zh}^{n+1}},\overline{H_{zh}^n}\right)-\frac{\mu_0}{2}\left(\overline{H_{zh}^n},\overline{H_{zh}^{n-1}}\right)+\left(\overline{\boldsymbol{D}_h^{n+1/2}}-\overline{\boldsymbol{D}_h^{n-1/2}},\overline{\overline{\boldsymbol{E}_h^n}}\right)=0. \tag{37}$$

The terms associated with spatial operators are canceled due to the specific choices of alternating fluxes in (28) or (29),

What remained is to analyze the third term in (37). To this end, based on (26d), one has

$$\begin{aligned}&\left(\overline{\boldsymbol{D}_h^{n+1/2}}-\overline{\boldsymbol{D}_h^{n-1/2}},\overline{\overline{\boldsymbol{E}_h^n}}\right)\\&=\frac{\epsilon_0\epsilon_\infty}{2}\left\|\overline{\boldsymbol{E}_h^{n+1/2}}\right\|^2-\frac{\epsilon_0\epsilon_\infty}{2}\left\|\overline{\boldsymbol{E}_h^{n-1/2}}\right\|^2+\frac{\epsilon_0 a(1-\theta)}{2}\left(\boldsymbol{Y}_h^{n+1}-\boldsymbol{Y}_h^{n-1},\overline{\overline{\boldsymbol{E}_h^n}}\right)\\&+\frac{\epsilon_0}{2}\left(\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1},\overline{\overline{\boldsymbol{E}_h^n}}\right)+\frac{\epsilon_0 a\theta}{2}\left(\boldsymbol{W}_h^{n+1}-\boldsymbol{W}_h^{n-1},\overline{\overline{\boldsymbol{E}_h^n}}\right).\end{aligned} \tag{38}$$

Here, $\|\cdot\|$ is the standard L^2 norm. For the Kerr term, using (26e) and the property of the interpolation operator $\mathcal{I}_h$ in Lemma 1, we have

$$\begin{aligned}&\left(\boldsymbol{Y}_h^{n+1}-\boldsymbol{Y}_h^{n-1},\overline{\overline{\boldsymbol{E}_h^n}}\right)\\&=\left(\mathcal{I}_h\left(\left(\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^2+\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2-\overline{\boldsymbol{E}_h^{n+1/2}}\cdot\overline{\boldsymbol{E}_h^{n-1/2}}\right)(\boldsymbol{E}_h^{n+1}-\boldsymbol{E}_h^{n-1})\right.\right.\\&\left.\left.+2\overline{\overline{\boldsymbol{E}_h^n}}\cdot(\boldsymbol{E}_h^{n+1}-\boldsymbol{E}_h^{n-1})\overline{\overline{\boldsymbol{E}_h^n}}\right),\overline{\overline{\boldsymbol{E}_h^n}}\right)\\&=\int_\Omega\mathcal{I}_h\left(\left(\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^2+\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2-\overline{\boldsymbol{E}_h^{n+1/2}}\cdot\overline{\boldsymbol{E}_h^{n-1/2}}\right)\left(\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^2-\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2\right)\right)\mathrm{d}\Omega\\&+\frac{1}{2}\int_\Omega\mathcal{I}_h\left(\left(\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^2-\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2\right)\left|\overline{\boldsymbol{E}_h^{n+1/2}}+\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2\right)\mathrm{d}\Omega\\&=\frac{3}{2}\int_\Omega\mathcal{I}_h\left(\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^4\right)\mathrm{d}\Omega-\frac{3}{2}\int_\Omega\mathcal{I}_h\left(\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^4\right)\mathrm{d}\Omega.\end{aligned} \tag{39}$$

For the Lorentz term in (38), by virtue of (26g), we have

$$
\begin{aligned}
&\left(\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1},\overline{\overline{\boldsymbol{E}_h^n}}\right)\\
&=\frac{1}{\omega_p^2}\left(\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1},\frac{\boldsymbol{P}_h^{n+1}-2\boldsymbol{P}_h^n+\boldsymbol{P}_h^{n-1}}{(\Delta t)^2}+\gamma\frac{\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1}}{2\Delta t}+\omega_0^2\overline{\overline{\boldsymbol{P}_h^n}}\right)\\
&=\frac{1}{\omega_p^2}\left\|\frac{\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^n}{\Delta t}\right\|^2-\frac{1}{\omega_p^2}\left\|\frac{\boldsymbol{P}_h^n-\boldsymbol{P}_h^{n-1}}{\Delta t}\right\|^2+\frac{2\gamma\Delta t}{\omega_p^2}\left\|\frac{\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1}}{2\Delta t}\right\|^2\\
&\quad+\frac{\omega_0^2}{\omega_p^2}\left\|\overline{\boldsymbol{P}_h^{n+1/2}}\right\|^2-\frac{\omega_0^2}{\omega_p^2}\left\|\overline{\boldsymbol{P}_h^{n-1/2}}\right\|^2.
\end{aligned}
\tag{40}
$$

For the Raman term, by (26f) and the property of the interpolation operator $\mathcal{I}_h$ in Lemma 1, we have

$$
\begin{aligned}
&\left(\boldsymbol{W}_h^{n+1}-\boldsymbol{W}_h^{n-1},\overline{\overline{\boldsymbol{E}_h^n}}\right)\\
&=\left(\mathcal{I}_h\left(\overline{Q_h^{n+1/2}}\,\overline{\boldsymbol{E}_h^{n+1/2}}-\overline{Q_h^{n-1/2}}\,\overline{\boldsymbol{E}_h^{n-1/2}}\right),\overline{\boldsymbol{E}_h^{n+1/2}}+\overline{\boldsymbol{E}_h^{n-1/2}}\right)\\
&=\int_\Omega\mathcal{I}_h\left(\overline{Q_h^{n+1/2}}\left|\overline{\boldsymbol{E}_h^{n+1/2}}\right|^2\right)\mathrm{d}\Omega-\int_\Omega\mathcal{I}_h\left(\overline{Q_h^{n-1/2}}\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2\right)\mathrm{d}\Omega\\
&\quad+\frac{1}{2}\left(Q_h^{n+1}-Q_h^{n-1},\mathcal{I}_h\left(\overline{\boldsymbol{E}_h^{n+1/2}}\cdot\overline{\boldsymbol{E}_h^{n-1/2}}\right)\right)
\end{aligned}
\tag{41}
$$

with the last term in (41) further formulated based on (26h),

$$
\begin{aligned}
&\left(Q_h^{n+1}-Q_h^{n-1},\mathcal{I}_h\left(\overline{\boldsymbol{E}_h^{n+1/2}}\cdot\overline{\boldsymbol{E}_h^{n-1/2}}\right)\right)\\
&=\frac{1}{\omega_v^2}\left(Q_h^{n+1}-Q_h^{n-1},\frac{Q_h^{n+1}-2Q_h^n+Q_h^{n-1}}{(\Delta t)^2}+\gamma_v\frac{Q_h^{n+1}-Q_h^{n-1}}{2\Delta t}+\omega_v^2\overline{\overline{Q_h^n}}\right)\\
&=\frac{1}{\omega_v^2}\left\|\frac{Q_h^{n+1}-Q_h^n}{\Delta t}\right\|^2-\frac{1}{\omega_v^2}\left\|\frac{Q_h^n-Q_h^{n-1}}{\Delta t}\right\|^2\\
&\quad+\frac{2\gamma_v\Delta t}{\omega_v^2}\left\|\frac{Q_h^{n+1}-Q_h^{n-1}}{2\Delta t}\right\|^2+\left\|\overline{Q_h^{n+1/2}}\right\|^2-\left\|\overline{Q_h^{n-1/2}}\right\|^2.
\end{aligned}
\tag{42}
$$

Combining (37)–(42), we obtain the energy relation (32) with the discrete energy $\mathcal{E}_h^{n-1/2}$ in (33).

Step 2: we will next derive a condition on the time step size, that will ensure the discrete energy $\mathcal{E}_h^{n-1/2}$ in (33) to be nonnegative, hence the energy relation (32) can lead to the energy stability. By (13) and (34), we get for any $\phi\in V_h^k$

$$\mu_0\left(\overline{H_{zh}^n}-\overline{H_{zh}^{n-1}},\phi\right)+\Delta t\sum_{j=1}^{N_y}\int_{x_a}^{x_b}\left(\widehat{\widehat{\overline{E_{xh}^{n-1/2}}}}[\phi]\right)(x,y_{j+\frac{1}{2}})\mathrm{d}x+\Delta t\left(\overline{E_{xh}^{n-1/2}},\partial_y\phi\right)$$
$$-\Delta t\sum_{i=1}^{N_x}\int_{y_a}^{y_b}\left(\widehat{\overline{E_{yh}^{n-1/2}}}[\phi]\right)(x_{i+\frac{1}{2}},y)\mathrm{d}y-\Delta t\left(\overline{E_{yh}^{n-1/2}},\partial_x\phi\right)=0. \tag{43}$$

Particularly, with $\phi=\overline{H_{zh}^n}$, we have

$$\begin{aligned}
&\mu_0\left\|\overline{H_{zh}^n}\right\|^2-\mu_0\left(\overline{H_{zh}^n},\overline{H_{zh}^{n-1}}\right)\\
&=\Delta t\sum_{i=1}^{N_x}\int_{y_a}^{y_b}\widehat{\overline{E_{yh}^{n-1/2}}}\left[\overline{H_{zh}^n}\right]_{x_{i+\frac{1}{2}}}\mathrm{d}y+\Delta t\left(\overline{E_{yh}^{n-1/2}},\partial_x\overline{H_{zh}^n}\right)\\
&\quad-\Delta t\sum_{j=1}^{N_y}\int_{x_a}^{x_b}\widehat{\widehat{\overline{E_{xh}^{n-1/2}}}}\left[\overline{H_{zh}^n}\right]_{y_{j+\frac{1}{2}}}\mathrm{d}x-\Delta t\left(\overline{E_{xh}^{n-1/2}},\partial_y\overline{H_{zh}^n}\right)\\
&\leqslant\frac{\Delta tC_\star}{h}\left(\left\|\overline{\boldsymbol{E}_h^{n-1/2}}\right\|^2+\left\|\overline{H_{zh}^n}\right\|^2\right),
\end{aligned} \tag{44}$$

hence,

$$\frac{\mu_0}{2}\left(\overline{H_{zh}^n},\overline{H_{zh}^{n-1}}\right)\geqslant\left(\frac{\mu_0}{2}-\frac{\Delta tC_\star}{2h}\right)\left\|\overline{H_{zh}^n}\right\|^2-\frac{\Delta tC_\star}{2h}\left\|\overline{\boldsymbol{E}_h^{n-1/2}}\right\|^2. \tag{45}$$

To get the last inequality in (44), we have applied some standard inverse inequalities associated with the discrete space V_h^k, see Lemma A.4 in [17]. The constant $C_\star$ depends on k and the mesh parameter δ. With the bound in (45) and the properties of the operator $\mathcal{I}_h$ in Lemma 1, we will find that if the time step size Δt is restricted by the following:

$$\frac{\mu_0}{2}-\frac{\Delta tC_\star}{2h}\geqslant 0,\quad \frac{\epsilon_0\epsilon_\infty}{2}-\frac{\Delta tC_\star}{2h}\geqslant 0,$$

that is, $\Delta t\leqslant\frac{\min(\mu_0,\epsilon_0\epsilon_\infty)}{C_\star}h$, along with the condition $\theta\in[0,\frac{3}{4}]$ as for the continuous model, we will have

$$\mathcal{E}_h^{n-1/2}\geqslant\left(\int_\Omega\frac{\epsilon_0 a\theta}{4}\mathcal{I}_h\left(\left(\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^2+\overline{Q_h^{n-1/2}}\right)^2\right)+\frac{\epsilon_0 a(3-4\theta)}{4}\mathcal{I}_h\left(\left|\overline{\boldsymbol{E}_h^{n-1/2}}\right|^4\right)\right)\mathrm{d}\Omega\geqslant 0.$$

Remark 1 For the governing equations of the mixed-order form considered here and that of the first-order form in [17], the PDE part is in the first-order form, and the time step condition for the energy stability of the proposed method here (see Theorem 2) is identical as that of the method in [17] (see Theorem 4 therein), with the same constant $C_\star$. Indeed, the time step condition is the same for the L^2 stability when the methods are applied to the model in the absence of the Kerr and Raman nonlinear effects and the linear Lorentz dispersion.

Remark 2 If we only apply nodal DG discretizations in space, with the time variable continuous, this will give the semi-discrete in space nodal DG methods. Following a similar analysis as in [17], one can establish a prior error estimates under some assumptions on the

strength of the nonlinearity when the exact solutions are sufficiently smooth. More specifically, the methods can be shown to be optimally accurate of the $(k+1)$th-order accuracy in the L^2 norm when V_h^k is used as the discrete space. The details of the proof are omitted.

Remark 3 In addition to alternating fluxes, other choices of numerical fluxes can be used, such as central or upwind type fluxes as in [4]. The resulting schemes still enjoy provable energy stability. They are not considered here, as the schemes with central fluxes can lead to sub-optimal accuracy, while the schemes with upwind-type fluxes are implicit and the associated nonlinear algebraic systems will no longer be local in nature. Interested readers can refer to [4] for more information.

2.3 On Implementation

The fully discrete schemes in (26) evolve the unknowns semi-explicitly. For the PDE part, with the leap-frog temporal strategy and the alternating numerical fluxes in space, $H_{zh}^{n+1/2}$ and $\boldsymbol{D}_h^{n+1}$ are updated explicitly, indeed locally in an element by element fashion. This renders high parallel efficiency as widely known for explicit DG methods. For the ODE part along with the constitutive law, the discretization is implicit. Over each time step, one needs to solve a system of nonlinear equations to update $\boldsymbol{E}_h^{n+1}$, $\boldsymbol{P}_h^{n+1}$, and Q_h^{n+1}, again locally for each element (i.e., K_{ij} when $d=2$). Moreover, benefited from the nodal treatment of the nonlinearities (see, e.g., (31)), the nonlinear system for each element is in fact a decoupled $(k+1)^{d_E}$ smaller nonlinear systems, with each associated with one interpolation point (i.e., (x_{im}, y_{jn}), $m,n=0,\cdots,k$ in the TE mode of the problem with $d=d_E=2$). Each of these small systems can be further reformulated and reduced to a $d_E \times d_E$ system of nonlinear equations, that only involves the unknown values of $\boldsymbol{E}_h^{n+1}$ at one nodal point, by eliminating $\boldsymbol{P}_h^{n+1}$ and Q_h^{n+1} using (26g)–(26h). All the algebraic nonlinear systems are cubic in nature due to the type of nonlinear effects considered in this work. They are entirely decoupled from each other, and can be numerically solved very efficiently in a parallel manner over one time step. The local nature of the entire schemes (in both the PDE and ODE discretizations) will be fully explored in our parallel scalability study in Sect. 3.1.2.

In terms of memory usage, to evolve the numerical solutions from t^{n-1}, t^n to t^{n+1}, we store u_h^{n-1} and u_h^n, with $u=\boldsymbol{E}, \boldsymbol{P}, Q$, as well as $H_{zh}^{n-1/2}$. We do not directly solve or store the auxiliary quantities $\boldsymbol{Y}_h^{n+1}$ and $\boldsymbol{W}_h^{n+1}$, as only their temporal differences are needed. In fact from (26d), we have

$$\begin{aligned}\boldsymbol{D}_h^{n+1}-\boldsymbol{D}_h^{n-1} = \epsilon_0\Big(&\epsilon_\infty(\boldsymbol{E}_h^{n+1}-\boldsymbol{E}_h^{n-1})+(\boldsymbol{P}_h^{n+1}-\boldsymbol{P}_h^{n-1})\\ &+a(1-\theta)(\boldsymbol{Y}_h^{n+1}-\boldsymbol{Y}_h^{n-1})+a\theta(\boldsymbol{W}_h^{n+1}-\boldsymbol{W}_h^{n-1})\Big).\end{aligned} \tag{46}$$

One can then apply (26e)–(26f) to eliminate $\boldsymbol{Y}_h^{n+1}-\boldsymbol{Y}_h^{n-1}$ and $\boldsymbol{W}_h^{n+1}-\boldsymbol{W}_h^{n-1}$, and only store $\boldsymbol{D}_h^n-\boldsymbol{D}_h^{n-1}$.

Note that the proposed schemes are two-step methods. In addition to the initialization at $t=0$ through interpolation, the one-step methods in [4, 17] are applied to obtain the numerical solution at $t=t^1$. The flow chart to implement the overall schemes for the TE mode of the system is summarized as Algorithm 1.

Algorithm 1: The flow chart to implement the proposed schemes for the TE mode

Initialization: Starting with u_h^0, $u = H_z$, $\boldsymbol{E}$, $\boldsymbol{P}$, $\partial_t \boldsymbol{P}$, Q, and $\partial_t Q$,

(i) compute $H_{zh}^{1/2}$ at $t = \Delta t/2$ using (26a), and

(ii) compute u_h^1, $u = \boldsymbol{E}$, $\boldsymbol{P}$, Q at $t = t^1$ using the one-step methods in [17].

Time marching: For $n = 1, \cdots, N_t$,

(**S.1**) update $H_{zh}^{n+1/2}$ locally in each element by solving

$$\mu_0 \left(\frac{H_{zh}^{n+1/2} - H_{zh}^{n-1/2}}{\Delta t}, \phi \right) + \mathcal{B}_h^{\boldsymbol{E}}(\boldsymbol{E}_h^n, \phi) = 0.$$

(**S.2**) update $\boldsymbol{D}_h^{n+1} - \boldsymbol{D}_h^n$, again locally in each element, based on (26b)–(26c).

(**S.3**) For $i = 1, \cdots, N_x, j = 1, \cdots, N_y$, $m, n = 0, \cdots, k$, in a decoupled fashion,

- compute $\boldsymbol{E}_h^{n+1}$ at (x_{im}, y_{jm}) by solving a $d_E \times d_E$ system of nonlinear equations, and then
- update $\boldsymbol{P}_h^{n+1}$ and Q_h^{n+1} at (x_{im}, y_{jm}) explicitly based on (26g)–(26h).

3 Numerical Experiments

In this section, we will demonstrate the performance of our proposed schemes in terms of accuracy, efficiency, and energy preservation through a set of 1D and 2D numerical examples. Comparison is made between the schemes proposed here and those in [4] (when $d = 1$) and [17] (when $d = d_E = 2$), that are based on the first-order form of the underlying equations and also satisfy provable energy stability relations. Moreover, a scalability study is performed to illustrate the parallel efficiency. All numerical tests are carried out using the nondimensionalized model, though some results are presented with physical units for better illustration. For each $d_E \times d_E$ nonlinear system $\boldsymbol{F}(\boldsymbol{u}) = \boldsymbol{0}$ associated with an interpolation point (with $d_E = 1, 2$), the classical Newton's iteration method is applied with $\|\boldsymbol{F}(\boldsymbol{u}^s)\|_\infty < \mathrm{Err}_{\mathrm{tol}}$ as the stopping criterion. Here $\mathrm{Err}_{\mathrm{tol}}$ is a given error tolerance. All simulations are performed on uniform meshes in double precision.

3.1 Accuracy, Energy Conservation, Parallel Scalability

3.1.1 1D Kink Shape Solution

In this example, we follow [4, 19] and consider the propagation of a 1D traveling wave of kink shape in $\Omega = [0, 6]$, governed by

$$\begin{aligned} &\partial_t H = \partial_x E, \quad \partial_t D = \partial_x H, \\ &\partial_{tt} P = -\omega_0^2 P + \omega_p^2 E, \\ &D = \epsilon_\infty E + aE^3 + P \end{aligned}$$

in the absence of the Ramen nonlinear effect as well as the damping in the linear dispersive Lorentz effect, i.e., with $\theta = 0, \gamma = 0$. The solution is 6-periodic, with the electric field in the form $E(x,t) = \Theta(x - vt)$ and satisfying

$$\frac{\mathrm{d}\Theta}{\mathrm{d}\xi} = \Phi,$$
$$\frac{\mathrm{d}\Phi}{\mathrm{d}\xi} = \frac{6av^2\Theta\Phi^2 + (\epsilon_\infty\omega_0^2 + \omega_p^2 - \omega_0^2/v^2)\Theta + a\omega_0^2\Theta^3}{1 - \epsilon_\infty v^2 - 3av^2\Theta^2},$$

where

$$\epsilon_\infty = 2.25,\ \epsilon_s = 5.25,\ \beta_1 = \epsilon_s - \epsilon_\infty,\ \omega_0 = 93.627\,179\,982\,222\,216,\ \omega_p = \omega_0\sqrt{\beta_1},$$
$$a = \epsilon_\infty/3,\ v = 0.654\,5/\sqrt{\epsilon_\infty},\ E(0) = 0,\ \Phi(0) = 0.249\,196\,667\,778\,658\,12.$$

The initial conditions for other quantities can be obtained as in [4].

This example will be used to illustrate the accuracy of the proposed methods and the energy conservation. In our simulation, the following two alternating numerical fluxes are used.

Alternating 1:

$$\widehat{E_h}(x_{i+\frac{1}{2}}) = E_h(x^-_{i+\frac{1}{2}}), \widetilde{H_h}(x_{i+\frac{1}{2}}) = H_h(x^+_{i+\frac{1}{2}}). \tag{47a}$$

Alternating 2:

$$\widehat{E_h}(x_{i+\frac{1}{2}}) = E_h(x^+_{i+\frac{1}{2}}), \widetilde{H_h}(x_{i+\frac{1}{2}}) = H_h(x^-_{i+\frac{1}{2}}). \tag{47b}$$

The final time is $T = 9/v$. To match the $(k+1)$th-order accuracy in space, we first set $\mathrm{d}t = Ch^{(k+1)/2}$ with $C = 0.2/v$ for $k = 1$, $C = 1$ for $k = 2$, and $C = 2$ for $k = 3$. The time step size is then taken as $\Delta t = \frac{T}{[T/\mathrm{d}t]+1}$, to render a uniform temporal mesh. Here $[x]$ represents the greatest integer less than or equal to x. The error tolerance in the Newton solve for each nonlinear equation is $\mathrm{Err}_{\mathrm{tol}}$=$10^{-10}$. This test is carried out sequentially using Fortran on Thinkpad-X1-Extreme with Intel Core i7-8850H CPU 2.70 GHz and 16 GB Memory.

In Tables 1 and 2 (left half), we report the L^2 and L^∞ errors and convergence orders of E computed by the proposed methods with $k = 1, 2, 3$ and two alternating fluxes (47a)–(47b), as well as the CPU time. The results confirm the optimal accuracy of $(k+1)$th order especially under the L^2 norm. The difference due to the two sets of fluxes is negligible. As a comparison, we also present in Tables 1 and 2 (right half) the results by the method in [4] based on the first-order form of the model, implemented in the nodal form. (Note the methods in [4] were described in their modal form.) Schemes of the same accuracy order from both families produce almost the same errors, with the comparable elapsed time. The nonlinear algebraic systems solved at each interpolation node in both families of methods are of the same size, with the proposed methods here involving relatively more algebraic evaluations to generate the nonlinear systems and hence requiring slightly more elapsed time.

For this example, we can define the discrete energy $\mathcal{E}_h^{n+1/2}$ as

Table 1 1D kink shape solution: numerical errors and convergence rates of E of the proposed method ("mixed-order form") and the method in [4] ("first-order form"), Alternating 1

N_x	Mixed-order form					First-order form				
	L^2		L^∞		Time/s	L^2		L^∞		Time/s
$k=1$										
100	1.28E−04	–	5.52E−04	–	0.01	1.28E−04	–	5.52E−04	–	0.01
200	4.19E−05	1.61	1.87E−04	1.56	0.05	4.19E−05	1.61	1.87E−04	1.56	0.04
400	1.21E−05	1.79	4.36E−05	2.10	0.28	1.21E−05	1.79	4.35E−05	2.10	0.22
800	3.30E−06	1.87	1.52E−05	1.52	0.63	3.31E−06	1.87	1.52E−05	1.52	0.58
1 600	7.37E−07	2.16	3.97E−06	1.94	2.30	7.37E−07	2.17	3.96E−06	1.94	1.99
$k=2$										
100	3.54E−05	–	1.70E−04	–	0.03	3.53E−05	–	1.70E−04	–	0.02
200	4.44E−06	3.00	2.24E−05	2.93	0.19	4.43E−06	3.00	2.23E−05	2.93	0.14
400	5.55E−07	3.00	2.71E−06	3.05	1.50	5.54E−07	3.00	2.70E−06	3.05	1.32
800	6.90E−08	3.01	3.46E−07	2.97	4.83	6.89E−08	3.01	3.45E−07	2.97	4.14
$k=3$										
100	8.37E−06	–	3.48E−05	–	0.07	8.35E−06	–	3.48E−05	–	0.06
200	5.23E−07	4.00	2.16E−06	4.01	0.59	5.22E−07	4.00	2.15E−06	4.01	0.58
400	3.25E−08	4.01	1.52E−07	3.83	6.91	3.24E−08	4.01	1.51E−07	3.83	6.04

Table 2 1D kink shape solution: numerical errors and convergence rates of E of the proposed method ("mixed-order form") and the method in [4] ("first-order form"), Alternating 2

N_x	Mixed-order form					First-order form				
	L^2		L^∞		Time/s	L^2		L^∞		Time/s
$k=1$										
100	1.42E−04	–	6.57E−04	–	0.01	1.42E−04	–	6.57E−04	–	0.01
200	3.79E−05	1.91	1.27E−04	2.38	0.04	3.79E−05	1.90	1.26E−04	2.38	0.03
400	1.03E−05	1.87	4.10E−05	1.63	0.23	1.03E−05	1.87	4.08E−05	1.63	0.21
800	3.23E−06	1.68	1.18E−05	1.80	0.59	3.23E−06	1.68	1.18E−05	1.80	0.54
1 600	9.10E−07	1.83	3.13E−06	1.91	2.13	9.11E−07	1.83	3.12E−06	1.91	1.92
$k=2$										
100	3.53E−05	–	1.53E−04	–	0.03	3.52E−05	–	1.53E−04	–	0.02
200	4.40E−06	3.00	1.93E−05	2.99	0.17	4.39E−06	3.00	1.92E−05	2.99	0.16
400	5.51E−07	3.00	2.42E−06	3.00	1.43	5.50E−07	3.00	2.41E−06	3.00	1.19
800	6.90E−08	3.00	3.12E−07	2.95	4.69	6.89E−08	3.00	3.11E−07	2.95	4.18
$k=3$										
100	8.36E−06	–	3.44E−05	–	0.07	8.35E−06	–	3.43E−05	–	0.08
200	5.23E−07	4.00	2.15E−06	4.00	0.56	5.22E−07	4.00	2.15E−06	4.00	0.51
400	3.25E−08	4.01	1.47E−07	3.87	6.39	3.24E−08	4.01	1.47E−07	3.87	6.04

$$\mathcal{E}_h^{n-1/2} = \int_\Omega \left(\frac{1}{2}\overline{H_h^n} \cdot \overline{H_h^{n-1}} + \frac{\epsilon_\infty}{2}\left(\overline{E_h^{n-1/2}}\right)^2 + \frac{1}{2\omega_p^2}\left(\frac{P_h^n - P_h^{n-1}}{\Delta t}\right)^2 + \frac{\omega_0^2}{2\omega_p^2}\left(\overline{P_h^{n-1/2}}\right)^2 + \frac{3a}{4}\mathcal{I}_h\left(\left(\overline{E_h^{n-1/2}}\right)^4\right)\right) \mathrm{d}\Omega.$$

Our analysis in Theorem 2 states that the discrete energy is conserved over time, namely, $\mathcal{E}_h^{n+1/2} - \mathcal{E}_h^{n-1/2} = 0, for\ all\ n$, and this is validated by the time history of the energy deviation $\mathcal{E}_h^{n+1/2} - \mathcal{E}_h^{1/2}$ in Fig. 1. One can see that the discrete energy is conserved up to the machine accuracy.

3.1.2 2D Manufactured Solution

In this section, we illustrate the accuracy, parallel scalability, and the energy conservation of the proposed methods in two dimensions.

Accuracy: To demonstrate the accuracy, we consider a manufactured solution

$$\begin{cases} H_z &= \mathrm{e}^{\cos(w(t+\alpha x+\beta y))}, \\ E_x &= \beta \mathrm{e}^{\cos(w(t+\alpha x+\beta y))}, \\ E_y &= -\alpha \mathrm{e}^{\cos(w(t+\alpha x+\beta y))}, \\ P_x &= E_x, \quad P_y = E_y, \quad Q = H_z, \\ D_x &= E_x, \quad D_y = E_y, \end{cases}$$

that satisfies the following 2D nondimensionalized system of the TE mode:

$$\begin{aligned} \partial_t H_z + \partial_x E_y - \partial_y E_x &= 0, \\ \partial_t D_x - \partial_y H_z &= 0, \\ \partial_t D_y + \partial_x H_z &= 0, \\ \boldsymbol{D} = \epsilon_\infty \boldsymbol{E} + \boldsymbol{P} + a(1-\theta)|\boldsymbol{E}|^2\boldsymbol{E} + a\theta Q\boldsymbol{E} + \boldsymbol{f_D}, \\ \partial_{tt}\boldsymbol{P} + \gamma\partial_t\boldsymbol{P} + \omega_0^2\boldsymbol{P} &= \omega_p^2\boldsymbol{E} + \boldsymbol{f_P}, \\ \partial_{tt}Q + \gamma_v\partial_t Q + \omega_v^2 Q &= \omega_v^2|\boldsymbol{E}|^2 + f_Q \end{aligned}$$

with suitable forcing terms $\boldsymbol{f_D}, \boldsymbol{f_P}, f_Q$ added to the constitutive relation and the ODE part. The computational domain is $\Omega = [0, 2\pi/(\alpha w)] \times [0, 2\pi/(\beta w)]$ with $\alpha = \cos(0.3\pi)$,

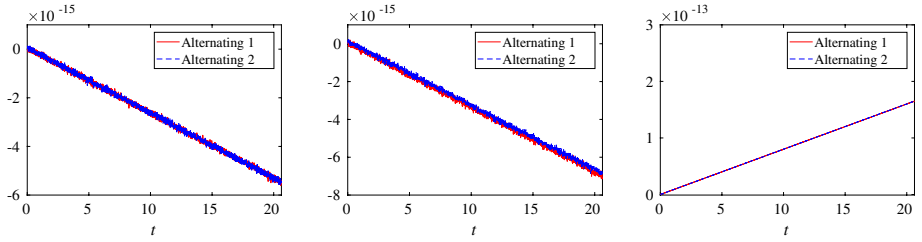

Fig. 1 1D kink shape solution: energy deviation $\mathcal{E}_h^{n+1/2} - \mathcal{E}_h^{1/2}$. Left: $k = 1$, middle: $k = 2$, right: $k = 3$. Mesh: $N_x = 400$

$\beta = \sin(0.3\pi)$, $w = 1.0$, and the boundary conditions are periodic. The model parameters are set as: $\epsilon_\infty = 1.0$, $\omega_0 = \omega_p = 1.0$, $\gamma = 0.05$, $a = 1/3$, $\theta = 0.5$, $\omega_v = 1.0$, $\gamma_v = 0.05$. The final time is $T = 1.0$.

Our numerical schemes (26) are adapted by adding the forcing term $\boldsymbol{f_D}^{n+1}$ to (26d), $\boldsymbol{f_P}^n$ to (26g), and f_Q^n to (26h). To match the $(k+1)$th-order accuracy in space, we first set $\mathrm{d}t = Ch^{(k+1)/2}$, with $h = \frac{0.5}{(\Delta x)^{-1}+(\Delta y)^{-1}}$, $C = 0.3$ for $k = 1$, $C = 1$ for $k = 2$, and $C = 2$ for $k = 3$. The time step size is then taken as $\Delta t = \frac{T}{[T/\mathrm{d}t]+1}$, to render a uniform temporal mesh. The error tolerance in the Newton solve for each nonlinear system is $\mathrm{Err}_{\mathrm{tol}} = 10^{-12}$. This experiment is carried out using MPI Fortran on the cluster LSSC-IV of the State Key Laboratory on Scientific and Engineering Computing, Chinese Academy of Sciences, with 16 processors.

In Tables 3 and 4 (left half), we report the L^2 and L^∞ errors and convergence orders of H_z computed by the proposed methods with $k = 1, 2, 3$ and two alternating fluxes (28)–(29), as well as the CPU time. The results confirm the optimal accuracy of $(k+1)$th order. As a comparison, we also present in Tables 3 and 4 (right half) the results by the method in [17] based on the first-order form of the model. Schemes of the same accuracy order from both families produce almost the same errors, with the slightly more elapsed time used by the proposed methods as we have seen in the 1D example with the kink shape solution.

Scalability: It is widely known that the local nature of (explicit) DG methods contributes to their high parallel efficiency. With the specific treatments of nonlinearity in the nodal DG setting (see Sect. 2.3), one can expect that our proposed algorithms will have the similar parallel efficiency as the standard DG methods applied to hyperbolic problems [3].

Table 3 2D manufactured solution: numerical errors and convergence rates of H_z of the proposed method ("mixed-order form") and the method in [17] ("first-order form"), Alternating 1

Mesh	Mixed-order form					First-order form				
	L^2		L^∞		Time/s	L^2		L^∞		Time/s
$k = 1$										
20×20	1.32E−02	–	9.06E−02	–	0.00	1.32E−02	–	9.08E−02	–	0.00
40×40	3.61E−03	1.87	2.30E−02	1.98	0.02	3.61E−03	1.87	2.30E−02	1.98	0.02
80×80	7.88E−04	2.20	5.56E−03	2.05	0.11	7.86E−04	2.20	5.56E−03	2.05	0.11
160×160	2.08E−04	1.92	1.45E−03	1.94	0.86	2.08E−04	1.92	1.45E−03	1.94	0.85
320×320	5.55E−05	1.91	3.81E−04	1.93	6.79	5.54E-05	1.91	3.77E-04	1.94	6.59
$k = 2$										
20×20	7.48E−04	–	6.43E−03	–	0.01	7.06E−04	–	6.12E−03	–	0.01
40×40	8.97E−05	3.06	6.02E−04	3.42	0.04	8.51E−05	3.05	5.46E−04	3.49	0.05
80×80	9.87E−06	3.18	6.83E−05	3.14	0.48	9.25E−06	3.20	6.52E−05	3.07	0.48
160×160	1.40E−06	2.81	1.36E−05	2.33	5.31	1.32E−06	2.81	1.27E−05	2.36	5.07
320×320	1.69E−07	3.05	1.27E−06	3.42	61.62	1.60E−07	3.05	1.16E−06	3.46	57.90
$k = 3$										
20×20	1.24E−04	–	4.02E−04	–	0.01	8.24E−05	–	3.58E−04	–	0.01
40×40	7.85E−06	3.98	2.54E−05	3.99	0.18	5.15E−06	4.00	2.42E−05	3.89	0.19
80×80	4.96E−07	3.98	1.78E−06	3.83	2.83	3.26E−07	3.98	1.60E−06	3.92	2.74
160×160	3.10E−08	4.00	1.06E−07	4.07	46.60	2.04E−08	3.99	9.56E−08	4.06	43.78
320×320	2.12E−09	3.87	7.14E−09	3.89	737.08	1.77E−09	3.53	6.72E−09	3.83	708.47

Table 4 2D manufactured solution: numerical errors and convergence rates of H_z of the proposed method ("mixed-order form") and the method in [17] ("first-order form"), Alternating 2

Mesh	Mixed-order form					First-order form				
	L^2		L^∞		Time/s	L^2		L^∞		Time/s
$k=1$										
20×20	1.54E−02	–	9.71E−02	–	0.01	1.54E−02	–	9.65E−02	–	0.01
40×40	3.30E−03	2.22	2.25E−02	2.11	0.02	3.30E−03	2.22	2.26E−02	2.10	0.02
80×80	8.34E−04	1.99	5.13E−03	2.13	0.11	8.32E−04	1.99	5.14E−03	2.13	0.11
160×160	2.18E−04	1.93	1.48E−03	1.79	0.85	2.18E−04	1.93	1.47E−03	1.81	0.85
320×320	5.36E−05	2.02	3.61E−04	2.04	6.77	5.35E−05	2.02	3.58E−04	2.04	6.61
$k=2$										
20×20	6.59E−04	–	3.91E−03	–	0.01	6.17E−04	–	3.42E−03	–	0.01
40×40	7.43E−05	3.15	3.61E−04	3.44	0.05	6.87E−05	3.17	3.72E−04	3.20	0.04
80×80	9.20E−06	3.01	5.50E−05	2.72	0.52	8.49E−06	3.02	5.65E−05	2.72	0.48
160×160	1.22E−06	2.92	8.82E−06	2.64	5.41	1.13E−06	2.91	7.94E−06	2.83	5.08
320×320	1.46E−07	3.06	9.69E−07	3.19	62.27	1.35E−07	3.06	9.05E−07	3.13	57.80
$k=3$										
20×20	1.24E−04	–	4.17E−04	–	0.01	8.13E−05	–	3.17E−04	–	0.01
40×40	7.84E−06	3.98	2.51E−05	4.05	0.18	5.10E−06	4.00	1.88E−05	4.08	0.19
80×80	4.94E−07	3.99	1.70E−06	3.88	2.81	3.21E−07	3.99	1.30E−06	3.85	2.74
160×160	3.09E−08	4.00	1.09E−07	3.96	46.01	2.01E−08	4.00	7.64E−08	4.09	43.68
320×320	2.12E−09	3.87	7.66E−09	3.83	734.70	1.75E−09	3.52	5.60E−09	3.77	713.91

We here will use our method (26) with $k=2$ and Alternating 2 flux as an example, applied to the 2D manufactured solution in the accuracy test, to illustrate the parallel scalability. To this end, we divide the computational domain into N_p subdomains, with N_p being the number of processors used in the simulation, and then time-march the methods in each subdomain on one processor, with additional communication in a surrounding ghost layer [2].

We first study the strong scalability of our scheme on a fixed mesh $N_x = N_y = 320$. In Table 5, we show the speed-up and parallel efficiency E_{strong} as the number of processors increases. The "speed-up" is the ratio of the computational time with one single processor to that with N_p processors, and $E_{\text{strong}} = \frac{\text{speed-up}}{N_p}$. As one can see, over 94% efficiency is achieved on up to 256 processors. If we further increase the computing resources (e.g., $N_p = 512, 1\,024$), the workload over each processor becomes smaller and smaller, and the parallel efficiency will be restricted by the communication between processors.

We further examine the weak scalability of our scheme by increasing the problem size and the number of processors N_p simultaneously while keeping the workload per processor a constant. In each simulation, the problem is solved over 200 time steps. As shown in Table 6, over 90% efficiency is achieved on up to 1 024 processors, exhibiting good scalability for large scale simulation, similarly to the standard DG methods applied to hyperbolic problems [3]. Here E_{weak} is the ratio of the computational time with one single processor to that with N_p processors.

Energy conservation: To validate the energy conserving property of our schemes in two dimensions, we use the same initial and boundary conditions as for the accuracy test in this section, while switching off the external sources $\boldsymbol{f_D}, \boldsymbol{f_P}, f_Q$, and setting the damping parameters $\gamma = \gamma_v = 0$ in the simulation. As predicted by Theorem 2, the discrete energy

Table 5 Strong scalability: $k = 2$, Alternating 2, $N_x = N_y = 320$

N_p	Time/s	Speed-up	E_{strong}
1	955.774	–	–
2	480.076	1.99	99.54%
4	241.309	3.96	99.02%
8	120.725	7.92	98.96%
16	60.631	15.76	98.52%
32	30.369	31.47	98.35%
64	15.221	62.79	98.11%
128	7.628	125.30	97.89%
256	3.959	241.40	94.30%
512	2.122	450.43	87.97%
1 024	2.564	372.77	36.40%

Table 6 Weak scalability: $k = 2$, Alternating 2, 200 time steps

N_p	$N_x \times N_y$	Time/s	E_{weak}
1	200×200	44.681	–
2	400×200	44.882	99.55%
4	400×400	45.014	99.26%
8	800×400	45.412	98.39%
16	800×800	45.572	98.04%
32	$1\,600 \times 800$	46.216	96.68%
64	$1\,600 \times 1\,600$	46.360	96.38%
128	$3\,200 \times 1\,600$	47.209	94.65%
256	$3\,200 \times 3\,200$	47.358	94.61%
512	$6\,400 \times 3\,200$	49.226	90.77%
1 024	$6\,400 \times 6\,400$	49.223	90.77%

is conserving. In Fig. 2, we plot the time history of energy deviation $\mathcal{E}_h^{n+1/2} - \mathcal{E}_h^{1/2}$ up to $T = 100$. As it shows, our schemes conserve the discrete energy quite well over long time simulation.

3.2 Physically Relevant Simulations

In this section, we apply our proposed schemes to simulate some physically relevant problems, including a 1D soliton-like wave propagation, 2D spatial-soliton propagation, as well as 2D two-beam interactions. With similarity in the results and to save space, we only present the results with Alternating 1 numerical fluxes, i.e., (47a) in one dimension and (28) in two dimensions. All 2D simulations are parallelly implemented in our codes. The initial configuration of the solutions are zero for all the examples considered in this section.

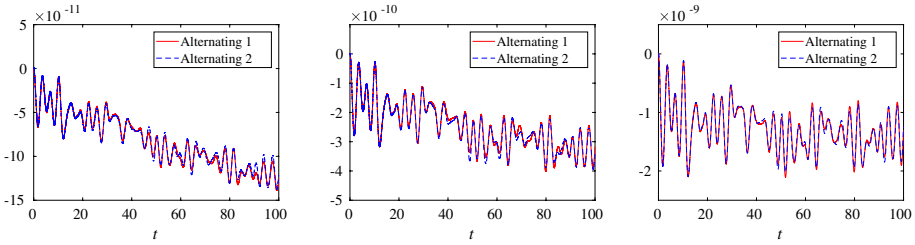

Fig. 2 2D example: energy deviation $\mathcal{E}_h^{n+1/2} - \mathcal{E}_h^{1/2}$. Left: $k = 1$, middle: $k = 2$, right: $k = 3$. Mesh: $N_x = N_y = 80$

3.2.1 1D Soliton-Like Wave Propagation

The propagation of a 1D soliton-like wave in the physical domain $\Omega = [0, 45]$ is governed by the following nondimensinalized model:

$$\begin{aligned} \partial_t H - \partial_x E &= 0, \\ \partial_t D - \partial_x H &= 0, \\ D = \epsilon_\infty E + P + a(1-\theta)E^3 + a\theta Q E, \\ \partial_{tt} P + \gamma \partial_t P + \omega_0^2 P &= \omega_p^2 E, \\ \partial_{tt} Q + \gamma_v \partial_t Q + \omega_v^2 Q &= \omega_v^2 E^2. \end{aligned}$$

Following [4, 11], the physical coefficients are

$$\epsilon_\infty = 2.25, \quad \epsilon_s = 5.25, \quad \beta_1 = \epsilon_s - \epsilon_\infty, \quad \gamma = 1.168 \times 10^{-5}, \quad \gamma_v = 29.2/32,$$
$$a = 0.07, \quad \theta = 0.3, \quad \Omega_0 = 12.57, \quad \omega_0 = 5.84, \quad \omega_v = 1.28, \quad \omega_p = \omega_0\sqrt{\beta_1}.$$

On the left boundary $x = 0$, an incident wave is injected as

$$E(x = 0, t) = f(t)\cos(\Omega_0 t), \quad f(t) = M\mathrm{sech}(t - 20), \tag{48}$$

where M physically characterizes the order of solitons. This boundary condition can be imposed through numerical flux straightforwardly with the use of Alternating 1 numerical flux in (47a). Interested readers can refer to [4] to see how Alternating 2 numerical flux in (47b) can be applied at $x = 0$, as for this case the magnetic field $H(x = 0, t)$ needs to be

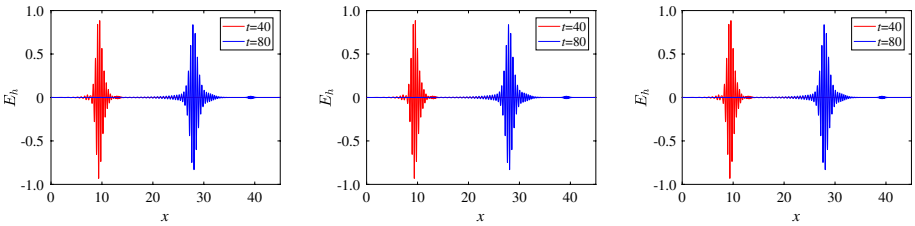

Fig. 3 1D soliton-like wave propagation: fundamental soliton with $M = 1$. Left: $k = 1$, middle: $k = 2$, right: $k = 3$

approximated first. At the right boundary $x = 45$, we follow the strategy in [4] and apply an approximate absorbing boundary condition by neglecting the nonlinear effects and the delayed response. The simulation is up to the final time $T = 80$. The mesh size h is equal to $45/6\,400$ with the time step size $\Delta t = 0.1h$, and $k = 1, 2, 3$. The error tolerance in the Newton solve is set as $\mathrm{Err}_{\mathrm{tol}} = 10^{-10}$.

In Figs. 3 and 4, we plot the electric fields recorded at $t = 40$ and $t = 80$ for the transient fundamental soliton with $M = 1$, and for the second-order soliton with $M = 2$, respectively, with $k = 1, 2, 3$. The results are comparable with those by the schemes in [4] based on the first-order form of the model. In Figs. 5 and 6, we further show the time evolution of the discrete energy $\mathcal{E}_h^{n+1/2}$ and the temporal spectrum of E_h, respectively. The spectrum is obtained through the discrete Fourier transform of the electric field $E_h(t)$ recorded at $x = 0.2$ from $t = 0$ to time $T_s = 2\pi/\Omega_0 \times 50$ over 50 time periods. One can see that the total energy decreases after the entire incident wave enters the domain, demonstrating the energy stability of the schemes. Moreover, in addition to the main soliton-like pulse with the fundamental frequency Ω_0, the third-order and fifth-order harmonic signals are generated and detected as in Fig. 6. The third-order harmonic signals correspond to the small daughter pulses shown in Figs. 3 and 4, while the fifth-order harmonic signals are too weak to observe in these plots.

3.2.2 Spatial-Soliton Propagation in two Dimensions

In this example, we simulate a 2D spatial-soliton propagation in optical glasses [13]. The governing equations are

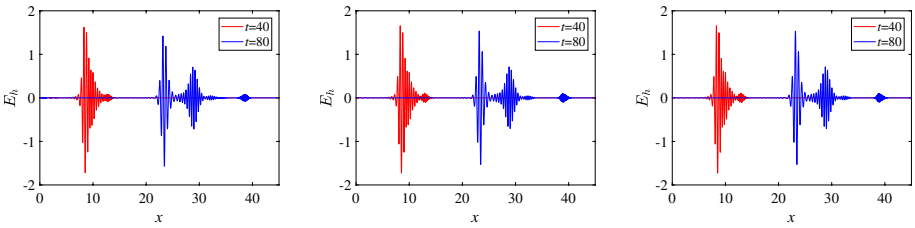

Fig. 4 1D soliton-like wave propagation: second-order soliton with $M = 2$. Left: $k = 1$, middle: $k = 2$, right: $k = 3$

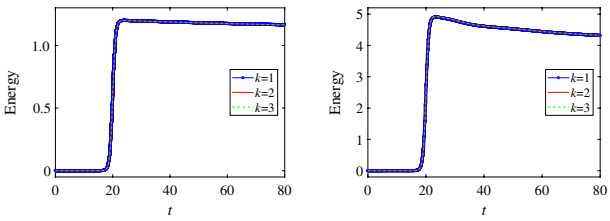

Fig. 5 1D soliton-like wave propagation: time history of the discrete energy. Left: $M = 1$, right: $M = 2$

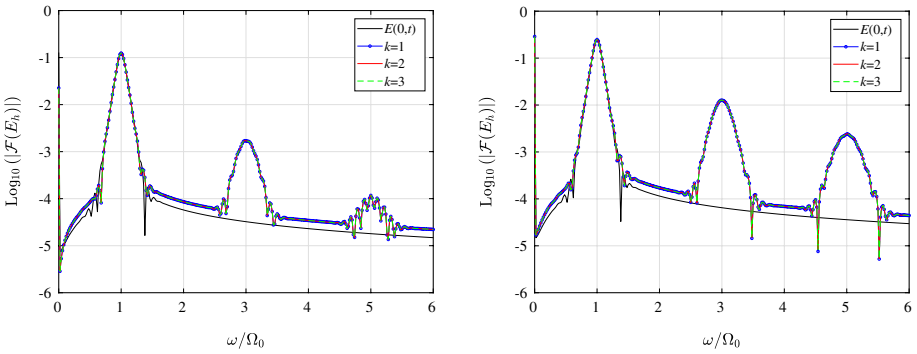

Fig. 6 1D soliton-like wave propagation: semi-log plot of the spectrum of E_h. Left: $M = 1$, right: $M = 2$

$$\mu_0 \partial_t H_z + \partial_x E_y - \partial_y E_x = 0, \tag{49a}$$

$$\partial_t D_x - \partial_y H_z = 0, \tag{49b}$$

$$\partial_t D_y + \partial_x H_z = 0, \tag{49c}$$

$$\boldsymbol{D} = \epsilon_0 \left(\epsilon_\infty \boldsymbol{E} + b \sum_{s=1}^{3} \boldsymbol{P}_s + a(1-\theta)|\boldsymbol{E}|^2 \boldsymbol{E} + a\theta Q \boldsymbol{E} \right), \tag{49d}$$

$$\partial_{tt} \boldsymbol{P}_s + \gamma_s \partial_t \boldsymbol{P}_s + \omega_{0s}^2 \boldsymbol{P}_s = \omega_{ps}^2 \boldsymbol{E}, \quad s = 1, 2, 3, \tag{49e}$$

$$\partial_{tt} Q + \gamma_v \partial_t Q + \omega_v^2 Q = \omega_v^2 |\boldsymbol{E}|^2, \tag{49f}$$

where

$$\omega_{01} = 2.753\,7 \times 10^{16}\,\text{rad/s},\ \omega_{02} = 1.620\,5 \times 10^{16}\,\text{rad/s},\ \omega_{03} = 1.903\,4 \times 10^{14}\,\text{rad/s},$$
$$\beta_1 = 0.696\,17,\ \beta_2 = 0.407\,94,\ \beta_3 = 0.897\,48,\ \omega_{ps} = \sqrt{\beta_s}\omega_{0s},\ \gamma_s = 0, s = 1, 2, 3,$$
$$\epsilon_\infty = 1.0,\ b = 1.0,\ a = 1.89 \times 10^{-22}\text{m}^2/\text{V}^2,\ \theta = 0.3,$$
$$\gamma_v = \frac{2}{\tau_2},\ \omega_v = \sqrt{\frac{\tau_1^2 + \tau_2^2}{\tau_1^2 \tau_2^2}},\ \tau_1 = 12.2\,\text{fs},\ \tau_2 = 32.0\,\text{fs}.$$

The physical domain is $\Omega_0 = [0, 38\,\mu\text{m}] \times [-3\,\mu\text{m}, 3\,\mu\text{m}]$. On the left boundary $x = 0$, the following signal is introduced:

$$H_z(x = 0, y, t) = H_0 \sin(\omega_c t)\text{sech}(y/w), \tag{50}$$

where $\omega_c = 4.35 \times 10^{15}\,\text{rad/s}$ is the carrier frequency. And w, H_0 are the width and the magnitude of the incident wave, respectively, which will be specified later. The boundary condition at $x = 0$ can be imposed directly by the Alternating 1 numerical flux (28).

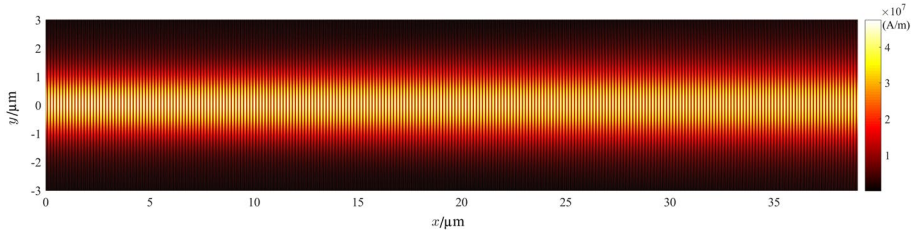

Fig. 7 2D spatial-soliton propagation: snapshot of $|H_z|$ at $T = 300$ fs for the fundamental soliton propagation, with $w = 667.0$ nm, $H_0 = 4.77 \times 10^7$ A/m, and $k = 2$

On the other three boundaries, approximate absorbing boundary conditions as in [17] are applied to suppress the unphysical reflections. These numerical boundary treatments are simple yet rough, and the actual domain used in our simulation is enlarged to $\Omega = [0, 60\,\mu\text{m}] \times [-4\,\mu\text{m}, 4\,\mu\text{m}]$ to compensate the possible insufficiency of the adopted strategy to control artificial reflections. We simulate the propagations up to the final time $T = 300$ fs with the time step size $\Delta t = 0.05\,h$ on a uniform spatial mesh with $h = 20$ nm in both the x and y directions. The discrete space V_h^k with $k = 2$ is used. The error tolerance in the Newton solve is set as $\text{Err}_{\text{tol}} = 10^{-8}$.

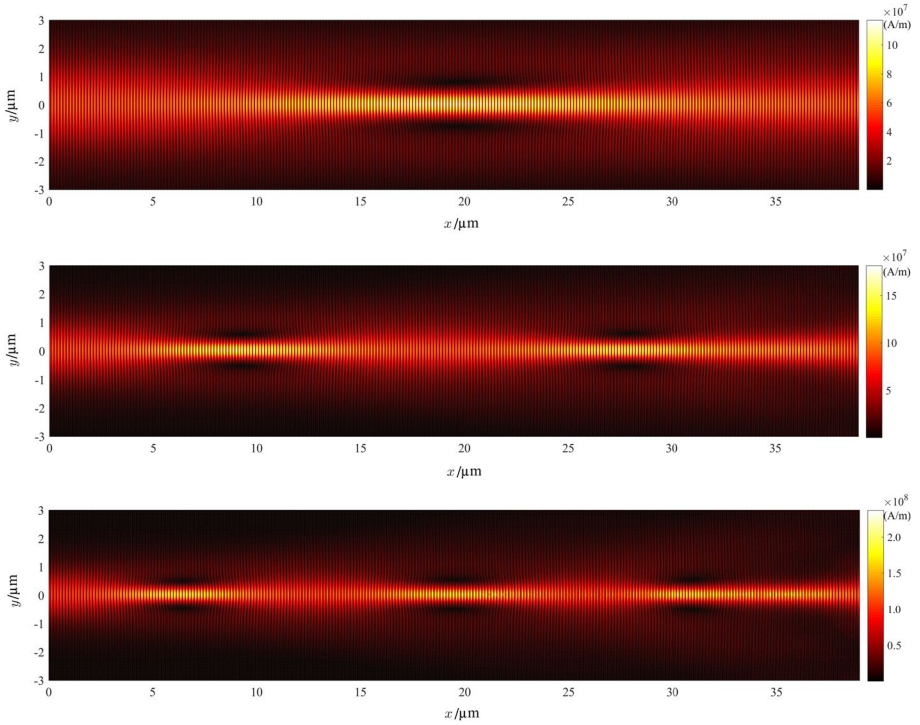

Fig. 8 2D spatial-soliton propagation: snapshot of $|H_z|$ at $T = 300$ fs for the second-order soliton propagation, with $w = r \times 667.0$ nm, $H_0 = 2/r \times 4.77 \times 10^7$ A/m, and $k = 2$. Top: **a** $r = 5.99/3.99$, middle: **b** $r = 3.99/3.99$, bottom: **c** $r = 3.26/3.99$

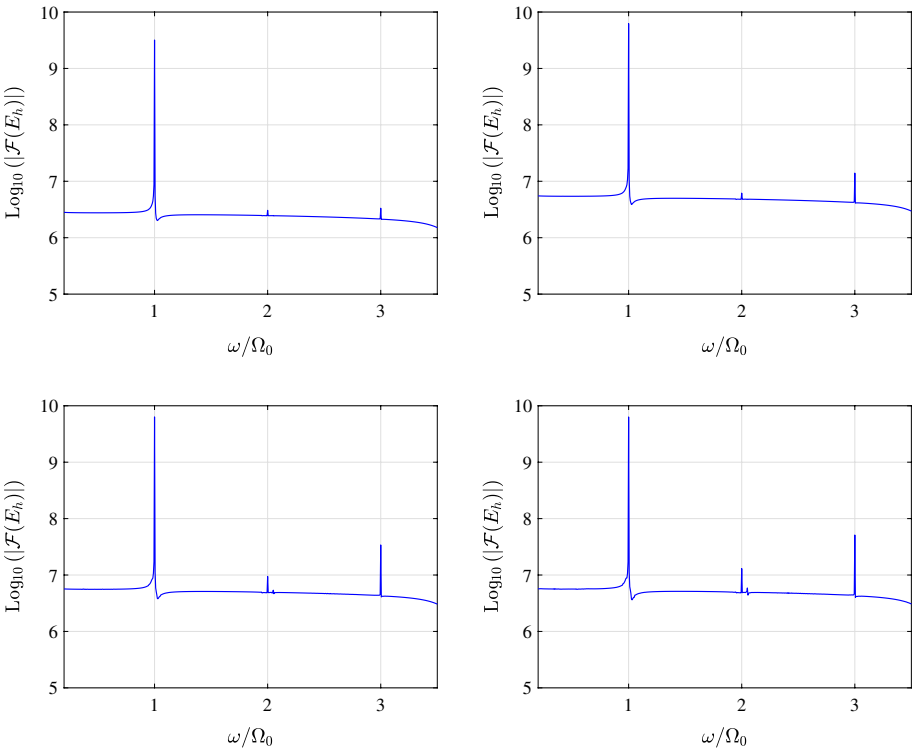

Fig. 9 2D spatial-soliton propagation: semi-log plot of the spectrum of the average of E_{yh} in the y direction at $x = 50$ nm and with $k = 2$. The top-left is for the fundamental soliton propagation, with $w = 667.0$ nm, $H_0 = 4.77 \times 10^7$ A/m. The rest is for the second-order soliton propagation, with $w = r \times 667.0$ nm, $H_0 = 2/r \times 4.77 \times 10^7$ A/m: top-right ($r = 5.99/3.99$), bottom-left ($r = 3.99/3.99$), bottom-right ($r = 3.26/3.99$)

Figure 7 shows the absolute value of the magnetic field $|H_{zh}|$ at $t = T$, when the parameters are taken as $w = 667.0$ nm and $H_0 = 4.77 \times 10^7$ A/m, corresponding to the fundamental soliton. As we can see, the shape and the magnitude remain the same as the wave propagates in the glass. When we enhance the magnitude of the incident waves properly, a second-order soliton is launched. This second-order soliton is the direct consequence of the interference between two fundamental solitons, leading to focusing and defocusing effects during the propagation, see Fig. 8 and its caption for three sets of parameter choices for ω and H_0.

To showcase the high-order harmonic generation, in Fig. 9 we present the discrete Fourier transform of $e_y(t)$ up to 200 time periods $T_s = 2\pi/\omega_c \times 200$ with $e_y(t)$ being the average in the y direction of the TE field $E_y(x, \cdot, t)$ at $x = 50$ nm. All four sets of parameter choices considered so far are examined. For all cases, one can clearly see that the third-order harmonic waves are generated in the interacting process. As the magnitude H_0 of the incident wave increases and its width ω decreases, the strength of the third-order harmonic also grows. From Fig. 9, we also observe small humps with magnitudes related to H_0 and w around the frequency of second-order harmonic waves. These second-order harmonic

waves, however, are not expected for isotropic media we consider here [7]. Based on a preliminary study by mesh refinement and domain enlargement, we find that those humps still persist even when the mesh is very refined, we hypothesize that the appearance of such a second-order harmonic wave is not a numerical artifact, but may be due to the incoming wave signal at the boundary.

3.2.3 Two-Beam Interactions in two Dimensions

Finally, we want to simulate another interesting example of two-beam interactions in the optical glasses [16], for which the same model and media property as in Sect. 3.2.2 are used.

The computational domain is $\Omega = [0, 50\,\mu\text{m}] \times [-15\,\mu\text{m}, 15\,\mu\text{m}]$. The final time is $T = 250\,\text{fs}$. The incident wave injected at the left boundary $x = 0$ is the superposition of two beams when $\eta = 1$,

$$\begin{aligned} \text{Beam}_1 &= \sin(\omega_c t)\text{sech}((y - y_1)/w), \\ \text{Beam}_2 &= \sin(\omega_c t + \varphi)\text{sech}((y - y_2)/w), \\ H(0, y, t) &= H_0\big(\text{Beam}_1 + \eta\ \text{Beam}_2\big). \end{aligned} \tag{51}$$

The amplitude $H_0 = 2.0 \times 10^8$ A/m, the characteristic width $w = 261$ nm, and $\omega_c = 2.35 \times 10^{15}$ rad/s. φ denotes the relative phase between the two beams. y_1 and y_2 are the locations of the peaks of the two beams, and they are taken as $y_1 = -1.317w \times d_n$ and $y_2 = -y_1$. Here d_n is the normalized separation distance and defined as the ratio of $|y_1 - y_2|$ and the full-width at half-maximum (FWHM). In our simulation, the time step size is $\Delta t = 0.05h$, and the spatial mesh is uniform with $h = 20$ nm in both the x and y directions. The discrete space V_h^k with $k = 2$ is used. The boundary condition at $x = 0$ is imposed through numerical fluxes, while on the rest of the domain boundaries, approximate absorbing boundary conditions as in [17] are applied. The error tolerance in the Newton solve is set as $\text{Err}_{\text{tol}} = 10^{-8}$.

We start with the case of a single beam, i.e., $\eta = 0$, initially located at $y_1 = 0$. Similar as the fundamental soliton examined in Sect. 3.2.2, from Fig. 10 one can observe that the narrow soliton preserves its profile quite well during the propagation.

We then consider the interactions between two beams, with $\eta = 1$. We firstly set the phase angle as $\varphi = 0$, and take different normalized separation distances, namely, $d_n = 3.5, 2.5, 1.5$. As one can see in Fig. 11, the initially separated beams rapidly merge into a single beam and exhibit periodic focusing and defocusing effects as the second-order spatial soliton shown in Sect. 3.2.2. And the distance when the two beams start to merge depends on the initial separation distance. Once two beams merge into one, they never

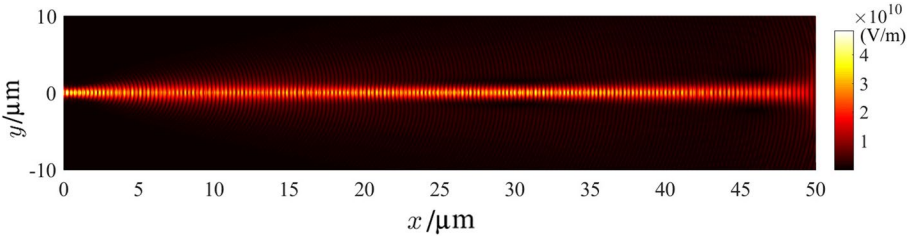

Fig. 10 Single-beam propagation, snapshot of $|\boldsymbol{E}|$ with $k = 2$ and $\varphi = 0$

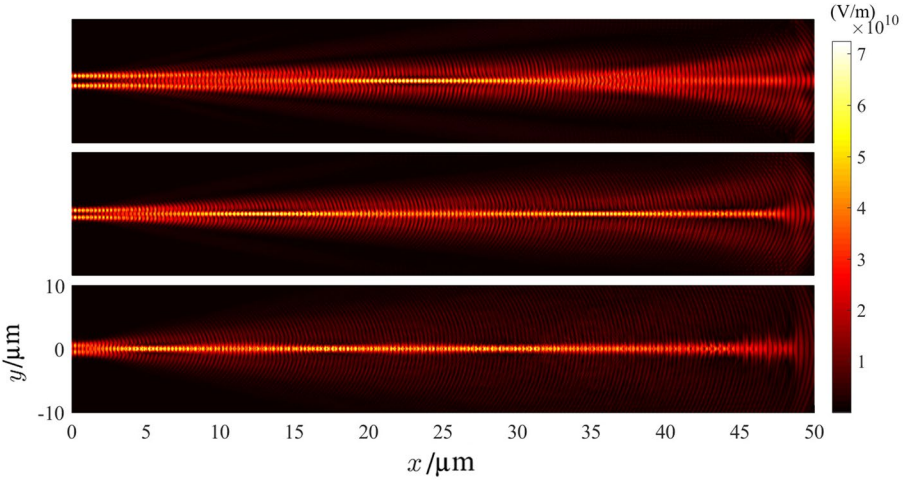

Fig. 11 Two-beam interactions, snapshot of $|\boldsymbol{E}|$ with $k = 2$ and $\varphi = 0$. From top to bottom: $d_n = 3.5, 2.5, 1.5$

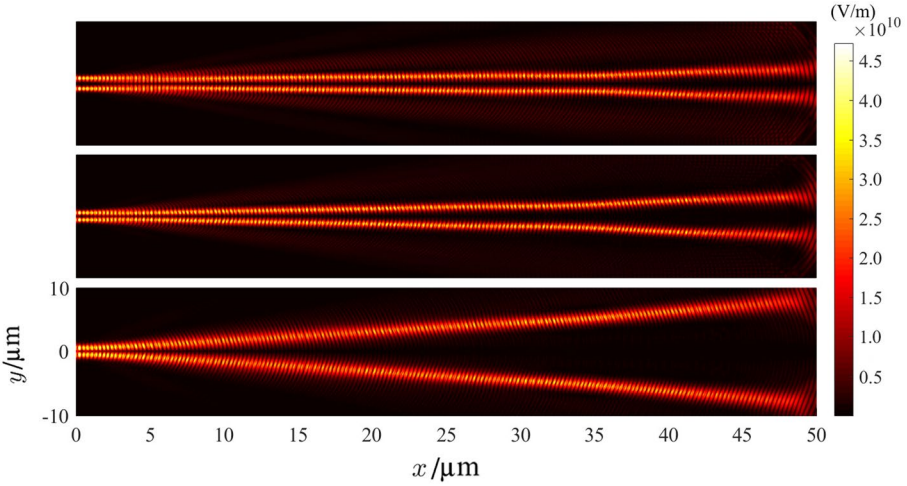

Fig. 12 Two-beam interactions, snapshot of $|\boldsymbol{E}|$ with $k = 2$ and $\varphi = \pi$. From top to bottom: $d_n = 3.5, 2.5, 1.5$

split again. On the right boundary, unphysical reflection occurs due to the relatively simple absorbing boundary conditions applied here.

When the phase angle is changed to $\varphi = \pi$, the interactions will be fairly different. Particularly, from Fig. 12, one can see that the two beams repel each other during the propagation and the onset of the mutual repulsion depend on the initial separation distance. In particular, when $d_n = 1.5$, the two beams begin to repel within a very short propagation distance, and stay divergent from each other.

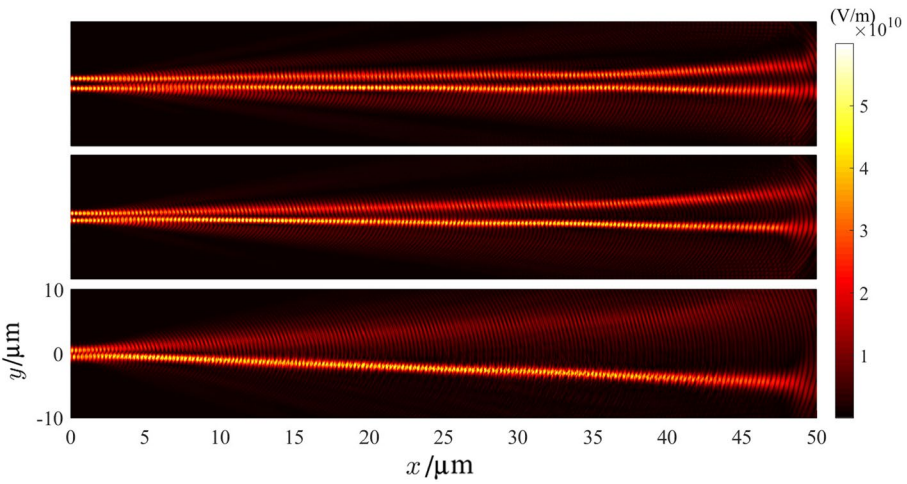

Fig. 13 Two-beam interactions, snapshot of $|\boldsymbol{E}|$ with $k = 2$ and $\varphi = \pi/2$. From top to bottom: $d_n = 3.5, 2.5, 1.5$

In the last case, we set the phase angle φ to be $\pi/2$. As shown in Fig. 13, the two beams also repel each other during the propagation. However, as the initial separation distance decreases, one of the two beams gets more energy at the expense of the other beam. All results above agree well with those reported in [16]. The magnetic field H_z, though not plotted here, shows similar behavior as the electric field $\boldsymbol{E}$ in each test.

4 Conclusions

In this work, a class of energy stable numerical schemes is proposed to simulate electromagnetic waves propagating in optical media where the media responses include the linear Lorentz dispersion, the instantaneous nonlinear cubic Kerr response, and the nonlinear delayed Raman molecular vibrational response. The underlying model is a mixed-order PDE-ODE system, and the methods are second-order accurate in time and arbitrary order accurate in space, with good efficiency especially in the parallel setting due to the decoupled nature of the nonlinear algebraic system over each time step. The main contribution to algorithmic development lies in the design of a new discretization of the nonlinear Kerr and Raman terms in a two-step temporal framework, to ensure a provable energy relation and stability. The methods are illustrated numerically for their accuracy, efficiency, energy stability, and parallel scalability through 1D and 2D examples with the scalar and vector-valued electric field. Together with the previous work [4, 17], the methods further enrich the energy stable numerical methods for nonlinear optics based on full Maxwell's equations. Though the mixed-order model consists of fewer unknowns, we observe numerically that the proposed methods are computational slightly more expensive than those in [4, 17] (with nodal implementation). For both types of methods, they solve the PDE part similarly, and the nonlinear algebraic systems solved at each interpolation node are essential of the same size (i.e., $d_E \times d_E$). The proposed methods here involve relatively more algebraic evaluations to generate the nonlinear systems. With the ODE part discretized by a two-step temporal method, our schemes also do not save in memory usage compared with those in [4, 17].

There are several future directions we plan to pursue. First, we will consider the model in its second-order form, that is, (1c)–(1e) with $\mu_0 \partial_{tt} \boldsymbol{D} + \nabla \times \nabla \times \boldsymbol{E} = \boldsymbol{0}$. The question is whether one can design an energy stable discretization for such model. Second, for (linear) models that are in first-order, mixed-order, or second-order forms, a detailed comparison study on the numerical dispersion relations for the designed schemes is necessary to compare their wave resolution properties. Finally, the simple absorbing boundary conditions adopted here from [17] have demonstrated some insufficiency to suppress unphysical reflections. More systematic investigation is needed for more robust and effective numerical boundary treatments in simulating optical phenomena in nonlinear and dispersive media. One can also explore the flexibility of DG methods to improve the efficiency through adaptive simulations. It remains an open question to design methods of higher than second-order temporal accuracy with provable energy stability and of similar local nature, hence the associated computational efficiency as our proposed methods.

Funding This work was supported by China Postdoctoral Science Foundation grant 2020TQ0344, by the NSFC grants 11871139 and 12101597, and by the NSF grants DMS-1720116, DMS-2012882, DMS-2011838, DMS-1719942, DMS-1913072.

Compliance with Ethical Standards

Conflict of Interest The authors have no competing interests to declare that are relevant to the content of this manuscript.

References

1. Agrawal, G.P.: Nonlinear fiber optics. In: Christiansen, P.L., Sørensen, M.P., Scott, A.C. (eds.) Nonlinear Science at the Dawn of the 21st Century, pp. 195–211. Springer (2000)
2. Bey, K.S., Oden, J.T., Patra, A.: A parallel *hp*-adaptive discontinuous Galerkin method for hyperbolic conservation laws. Appl. Numer. Math. **20**(4), 321–336 (1996)
3. Biswas, R., Devine, K.D., Flaherty, J.E.: Parallel, adaptive finite element methods for conservation laws. Appl. Numer. Math. **14**(1/2/3), 255–283 (1994)
4. Bokil, V.A., Cheng, Y., Jiang, Y., Li, F.: Energy stable discontinuous Galerkin methods for Maxwell's equations in nonlinear optical media. J. Comput. Phys. **350**, 420–452 (2017)
5. Bokil, V.A., Cheng, Y., Jiang, Y., Li, F., Sakkaplangkul, P.: High spatial order energy stable FDTD methods for Maxwell's equations in nonlinear optical media in one dimension. J. Sci. Comput. **77**(1), 330–371 (2018)
6. Boyd, R.W.: Nonlinear Optics. Academic Press, New York (2003)
7. Chen, C.-L.: Optical Solitons in Optical Fibers. Wiley, New York (2005)
8. Cockburn, B., Singler, J.R., Zhang, Y.: Interpolatory HDG method for parabolic semilinear PDEs. J. Sci. Comput. **79**(3), 1777–1800 (2019)
9. Fezoui, L., Lanteri, S.: Discontinuous Galerkin methods for the numerical solution of the nonlinear Maxwell equations in 1d. PhD thesis, INRIA (2015)
10. Fisher, A., White, D., Rodrigue, G.: An efficient vector finite element method for nonlinear electromagnetic modeling. J. Comput. Phys. **225**(2), 1331–1346 (2007)
11. Gilles, L., Hagness, S.C., Vázquez, L.: Comparison between staggered and unstaggered finite-difference time-domain grids for few-cycle temporal optical soliton propagation. J. Comput. Phys. **161**(2), 379–400 (2000)
12. Greene, J., Taflove, A.: Scattering of spatial optical solitons by subwavelength air holes. Microw. Wirel. Compon. Lett. IEEE **17**, 760–762 (2007)
13. Greene, J.H., Taflove, A.: General vector auxiliary differential equation finite-difference time-domain method for nonlinear optics. Opt. Express **14**(18), 8305–8310 (2006)
14. Hesthaven, J.S., Warburton, T.: Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications. Springer Science & Business Media, Berlin (2007)
15. Jackson, J.D.: Classical Electrodynamics. Wiley, New York (1999)

16. Lubin, Z., Greene, J.H., Taflove, A.: FDTD computational study of ultra-narrow TM non-paraxial spatial soliton interactions. IEEE Microw. Wirel. Compon. Lett. **21**(5), 228–230 (2011)
17. Lyu, M., Bokil, V.A., Cheng, Y., Li, F.: Energy stable nodal discontinuous Galerkin methods for nonlinear Maxwell's equations in multi-dimensions. J. Sci. Comput. **89**, 1–42 (2021)
18. Patel, C.K.N.: Efficient phase-matched harmonic generation in Tellurium with a CO_2 laser at 10.6 μ. Phys. Rev. Lett. **15**(26), 1027–1030 (1965)
19. Sørensen, M.P., Webb, G.M., Brio, M., Moloney, J.V.: Kink shape solutions of the Maxwell-Lorentz system. Phys. Rev. E **71**(3), 036602 (2005)

Authors and Affiliations

Maohui Lyu[1] · Vrushali A. Bokil[2] · Yingda Cheng[3,4] · Fengyan Li[5]

Maohui Lyu
mlyu@lsec.cc.ac.cn

Vrushali A. Bokil
bokilv@math.oregonstate.edu

Yingda Cheng
ycheng@msu.edu

[1] State Key Laboratory of Scientific and Engineering Computing (LSEC), Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100190, China

[2] Department of Mathematics, College of Science, Oregon State University, Corvallis, OR 97331, USA

[3] Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

[4] Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

[5] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA