Journal of Computational Physics 402 (2020) 109101

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

# Asymptotic and positivity preserving methods for Kerr-Debye model with Lorentz dispersion in one dimension

Zhichao Peng [a], Vrushali A. Bokil [b,1], Yingda Cheng [c,2], Fengyan Li [a,*,3]

[a] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[b] Department of Mathematics, Oregon State University, Corvallis, OR 97331, USA
[c] Department of Mathematics, Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

*Article history:*
Received 2 May 2019
Received in revised form 1 October 2019
Accepted 2 November 2019
Available online 12 November 2019

*Keywords:*
Full Maxwell's equations
Nonlinear media
Kerr-Debye with linear Lorentz
Asymptotic preserving
Positivity preserving
Energy stable

ABSTRACT

In this paper, we continue our recent developments in [4,5] to devise numerical methods that have important provable properties to simulate electromagnetic wave propagation in nonlinear optical media. Particularly, we consider the one dimensional Kerr-Debye model with the Lorentz dispersion, termed as the Kerr-Debye-Lorentz model, where the nonlinearity in the polarization is a relaxed cubic Kerr type effect. The polarization also includes the linear Lorentz dispersion. As the relaxation time $\varepsilon$ goes to zero, the model will approach the Kerr-Lorentz model.

The objective of this work is to devise and analyze asymptotic preserving (AP) and positivity preserving (PP) methods for the Kerr-Debye-Lorentz model. Being AP, the methods address the stiffness of the model associated with small $\varepsilon$, while capturing the correct Kerr-Lorentz limit as $\varepsilon \to 0$ on under-resolved meshes. Being PP, the third-order nonlinear susceptibility will stay non-negative and this is important for the energy stability. In the proposed methods, the nodal discontinuous Galerkin (DG) discretizations of arbitrary order accuracy are applied in space to effectively handle nonlinearity; in time, several first and second order methods are developed. We prove that the first order in time fully discrete schemes are AP, PP and also energy stable. For the second order temporal accuracy, a novel modified exponential time integrator is proposed for the stiff part of the auxiliary differential equations modeling the electric polarization, and this is a key ingredient for the methods to be both AP and PP. In addition to a straightforward discretization of the constitutive law, we further propose a non-trivial energy-based approximation, with which the energy stability is also established mathematically. Numerical examples are presented that include an ODE example, a manufactured solution, the soliton-like propagation and the propagation of *Sech* signal in fused bulk silica, to compare the proposed methods and to demonstrate the accuracy, AP and PP property. The effect of the finite relaxation time $\varepsilon$ in the model is also examined numerically.



* Corresponding author.
*E-mail addresses:* pengz2@rpi.edu (Z. Peng), bokilv@math.oregonstate.edu (V.A. Bokil), ycheng@msu.edu (Y. Cheng), lif@rpi.edu (F. Li).

[1] Research is supported by NSF grant DMS-1720116.
[2] Research is supported by NSF grants DMS-1453661 and DMS-1720023.
[3] Research is supported by NSF grant DMS-1719942.

https://doi.org/10.1016/j.jcp.2019.109101
0021-9991/

## 1. Introduction

In this paper, we continue our recent developments in [4,5] to devise numerical methods that have important *provable* properties to simulate electromagnetic wave propagation in nonlinear optical media. In such media, the material responds to light nonlinearly, and this contributes to many interesting optical effects including frequency mixing, second and third-harmonic generation, and self-steepening etc [3,6,26].

Instead of working with approximated models such as asymptotic or paraxial wave models [3,6], we focus on the full Maxwell's equations for the electromagnetic wave propagation along with a constitutive law that models linear and nonlinear electric polarizations. The polarization model examined here is phenomenological at a macroscopic scale, and it is given mathematically in the form of auxiliary differential equations (ADEs) appended to the Maxwell part. Full Maxwell models are important to simulate and capture optical phenomena at small length scales or with multiple spatial and temporal scales. Numerical simulation of Maxwell's equations is a well-studied area. In the literature, finite difference time domain (FDTD) methods [19,28,16], finite element (FEM) methods [15], pseudospectral methods [24], finite volume (FV) methods [2,10] are available for the integration of the Maxwell's equations in nonlinear media, along with ADEs for the material response with various effects. We want to mention our recent efforts in [4,5] to design energy stable FDTD and discontinuous Galerkin (DG) methods for Maxwell's equations with the nonlinear Kerr and Raman effects together with the linear Lorentz polarization, and the related references within.

In this paper, we consider the one dimensional Kerr-Debye model with the Lorentz dispersion. Here the linear effect in the polarization includes a Lorentz oscillator driven by the electric field $E$. The classical Kerr effect is a phenomenon in which the refractive index of a material changes proportionally to the square of the applied electric field. This cubic nonlinearity is often modeled as an instantaneous effect, and it is relevant when the pulse widths are sufficiently short. The Kerr-Debye model considered here takes into account a relaxation process in the Kerr effect. The relaxation time of the nonlinear Kerr effect is at the sub-femtosecond (fs) level, and when it is neglected, one will recover the classical instantaneous Kerr model. However, to describe ultrashort pulses (e.g. a few cycles), the effect of the finite relaxation time can be important [30,17,1]. For instance in [17], the authors investigated the propagation of ultrashort pulses through a dispersive medium with the linear effect via three linear Lorentz (resonance) poles and the nonlinear Kerr effect with a finite relaxation time. The effects of a finite relaxation time, such as the reduction of the production of phase matched harmonics and slowing down of the self-steepening of the electric field envelope, are observed.

Beyond the common issues one encounters to numerically simulate wave propagations in dispersive media, there are additional challenges and considerations associated with simulating Kerr-Debye type models. In the Kerr-Debye-Lorentz model, the relaxation of the Kerr effect is modeled by an ADE for the third-order nonlinear susceptibility $\chi$, given as $\partial_t \chi = -\frac{1}{\varepsilon}(\chi - E^2)$, that enters into the constitutive law for the electric displacement (see equation (2.1c)). The parameter $\varepsilon$ is the relaxation time. When $\varepsilon$ goes to 0, $\chi$ approaches $E^2$, and the model subsequently becomes the classical Kerr-Lorentz model. On the other hand, physically $\chi$ is always positive (or non-negative). When the relaxation time $\varepsilon$ is small, the model is stiff and standard numerical methods may suffer from a restricted time step condition, $\Delta t = O(\varepsilon)$, for numerical stability. Numerically, we also want to preserve the positivity of $\chi$, and this is important for the energy stability property. Numerical methods that preserve the non-negative property of $\chi$ are referred to as positivity preserving (PP) schemes.

The objective of this paper is to devise and analyze numerical methods that maintain at the discrete level some important properties of the continuous Kerr-Debye-Lorentz model. We also want to numerically examine the effect of the finite relaxation time in the underlying model. In particular, we will propose computational methods that are *simultaneously* asymptotic preserving (AP) and PP, and possibly also energy stable. AP methods are a class of numerical methods that are designed to work with a wide range of the relaxation time $\varepsilon$. In addition, as $\varepsilon \to 0$ and on the under-resolved meshes, the schemes will become consistent and stable discretizations of the limiting Kerr-Lorentz model that includes both the classical instantaneous nonlinear Kerr effect and the linear Lorentz dispersion. In other words, AP methods can address the stiffness of the model associated with the small $\varepsilon$, while capturing the correct asymptotic limit as $\varepsilon \to 0$ without resolving the small $\varepsilon$ scale. The interested readers are referred to [23,11,12] for some reviews on AP methods. Besides being AP, the proposed methods will be PP and produce non-negative approximation for the third-order nonlinear susceptibility $\chi$. The variable $\chi$ being non-negative can ensure the energy being nonnegative. Whenever possible, we also want to achieve an energy stability property of the proposed methods.

For the proposed methods, we first discretize in space using nodal DG methods that allow arbitrary spatial accuracy. The nodal DG spatial discretizations provide an efficient and natural way to handle the nonlinearity and to (potentially) preserve the positivity of $\chi$. They equivalently can be rewritten in an interpolatory weak form [14,8], that are more compact and will be used in most part of the paper. For the semi-discrete in space schemes, we further discretize in time. We first introduce first order temporal discretizations, that involve either forward or backward Euler methods for the Maxwell part, and the backward Euler method for discretizing the ADEs. We prove that the resulting fully discrete schemes are simultaneously AP and PP, and moreover they satisfy the energy stability property.

Even though possessing theoretically desirable properties, the first order in time methods, as expected, are too dissipative to be computationally efficient to simulate wave propagation (see numerical results in Section 6). To improve the accuracy and especially the dispersion property, we then propose several second order temporal discretizations. The resulting fully discrete schemes are all AP and PP, with one also being energy stable. In these methods, second order symplectic time integrators, including the staggered leapfrog and implicit trapezoidal methods are applied to the non-stiff Maxwell part (see

equations (2.1a)-(2.1b)). For the stiff part of the ADEs that evolves the third-order nonlinear susceptibility $\chi$ (see equation (2.1d)), a novel modified exponential time integrator is proposed. This is inspired by the work in [9] and is based on a suitably designed $O(\Delta t^2)$-perturbed form of the stiff equation. This modified exponential time integrator is the key for the resulting schemes to be simultaneously AP and PP. For the non-stiff part of the ADEs (see equations (2.1e)-(2.1f)) that models the linear Lorentz dispersion effect, implicit trapezoidal time integrator is applied. If the strategies described above are coupled with the constitutive law (2.1c) directly, the resulting second order fully discrete methods will be AP and PP. In order to ultimately get a second order in time method that is AP and PP and additionally satisfies energy stability, we also propose a nontrivial second order energy-based numerical approximation to the constitutive law (2.1c). Though extra computational effort is needed (see Sections 5.6.2 and 6.2.2), this reformulation leads to provable energy stability, which can not be achieved by directly imposing the constitutive law. We want to point out that the temporal discretization strategies introduced in this work together with the energy-based numerical constitutive relation can be coupled with other spatial discretizations, such as finite difference methods, to achieve AP and PP properties, and possibly also the energy stability.

To put our work into perspective, we briefly review some recent developments for AP and PP schemes of second order accuracy for nonlinear optical models and some other problems. In [7], semi-implicit Runge-Kutta (RK) schemes with a correction step preserving the positivity and the steady state are proposed for a shallow water model. In [22] and [21], implicit-explicit RK (IMEX-RK) schemes with a correction step preserving the positivity and the asymptotic limit are proposed for the Kerr-Debye model ([22]) and for the stiff BGK model ([21]). Both the semi-implicit RK schemes and the second order accurate IMEX-RK schemes above start with a convex combination of first order schemes with AP and PP properties. Due to the fact that there is no second or higher order implicit strong stability preserving (SSP) RK schemes [18], an additional correction step is introduced, based on the Fréchet derivative of the stiff part, to achieve the second order accuracy while keeping the PP property. In [13] and [20], AP and PP exponential time integrators are designed for some stiff kinetic model, by utilizing the structure of the collision operator of the underlying model. One can see Section 5.2.3 for more comments and discussions.

The rest of the paper is organized as follows. In Section 2, the physical model is introduced, and positivity preservation property and energy stability are derived. In Section 3, the nodal DG spatial discretization is formulated, together with its equivalent interpolatory form. We then define first order temporal discretizations in Section 4, and show the resulting semi-implicit or implicit fully discrete scheme is AP, PP and energy stable. In Section 5, the formulation of the second order in time AP and PP schemes are presented. A modified second order exponential time integrator is designed for the stiff part of the ADEs, and this is the key to achieve the AP and PP property of the resulting schemes. To further achieve energy stability, an energy-based numerical constitutive relation is designed. The AP, PP and possibly the energy stable properties are established, and we also discuss some aspects of the nonlinear algebraic solvers. The performance of the proposed schemes are numerically demonstrated in Section 6, which is followed by concluding remarks and future work in Section 7.

## 2. The model

In this paper, we consider the one dimensional Kerr-Debye model with the linear Lorentz dispersion represented by the following first order system.

$$\mu_0 \partial_t H = \partial_x E, \tag{2.1a}$$

$$\partial_t D = \partial_x H, \tag{2.1b}$$

$$D = \epsilon_0 \left( (\epsilon_\infty + a\chi) E + P \right), \tag{2.1c}$$

$$\partial_t \chi = -\frac{1}{\varepsilon}(\chi - E^2), \tag{2.1d}$$

$$\partial_t P = J, \tag{2.1e}$$

$$\partial_t J + \frac{1}{\tau} J + \omega_0^2 P = \omega_p^2 E, \tag{2.1f}$$

in $\Omega \times [0, T]$, where $E$ is the electric field, $H$ is the magnetic field, $D$ is the electric displacement, $\chi$ is the third-order nonlinear susceptibility, $P$ is the linear polarization due to the Lorentz dispersion, and $J$ is the linear polarization current density. The parameters are the electric permittivity of free space $\epsilon_0$, the relative electric permittivity in the limit of the infinite frequency $\epsilon_\infty$, the magnetic permeability of free space $\mu_0$, the resonance frequency of the medium $\omega_0$, the plasma frequency of the medium $\omega_p$, a damping constant $\tau$, and a relaxation time $\varepsilon$. The equations (2.1a)-(2.1b) are the Maxwell's equations, (2.1c) is the constitutive law, while (2.1d)-(2.1f) are auxiliary differential equations (ADEs) modeling the evolution in time of the nonlinear susceptibility and linear polarization. The system (2.1) will be referred to as the Kerr-Debye-Lorentz model.

Formally, as the relaxation time $\varepsilon$ goes to 0, (2.1d) will become $\chi = E^2$, and we will get the asymptotic limit of the system (2.1),

$$\mu_0 \partial_t H = \partial_x E, \tag{2.2a}$$

$$\partial_t D = \partial_x H, \tag{2.2b}$$

$$D = \epsilon_0\big((\epsilon_\infty + aE^2)E + P\big), \tag{2.2c}$$

$$\partial_t P = J, \tag{2.2d}$$

$$\partial_t J + \frac{1}{\tau} J + \omega_0^2 P = \omega_p^2 E, \tag{2.2e}$$

which is the Kerr model with the linear Lorentz dispersion. Unlike in the Kerr-Debye-Lorentz model, the Kerr effect in the limiting Kerr-Lorentz model (2.2) is modeled as an instantaneous response.

Below we gather some important properties of systems (2.1) and (2.2).

**Theorem 2.1** *(Positivity preserving property of the susceptibility $\chi$). For the system (2.1), suppose $\chi(x,0) \geq 0$, then $\chi(x,t) \geq 0$, $\forall t \geq 0$.*

**Proof.** From (2.1d), $\chi(x,t) = e^{-\frac{t}{\varepsilon}}\chi(x,0) + \int_0^t e^{\frac{\tau - t}{\varepsilon}} E(x,\tau)^2 d\tau$. Hence, the theorem holds. □

**Theorem 2.2** *(Energy stability for system (2.1)). Under periodic boundary conditions, the energy for (2.1)*

$$\mathcal{E}(t) = \frac{1}{2}\int_\Omega \left(\epsilon_0(\epsilon_\infty + a\chi)E^2 + \mu_0 H^2 + \frac{\epsilon_0 a}{2}\chi^2 + \frac{\epsilon_0\omega_0^2}{\omega_p^2}P^2 + \frac{\epsilon_0}{\omega_p^2}J^2\right)dx \tag{2.3}$$

*satisfies the following relation,*

$$\frac{d\mathcal{E}(t)}{dt} = -\int_\Omega \left(\frac{\epsilon_0 a}{2\varepsilon}(\chi - E^2)^2 + \frac{\epsilon_0}{\tau\omega_p^2}J^2\right)dx \leq 0. \tag{2.4}$$

**Proof.** Multiply $H$ to (2.1a), $E$ to (2.1b), $(\chi - E^2)$ to (2.1d), $P$ to (2.1e) and $J$ to (2.1f), then integrate in $x$ and sum the resulting equations up with suitable weights. Based on (2.1c), the energy stability (2.4) follows. □

**Remark 2.3.** The fact $\chi \geq 0$ guarantees the energy $\mathcal{E}$ is non-negative. Hence, (2.4) gives the energy stability of (2.1).

**Theorem 2.4** *(Energy stability for system (2.2)). Under periodic boundary conditions, the energy for (2.2)*

$$\mathcal{E}_{Kerr}(t) = \frac{1}{2}\int_\Omega \left(\epsilon_0\epsilon_\infty E^2 + \frac{3\epsilon_0 a}{2}E^4 + \mu_0 H^2 + \frac{\epsilon_0\omega_0^2}{\omega_p^2}P^2 + \frac{\epsilon_0}{\omega_p^2}J^2\right)dx \tag{2.5}$$

*satisfies the following relation*

$$\frac{d\mathcal{E}_{Kerr}(t)}{dt} = -\frac{\epsilon_0}{\tau\omega_p^2}\int_\Omega J^2 dx \leq 0. \tag{2.6}$$

**Proof.** Multiply $H$ to (2.2a), $E$ to (2.2b), $P$ to (2.2d) and $J$ to (2.2e), then integrate in $x$ and sum the resulting equations up with suitable weights. Based on (2.2c), the energy stability is obtained. □

**Remark 2.5.** Formally, as $\varepsilon \to 0$, we have $\chi \to E^2$ and $\mathcal{E} \to \mathcal{E}_{Kerr}$.

The goal of this paper is to design efficient numerical schemes for the Kerr-Debye-Lorentz system (2.1) that can maintain the important physical properties discussed above, such as the asymptotic limit when $\varepsilon \to 0$, positivity of $\chi$ and energy stability. We will discuss the spatial discretization first in Section 3, followed by first and second order temporal discretizations in Sections 4 and 5.

## 3. Semi-discrete DG spatial discretization

In this section, we formulate DG spatial discretizations for system (2.1). The nonlinearity is treated via the nodal form, or equivalently the interpolatory form, of the DG method. Other spatial discretizations may also be viable, such as finite difference methods, which will be investigated in a future work.

We assume the computational domain is $\Omega = [x_{\min}, x_{\max}]$, with the mesh defined as $x_{\min} = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N_x+\frac{1}{2}} = x_{\max}$. Let $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})$, $\Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$, for $1 \leq j \leq N_x$, and $h = \min_{j=1}^{N_x}\Delta x_j$. The mesh is

assumed to be quasi-uniform, namely, there exists a positive constant $\delta > 0$, such that $h/\Delta x_j \le \delta, \forall j$ as the mesh is refined. We use the following discrete space

$$U_h^K = \left\{ p \in L^2(\Omega_x) : p|_{I_j} \in P^K(I_j), \forall j \right\},$$

where $P^K(I_j)$ is the set of polynomials with degree at most $K$ on $I_j$. The vector version of the space is denoted as

$$(U_h^K)^m = \left\{ p = (p_1, \dots, p_m)^T : p_l \in U_h^K, \forall 1 \le l \le m \right\}.$$

Let $v(x^{\pm}) = \lim_{\Delta x \to 0^{\pm}} v(x + \Delta x)$, $v^{\pm}_{j\pm\frac{1}{2}} = v(x^{\pm}_{j\pm\frac{1}{2}})$, and the average and jump be $\{v\}_{j+\frac{1}{2}} = \frac{1}{2}(v^+_{j+\frac{1}{2}} + v^-_{j+\frac{1}{2}})$ and $[v]_{j+\frac{1}{2}} = v^+_{j+\frac{1}{2}} - v^-_{j+\frac{1}{2}}$.

To introduce the nodal DG formulation, we consider the Gaussian-Legendre quadrature points $\{\xi_k\}_{k=0}^K$ on $[-1, 1]$ with the weights $\{\widehat{\omega}_k\}_{k=0}^K$, satisfying $\sum_{k=0}^K \widehat{\omega}_k = 2$. Let $x_{jk} = x_j + \frac{\Delta x_j}{2}\xi_k$, and the associated interpolant is $L_{jk}(x) \in U_h^K$, satisfying

$$L_{jk}(x_{il}) = \delta_{ij}\delta_{kl},$$

where $\delta_{ij}$ is the Kronecker delta. The collection of $\{L_{jk}\}_{j=1,k=0}^{N_x,K}$ forms a set of Lagrange basis of $U_h^K$. For any $p \in U_h^K$, it can be expressed as $p(x) = \sum_{j=1}^{N_x}\sum_{k=0}^K p_{jk}L_{jk}(x)$, where $p_{jk} = p(x_{jk})$. Such expansion will be used throughout the paper.

The semi-discrete nodal DG method for system (2.1) is: we look for $\mathbf{u_h} = (H_h, D_h, E_h, P_h, J_h, \chi_h)^T$, with $\mathbf{u_h}(x,t) = \sum_{j=1}^{N_x}\sum_{k=0}^K \mathbf{u_{jk}}(t)L_{jk}(x)$ and $\mathbf{u_{jk}} = \left(H_{jk}, D_{jk}, E_{jk}, P_{jk}, J_{jk}, \chi_{jk}\right)^T$ such that

$$\frac{\Delta x_j}{2}\mu_0\widehat{\omega}_k\partial_t H_{jk} + \frac{\Delta x_j}{2}\sum_{l=0}^K \widehat{\omega}_l E_{jl}\partial_x L_{jk}(x_{jl}) - (\widehat{E_h})_{j+\frac{1}{2}}\left(L_{jk}\right)^-_{j+\frac{1}{2}} + (\widehat{E_h})_{j-\frac{1}{2}}\left(L_{jk}\right)^+_{j-\frac{1}{2}} = 0, \tag{3.7a}$$

$$\frac{\Delta x_j}{2}\widehat{\omega}_k\partial_t D_{jk} + \frac{\Delta x_j}{2}\sum_{l=0}^K \widehat{\omega}_l H_{jl}\partial_x L_{jk}(x_{jl}) - (\widetilde{H_h})_{j+\frac{1}{2}}\left(L_{jk}\right)^-_{j+\frac{1}{2}} + (\widetilde{H_h})_{j-\frac{1}{2}}\left(L_{jk}\right)^+_{j-\frac{1}{2}} = 0, \tag{3.7b}$$

$$D_{jk} = \epsilon_0\left((\epsilon_\infty + a\chi_{jk})E_{jk} + P_{jk}\right), \tag{3.7c}$$

$$\partial_t \chi_{jk} = -\frac{1}{\varepsilon}\left(\chi_{jk} - E^2_{jk}\right), \tag{3.7d}$$

$$\partial_t P_{jk} = J_{jk}, \tag{3.7e}$$

$$\partial_t J_{jk} + \frac{1}{\tau}J_{jk} + \omega_0^2 P_{jk} = \omega_p^2 E_{jk}, \qquad k = 0, 1, \cdots K, \quad j = 1, 2, \cdots N_x, \tag{3.7f}$$

where $\widehat{E_h}$ and $\widetilde{H_h}$ are numerical fluxes, which can be chosen as one of the following,

$$\widetilde{H_h} = H_h^+ \quad \text{and} \quad \widehat{E_h} = E_h^-, \qquad \text{(alternating flux 1)} \tag{3.8a}$$

$$\widetilde{H_h} = H_h^- \quad \text{and} \quad \widehat{E_h} = E_h^+, \qquad \text{(alternating flux 2)} \tag{3.8b}$$

$$\widetilde{H_h} = \{H_h\} + \frac{1}{2}\sqrt{\frac{\epsilon_0\epsilon_\infty}{\mu_0}}[E_h] \quad \text{and} \quad \widehat{E_h} = \{E_h\} + \frac{1}{2}\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_\infty}}[H_h], \qquad \text{(upwind flux)}. \tag{3.8c}$$

The "upwind" flux in (3.8c) is actually the upwind flux for the Maxwell's equations in free space. Without confusion, we will refer to it as the upwind flux in this paper.

The nodal DG method above can be rewritten more compactly in an interpolatory weak form [14,8]. Define a local interpolation operator $I_h^j : H^1(I_j) \to P^K(I_j)$ as $(I_h^j u)(x) = \sum_{k=0}^K u(x_{jk})L_{jk}(x)$, and the global one $I_h : H_h^1(\Omega) = \left\{p(x) : p(x)|_{I_j} \in H^1(I_j), \forall 1 \le j \le N_x\right\} \to U_h^K$ as $I_h|_{I_j} = I_h^j$, $\forall 1 \le j \le N_x$. The nodal DG method (3.7) is equivalent to: we look for $\mathbf{u_h} = (H_h, D_h, E_h, P_h, J_h, \chi_h)^T \in (U_h^k)^6$ such that for any $\phi, \varphi \in U_h^K$ and $1 \le j \le N_x$:

$$\mu_0\int_{I_j}\partial_t H_h\phi dx + \int_{I_j} E_h\partial_x\phi dx - (\widehat{E_h})_{j+\frac{1}{2}}\phi^-_{j+\frac{1}{2}} + (\widehat{E_h})_{j-\frac{1}{2}}\phi^+_{j-\frac{1}{2}} = 0, \tag{3.9a}$$

$$\int_{I_j}\partial_t D_h\varphi dx + \int_{I_j} H_h\partial_x\varphi dx - (\widetilde{H_h})_{j+\frac{1}{2}}\varphi^-_{j+\frac{1}{2}} + (\widetilde{H_h})_{j-\frac{1}{2}}\varphi^+_{j-\frac{1}{2}} = 0, \tag{3.9b}$$

$$D_h = \epsilon_0\left(\epsilon_\infty E_h + aI_h\left(\chi_h E_h\right)\right) + P_h\right), \tag{3.9c}$$

$$\partial_t \chi_h = -\frac{1}{\varepsilon}\left(\chi_h - I_h(E_h^2)\right), \tag{3.9d}$$

$$\partial_t P_h = J_h, \tag{3.9e}$$

$$\partial_t J_h + \frac{1}{\tau} J_h + \omega_0^2 P_h = \omega_p^2 E_h. \tag{3.9f}$$

The equivalency can be established by the exactness of the $K+1$ points Gaussian quadrature for polynomials with degree up to $2K+1$, which is stated in the following lemma.

**Lemma 3.1.** *For any $v \in P^{2K+1}(I_j)$, the $K+1$ points Gaussian quadrature is exact, namely,*

$$\int_{I_j} v(x)dx = \frac{\Delta x_j}{2} \sum_{k=0}^{K} \widehat{\omega}_k v(x_{jk}). \tag{3.10}$$

The following "commutative property" of the interpolation operator will be used later in the analysis.

**Lemma 3.2.** *For any $v_1, v_2, v_3 \in U_h^K$, there holds*

$$\int_{\Omega} \Big(I_h(v_1 v_2)\Big)(x) v_3(x)dx = \int_{\Omega} \Big(I_h(v_2 v_3)\Big)(x) v_1(x)dx. \tag{3.11}$$

**Proof.** Note that $I_h(v_1 v_2) \in U_h^K$ and $I_h(v_2 v_3) \in U_h^K$, based on Lemma 3.1, we have

$$\int_{\Omega} \Big(I_h(v_1 v_2)\Big)(x) v_3(x)dx = \sum_{j=1}^{N_x} \sum_{k}^{K} \frac{\Delta x_j}{2} \widehat{\omega}_k v_1(x_{jk}) v_2(x_{jk}) v_3(x_{jk}) = \int_{\Omega} \Big(I_h(v_2 v_3)\Big)(x) v_1(x)dx. \quad \square \tag{3.12}$$

Standard inverse inequalities for polynomial spaces will also be used frequently: there exist constants $C_{inv}$ and $\widehat{C}_{inv}$, dependent of $K$, such that

$$\Delta x_j w(x_{j\pm\frac{1}{2}})^2 \le C_{inv} ||w||^2_{L^2(I_j)}, \quad \forall w \in P^K(I_j), \tag{3.13a}$$

$$\Delta x_j^2 ||w_x||^2_{L^2(I_j)} \le \widehat{C}_{inv} ||w||^2_{L^2(I_j)}, \quad \forall w \in P^K(I_j). \tag{3.13b}$$

Here $||\cdot||_{L^2(I_j)}$ denotes the standard $L^2$ norm on $I_j$. It is easy to see $C_{inv} = 1$ when $K = 0$.

## 4. First order in time AP, PP and energy stable fully discrete schemes

In this section, we will further introduce first order temporal discretizations to get fully discrete schemes. Particularly, we apply the forward or backward Euler method to the Maxwell part and the backward Euler method to the ADEs. The resulting fully discrete schemes are called the semi-implicit first order scheme (SI1) and the fully implicit first order scheme (FI1), respectively, and they will be shown to be AP, PP, and satisfy energy stability.

For simplicity, we present the fully discrete scheme in the interpolatory form. The semi-implicit scheme (SI1) is defined as follows. Given $\mathbf{u_h^n} \in (U_h^K)^6$ at $t^n$, we look for $\mathbf{u_h^{n+1}} \in (U_h^K)^6$ at $t^{n+1} = t^n + \Delta t$, such that for any $\phi, \varphi \in U_h^K$ and $1 \le j \le N_x$,

$$\int_{I_j} \mu_0 \frac{H_h^{n+1} - H_h^n}{\Delta t} \phi dx + \int_{I_j} E_h^n \partial_x \phi dx - (\widehat{E_h^n})_{j+\frac{1}{2}} \phi^-_{j+\frac{1}{2}} + (\widehat{E_h^n})_{j-\frac{1}{2}} \phi^+_{j-\frac{1}{2}} = 0, \tag{4.14a}$$

$$\int_{I_j} \frac{D_h^{n+1} - D_h^n}{\Delta t} \varphi dx + \int_{I_j} H_h^n \partial_x \varphi dx - (\widetilde{H_h^n})_{j+\frac{1}{2}} \varphi^-_{j+\frac{1}{2}} + (\widetilde{H_h^n})_{j-\frac{1}{2}} \varphi^+_{j-\frac{1}{2}} = 0, \tag{4.14b}$$

$$D_h^{n+1} = \epsilon_0 \Big(\epsilon_\infty E_h^{n+1} + a I_h(\chi_h^{n+1} E_h^{n+1}) + P_h^{n+1}\Big), \tag{4.14c}$$

$$\frac{\chi_h^{n+1} - \chi_h^n}{\Delta t} = -\frac{1}{\varepsilon} \left(\chi_h^{n+1} - I_h((E_h^{n+1})^2)\right), \tag{4.14d}$$

$$\frac{P_h^{n+1} - P_h^n}{\Delta t} = J_h^{n+1}, \tag{4.14e}$$

$$\frac{J_h^{n+1} - J_h^n}{\Delta t} + \frac{1}{\tau} J_h^{n+1} + \omega_0^2 P_h^{n+1} = \omega_p^2 E_h^{n+1}, \tag{4.14f}$$

where the numerical fluxes are chosen as one of the options in (3.8). Similarly, the fully implicit scheme (I1) can be defined by coupling

$$\int_{I_j} \mu_0 \frac{H_h^{n+1}-H_h^n}{\Delta t}\phi dx + \int_{I_j} E_h^{n+1}\partial_x\phi dx - (\widehat{E_h^{n+1}})_{j+\frac{1}{2}}\phi^-_{j+\frac{1}{2}} + (\widehat{E_h^{n+1}})_{j-\frac{1}{2}}\phi^+_{j-\frac{1}{2}} = 0, \tag{4.15a}$$

$$\int_{I_j} \frac{D_h^{n+1}-D_h^n}{\Delta t}\varphi dx + \int_{I_j} H_h^{n+1}\partial_x\varphi dx - (\widetilde{H_h^{n+1}})_{j+\frac{1}{2}}\varphi^-_{j+\frac{1}{2}} + (\widetilde{H_h^{n+1}})_{j-\frac{1}{2}}\varphi^+_{j-\frac{1}{2}} = 0, \tag{4.15b}$$

with (4.14c)-(4.14f).

### 4.1. AP and PP properties

As $\varepsilon$ goes to 0, formally, we can get the limiting scheme of the semi-implicit scheme (4.14): given $\mathbf{v_h^n} = (H_h^n, D_h^n, E_h^n, P_h^n, J_h^n)^T \in (U_h^K)^5$, look for $\mathbf{v_h^{n+1}} = \left(H_h^{n+1}, D_h^{n+1}, E_h^{n+1}, P_h^{n+1}, J_h^{n+1}\right)^T \in (U_h^K)^5$ such that for any $\phi, \varphi \in U_h^K$ and $1 \le j \le N_x$,

$$\int_{I_j} \mu_0 \frac{H_h^{n+1}-H_h^n}{\Delta t}\phi dx + \left(E_h^n, \partial_x\phi\right) - \widehat{(E_h^n)}_{j+\frac{1}{2}}\phi^-_{j+\frac{1}{2}} + \widehat{(E_h^n)}_{j-\frac{1}{2}}\phi^+_{j-\frac{1}{2}} = 0, \tag{4.16a}$$

$$\int_{I_j} \frac{D_h^{n+1}-D_h^n}{\Delta t}\varphi dx + \left(H_h^n, \partial_x\varphi\right) - \widetilde{(H_h^n)}_{j+\frac{1}{2}}\varphi^-_{j+\frac{1}{2}} + \widetilde{(H_h^n)}_{j-\frac{1}{2}}\varphi^+_{j-\frac{1}{2}} = 0, \tag{4.16b}$$

$$D_h^{n+1} = \epsilon_0\left(\epsilon_\infty E_h^{n+1} + aI_h((E_h^{n+1})^3) + P_h^{n+1}\right), \tag{4.16c}$$

$$\frac{P_h^{n+1}-P_h^n}{\Delta t} = J_h^{n+1}, \tag{4.16d}$$

$$\frac{J_h^{n+1}-J_h^n}{\Delta t} + \frac{1}{\tau}J_h^{n+1} + \omega_0^2 P_h^{n+1} = \omega_p^2 E_h^{n+1}. \tag{4.16e}$$

Similarly, the limiting scheme of the fully implicit scheme is (4.15) coupled with (4.16c)-(4.16e). Both limiting schemes are first order in time nodal DG methods for the Kerr model with the Lorentz dispersion in (2.2). This shows the proposed schemes are formally AP.

The following theorem confirms the PP property.

**Theorem 4.1.** *Both first order in time fully discrete schemes are PP,* in the sense *that if* $\chi_{jk}^n \ge 0$, *then* $\chi_{jk}^{n+1} \ge 0$, *for any* $1 \le j \le N_x$, $0 \le k \le K$, *and* $\forall n \ge 0$.

**Proof.** Note that (4.14d) is equivalent to

$$\frac{\chi_{jk}^{n+1}-\chi_{jk}^n}{\Delta t} = -\frac{1}{\varepsilon}\left(\chi_{jk}^{n+1} - (E_{jk}^{n+1})^2\right), \qquad 1 \le j \le N_x, \quad 0 \le k \le K. \tag{4.17}$$

This leads to $\chi_{jk}^{n+1} = \frac{\varepsilon}{\varepsilon+\Delta t}\chi_{jk}^n + \frac{\Delta t}{\varepsilon+\Delta t}(E_{jk}^{n+1})^2 \ge \min\left(\chi_{jk}^n, (E_{jk}^{n+1})^2\right)$. From here, one can conclude the PP property. □

### 4.2. Energy stability

We now state and perform the energy stability analysis for the proposed first order fully discrete schemes. The proof can be found in Appendix A.

**Theorem 4.2.** *Given* $T > 0$. *Assume* $\chi(x,0) \ge 0$ *and periodic boundary conditions, for the proposed first order fully discrete schemes, we introduce*

$$\mathcal{E}_{h,1}^n = \frac{1}{2}\left(\epsilon_0\epsilon_\infty||E_h^n||^2 + a\epsilon_0\left(\chi_h^n, I_h((E_h^n)^2)\right)\right) + \mu_0||H_h^n||^2 + \frac{\epsilon_0 a}{2}||\chi_h^n||^2 + \frac{\epsilon_0\omega_0^2}{\omega_p^2}||P_h^n||^2 + \frac{\epsilon_0}{\omega_p^2}||J_h^n||^2\Bigg), \tag{4.18}$$

*then* $\mathcal{E}_{h,1}^n \ge 0$, $\forall n$, *and hence it defines a discrete energy. Moreover, the following energy stability results hold.*

1.) *For the semi-implicit scheme (4.14), with either the upwind or alternating flux defined in (3.8), we have* $\mathcal{E}_{h,1}^n \le e^{C'T}\mathcal{E}_{h,1}^0, \forall n \le \frac{T}{\Delta t}$, *under the time step condition* $\Delta t \le Ch^2$. *Here* $C$ *can be taken as any fixed generic constant, while* $C'$ *is a constant dependent of* $K$, $\epsilon_0$, $\epsilon_\infty$, $\mu_0$ *and the quasi-uniform mesh parameter* $\delta$. *Furthermore, when* $K = 0$ *with the upwind flux, the scheme satisfies a strong energy stability result, namely,*

$$\mathcal{E}_{h,1}^{n+1} \le \mathcal{E}_{h,1}^n, \quad \forall n \ge 0,$$

*under a hyperbolic type time step condition* $\Delta t \le \frac{\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{4\delta}h$.

2.) *For the fully implicit scheme (4.15), with either the upwind or alternating flux defined in (3.8), the strong energy stability holds, namely,* $\mathcal{E}_{h,1}^{n+1} \le \mathcal{E}_{h,1}^n, \forall n \ge 0$ *for any* $\Delta t > 0$. *In other words, the scheme (4.15) is unconditionally energy stable.*

*Here in (4.18),* $||\cdot||$ *is the* $L^2$ *norm for* $L^2(\Omega)$, *induced from the standard* $L^2$ *inner product* $(\cdot,\cdot)$.

In summary, the first order semi-implicit scheme SI1 in (4.14) is AP, PP and energy stable under a suitable CFL condition. The first order fully implicit scheme FI1 in (4.15) is AP, PP, and unconditionally energy stable.

## 5. Second order in time AP and PP fully discrete schemes

In this section, we propose several second order AP and PP schemes for (2.1). All these schemes include three main ingredients:

1. the staggered leapfrog or trapezoidal nodal DG method for the Maxwell (also the PDE) part (2.1a)-(2.1b);
2. a modified exponential time integrator, applied to the stiff part of the ADEs, namely (2.1d), coupled with the trapezoidal time integrator, applied to (2.1e)-(2.1f), the non-stiff part of the ADEs;
3. an algebra-based or energy-based numerical constitutive law approximating the relation (2.1c).

For the PDE part, we use the second order staggered leapfrog or trapezoidal temporal discretization coupled with the nodal DG spatial discretization. To obtain AP and PP properties, we design a new modified exponential time integrator for the stiff part (2.1d), combined with the trapezoidal time integrator for (2.1e)-(2.1f), the non-stiff part of the ADEs. For the constitutive law, two different approximations are considered. The first one is based on the algebraic relation (2.1c). This algebra-based approximation is straightforward, yet it does not lead to a provable energy stability. In order to achieve such provable result, we propose an energy-based numerical constitutive law.

As all the proposed schemes employ a modified exponential time integrator for (2.1f), we call them the modified exponential leapfrog (or trapezoidal) method with an algebra-based (or energy-based) approximation for the constitutive law. Particularly, we consider three methods: the modified exponential trapezoidal method with algebra-based numerical constitutive law (META), the modified exponential leapfrog method with algebra-based numerical constitutive law (MELA), and the modified exponential leapfrog method with energy-based numerical constitutive law (MELE). All three schemes preserve the correct asymptotic limit as $\varepsilon \to 0$ and the positivity of $\chi$. The MELE scheme can further be proved to preserve an energy relation on the discrete level. In actual numerical implementations, all three schemes require to solve some nonlinear algebraic equations, for which Newton-type nonlinear solvers are used. These solvers performs relatively more robust for the algebra-based schemes (META and MELA). When the MELE scheme is applied to some examples, we observe that a good initial guess for the Newton-type solver is crucial to capture the physically relevant solution (see Section 5.6 for more details). We do not consider the method that combines the trapezoidal method for the PDE part and the energy-based numerical constitutive law, as this method is observed to be too sensitive to the initial guesses in the Newton solvers. In the following, we will discuss each ingredient of the proposed methods, see Sections 5.1-5.3. The PP property is established in Section 5.2.1 and the AP property is in Section 5.4. In Section 5.5, energy stability result is established for the MELE scheme. In Section 5.6, we discuss the solvers for the nonlinear algebraic equations resulted from the proposed schemes.

### 5.1. Discretization of the Maxwell part

For the Maxwell part (2.1a)-(2.1b), we apply the second order staggered leapfrog or trapezoidal time integrator. With the leapfrog method, the fully discrete scheme is: given $\mathbf{u_h^n} \in (U_h^K)^6$, we look for $\mathbf{u_h^{n+1}} \in (U_h^K)^6$, such that for any $\phi, \varphi \in U_h^K$ and $1 \le j \le N_x$,

$$\mu_0 \int_{I_j} \frac{H_h^{n+\frac{1}{2}} - H_h^n}{\Delta t/2}\phi dx + \int_{I_j} E_h^n \partial_x \phi dx - \widehat{(E_h^n)}_{j+\frac{1}{2}}\phi_{j+\frac{1}{2}}^- + \widehat{(E_h^n)}_{j-\frac{1}{2}}\phi_{j-\frac{1}{2}}^+ = 0, \tag{5.19a}$$

$$\int_{I_j} \frac{D_h^{n+1} - D_h^n}{\Delta t}\varphi dx + \int_{I_j} H_h^{n+\frac{1}{2}} \partial_x \varphi dx - \widetilde{(H_h^{n+\frac{1}{2}})}_{j+\frac{1}{2}}\varphi_{j+\frac{1}{2}}^- + \widetilde{(H_h^{n+\frac{1}{2}})}_{j-\frac{1}{2}}\varphi_{j-\frac{1}{2}}^+ = 0, \tag{5.19b}$$

$$\mu_0 \int_{I_j} \frac{H_h^{n+1} - H_h^{n+\frac{1}{2}}}{\Delta t/2} \psi dx + \int_{I_j} E_h^{n+1} \partial_x \psi dx - \widehat{(E_h^{n+1})}_{j+\frac{1}{2}} \psi^-_{j+\frac{1}{2}} + \widehat{(E_h^{n+1})}_{j-\frac{1}{2}} \psi^+_{j-\frac{1}{2}} = 0, \tag{5.19c}$$

where $\widehat{E_h^n}$ and $\widetilde{H_h^{n+\frac{1}{2}}}$ are numerical fluxes. Even though both the alternating and upwind fluxes in (3.8) can be used here, we choose to only work with the alternating flux due to the consideration for cost efficiency. When the upwind flux is applied in the DG discretization with the leapfrog scheme in time, the term $\frac{1}{2}(E_h^{n+1} + E_h^n)$ is needed to evaluate $\widetilde{H_h^{n+\frac{1}{2}}}$. This will lead to a globally coupled nonlinear system to solve. Meanwhile, the time step condition of the leapfrog scheme for numerical stability is restricted by a CFL condition. Hence the leapfrog scheme with the upwind DG method in space is computationally much more expensive than that with the alternating flux. With the alternating flux, the nonlinear system is defined locally at each nodal point $x_{jk}$ and this results in good parallel efficiency.

With the trapezoidal method, which is implicit, the fully discrete scheme is: given $\mathbf{u_h^n} \in (U_h^K)^6$, we look for $\mathbf{u_h^{n+1}} \in (U_h^K)^6$, such that for any $\phi, \varphi \in U_h^K$ and $1 \le j \le N_x$,

$$\mu_0 \int_{I_j} \frac{H_h^{n+1} - H_h^n}{\Delta t} \phi dx + \int_{I_j} \frac{E_h^n + E_h^{n+1}}{2} \partial_x \phi dx - \left( \widehat{\frac{E_h^{n+1} + E_h^n}{2}} \right)_{j+\frac{1}{2}} \phi^-_{j+\frac{1}{2}} + \left( \widehat{\frac{E_h^{n+1} + E_h^n}{2}} \right)_{j-\frac{1}{2}} \phi^+_{j-\frac{1}{2}} = 0, \tag{5.20a}$$

$$\int_{I_j} \frac{D_h^{n+1} - D_h^n}{\Delta t} \varphi dx + \int_{I_j} \frac{H_h^n + H_h^{n+1}}{2} \partial_x \varphi dx - \left( \widetilde{\frac{H_h^n + H_h^{n+1}}{2}} \right)_{j+\frac{1}{2}} \varphi^-_{j+\frac{1}{2}} + \left( \widetilde{\frac{H_h^n + H_h^{n+1}}{2}} \right)_{j-\frac{1}{2}} \varphi^+_{j-\frac{1}{2}} = 0. \tag{5.20b}$$

Both the alternating and upwind fluxes in (3.8) are considered. Although a global nonlinear system needs to be solved, this fully implicit scheme allows larger time step size.

### 5.2. Discretization of the ADE part: a modified exponential time integrator

Our second order approximation for the ADE part (2.1d)-(2.1f) is given as follows.

$$(\frac{1}{2} + \frac{1}{2} e^{-\frac{\Delta t}{\varepsilon}} + \frac{\Delta t}{2\varepsilon} e^{-\frac{\Delta t}{2\varepsilon}}) \chi_h^{n+1} = e^{-\frac{\Delta t}{\varepsilon}} \chi_h^n + \frac{1}{2}(1 - e^{\frac{-\Delta t}{\varepsilon}}) I_h \left( (E_h^{n+1})^2 \right) + \frac{\Delta t}{2\varepsilon} e^{-\frac{\Delta t}{2\varepsilon}} I_h \left( (E_h^n)^2 \right), \tag{5.21a}$$

$$\frac{P_h^{n+1} - P_h^n}{\Delta t} = \frac{1}{2} \left( J_h^{n+1} + J_h^n \right), \tag{5.21b}$$

$$\frac{J_h^{n+1} - J_h^n}{\Delta t} = -\frac{1}{2} \left( \frac{1}{\tau} (J_h^{n+1} + J_h^n) + \omega_0^2 (P_h^{n+1} + P_h^n) - \omega_p^2 (E_h^{n+1} + E_h^n) \right). \tag{5.21c}$$

Note that the standard trapezoidal time integrator is applied in (5.21b)-(5.21c), while a modified exponential time integrator is proposed in (5.21a). This new exponential integrator is to ensure that our schemes simultaneously preserve the correct asymptotic limit as $\varepsilon \to 0$ and the positivity of $\chi$.

Next in Section 5.2.1 we will elaborate how (5.21a) is derived. We then extend the underlying idea to more general ODE systems in Section 5.2.2. Additional comments and discussions will be made in Section 5.2.3 to further understand the proposed strategy.

#### 5.2.1. Derivation of the modified exponential time integrator in (5.21a)

There are two steps to derive (5.21a): (i) modify the ODE (2.1d), and (ii) discretize the modified ODE to preserve the correct asymptotic limit and the positivity of $\chi$. Each step is second order accurate. The modification step is inspired by [9]. The AP and PP properties of the resulting discretization are due to the positivity of exponential functions as well as their fast decay property (surely when the functions decay).

The following estimate will be used repeatedly. Here the big-O notation is with respect to $\Delta t$.

$$\left| \frac{\Delta t}{\varepsilon} e^{-\frac{\Delta t}{2\varepsilon}} - (1 - e^{\frac{-\Delta t}{\varepsilon}}) \right| = \left| \frac{\Delta t}{\varepsilon} e^{-\frac{\Delta t}{2\varepsilon}} - \int_0^{\Delta t} \frac{1}{\varepsilon} e^{\frac{-\tau}{\varepsilon}} d\tau \right| = O(\Delta t^3). \tag{5.22}$$

The spatial dependence of the function is irrelevant, and it is suppressed in the following derivation. As the first step, we modify the ODE. Note that the exact solution of (2.1d) satisfies

$$\chi(t^{n+1}) = e^{-\frac{\Delta t}{\varepsilon}} \chi(t^n) + \int_0^{\Delta t} \frac{1}{\varepsilon} e^{\frac{\tau - \Delta t}{\varepsilon}} \left( E(t^n + \tau) \right)^2 d\tau = e^{-\frac{\Delta t}{\varepsilon}} \chi(t^n) + \frac{\Delta t}{\varepsilon} e^{-\frac{\Delta t}{2\varepsilon}} \left( E(t^{n+\frac{1}{2}}) \right)^2 + O(\Delta t^3), \tag{5.23}$$

or equivalently,

$$\frac{\chi(t^{n+1})-\chi(t^n)}{\Delta t}=\frac{e^{-\frac{\Delta t}{\varepsilon}}-1}{\Delta t}\chi(t^n)+\frac{1}{\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^{n+\frac{1}{2}})\right)^2+O(\Delta t^2). \tag{5.24}$$

Here $t^{n+\frac{1}{2}}=t^n+\frac{\Delta t}{2}$. With the Taylor series expansion, (5.24) becomes

$$\begin{aligned}\frac{\chi(t^{n+1})-\chi(t^n)}{\Delta t}&=\frac{e^{-\frac{\Delta t}{\varepsilon}}-1}{\Delta t}\chi(t^n)+\frac{1}{\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^{n+\frac{1}{2}})\right)^2+O(\Delta t^2),\\ \partial_t\chi(t^{n+\frac{1}{2}})&=\frac{e^{\frac{-\Delta t}{\varepsilon}}-1}{\Delta t}\left(\chi(t^{n+\frac{1}{2}})-\frac{\Delta t}{2}\partial_t\chi(t^{n+\frac{1}{2}})\right)+\frac{1}{\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^{n+\frac{1}{2}})\right)^2+O(\Delta t^2),\\ (\frac{1}{2}+\frac{1}{2}e^{-\frac{\Delta t}{\varepsilon}})\partial_t\chi(t^{n+\frac{1}{2}})&=\frac{e^{-\frac{\Delta t}{\varepsilon}}-1}{\Delta t}\chi(t^{n+\frac{1}{2}})+\frac{1}{\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^{n+\frac{1}{2}})\right)^2+O(\Delta t^2).\end{aligned} \tag{5.25}$$

The equation (5.25) can be regarded as a $O(\Delta t^2)$ perturbation to the original ODE (2.1d). Next we want to discretize the modified ODE (5.25) with a second order approximation. To achieve this, applying central differencings to each term, we get

$$\begin{aligned}(\frac{1}{2}+\frac{1}{2}e^{-\frac{\Delta t}{\varepsilon}})\left(\frac{\chi(t^{n+1})-\chi(t^n)}{\Delta t}\right)&=\left(\frac{e^{-\frac{\Delta t}{\varepsilon}}-1}{\Delta t}\right)\frac{\chi(t^{n+1})+\chi(t^n)}{2}+\frac{1}{\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\frac{\left(E(t^{n+1})\right)^2+\left(E(t^n)\right)^2}{2}+O(\Delta t^2),\\ (\frac{1}{2}+\frac{1}{2}e^{\frac{-\Delta t}{\varepsilon}}+\frac{1-e^{-\frac{\Delta t}{\varepsilon}}}{2})\chi(t^{n+1})&=e^{-\frac{\Delta t}{\varepsilon}}\chi(t^n)+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^{n+1})\right)^2+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^n)\right)^2+O(\Delta t^3).\end{aligned} \tag{5.26}$$

We then apply (5.22) selectively to some coefficients in (5.26), and reach

$$(\frac{1}{2}+\frac{1}{2}e^{\frac{-\Delta t}{\varepsilon}}+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}})\chi(t^{n+1})=e^{-\frac{\Delta t}{\varepsilon}}\chi(t^n)+\frac{1-e^{-\frac{\Delta t}{\varepsilon}}}{2}\left(E(t^{n+1})\right)^2+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}\left(E(t^n)\right)^2+O(\Delta t^3). \tag{5.27}$$

The derivation above justifies the second order discretization (5.21a) for the equation (2.1d).

As $\varepsilon\to 0$, (5.21a) becomes $\chi_h^{n+1}=I_h((E_h^{n+1})^2)$, or equivalently, $\chi_{jk}^{n+1}\to(E_{jk}^{n+1})^2,\forall j,k$. Hence, the correct asymptotic limit is captured. We will state the asymptotic preserving property for the entire algorithm in Section 5.4. The modified exponential time integrator also ensures the PP property of $\chi_h$.

**Theorem 5.1.** *If $\chi_{jk}^n\geq 0$, then $\chi_{jk}^{n+1}\geq 0$.*

**Proof.** Equivalent to (5.21a), we have

$$(\frac{1}{2}+\frac{1}{2}e^{-\frac{\Delta t}{\varepsilon}}+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}})\chi_{jk}^{n+1}=e^{-\frac{\Delta t}{\varepsilon}}\chi_{jk}^n+\frac{1}{2}(1-e^{\frac{-\Delta t}{\varepsilon}})(E_{jk}^{n+1})^2+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}(E_{jk}^n)^2,$$

where $e^{-\frac{\Delta t}{\varepsilon}}>0$, $\frac{1}{2}(1-e^{\frac{-\Delta t}{\varepsilon}})>0$, and $\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}>0$. One can then conclude $\chi_{jk}^{n+1}\geq 0$ provided that $\chi_{jk}^n\geq 0$. □

**Remark 5.2.** If the exponential time integrator is defined based on (5.26) instead of (5.27), it will lead to $\chi_h^{n+1}=0$ as $\varepsilon\to 0$, an inconsistent discretization for $\chi=E^2$. The corresponding numerical method will no longer be AP, though it is still positivity-preserving.

*5.2.2. Extension to general ODE systems*

The underlying idea of the modified exponential time integrator in Section 5.2 indeed can be applied to more general ODE systems to get second order AP and PP methods. Consider an ODE system of the following form,

$$\frac{d\boldsymbol{u}}{dt}=\boldsymbol{f}(\boldsymbol{u},v,t), \tag{5.28a}$$

$$\frac{dv}{dt}=-\frac{1}{\varepsilon}(v-g(\boldsymbol{u},v,t)), \tag{5.28b}$$

where $\boldsymbol{u}\in\mathbb{R}^n$. If $g(\boldsymbol{u},v,t)\geq 0$, the exact solution $v(t)=e^{-\frac{t}{\varepsilon}}v(0)+\int_0^t\frac{e^{\frac{\tau-t}{\varepsilon}}}{\varepsilon}g(\boldsymbol{u},v,\tau)d\tau$ is also non-negative. Moreover, as $\varepsilon\to 0$, $v$ formally goes to $g(\boldsymbol{u},v,t)$.

Following the similar idea in Section 5.2, a second order approximation for (5.28b) is given as follows,

$$(\frac{1}{2}+\frac{1}{2}e^{\frac{-\Delta t}{\varepsilon}}+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}})v^{n+1}=e^{-\frac{\Delta t}{\varepsilon}}v^n+\frac{1}{2}(1-e^{-\frac{\Delta t}{\varepsilon}})g(\boldsymbol{u}^{n+1},v^{n+1},t^{n+1})+\frac{\Delta t}{2\varepsilon}e^{-\frac{\Delta t}{2\varepsilon}}g(\boldsymbol{u}^n,v^n,t^n). \tag{5.29}$$

Just as in Theorem 5.1, for the numerical solution of (5.29), we can prove $v^{n+1} \geq 0$ provided $v^n \geq 0$. Moreover, as $\varepsilon \to 0$, $v^{n+1} \to g(\boldsymbol{u}^{n+1}, v^{n+1}, t^{n+1})$.

In general we can discretize (5.28a) using the trapezoidal method,

$$\frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^n}{\Delta t} = \frac{1}{2}\left(f(\boldsymbol{u}^n, v^n, t^n) + f(\boldsymbol{u}^{n+1}, v^{n+1}, t^{n+1})\right). \tag{5.30}$$

When $n = 2m$ and (5.28a) is of the following special form,

$$\boldsymbol{u} = (\boldsymbol{u}_1, \boldsymbol{u}_2)^T, \quad \boldsymbol{f}(\boldsymbol{u}, v, t) = \left(\boldsymbol{f}_1(\boldsymbol{u}_2, t), \boldsymbol{f}_2(\boldsymbol{u}_1, t)\right)^T, \quad \text{with } \boldsymbol{u}_i \in \mathbb{R}^m, i = 1, 2,$$

one can also apply the staggered leapfrog method and get

$$\frac{\boldsymbol{u}_1^{n+\frac{1}{2}} - \boldsymbol{u}_1^n}{\Delta t/2} = \boldsymbol{f}_1(\boldsymbol{u}_2^n, t^n), \tag{5.31a}$$

$$\frac{\boldsymbol{u}_2^{n+1} - \boldsymbol{u}_2^n}{\Delta t} = \boldsymbol{f}_2(\boldsymbol{u}_1^{n+\frac{1}{2}}, t^{n+\frac{1}{2}}), \tag{5.31b}$$

$$\frac{\boldsymbol{u}_1^{n+1} - \boldsymbol{u}_1^{n+\frac{1}{2}}}{\Delta t/2} = \boldsymbol{f}_1(\boldsymbol{u}_2^{n+1}, t^{n+1}). \tag{5.31c}$$

*5.2.3. More discussions*

Here we want to briefly compare our proposed strategy with some available AP and PP schemes of second order accuracy in the literature. In [7], semi-implicit Runge-Kutta (RK) schemes with a correction step preserving the positivity and the steady state are proposed for a shallow water model. In [22] and [21], implicit-explicit RK (IMEX-RK) schemes with a correction step preserving the positivity and the asymptotic limit are proposed for the Kerr-Debye model (a nonlinear hyperbolic system with stiff relaxation, [22]) and for the stiff BGK model ([21]). Either the semi-implicit RK schemes or the IMEX-RK schemes above start with a convex combination of first order schemes with AP and PP properties. Due to the fact that there is no second or higher order implicit strong stability preserving (SSP) RK schemes [18], an additional correction step is introduced, based on the Fréchet derivative of the stiff part, to achieve the second order accuracy while keeping the PP property.

What we propose here follows a different strategy. While the non-stiff part of the model is discretized using the leapfrog or trapezoidal method, for the stiff part, we design a modified exponential time integrator, with which second order accuracy is achieved together with the positivity and AP properties. On the implementation level, using the proposed strategy, a nonlinear algebraic equation needs to be solved once per time step. If the strategy in [22,21] is adapted to our present model, a nonlinear algebraic equation would need to be solved in each inner stage of the RK method and in the correction step.

We also want to comment that using our proposed strategy, the error in (5.22) is no longer negligible when $\Delta t/\varepsilon = O(1)$. Hence one may observe some order reduction as many AP schemes encounter in this intermediate regime (also see numerical examples in Section 6). One way to alleviate this is to implement the modified exponential time integrator adaptively, namely, it is only used when some indicator implies the positivity of $\chi$ is lost at some nodal points. Elsewhere, one can use other second order discretizations which is only responsible for the AP property. With this, the modified exponential time integrator we proposed will be applied like a PP limiter. Such adaptive strategy is not explored in the present work.

*5.3. Numerical constitutive laws*

To impose the constitutive law (2.1c), we require

$$D_h^{n+1} = \epsilon_0((\epsilon_\infty + a\chi_h^{n+1})E_h^{n+1} + P_h^{n+1}), \tag{5.32}$$

and this will be referred to as the algebra-based numerical constitutive law. This strategy, though very natural, does not lead to a provable energy relation or energy stability. To achieve such a provable result, we also propose an energy-based numerical constitutive law:

$$\begin{aligned}
&(D_h^{n+1} - D_h^n)(E_h^{n+1} + E_h^n) - \frac{a\epsilon_0\Delta t}{\varepsilon}\left(\frac{\chi_h^{n+1} + \chi_h^n}{2} - I_h\left(\frac{(E_h^{n+1})^2 + (E_h^n)^2}{2}\right)\right)^2 \\
&= \epsilon_0\epsilon_\infty\left((E_h^{n+1})^2 - (E_h^n)^2\right) + \frac{\epsilon_0 a}{2}\left((\chi_h^{n+1})^2 - (\chi_h^n)^2\right) + \epsilon_0 a\left(\chi_h^{n+1} I_h((E_h^{n+1})^2) - \chi_h^n I_h((E_h^n)^2)\right) \\
&\quad + \epsilon_0(P_h^{n+1} - P_h^n)(E_h^{n+1} + E_h^n).
\end{aligned} \tag{5.33}$$

In order to understand this energy-based approximation, note that the exact solution satisfies

$$D(t^{n+1}) - D(t^n) = \epsilon_0\epsilon_\infty\big(E(t^{n+1}) - E(t^n)\big) + a\epsilon_0\big(Y(t^{n+1}) - Y(t^n)\big) + \epsilon_0\big(P(t^{n+1}) - P(t^n)\big), \tag{5.34}$$

where $Y(t) = \chi(t)E(t)$. Use Taylor series expansion, and assume $E$ is nonzero for the time being, we have

$$\begin{aligned} Y(t^{n+1}) - Y(t^n) &= Y_t(t^{n+\frac{1}{2}})\Delta t + O(\Delta t^3) = \frac{\chi_t E^2 + (\chi E^2)_t}{2E}\Big|_{t=t^{n+\frac{1}{2}}}\Delta t + O(\Delta t^3) \\ &= \frac{1}{E(t^{n+1}) + E(t^n)}\Big(\Delta t(\chi_t E^2)(t^{n+\frac{1}{2}}) + \chi(t^{n+1})E^2(t^{n+1}) - \chi(t^n)E^2(t^n)\Big) + O(\Delta t^3). \end{aligned} \tag{5.35}$$

On the other hand, multiply $\chi - E^2$ to (2.1d) and use Taylor series expansion, we get

$$\begin{aligned} \chi_t E^2\Big|_{t=t^{n+\frac{1}{2}}} &= \Big(\frac{1}{\varepsilon}(\chi - E^2)^2 + \chi_t\chi\Big)\Big|_{t=t^{n+\frac{1}{2}}} \\ &= \frac{1}{\varepsilon}\Big(\frac{\chi(t^{n+1}) + \chi(t^n)}{2} - \frac{E^2(t^{n+1}) + E^2(t^n)}{2}\Big)^2 + \frac{1}{2\Delta t}\big(\chi^2(t^{n+1}) - \chi^2(t^n)\big) + O(\Delta t^2), \end{aligned} \tag{5.36}$$

hence (5.35) becomes

$$\begin{aligned} Y(t^{n+1}) - Y(t^n) = \frac{1}{E(t^{n+1}) + E(t^n)}\Big(&\frac{\Delta t}{\varepsilon}\Big(\frac{\chi(t^{n+1}) + \chi(t^n)}{2} - \frac{E^2(t^{n+1}) + E^2(t^n)}{2}\Big)^2 \\ &+ \frac{\chi^2(t^{n+1})}{2} - \frac{\chi^2(t^n)}{2} + \chi(t^{n+1})E^2(t^{n+1}) - \chi(t^n)E^2(t^n)\Big) + O(\Delta t^3). \end{aligned} \tag{5.37}$$

With (5.37) plugged into (5.34), one can find that the energy-based numerical constitutive law in (5.33) is simply a second order (in time) approximation for the relation (5.34) multiplied by $E(t^{n+1}) + E(t^n)$. The reason to multiply (5.34) by $E(t^{n+1}) + E(t^n)$ is for a more robust implementation of (5.33).

Now we are ready to define three fully discrete schemes. The META scheme is defined as (5.20), (5.21), and (5.32). The MELA scheme is defined as (5.19), (5.21), and (5.32). And the MELE scheme is defined as (5.19), (5.21), and (5.33).

### 5.4. AP property

When $\Delta t$ (and the spatial mesh) is fixed, $e^{-\frac{\Delta t}{\varepsilon}} = O(\varepsilon^p)$ holds for any $p > 0$. Based on this and with dominance balance argument, one can formally derive

$$\chi_h^{n+1} - I_h((E_h^{n+1})^2) = O(\varepsilon^p), \quad \forall p > 0,\ \forall n \geq 0, \tag{5.38}$$

from the scheme (5.21a), provided $E_h^n = O(1)$.

With (5.38) and as $\varepsilon \to 0$, the limiting scheme of the MELA scheme is formally defined as follows: given $\mathbf{v_h^n} = (E_h^n, H_h^n, D_h^n, \chi_h^n, P_h^n)^T \in (U_h^K)^5$, we look for $\mathbf{v_h^{n+1}} \in (U_h^K)^5$ satisfying (5.19), (5.21b)-(5.21c), and

$$D_h^{n+1} = \epsilon_0\big(\epsilon_\infty E_h^{n+1} + aI_h\Big((E_h^{n+1})^3\Big) + P_h^{n+1}\big). \tag{5.39}$$

And the limiting scheme of the META scheme is given by (5.20), (5.21b)-(5.21c), and (5.39). Both limiting schemes are second order approximations for the Kerr-Lorentz model in (2.2), and this shows the AP property of the proposed MELA and META schemes.

To study the limiting behavior of the MELE scheme, we need to take a closer look at the energy-based numerical constitutive law (5.33) and examine the influence of the initial data. If the initial data satisfies

$$\frac{1}{\varepsilon}\Big(\chi_h^0 - I_h\big((E_h^0)^2\big)\Big) \to 0, \quad \varepsilon \to 0, \tag{5.40}$$

then based on (5.38), the relation (5.33) will become

$$D_h^{n+1} - D_h^n = \epsilon_0\epsilon_\infty\Big(E_h^{n+1} - E_h^n\Big) + \frac{3\epsilon_0 a}{2}I_h\Big(\Big((E_h^{n+1})^2 + (E_h^n)^2\Big)(E_h^{n+1} - E_h^n)\Big) + \epsilon_0(P_h^{n+1} - P_h^n), \tag{5.41}$$

as $\varepsilon \to 0$, hence the limiting scheme of the MELE scheme is defined as (5.19), (5.21b)-(5.21c), and (5.41). This is a second order discretization of the Kerr-Lorentz model (2.2).

When the initial data is not as prepared as (5.40), then the term $\frac{1}{\varepsilon}\big(\chi_h^0 - I_h\big((E_h^0)^2\big)\big)$ in (5.33) will have a lingering effect to later time steps as $\varepsilon \to 0$. This will either lead to order reduction or even inaccurate limiting behavior (see [27] for some discussion and analysis for a similar phenomenon in simulating multi-scale kinetic transport problems). To avoid this, one can modify the MELE scheme as follows: when $n = 0$, use the algebra-based numerical constitutive law (5.32); When

$n \geq 1$, we switch back to the energy based constitutive law (5.33). This modified MELE scheme still enjoys the provable energy relation as the original MELE scheme (this will proved in next section), while its limiting scheme is given as (5.19), (5.21b)-(5.21c), together with (5.39) when $n = 0$ and (5.41) when $n \geq 1$. This is still a consistent discretization of the limiting Kerr-Lorentz model.

### 5.5. Energy stability for the MELE scheme

**Theorem 5.3.** *Assume $\chi(x, 0) \geq 0$ and periodic boundary conditions. For the proposed MELE scheme defined as (5.19), (5.21), and (5.33), together with the alternating flux in (3.8), we introduce*

$$\mathcal{E}_{h,2}^{n} = \frac{1}{2}\left(\epsilon_0\epsilon_\infty||E_h^n||^2 + a\epsilon_0\left(\chi_h^n, I_h\left((E_h^n)^2\right)\right) + \mu_0\left(H_h^{n+\frac{1}{2}}, H_h^{n-\frac{1}{2}}\right) + \frac{\epsilon_0 a}{2}||\chi_h^n||^2 + \frac{\epsilon_0\omega_0^2}{\omega_p^2}||P_h^n||^2 + \frac{\epsilon_0}{\omega_p^2}||J_h^n||^2\right), \tag{5.42}$$

*then $\mathcal{E}_{h,2}^{n+1} \leq \mathcal{E}_{h,2}^{n}, \forall n \geq 1$. Moreover, under a hyperbolic type time step condition,*

$$\Delta t \leq \frac{2\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{\delta\left(\sqrt{\widehat{C}_{inv}} + 2C_{inv}\right)}h, \tag{5.43}$$

*there holds $\mathcal{E}_{h,2}^{n} \geq 0, \forall n \geq 1$, hence $\mathcal{E}_{h,2}^{n}$ defines a discrete energy at $t^n$.*

**Proof.** Apply two time steps of (5.19a) and (5.19c) with $\psi = \phi$, we have

$$\begin{aligned}&\mu_0\int_{I_j}\frac{H_h^{n+\frac{3}{2}} - H_h^{n-\frac{1}{2}}}{\Delta t}\phi dx + \int_{I_j}(E_h^{n+1} + E_h^n)\partial_x\phi dx - \left(\widehat{E_h^n} + \widehat{E_h^{n+1}}\right)_{j+\frac{1}{2}}\phi_{j+\frac{1}{2}}^{-} + \left(\widehat{E_h^n} + \widehat{E_h^{n+1}}\right)_{j-\frac{1}{2}}\phi_{j-\frac{1}{2}}^{+}\\ &= 0, \qquad \forall\phi \in U_h^K.\end{aligned} \tag{5.44}$$

Take $\varphi = E_h^n + E_h^{n+1}$ in (5.19b) and $\phi = H_h^{n+\frac{1}{2}}$ in (5.44), sum them up over all elements, use the definition of the alternating flux and periodic boundary conditions, we obtain

$$\mu_0\left(H_h^{n+\frac{3}{2}} - H_h^{n-\frac{1}{2}}, H_h^{n+\frac{1}{2}}\right) + \left(D_h^{n+1} - D_h^n, E_h^{n+1} + E_h^n\right) = 0. \tag{5.45}$$

Multiply $P_h^{n+1} + P_h^n$ to (5.21b) and integrate in $x$, we get

$$||P_h^{n+1}||^2 - ||P_h^n||^2 = \frac{\Delta t}{2}\left(J_h^{n+1} + J_h^n, P_h^{n+1} + P_h^n\right). \tag{5.46}$$

Multiply $\frac{1}{\omega_p^2}(J_h^{n+1} + J_h^n)$ to (5.21c), integrate in $x$, and rewrite the result by applying (5.21b) and (5.46), we get

$$\frac{1}{\omega_p^2}\left(||J_h^{n+1}||^2 - ||J_h^n||^2\right) + \frac{\Delta t}{2\omega_p^2\tau}||J_h^{n+1} + J_h^n||^2 + \frac{\omega_0^2}{\omega_p^2}\left(||P_h^{n+1}||^2 - ||P_h^n||^2\right) = \left(E_h^{n+1} + E_h^n, P_h^{n+1} - P_h^n\right). \tag{5.47}$$

Now combine (5.33), (5.45) and (5.47), we reach

$$\begin{aligned}&\mu_0\left(H_h^{n+\frac{3}{2}}, H_h^{n+\frac{1}{2}}\right) - \mu_0\left(H_h^{n+\frac{1}{2}}, H_h^{n-\frac{1}{2}}\right) + \epsilon_0\epsilon_\infty\left(||E_h^{n+1}||^2 - ||E_h^n||^2\right)\\ &\quad + \epsilon_0 a\left(\left(\chi_h^{n+1}, I_h((E_h^{n+1})^2)\right) - \left(\chi_h^n, I_h((E_h^n)^2)\right)\right)\\ &\quad + \frac{a\epsilon_0}{2}\left(||\chi_h^{n+1}||^2 - ||\chi_h^n||^2\right) + \frac{\epsilon_0 a\Delta t}{\varepsilon}\left|\left|\frac{\chi_h^{n+1} + \chi_h^n}{2} - I_h\left(\frac{(E_h^{n+1})^2 + (E_h^n)^2}{2}\right)\right|\right|^2\\ &\quad + \frac{\epsilon_0}{\omega_p^2}\left(||J_h^{n+1}||^2 - ||J_h^n||^2\right) + \frac{\epsilon_0\omega_0^2}{\omega_p^2}\left(||P_h^{n+1}||^2 - ||P_h^n||^2\right) + \frac{\epsilon_0\Delta t}{2\omega_p^2\tau}||J_h^{n+1} + J_h^n||^2 = 0.\end{aligned} \tag{5.48}$$

In terms of $\mathcal{E}_{h,2}^{n}$, (5.48) can also be written as

$$\mathcal{E}_{h,2}^{n+1}-\mathcal{E}_{h,2}^{n}=-\frac{\epsilon_0 a\Delta t}{2\varepsilon}\left|\left|\frac{\chi_h^{n+1}+\chi_h^{n}}{2}-I_h\left(\frac{(E_h^{n+1})^2+(E_h^{n})^2}{2}\right)\right|\right|^2-\frac{\omega_0^2\Delta t}{\omega_p^2\tau}\left|\left|\frac{J_h^{n+1}+J_h^{n}}{2}\right|\right|^2, \tag{5.49}$$

and this implies $\mathcal{E}_{h,2}^{n+1}\le\mathcal{E}_{h,2}^{n},\ \ \forall n\ge 0$.

What remained is to figure out when the term $\mathcal{E}_{h,2}^{n}$ is non-negative, hence when it defines a discrete energy. Based on (A.66) and Theorem 5.1, we know $(\chi_h^n, I_h((E_h^n)^2))\ge 0$. We now only need to examine $(H_h^{n+\frac{1}{2}}, H_h^{n-\frac{1}{2}})$. Following the same argument as in [4] (see Theorem 5.1, particularly (5.22)-(5.23)), we find

$$\left(H_h^{n+\frac{1}{2}}, H_h^{n-\frac{1}{2}}\right)=||H_h^n||^2-||H_h^n-H_h^{n-\frac{1}{2}}||^2. \tag{5.50}$$

From the scheme (5.19c) and inverse inequalities, one gets

$$\begin{aligned}\frac{\mu_0}{2}||H_h^n-H_h^{n-\frac{1}{2}}||^2&=\frac{\Delta t}{4}\sum_{j=1}^{N_x}\left(-\int_{I_j}E_h^n\partial_x(H_h^n-H_h^{n-\frac{1}{2}})dx+(\widehat{E}_h^n)_{j+\frac{1}{2}}(H_h^n-H_h^{n-\frac{1}{2}})^-_{j+\frac{1}{2}}-(\widehat{E}_h^n)_{j-\frac{1}{2}}(H_h^n-H_h^{n-\frac{1}{2}})^+_{j-\frac{1}{2}}\right)\\&\le||H_h^n-H_h^{n-\frac{1}{2}}||\left(\sqrt{\widehat{C}_{inv}}+2C_{inv}\right)\frac{\delta\Delta t}{4h}||E_h^n||,\end{aligned}$$

hence

$$\frac{\mu_0}{2}||H_h^n-H_h^{n-\frac{1}{2}}||^2\le\left(\sqrt{\widehat{C}_{inv}}+2C_{inv}\right)^2\frac{(\delta\Delta t)^2}{8\mu_0h^2}||E_h^n||^2. \tag{5.51}$$

Finally, we combine (5.50) and (5.51), and conclude

$$\frac{\mu_0}{2}\left(H_h^{n+\frac{1}{2}}, H^{n-\frac{1}{2}}\right)+\frac{\epsilon_0\epsilon_\infty}{2}||E_h^n||^2\ge\frac{\mu_0}{2}||H_h^n||^2+\left(\frac{\epsilon_0\epsilon_\infty}{2}-\left(\sqrt{\widehat{C}_{inv}}+2C_{inv}\right)^2\frac{(\delta\Delta t)^2}{8\mu_0h^2}\right)||E_h^n||^2\ge 0, \tag{5.52}$$

under the time step condition

$$\Delta t\le\frac{2\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{\delta\left(\sqrt{\widehat{C}_{inv}}+2C_{inv}\right)}h. \tag{5.53}$$

This condition also ensures $\mathcal{E}_{h,2}^{n}\ge 0$. □

### 5.6. Nonlinear algebraic solvers

Due to the implicit treatments in the proposed numerical methods, nonlinear algebraic equations need to be solved in actual implementation. In this section, we will discuss some aspects of the nonlinear solvers.

#### 5.6.1. Algorithm

All the proposed methods are implemented in their nodal form. For the leapfrog schemes, namely, the MELA scheme and the MELE scheme, the implicitness comes from the discretizations of the ADE part and the constitutive laws. A small size nonlinear algebraic system needs to be solved locally at each $x_{jk}$. The standard Newton's method is applied. For the trapezoidal META scheme, the implicitness also comes from the PDE discretization. The resulting nonlinear algebraic system is globally coupled, and it is solved using a Jacobian-free Newton-Krylov solver [25]. The solver is implemented within the framework of the KINSOL package [29].

The nonlinear solver is described next as an algorithm. Suppose $\mathbf{u}^n$ is the data representing the numerical solution from the time $t^n$, and we need to solve an algebraic system $\mathbf{f}(\mathbf{u})=0$ (which is either locally or globally defined; and $\mathbf{f}$ generally depends on $n$) to get the data representing the numerical solution $\mathbf{u}^{n+1}$ at $t^{n+1}$.

**Algorithm:** Given $\mathbf{u}^n$, we want to solve $\mathbf{f}(\mathbf{u})=\mathbf{0}$ for $\mathbf{u}^{n+1}$.

Initialization: Let $\mathbf{u_{-1}}=\mathbf{u}^n$. Let $\mathbf{u_0}$ be an initial guess, and $k=0$.

**while** $\left(k<k_{\max})\ \textbf{\textit{and}}\ \left((||\mathbf{f}(\mathbf{u_k})||>Tol)\ \textbf{\textit{or}}\ (||\mathbf{u_k}-\mathbf{u_{k-1}}||>Tol)\right)\right.$ **do**

  $\mathbf{u_{k+1}}=\mathbf{u_k}-J_f^{-1}\mathbf{f}(\mathbf{u_k})$, and $k:=k+1$.

**end**

Output: $\mathbf{u}^{n+1}=\mathbf{u_k}$.

Here $J_f$ is the Jacobian of $\mathbf{f}(\mathbf{u})$ or its approximation, depending on which solver is used. In our implementation, the error tolerance is taken as $\text{Tol} = 10^{-10}$ and the maximum iteration number is $k_{\max} = 200$. Unless otherwise specified, the initial guess $\mathbf{u_0}$ is chosen as the data from the previous time step, namely, $\mathbf{u_0} = \mathbf{u}^n$. With the local nature of the leapfrog schemes, $J_f$ and its inverse can be easily obtained analytically for each nodal point $x_{jk}$. When the Jacobi-free Newton Krylov solver is applied for the META scheme, the GMRES method is used as the linear solver to obtain $\mathbf{u_{k+1}}$.

*5.6.2. Initial guess for the MELE scheme*

For the schemes with the algebra-based constitutive law, including the MELA and META schemes, they perform quite well when the initial guess in the Newton's method is based on the data from the previous time step, namely, $\mathbf{u_0} = \mathbf{u}^n$. This initial guess however, when used in the MELE scheme that involves the energy-based constitutive law, may result in the convergence of the Newton solver to a nonphysical solution. Next, we will elaborate the reason and propose strategies for possibly finding a better initial guess.

Recall the nonlinearity of the underlying model is cubic. When the energy-based constitutive law (5.33) is used, the resulting nonlinear algebraic problem is quartic. It is one degree higher than that from the MELA or META scheme, and allows one more solution which can be non-physical. This can be explained using the soliton propagation problem in Section 6.2.2. In this example, the initial conditions for all fields are zero, and a non-zero signal is introduced to the domain as a left boundary condition. For the MELE scheme, the signal from the left boundary changes the $D$ term through (5.19), and then updates the $E$ term through (5.33) together with (5.21). When the solutions at $t^n$ are zero, (5.21) and (5.33) allow a zero solution $E_h^{n+1}$ regardless of the boundary condition. This zero solution is not physical, yet it will be captured by the Newton's solver if the initial guess is set as the zero numerical solution at the previous time. As a result, we will observe that $E_h^n = 0,\ \forall n \geq 1$.

We next propose several strategies to assign an initial guess $\mathbf{u_0}$ for the Newton solver in the MELE scheme. Assume the numerical solutions at $t^n$ are known, and $D_h^{n+1}$ has already been computed from (5.19). We will compute $E^*, \chi^*, P^*, J^* \in U_h^K$ based on one of the following strategies, then use them as the initial guess for the Newton solver in the MELE scheme to get $E_h^{n+1}, \chi_h^{n+1}, P_h^{n+1}, J_h^{n+1} \in U_h^K$.

1.) "*backward Euler initial guess*": We solve for $E^*, \chi^*, P^*, J^* \in U_h^K$ based on a back Euler discretization of the ADE part and a direct discretization of the constitutive relation, namely,

$$D_h^{n+1} = \epsilon_0 \left(\epsilon_\infty E^* + aI_h\left(\chi^* E^*\right) + P^*\right), \tag{5.54a}$$

$$\frac{\chi^* - \chi_h^n}{\Delta t} = -\frac{1}{\varepsilon}\Big(\chi^* - I_h\left((E^*)^2\right)\Big), \tag{5.54b}$$

$$\frac{P^* - P_h^n}{\Delta t} = J^*, \tag{5.54c}$$

$$\frac{J^* - J_h^n}{\Delta t} + \frac{1}{\tau} J^* + \omega_0^2 P^* = \omega_p^2 E^*. \tag{5.54d}$$

2.) "*trapezoidal initial guess*": We solve for $E^*, \chi^*, P^*, J^* \in U_h^K$ based on a trapezoidal discretization of the ADE part and a direct discretization of the constitutive law, namely,

$$D_h^{n+1} = \epsilon_0 \left(\epsilon_\infty E_h^* + aI_h\left(\chi^* E^*\right) + P^*\right), \tag{5.55a}$$

$$\frac{\chi^* - \chi_h^n}{\Delta t} = -\frac{1}{2\varepsilon}\Big(\chi^* - I_h((E^*)^2) + \chi_h^n - I_h((E_h^n)^2)\Big), \tag{5.55b}$$

$$\frac{P^* - P_h^n}{\Delta t} = \frac{1}{2}\left(J^* + J_h^n\right), \tag{5.55c}$$

$$\frac{J^* - J_h^n}{\Delta t} + \frac{1}{2\tau}\left(J^* + J_h^n\right) + \frac{\omega_0^2}{2}\left(P^* + P_h^n\right) = \frac{\omega_p^2}{2}(E^* + E_h^n). \tag{5.55d}$$

3.) "*limiting initial guess*": based on a second order discretization of the ADEs of the limiting Kerr-Lorentz model and the difference form of the constitutive law, we solve for $E^*, \chi^*, P^*, J^* \in U_h^K$ as follows,

$$D_h^{n+1} - D_h^n = \epsilon_0 \left(\epsilon_\infty (E^* - E_h^n) + a(Y^* - Y_h^n) + P^* - P_h^n\right), \tag{5.56a}$$

$$Y^* - Y_h^n = \frac{3}{2} I_h\Big(\left((E^*)^2 + (E_h^n)^2\right)\left(E^* - E_h^n\right)\Big), \tag{5.56b}$$

$$\chi^* = I_h((E^*)^2), \tag{5.56c}$$

$$\frac{P^* - P_h^n}{\Delta t} = \frac{1}{2}\left(J^* + J_h^n\right), \tag{5.56d}$$

**Table 5.1**
Properties of the proposed second order MELE, MELA and META schemes.

| | MELE | MELA | META |
|---|---|---|---|
| Maxwell part | leapfrog | leapfrog | trapezoidal |
| Numerical fluxes | alternating | alternating | alternating, upwind |
| Constitutive law | energy-based | algebra-based | algebra-based |
| Stability | conditional | conditional | unconditional |
| Positivity and asymptotic preserving | ✓ | ✓ | ✓ |
| Provable energy relation | ✓ | ✗ | ✗ |
| Nonlinear algebraic equations | local | local | global |
| Initial guess | special choice | previous step | previous step |

$$\frac{J^* - J_h^n}{\Delta t} + \frac{1}{2\tau}\left(J^* + J_h^n\right) + \frac{\omega_0^2}{2}\left(P^* + P_h^n\right) = \frac{\omega_p^2}{2}\left(E^* + E_h^n\right). \tag{5.56e}$$

For each strategy above to get $\mathbf{u_0}$, a nonlinear algebraic system needs to be solved. For this we apply Newton solver with $\mathbf{u}^n$ as an initial guess. The performance of these strategies will be investigated in the numerical section based on the soliton propagation example, see Section 6.2.2. The study suggests the use of the trapezoidal initial guess strategy in the MELE scheme, based on its overall cost efficiency for different $\varepsilon$'s. Finally in Table 5.1, we summarize the features and properties of the three second order in time schemes proposed in this section.

## 6. Numerical examples

In this section, we will present a set of numerical examples to demonstrate the performance of the proposed schemes. We will start with an ODE example before moving to the Kerr-Debye-Lorentz model.

### 6.1. An ODE example

We will use the following ODE example

$$\frac{du_1}{dt} = -u_2^2, \tag{6.57a}$$

$$\frac{du_2}{dt} = -u_1^2, \tag{6.57b}$$

$$\frac{dv}{dt} = -\frac{1}{\varepsilon}\left(v - \frac{u_1^2 + u_2^2}{v^2}\right), \tag{6.57c}$$

to demonstrate the performance of the time integrator in Section 5.2.2. For the non-stiff part, the trapezoidal or the staggered leapfrog method is applied. For the stiff part, the modified exponential time integrator is applied. We consider the initial data $u_1(0) = 2$, $u_2(0) = 1$, and $v(0) = 1$.

We first compute a reference solution, $u^{ref}$ and $v^{ref}$, up to $T = 1$, using the RK4 method with $\max(50000, \text{floor}(\frac{100}{\varepsilon}))$ uniform time steps. The errors at $T = 1$ are shown in Table 6.2 for $\Delta t = \frac{1}{40}, \frac{1}{80}, \frac{1}{160}, \frac{1}{320}$, and $\varepsilon = 1, 10^{-2}, 10^{-6}$. For the trapezoidal method, the second order accuracy is observed for $u_1$ and $u_2$ with all $\varepsilon$. When $\Delta t \ll \varepsilon$ or $\varepsilon \ll \Delta t$, second order accuracy is observed for $v$. When $\Delta t/\varepsilon = O(1)$, we can see that the convergence rate for $v$ slows down. When the mesh is further refined, the second order accuracy will come back again. This phenomenon is the so-called order reduction of AP schemes in the intermediate regime. This can be explained by the large error in (5.22) when $\Delta t$ and $\varepsilon$ are of the same order of magnitude. The observations for the leapfrog method are very similar.

In Fig. 6.1, we further show the convergence rates and errors of the numerical solutions by uniformly sampling $\varepsilon$ in its logarithmic scale. The convergence rate is calculated based on the numerical errors with $\Delta t = \frac{1}{100}$ and $\Delta t = \frac{1}{200}$. Errors in $\mathbf{u} = (u_1, u_2)^T$ are computed based on its $l_2$ norm. Second order accuracy for $v$ is observed when $\varepsilon \ll \Delta t$ or $\varepsilon \gg \Delta t$. For the intermediate regime $\varepsilon \approx O(\Delta t)$, a wiggle in the convergence rate of $v$ is observed due to the order reduction, while the numerical errors in the solutions are still fairly small.

### 6.2. Kerr-Debye-Lorentz model

In this section, we will present the numerical results for the Kerr-Debye-Lorentz model. For the proposed schemes with the first order temporal accuracy, the piecewise constant space $U_h^K$ with $K = 0$ is used; for the proposed schemes with the second order temporal accuracy, the piecewise linear space $U_h^K$ with $K = 1$ is used. Increasing spatial orders can improve the numerical resolution significantly. We here choose to present the schemes with the same formal accuracy orders in space and time to better reveal and compare the properties of the different methods. Altogether, we will examine the performance

**Table 6.2**
Errors and orders at $T = 1$ for the ODE example in Section 6.1.

| Scheme | $\varepsilon$ | $\Delta t$ | $\lvert u_{1,h}(T) - u_1^{ref}(T)\rvert$ | Order | $\lvert u_{2,h}(T) - u_2^{ref}(T)\rvert$ | Order | $\lvert v_h(T) - v^{ref}(T)\rvert$ | Order |
|---|---|---|---|---|---|---|---|---|
| Trapezoidal | 1 | $\frac{1}{40}$ | 1.182e−03 | – | 8.404e−04 | – | 2.715e−04 | – |
| | | $\frac{1}{80}$ | 2.958e−04 | 2.00 | 2.100e−04 | 2.00 | 6.791e−05 | 2.00 |
| | | $\frac{1}{160}$ | 7.398e−05 | 2.00 | 5.250e−05 | 2.00 | 1.698e−05 | 2.00 |
| | | $\frac{1}{320}$ | 1.850e−05 | 2.00 | 1.312e−05 | 2.00 | 4.245e−06 | 2.00 |
| | $10^{-2}$ | $\frac{1}{40}$ | 1.182e−03 | – | 8.404e−04 | – | 2.768e−04 | – |
| | | $\frac{1}{80}$ | 2.958e−04 | 2.00 | 2.100e−04 | 2.00 | 1.188e−04 | 1.22 |
| | | $\frac{1}{160}$ | 7.398e−05 | 2.00 | 5.250e−05 | 2.00 | 3.639e−05 | 1.71 |
| | | $\frac{1}{320}$ | 1.850e−05 | 2.00 | 1.312e−05 | 2.00 | 9.955e−06 | 1.87 |
| | $10^{-6}$ | $\frac{1}{40}$ | 1.182e−03 | – | 8.404e−04 | – | 2.012e−04 | – |
| | | $\frac{1}{80}$ | 2.958e−04 | 2.00 | 2.100e−04 | 2.00 | 5.034e−05 | 2.00 |
| | | $\frac{1}{160}$ | 7.398e−05 | 2.00 | 5.250e−05 | 2.00 | 1.268e−05 | 1.99 |
| | | $\frac{1}{320}$ | 1.850e−05 | 2.00 | 1.312e−05 | 2.00 | 3.273e−06 | 1.95 |
| Leapfrog | 1 | $\frac{1}{40}$ | 6.002e−04 | – | 4.655e−04 | – | 1.822e−04 | – |
| | | $\frac{1}{80}$ | 1.503e−04 | 2.00 | 1.165e−04 | 2.00 | 4.559e−05 | 2.00 |
| | | $\frac{1}{160}$ | 3.759e−05 | 2.00 | 2.913e−05 | 2.00 | 1.140e−05 | 2.00 |
| | | $\frac{1}{320}$ | 9.399e−06 | 2.00 | 7.284e−06 | 2.00 | 2.850e−06 | 2.00 |
| | $10^{-2}$ | $\frac{1}{40}$ | 6.002e−04 | – | 4.655e−04 | – | 3.601e−04 | – |
| | | $\frac{1}{80}$ | 1.503e−04 | 2.00 | 1.165e−04 | 2.00 | 1.394e−04 | 1.37 |
| | | $\frac{1}{160}$ | 3.759e−05 | 2.00 | 2.913e−05 | 2.00 | 4.152e−05 | 1.75 |
| | | $\frac{1}{320}$ | 9.399e−06 | 2.00 | 7.284e−06 | 2.00 | 1.123e−05 | 1.89 |
| | $10^{-6}$ | $\frac{1}{40}$ | 6.002e−04 | - | 4.655e−04 | – | 1.215e−04 | – |
| | | $\frac{1}{80}$ | 1.503e−04 | 2.00 | 1.165e−04 | 2.00 | 3.051e−05 | 1.99 |
| | | $\frac{1}{160}$ | 3.759e−05 | 2.00 | 2.913e−05 | 2.00 | 7.732e−06 | 1.98 |
| | | $\frac{1}{320}$ | 9.399e−06 | 2.00 | 7.284e−06 | 2.00 | 2.035e−06 | 1.93 |

of the first order semi-implicit SI1-al and SI1-up schemes, the first order fully implicit FI1-al and FI-up schemes, the second order MELA and MELE schemes, and the second order META-al and META-up schemes. Here "-al" ("-up") is to indicate the use of the alternating (upwind) flux. As discussed before, the MELA and MELE schemes are only coupled with the alternating fluxes.

All simulations are done on uniform meshes. Let $h$ be the mesh size in space. The time step size $\Delta t$ is determined by $\Delta t = 0.75h$ for the META and FI1 schemes. Based on the energy analysis in Section 5.5, $\Delta t = 0.9 \times \frac{2\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{\sqrt{12}+8}h$ is taken for the MELE scheme, and this time step size formula is also used for the MELA scheme. For the SI1 scheme, we take $\Delta t = \frac{\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{4}h$ for the upwind flux, and $\Delta t = 0.9h^2$ for the alternating flux.

The actual simulation is carried out based on the dimensionless form of the equations. Consider the Kerr-Debye system with $N_\omega$ linear Lorentz responses

$$\mu_0 \partial_t H = \partial_x E, \tag{6.58a}$$

$$\partial_t D = \partial_x H, \tag{6.58b}$$

$$D = \epsilon_0 \left( (\epsilon_\infty + a\chi)E + \sum_{k=1}^{N_\omega} P_k \right), \tag{6.58c}$$

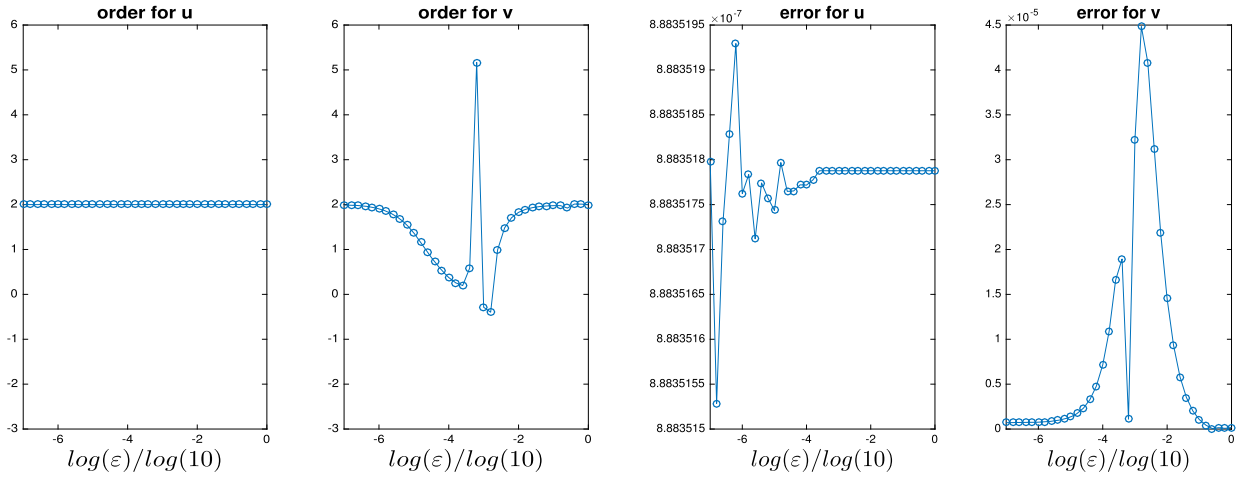
$$\partial_t \chi = -\frac{1}{\varepsilon}(\chi - E^2), \tag{6.58d}$$

$$\partial_t P_k = J_k, \quad k = 1, 2, \ldots, N_\omega, \tag{6.58e}$$

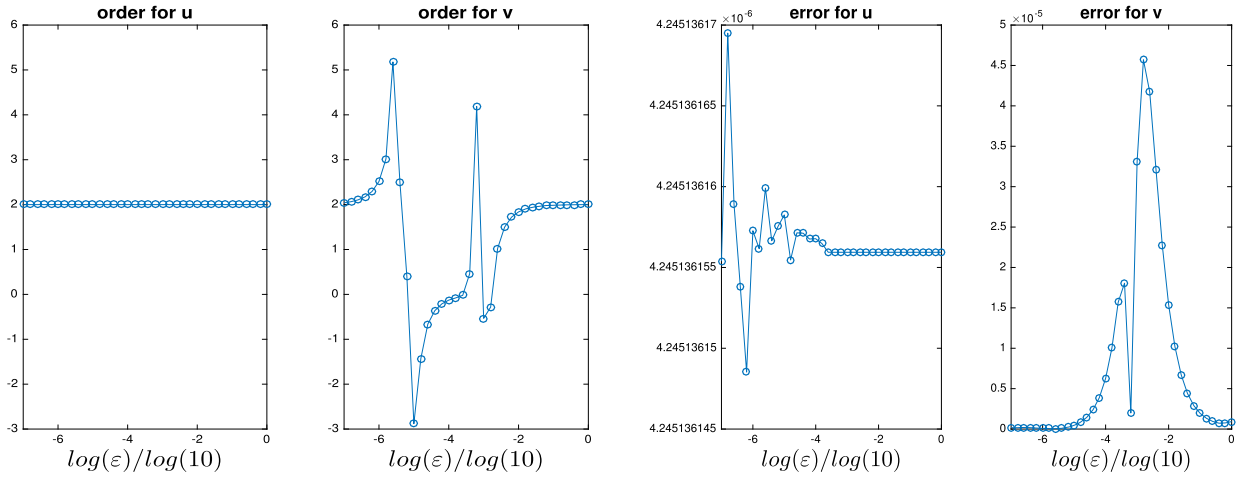
$$\partial_t J_k + \frac{1}{\tau_k} J_k + \omega_k^2 P_k = (\omega_p^{(k)})^2 E, \quad k = 1, 2, \ldots, N_\omega. \tag{6.58f}$$

Let $\bar{\tau}$ be the reference time scale, $c\bar{\tau}$ be the reference space scale, where $c = \frac{1}{\sqrt{\mu_0\epsilon_0}}$ is the speed of light, and let $\bar{E}$ be the reference electric field. By setting

$$t/\bar{\tau} \to t, \quad x/(c\bar{\tau}) \to x,$$




(a) Convergence rates, trapezoidal method (b) Errors, trapezoidal method, $\Delta t = \frac{1}{200}$




(c) Convergence rates, leapfrog method (d) Errors, leapfrog method, $\Delta t = \frac{1}{200}$

**Fig. 6.1.** Convergence rate vs $\varepsilon$ (left column) and error vs $\varepsilon$ (right column) for the ODE example in Section 6.1.

$$(H/\bar{E})\sqrt{\mu_0/\epsilon_0} \to H, \quad D/(\epsilon_0\bar{E}) \to D, \quad P/\bar{E} \to P,$$
$$(J/\bar{E})\bar{\tau} \to J, \quad E/\bar{E} \to E, \quad \chi/\bar{E}^2 \to \chi, \quad a\bar{E}^2 \to a,$$
$$\omega_k\bar{\tau} \to \omega_k, \quad (1/\tau_k)\bar{\tau} \to 1/\tau_k, \quad \bar{\tau}\omega_p^{(k)} \to \omega_p^{(k)}, \quad \frac{1}{\varepsilon}\bar{\tau} \to \frac{1}{\varepsilon}, \tag{6.59}$$

we obtain the dimensionless system,

$$\partial_t H = \partial_x E, \tag{6.60a}$$
$$\partial_t D = \partial_x H, \tag{6.60b}$$
$$D = (\epsilon_\infty + a\chi)E + \sum_{k=1}^{N_\omega} P_k, \tag{6.60c}$$
$$\partial_t \chi = -\frac{1}{\varepsilon}(\chi - E^2), \tag{6.60d}$$
$$\partial_t P_k = J_k, \quad k = 1, 2, \ldots, N_\omega, \tag{6.60e}$$
$$\partial_t J_k + \frac{1}{\tau_k}J_k + \omega_k^2 P_k = (\omega_p^{(k)})^2 E, \quad k = 1, 2, \ldots, N_\omega. \tag{6.60f}$$

Here without confusion, we use the same notation for the scaled and the original variables.

**Table 6.3**
$L^2$ errors and convergence orders for the example in Section 6.2.1, $T = 1$, the SI1 method.

| Flux | $\varepsilon$ | $N_x$ | $L^2$ error of $D$ | Order | $L^2$ error of $H$ | Order | $L^2$ error of $\chi$ | Order |
|---|---|---|---|---|---|---|---|---|
| Alternating 1 | 1 | 100 | 6.047E−03 | – | 1.506E−02 | – | 5.565E−03 | – |
| | | 200 | 3.008E−03 | 1.01 | 7.518E−03 | 1.00 | 2.662E−03 | 1.06 |
| | | 400 | 1.502E−03 | 1.00 | 3.757E−03 | 1.00 | 1.315E−03 | 1.02 |
| | | 800 | 7.508E−04 | 1.00 | 1.878E−03 | 1.00 | 6.556E−04 | 1.00 |
| | $10^{-2}$ | 100 | 5.376E−03 | – | 1.848E−02 | – | 1.120E−02 | – |
| | | 200 | 2.674E−03 | 1.01 | 9.227E−03 | 1.00 | 5.568E−03 | 1.01 |
| | | 400 | 1.335E−03 | 1.00 | 4.612E−03 | 1.00 | 2.780E−03 | 1.00 |
| | | 800 | 6.674E−04 | 1.00 | 2.305E−03 | 1.00 | 1.390E−03 | 1.00 |
| | $10^{-6}$ | 100 | 5.374E−03 | 1.00 | 1.854E−02 | 1.01 | 1.131E−02 | 1.01 |
| | | 200 | 2.673E−03 | 1.01 | 9.257E−03 | 1.00 | 5.625E−03 | 1.01 |
| | | 400 | 1.335E−03 | 1.00 | 4.627E−03 | 1.00 | 2.809E−03 | 1.00 |
| | | 800 | 6.671E−04 | 1.00 | 2.313E−03 | 1.00 | 1.404E−03 | 1.00 |
| Upwind | 1 | 100 | 5.627E−03 | – | 1.200E−02 | – | 1.001E−02 | – |
| | | 200 | 2.856E−03 | 0.98 | 6.042E−03 | 0.99 | 5.040E−03 | 0.99 |
| | | 400 | 1.439E−03 | 0.99 | 3.032E−03 | 0.99 | 2.529E−03 | 0.99 |
| | | 800 | 7.226E−04 | 0.99 | 1.519E−03 | 1.00 | 1.267E−03 | 1.00 |
| | $10^{-2}$ | 100 | 5.908E−03 | – | 1.247E−02 | – | 1.135E−02 | – |
| | | 200 | 3.003E−03 | 0.98 | 6.273E−03 | 0.99 | 5.755E−03 | 0.98 |
| | | 400 | 1.515E−03 | 0.99 | 3.147E−03 | 1.00 | 2.899E−03 | 0.99 |
| | | 800 | 7.606E−04 | 0.99 | 1.576E−03 | 1.00 | 1.455E−03 | 0.99 |
| | $10^{-6}$ | 100 | 5.926E−03 | – | 1.248E−02 | – | 1.137E−02 | – |
| | | 200 | 3.013E−03 | 0.98 | 6.279E−03 | 0.99 | 5.767E−03 | 0.98 |
| | | 400 | 1.519E−03 | 0.99 | 3.150E−03 | 1.00 | 2.905E−03 | 0.99 |
| | | 800 | 7.629E−04 | 0.99 | 1.578E−03 | 1.00 | 1.458E−03 | 0.99 |

*6.2.1. Accuracy test: a manufactured solution*

A manufactured solution is used to demonstrate the accuracy of the scheme. Consider the following dimensionless system:

$$\begin{aligned} &\partial_t H = \partial_x E + F_1, \qquad \partial_t D = \partial_x H + F_2, \\ &D = (1+\chi)E + P, \qquad \chi_t = -\frac{1}{\varepsilon}(\chi - E^2), \\ &\partial_t P = J, \qquad \partial_t J + J + P = E + F_3, \end{aligned} \tag{6.61}$$

on $\Omega = [0, 2\pi]$, with the exact solutions $E(x,t) = \sqrt{\frac{1}{2}t - \frac{1}{4}\sin(2t) + \sin^2(x) + \frac{1}{2} + \varepsilon\sin^2(t)}$, $H(x,t) = \sin(x)\sin(t)$, $\chi(x,t) = \frac{1}{2}t - \frac{1}{4}\sin(2t) + \sin^2(x) + \frac{1}{2}$, $P(x,t) = \cos(x)\sin(t)$, $D(x,t) = (1+\chi(x,t))E(x,t) + P(x,t)$. $F_1$, $F_2$, $F_3$ are the forcing terms to make (6.61) hold. Periodic boundary conditions are considered.

In Tables 6.3-6.6, $L^2$ errors and convergence orders are reported for the proposed first and second order schemes, when $\varepsilon = 1, 10^{-2}, 10^{-6}$. The alternating fluxes 1 and 2 lead to the same leading digits in errors, and only the results from schemes with the alternating flux 1 are shown. The results of the MELA and the MELE scheme also coincide (at least in their leading few digits, see Table 6.5). When $\varepsilon = 1.0, 10^{-2}, 10^{-6}$, one can see the first order accuracy for the SI1 and FI1 schemes, and the second order accuracy for the MELE and MELA schemes. For the META scheme, second order accuracy is observed for $\varepsilon = 1, 10^{-6}$, yet when $\varepsilon = 10^{-2}$, an order reduction is observed with $N_x = 200, 400$. In this case, $\Delta t$ and $\varepsilon$ are comparable, and the time integrator suffers from the order reduction in this intermediate regime. As we further refine the mesh, the second order accuracy will be regained.

*6.2.2. Soliton-like propagation*

Here, we consider a soliton-like propagation problem with a setup similar to [4,5,16]. A dimensionless system (6.58) is considered with $N_\omega = 1$, and $P = P_1$, $J = J_1$. The parameters are taken as

$$\epsilon_\infty = 2.25, \quad \epsilon_s = 5.25, \quad \beta_1 = \epsilon_s - \epsilon_\infty, \quad \tau^{-1} = 1.168 \times 10^{-5}, \quad \omega_0 = 5.84, \quad \omega_p = \sqrt{\beta_1}\omega_0.$$

Initially, all the fields are zero. From the left boundary $x = 0$, a *sech* signal will enter the domain,

$$E(x=0,t) = sech(t-20)\cos(\Omega_0 t), \quad \Omega_0 = 12.57. \tag{6.62}$$

As discussed in Section 5.1 of [16], if the soliton condition

$$a_{circle} \approx 3.78 \times 10^{-2} \quad \text{and} \quad a = \frac{4}{3}a_{circle}$$

**Table 6.4**
$L^2$ errors and convergence orders for the example in Section 6.2.1, $T = 1$, the FI1 method.

| Flux | $\varepsilon$ | $N_x$ | $L^2$ error of $E$ | Order | $L^2$ error of $H$ | Order | $L^2$ error of $\chi$ | Order |
|---|---|---|---|---|---|---|---|---|
| Alternating 1 | 1 | 100 | 9.966E−03 | - | 1.540E−02 | – | 1.089E−02 | – |
| | | 200 | 5.049E−03 | 0.98 | 7.794E−03 | 0.98 | 5.484E−03 | 0.99 |
| | | 400 | 2.554E−03 | 0.98 | 3.921E−03 | 0.99 | 2.762E−03 | 0.99 |
| | | 800 | 1.280E−03 | 1.00 | 1.966E−03 | 1.00 | 1.383E−03 | 1.00 |
| | $10^{-2}$ | 100 | 7.641E−03 | – | 1.839E−02 | – | 1.557E−02 | – |
| | | 200 | 3.842E−03 | 0.99 | 9.297E−03 | 0.98 | 7.813E−03 | 0.99 |
| | | 400 | 1.931E−03 | 0.99 | 4.675E−03 | 0.99 | 3.917E−03 | 1.00 |
| | | 800 | 9.664E−04 | 1.00 | 2.344E−03 | 1.00 | 1.960E−03 | 1.00 |
| | $10^{-6}$ | 100 | 7.656E−03 | – | 1.844E−02 | – | 1.566E−02 | – |
| | | 200 | 3.848E−03 | 0.99 | 9.325E−03 | 0.98 | 7.862E−03 | 0.99 |
| | | 400 | 1.932E−03 | 0.99 | 4.689E−03 | 0.99 | 3.945E−03 | 0.99 |
| | | 800 | 9.672E−04 | 1.00 | 2.351E−03 | 1.00 | 1.974E−03 | 1.00 |
| Upwind | 1 | 100 | 6.820E−03 | – | 1.640E−02 | – | 9.889E−03 | – |
| | | 200 | 3.516E−03 | 0.96 | 8.216E−03 | 1.00 | 5.084E−03 | 0.96 |
| | | 400 | 1.800E−03 | 0.97 | 4.113E−03 | 1.00 | 2.590E−03 | 0.97 |
| | | 800 | 9.065E−04 | 0.99 | 2.057E−03 | 1.00 | 1.303E−03 | 0.99 |
| | $10^{-2}$ | 100 | 5.497E−03 | – | 1.519E−02 | – | 1.132E−02 | – |
| | | 200 | 2.865E−03 | 0.94 | 7.552E−03 | 1.01 | 5.880E−03 | 0.94 |
| | | 400 | 1.468E−03 | 0.96 | 3.766E−03 | 1.00 | 3.002E−03 | 0.97 |
| | | 800 | 7.415E−04 | 0.99 | 1.880E−03 | 1.00 | 1.515E−03 | 0.99 |
| | $10^{-6}$ | 100 | 5.544E−03 | – | 1.517E−02 | – | 1.134E−02 | – |
| | | 200 | 2.889E−03 | 0.94 | 7.543E−03 | 1.01 | 5.891E−03 | 0.94 |
| | | 400 | 1.479E−03 | 0.97 | 3.761E−03 | 1.00 | 3.011E−03 | 0.97 |
| | | 800 | 7.470E−04 | 0.99 | 1.878E−03 | 1.00 | 1.520E−03 | 0.99 |

**Table 6.5**
$L^2$ errors and convergence orders for the example in Section 6.2.1, $T = 1$, the MELE (MELA) method.

| Flux | $\varepsilon$ | $N_x$ | $L^2$ error of $E$ | Order | $L^2$ error of $H$ | Order | $L^2$ error of $\chi$ | Order |
|---|---|---|---|---|---|---|---|---|
| Alternating 1 | 1 | 100 | 1.182E−04 | – | 1.358E−04 | – | 1.689E−04 | – |
| | | 200 | 2.752E−05 | 2.10 | 3.451E−05 | 1.98 | 4.222E−05 | 2.00 |
| | | 400 | 7.574E−06 | 1.86 | 8.341E−06 | 2.05 | 1.056E−05 | 2.00 |
| | | 800 | 1.753E−06 | 2.11 | 2.028E−06 | 2.04 | 2.638E−06 | 2.00 |
| | $10^{-2}$ | 100 | 1.357E−04 | – | 1.136E−04 | – | 3.248E−04 | – |
| | | 200 | 3.918E−05 | 1.79 | 2.835E−05 | 2.00 | 9.495E−05 | 1.77 |
| | | 400 | 1.054E−05 | 1.89 | 7.084E−06 | 2.00 | 2.556E−05 | 1.89 |
| | | 800 | 2.916E−06 | 1.85 | 1.771E−06 | 2.00 | 7.027E−06 | 1.86 |
| | $10^{-6}$ | 100 | 1.685E−04 | – | 2.139E−04 | – | 3.910E−04 | – |
| | | 200 | 3.726E−05 | 2.18 | 5.530E−05 | 1.95 | 7.388E−05 | 2.40 |
| | | 400 | 1.035E−05 | 1.85 | 1.295E−05 | 2.09 | 2.175E−05 | 1.76 |
| | | 800 | 2.310E−06 | 2.16 | 3.293E−06 | 1.98 | 5.084E−06 | 2.10 |

holds, a soliton structure is supported by the Kerr media ($\varepsilon = 0$). Based on this, we set $a = 0.0504$. The relaxation time in our Kerr-Debye-Lorentz model is taken as $\varepsilon = 10^{-8}$ and $\varepsilon = 0.1$.

In our simulation, we take the computational domain as $\Omega = [0, 50]$, and the final time as $T = 40$ or 80. With these choices, the signal will not reach the right boundary by the final time, thus zero boundary conditions are used for $E$ and $H$ at $x = 50$. Depending on the numerical fluxes, the numerical boundary condition may be needed for the magnetic field $H$ at the left boundary. Just as in [4,5,16], based on the dispersion relation of the linear model with the Lorentz response, namely

$$k(\omega) = \omega\sqrt{\epsilon_\infty - \frac{\omega_p^2}{\omega^2 - i\omega/\tau - \omega_0^2}}, \tag{6.63}$$

we derive and propose the following numerical boundary condition

$$H(x = 0, t) \approx \frac{1}{2}\left(\sum_{m=0}^{2} \frac{(-i)^m}{m!}(\frac{1}{Z})^{(m)}\Big|_{\omega=\Omega_0} f^{(m)}(t)\right) e^{i\Omega_0 t} + c.c., \tag{6.64}$$

**Table 6.6**
$L^2$ errors and convergence orders for the example in Section 6.2.1, $T = 1$, the META method.

| Flux | $\varepsilon$ | $N_x$ | $L^2$ error of $E$ | Order | $L^2$ error of $H$ | Order | $L^2$ error of $\chi$ | Order |
|---|---|---|---|---|---|---|---|---|
| Alternating 1 | 1 | 100 | 1.303E−04 | – | 1.709E−04 | – | 1.707E−04 | – |
| | | 200 | 3.192E−05 | 2.03 | 4.136E−05 | 2.05 | 4.292E−05 | 1.99 |
| | | 400 | 8.394E−06 | 1.93 | 1.122E−05 | 1.88 | 1.083E−05 | 1.99 |
| | | 800 | 2.358E−06 | 1.83 | 2.554E−06 | 2.14 | 2.699E−06 | 2.00 |
| | $10^{-2}$ | 100 | 1.651E−04 | – | 1.373E−04 | – | 5.476E−04 | – |
| | | 200 | 1.355E−04 | 0.29 | 4.537E−05 | 1.60 | 3.687E−04 | 0.57 |
| | | 400 | 1.429E−04 | −0.07 | 1.385E−05 | 1.71 | 3.170E−04 | 0.22 |
| | | 800 | 3.804E−05 | 1.91 | 3.819E−06 | 1.86 | 8.455E−05 | 1.91 |
| | | 1600 | 1.087E−05 | 1.81 | 1.002E−06 | 1.93 | 2.397E−05 | 1.82 |
| | $10^{-6}$ | 100 | 1.558E−04 | – | 2.559E−04 | – | 3.105E−04 | – |
| | | 200 | 4.736E−05 | 1.72 | 6.276E−05 | 2.03 | 9.966E−05 | 1.64 |
| | | 400 | 1.124E−05 | 2.08 | 1.404E−05 | 2.16 | 2.503E−05 | 1.99 |
| | | 800 | 2.542E−06 | 2.14 | 3.681E−06 | 1.93 | 4.986E−06 | 2.33 |
| Upwind | 1 | 100 | 1.076E−04 | – | 1.067E−04 | – | 1.169E−04 | – |
| | | 200 | 2.776E−05 | 1.95 | 2.668E−05 | 2.00 | 3.031E−05 | 1.95 |
| | | 400 | 7.058E−06 | 1.98 | 6.676E−06 | 2.00 | 7.744E−06 | 1.97 |
| | | 800 | 1.779E−06 | 1.99 | 1.669E−06 | 2.00 | 1.954E−06 | 1.99 |
| | $10^{-2}$ | 100 | 1.427E−04 | – | 1.443E−04 | – | 5.295E−04 | – |
| | | 200 | 1.348E−04 | 0.08 | 4.720E−05 | 1.61 | 3.686E−04 | 0.52 |
| | | 400 | 1.429E−04 | -0.08 | 1.421E−05 | 1.73 | 3.170E−04 | 0.22 |
| | | 800 | 3.804E−05 | 1.91 | 3.899E−06 | 1.87 | 8.455E−05 | 1.91 |
| | | 1600 | 1.087E−05 | 1.81 | 1.020E−06 | 1.93 | 2.397E−05 | 1.82 |
| | $10^{-6}$ | 100 | 1.168E−04 | – | 1.319E−04 | – | 2.813E−04 | – |
| | | 200 | 3.007E−05 | 1.96 | 3.296E−05 | 2.00 | 7.223E−05 | 1.96 |
| | | 400 | 7.671E−06 | 1.97 | 8.239E−06 | 2.00 | 1.822E−05 | 1.99 |
| | | 800 | 1.985E−06 | 1.95 | 2.059E−06 | 2.00 | 4.508E−06 | 2.02 |

where *c.c.* stands for the complex conjugate of the first term, $f^{(m)}$ is the $m$-th derivative of $f(t) = sech(t-20)$, $Z(\omega) = -\frac{\omega}{k(\omega)}$ and $(\frac{1}{Z})^{(m)}$ is the $m$-th derivative of $\frac{1}{Z(\omega)}$ with respect to $\omega$.

**1.) Initial guess of the MELE scheme**

As discussed in Section 5.6.2, when the MELE scheme is applied, special initial guesses may be needed for the Newton solver to capture the physically relevant solution. In this section, we want to use this soliton-like propagation as an example to numerically examine the effectiveness and efficiency of the various initial guesses proposed in Section 5.6.2.

Recall that to update the solution at $t^{n+1}$ based on the solution at $t^n$, the MELE scheme at each nodal point $x_{jk}, \forall j,k$ consists of two steps:

- Initial-guess step: to get the initial guess $\mathbf{u_0}$, by applying Newton solver to one of the strategies in (5.54), (5.55), (5.56). For the Newton solver in this step, it also requires an initial guess which is set based on $\mathbf{u_h^n}$.
- Solution step: to apply Newton solver with the initial guess $\mathbf{u_0}$ to solve (5.21) and (5.33), and obtain $\mathbf{u_h^{n+1}}$.

For the initial-guess step, we will count the total number of Newton iterations, denoted as $Tot_{guess}$, used at all nodal points $x_{jk}, \forall j,k$ up to the final time $T$. During the $i$-th time step, we also count the total number of nodal points, denoted as $N^i_{guess}$, for each of which at least one iteration is needed in the Newton solver. The averaged number of iterations is then defined as

$$Avg_{guess} = \frac{Tot_{guess}}{\sum_{i:i\Delta t\le T} N^i_{guess}}.$$

Similarly, for the solution step, we can count and define $Tot_{sol}$ and $Avg_{sol}$.

Finally, we define $Tot_{total} = Tot_{guess} + Tot_{sol}$ and $Avg_{total} = Avg_{guess} + Avg_{sol}$. We want to point out that for each step, it is possible that zero iteration is needed in the Newton solver (see the flow chart of the algorithm in Section 5.6.1).

We implement the algorithm up to $T = 80$ on a uniform mesh with $N_x = 3500$, and the iteration numbers defined above are reported in Table 6.7 when the relaxation time is $\varepsilon = 10^{-8}$ or 0.1. The error tolerance is set as $\mathrm{Tol} = 10^{-10}$ for both the initial-guess step and the solution step (see the algorithm in Section 5.6.1). For the simulations with a given $\varepsilon$, the numerical solutions based on all three initial guess strategies share the same leading digits and visually coincide with each other. From Table 6.7, we can see that the *limiting initial guess strategy* needs the fewest iterations when $\varepsilon = 10^{-8}$ and the most iterations when $\varepsilon = 0.1$. The efficiency seems to depend on the value of $\varepsilon$. The *trapezoidal initial guess strategy* performs well for both $\varepsilon = 0.1$ and $\varepsilon = 10^{-8}$. When $\varepsilon = 10^{-8}$, both the *limiting and trapezoidal initial guess strategies* are

**Table 6.7**
Comparison of various initial guess strategies based on the soliton-like propagation.

| $\varepsilon$ | Initial guess | $Tot_{guess}$ | $Avg_{guess}$ | $Tot_{sol}$ | $Avg_{sol}$ | $Tot_{total}$ | $Avg_{total}$ |
|---|---|---|---|---|---|---|---|
| 0.1 | Backward Euler | 248639720 | 2.18 | 217314553 | 1.95 | 465954273 | 4.13 |
| | Trapezoidal | 261425221 | 2.16 | 147253639 | 1.26 | 408678860 | 3.42 |
| | Limiting | 265499559 | 2.20 | 279257430 | 2.38 | 544756989 | 4.58 |
| $10^{-8}$ | Backward Euler | 251777821 | 2.20 | 227013893 | 2.03 | 478791714 | 4.23 |
| | Trapezoidal | 265115776 | 2.18 | 140879982 | 1.20 | 405995758 | 3.38 |
| | Limiting | 263228932 | 2.18 | 117477103 | 1.00 | 380706035 | 3.18 |

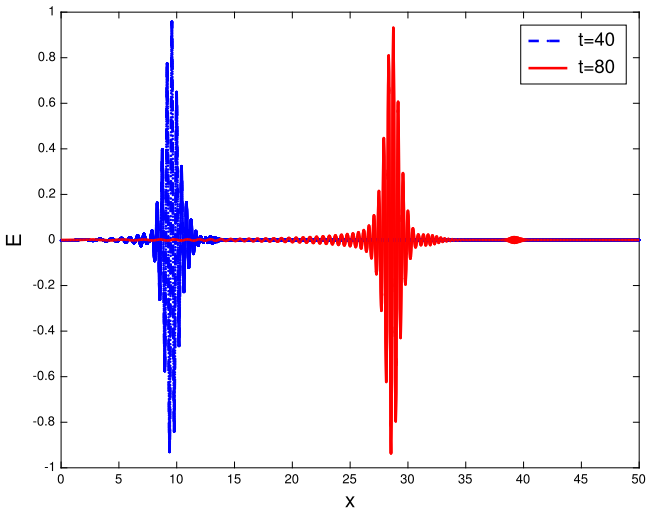



**Fig. 6.2.** Reference solution of the soliton-like propagation in Section 6.2.2, $\varepsilon = 10^{-8}$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

comparably efficient. Compared with the trapezoidal initial guess strategy, the *backward Euler initial guess strategy* needs more iterations. Based on this study, we recommend and will use the *trapezoidal initial guess* strategy for the MELE scheme in all the remaining numerical experiments.

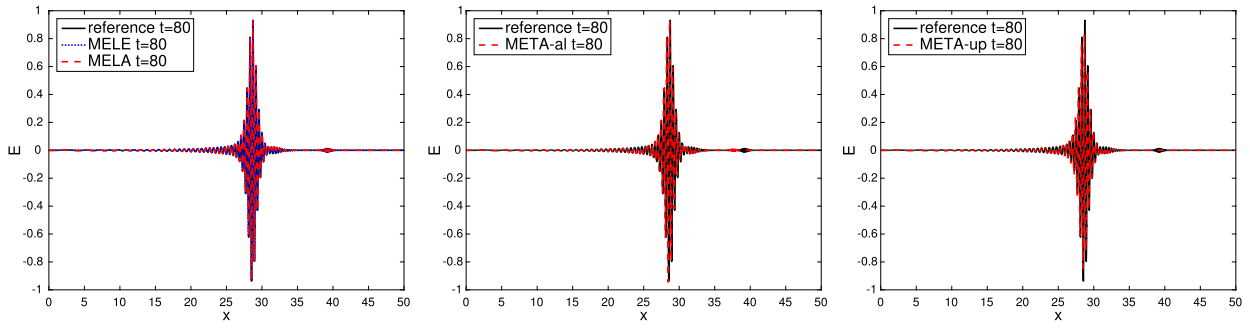
**2.) Numerical results for the soliton-like propagation**

Next we want to present the numerical results from simulating the soliton-like propagation using the proposed methods. With this example, we want to illustrate the AP and PP properties, and compare the performance of different schemes. We also want to examine the influence of the finite relaxation time on the soliton propagation.

First of all, no negative value of $\chi_{jk}$ is observed throughout the simulation of each scheme. This confirms the PP property of the methods. Numerically, we deem $\chi_h$ to be nonnegative if $\chi_{jk} > \text{Tol}_*, \forall j, k$. We set $\text{Tol}_* = -5 \times 10^{-13}$ in all the simulations.

When $\varepsilon = 10^{-8}$, the reference solution is computed by applying (the nodal version of) the leapfrog DG scheme in [4], with an alternating flux, to the limiting Kerr-Lorenz model on a uniform mesh with $h = \frac{1}{400}$, and it is plotted in Fig. 6.2. One can see the soliton propagating in time, and moreover, a daughter pulse is observed at $t = 80$ due to the harmonic generation.

- In Figs. 6.3 and 6.4, we show the computed electric field $E$ by all the second order schemes with $h = \frac{1}{70}$ and $\varepsilon = 10^{-8}$ at $t = 80$. The numerical results of the MELA and the MELE schemes are visually indistinguishable. From Fig. 6.3, one can see that the result of each method matches well the overall structure of the reference solution, indicating the AP property of the proposed methods. In Fig. 6.4, the plots are zoomed-in to compare the detailed structures of the main pulse and the daughter pulse due to the harmonic generation. The MELE and MELA schemes perform the best among all, in terms of capturing the location and amplitude of the main and daughter pulses. The META-al scheme captures the amplitude of both pulses yet not their location, especially of the daughter pulse, due to the numerical dispersive errors. As a more diffusive method with the use of the upwind flux, the META-up scheme captures the main pulse with slightly reduced amplitude and shifted location, while it is too dissipative to capture the daughter pulse.
- In Figs. 6.5 and 6.6, we show the zoomed-in electric field $E$ by the first order semi-implicit SI1 scheme and the fully implicit FI1 scheme with $h = \frac{1}{70}$ and $\varepsilon = 10^{-8}$ at $t = 80$. The SI1-al scheme captures both the main and daughter pulses well; the SI1-up schemes captures the main pulse reasonably well with the amplitude slighted reduced, yet it does not




(a) $\varepsilon = 10^{-8}$, MELE (MELA), $t = 80$ (b) $\varepsilon = 10^{-8}$, META-al, $t = 80$ (c) $\varepsilon = 10^{-8}$, META-up, $t = 80$

**Fig. 6.3.** The electric field $E$ in the soliton-like propagation at $t = 80$, computed by the second order MELE, MELA and META schemes, $\varepsilon = 10^{-8}$.

**Table 6.8**
Normalized CPU times for the example in Section 6.2.2 with $T = 40$ and $h = \frac{1}{70}$.

| | MELA | MELE | META-al | META-up |
|---|---|---|---|---|
| $\varepsilon = 0.1$ | 1.00 | 1.28 | 606.42 | 457.71 |
| $\varepsilon = 10^{-8}$ | 1.53 | 1.80 | 777.45 | 574.24 |

capture the daughter pulse with its numerical dissipation. The fully implicit Fl1 scheme is too dissipative to capture either pulse.

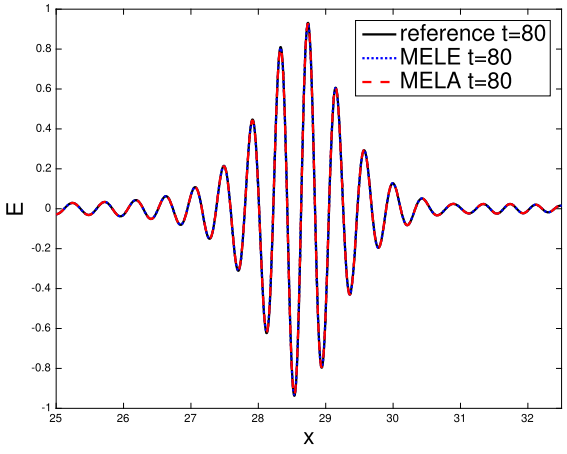
When $\varepsilon = 0.1$, the reference solution is obtained by the MELA scheme with $h = \frac{1}{400}$ and shown in Fig. 6.7. On such refined mesh, the numerical solutions are not sensitive to the choice of the schemes. Compared with the solution for $\varepsilon = 10^{-8}$, the main pulse now is broadened and the soliton structure is broken. In Fig. 6.8, we present the zoomed-in plots of the computed electric field $E$ by the proposed second order schemes at $t = 80$ with $h = \frac{1}{70}$, together with the reference solution. Similar observations can be made as for the case with $\varepsilon = 10^{-8}$: the MELA and the MELE schemes perform the best, and there is no visual difference between their results; the META-al method gives a slight different location of the main pulse and fails to capture the correct location of the daughter pulse. The META-up scheme still does not capture the daughter pulse, due to its numerical dissipation. When the proposed first order schemes are applied, one can make observations and comments similarly as when these methods are applied to the case of $\varepsilon = 10^{-8}$. The numerical results are omitted here.

From the numerical experiments for the model with $\varepsilon = 0$ (see the reference solutions in Fig. 6.2), $\varepsilon = 10^{-8}$ and $\varepsilon = 0.1$, we learn that as the relaxation time increases, the main pulse of the soliton-like structure is broadened and the amplitude of the daughter pulse due to the harmonic generation is reduced.

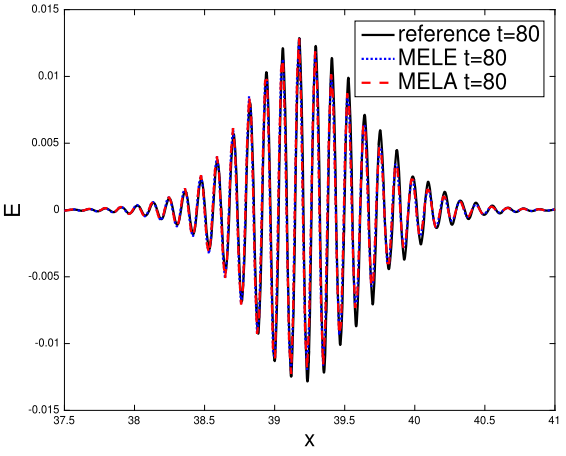
**Remark 6.1.** Note that for all the methods we implemented, the same spatial mesh size is used, while the time step size is chosen for each method according to the beginning of Section 6. In particular, for the second order methods, the trapezoidal schemes (namely, META-a and META-up) use a time step size that is about five times of that for the leapfrog schemes (namely, MELE and MELA), yet the trapezoidal schemes still take much longer time to run with their nonlinear algebraic being global. To see this, we report in Table 6.8 the normalized CPU times of various second order methods on a mesh of $h = \frac{1}{70}$ up to the final time $T = 40$ with $\varepsilon = 0.1, 10^{-8}$. The normalization is with respect to the CPU time of the MELA scheme with $\varepsilon = 0.1$. The simulations are performed on a Dell precision T5500 machine with 2.27 GHz Intel Xeon E5607 CPU and 12300 MB memory. As mentioned before, the nonlinear algebraic solvers for the trapezoidal and leapfrog schemes are different. Specifically, we use a Jacobian-free Newton Krylov solver for the trapezoidal schemes within the framework of the KINSOL package [29] and an in-house Newton solver for the leapfrog schemes. Even with the difference in time step sizes and nonlinear algebraic solvers, Table 6.8 qualitatively shows that the leapfrog schemes have much better computational efficiency than the trapezoidal ones. According to the earlier on discussion in this section, the cost-efficient MELE and MELA methods also better resolve the solution than the META-a and META-up methods. The resolution of these trapezoidal methods can be improved if the time step size is reduced, and this will come at the price of a further increased computational time. Similar comment can be made for the first order semi-implicit and full implicit schemes.

### *6.2.3. Sech signal in fused bulk silica*

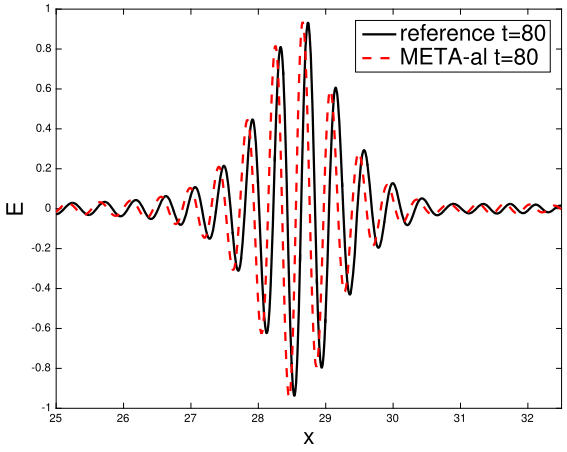
Our final example concerns a signal propagating in the fused bulk silica that follows a similar setting as in [17]. The model involves three linear dispersion responses. Before non-dimensionalization, the physical parameters in (6.58) are taken as



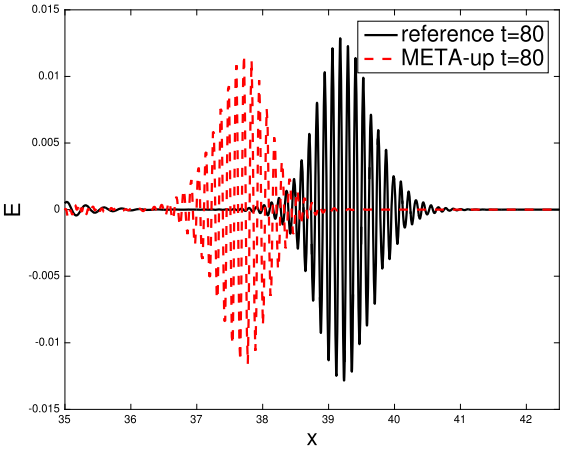

(a) $\varepsilon = 10^{-8}$, MELE (MELA), main pulse



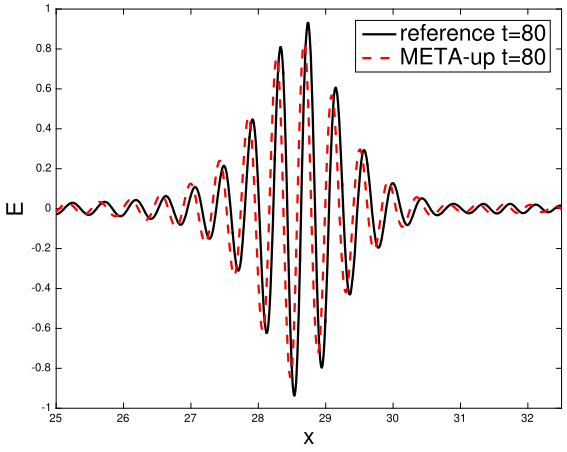

(b) $\varepsilon = 10^{-8}$, MELE (MELA), daughter pulse



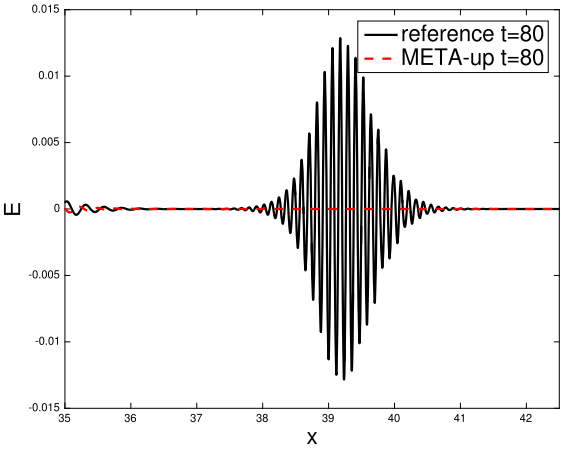

(c) $\varepsilon = 10^{-8}$, META-al, main pulse




(d) $\varepsilon = 10^{-8}$, META-al, daughter pulse

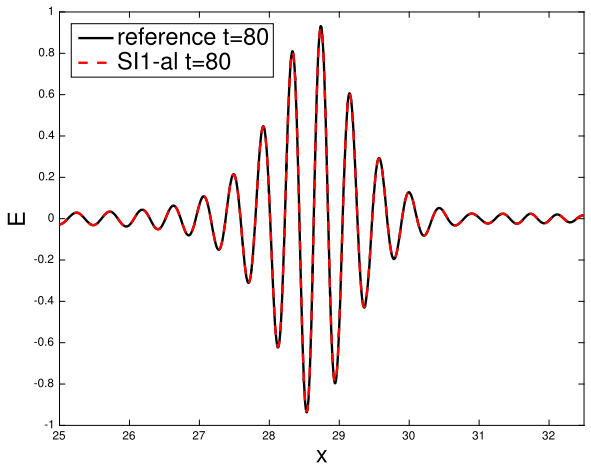



(e) $\varepsilon = 10^{-8}$, META-up, main pulse

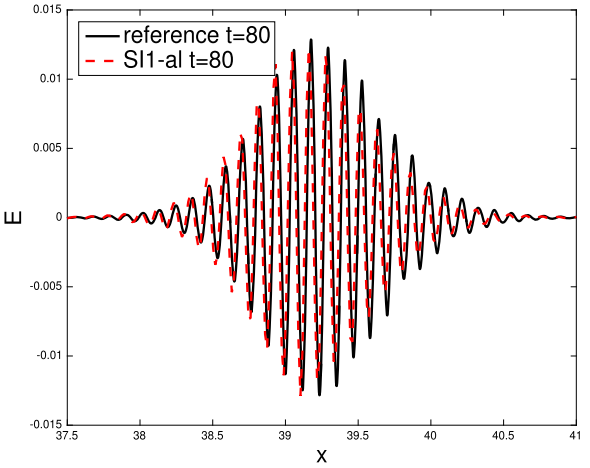



(f) $\varepsilon = 10^{-8}$, META-up, daughter pulse

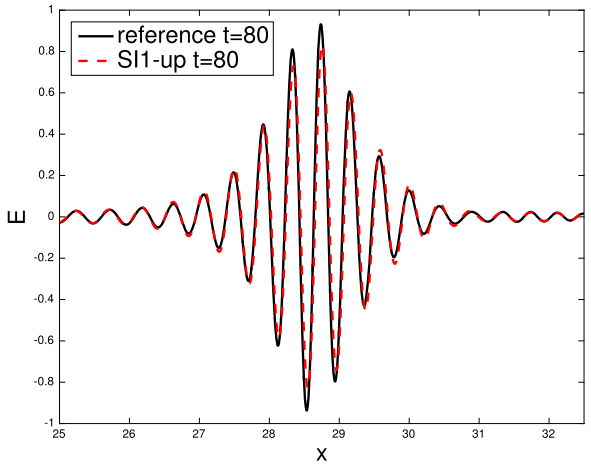
**Fig. 6.4.** The zoomed-in electric field $E$ in the soliton-like propagation at $t = 80$, computed by the second order MELE, MELA and META schemes, $\varepsilon = 10^{-8}$. Left: main pulse; Right: daughter pulse.



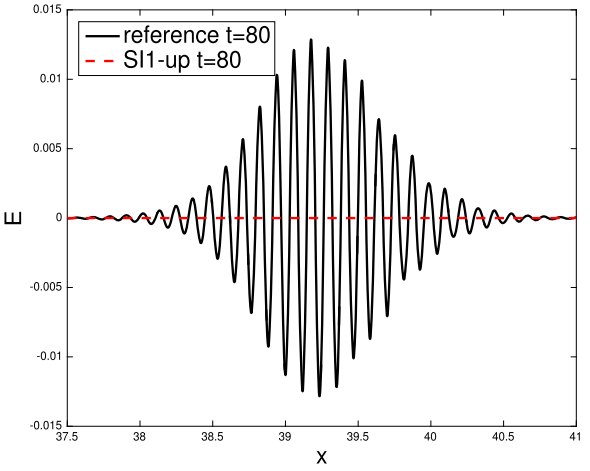

(a) $\varepsilon = 10^{-8}$, SI1-al, main pulse




(b) $\varepsilon = 10^{-8}$, SI1-al, daughter pulse




(c) $\varepsilon = 10^{-8}$, SI1-up, main pulse




(d) $\varepsilon = 10^{-8}$, SI1-up, daughter pulse

**Fig. 6.5.** The zoomed-in electric field $E$ in the soliton-like propagation at $t = 80$, computed by the first order semi-implicit SI1 scheme, $\varepsilon = 10^{-8}$. Left: main pulse; Right: daughter pulse.

$$\epsilon_\infty = 1, \quad \mu_0 = 4\pi \times 10^{-7}\ \text{H/m}, \quad \epsilon_0 = \frac{1}{\mu_0 c^2},$$
$$\tau_1 = \tau_2 = \tau_3 = 10\ \mu\text{s},$$
$$\omega_k = \frac{2\pi\lambda_k}{c}, \quad \omega_p^{(k)} = \sqrt{\beta_k}\omega_k, \quad k = 1, 2, 3,$$

where

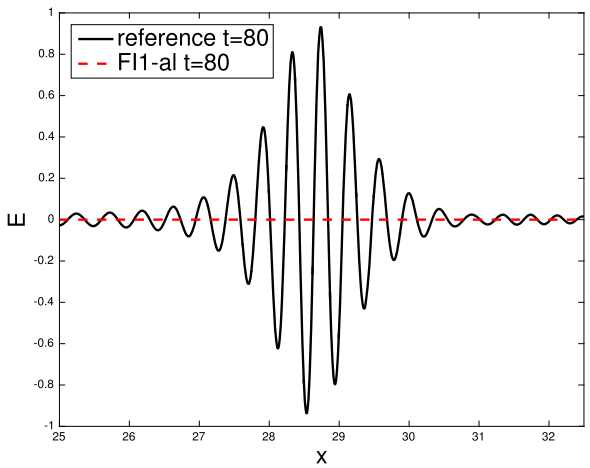
$$\beta_1 = 0.7, \quad \beta_2 = 0.4, \quad \beta_3 = 0.9, \qquad \lambda_1 = 0.068\ \mu\text{m}, \quad \lambda_2 = 0.116\ \mu\text{m}, \quad \lambda_3 = 9.9\ \mu\text{m}.$$

Here again $c$ is the speed of light.

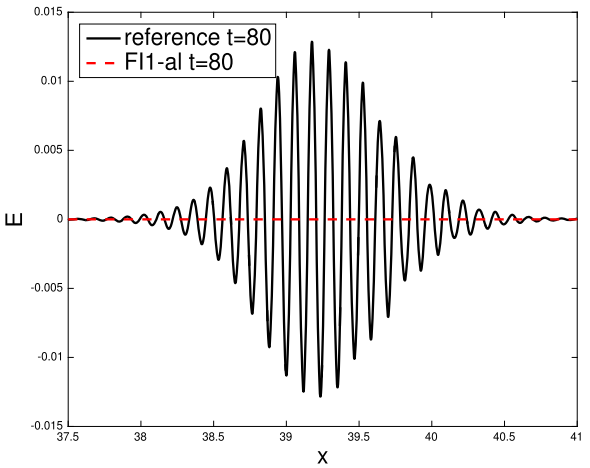
Initially, all the fields are zero. From the left boundary $x = 0$, a *sech* signal enters the domain

$$E(x = 0, t) = E_0 sech\left(\frac{t - 50fs}{t_0}\right)\cos(\Omega_0 t), \tag{6.65}$$

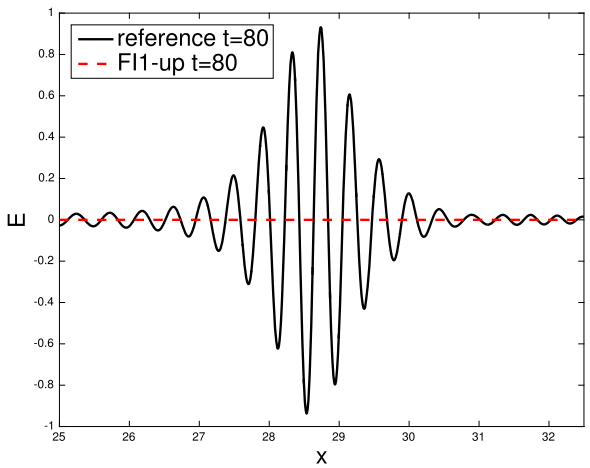
where $t_0 = 5$ fs, $\Omega_0 = 1520$ THz, and $E_0$ satisfies $aE_0^2 = 0.06$. The final time is $T = 40\pi/\Omega_0$, and the spatial domain is $[0, 15ct_0]$. The relaxation time takes the value of $\varepsilon = 0.5$ fs, 0.1 fs and $10^{-4}$ fs. To get a dimensionless system, we set $\bar{\tau} = 5$ fs $= 5 \times 10^{-15}$ s and $\bar{E} = E_0$ in (6.59). The numerical boundary condition for $H$ at the left boundary is obtained similarly as for the example in Section 6.2.2. The signal will not reach the right boundary at the final time, so zero boundary conditions are used at the right boundary for $E$ and $H$. With the physical parameters specified above, the nonlinear effect is



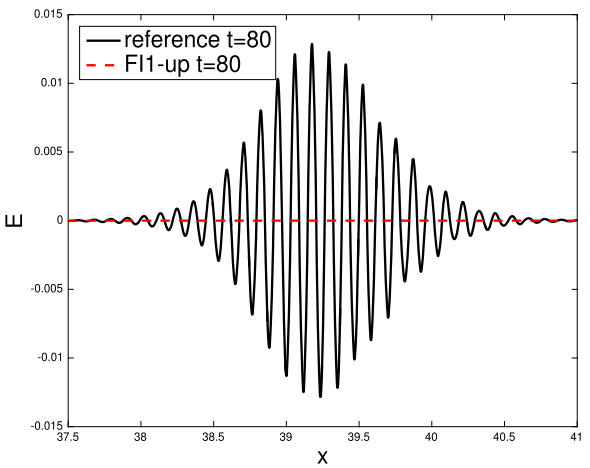

(a) $\varepsilon = 10^{-8}$, FI1-al, main pulse




(b) $\varepsilon = 10^{-8}$, FI1-al, daughter pulse




(c) $\varepsilon = 10^{-8}$, FI1-up, main pulse




(d) $\varepsilon = 10^{-8}$, FI1-up, daughter pulse

**Fig. 6.6.** The zoomed-in electric field $E$ in the soliton-like propagation at $t = 80$, computed by the first order fully implicit FI1 scheme, $\varepsilon = 10^{-8}$. Left: main pulse; Right: daughter pulse.

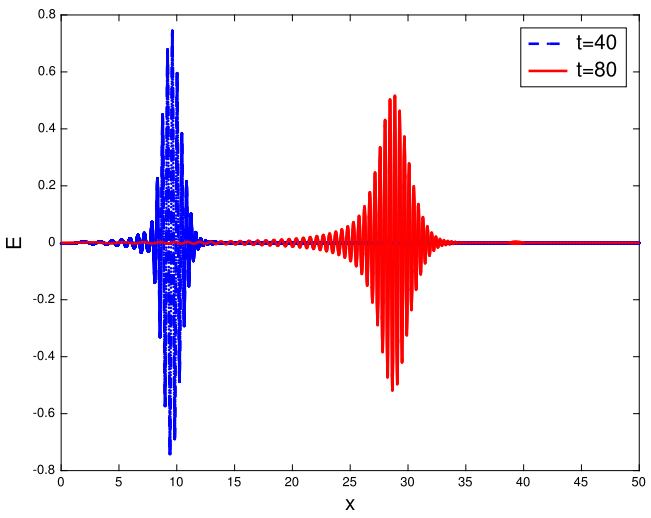
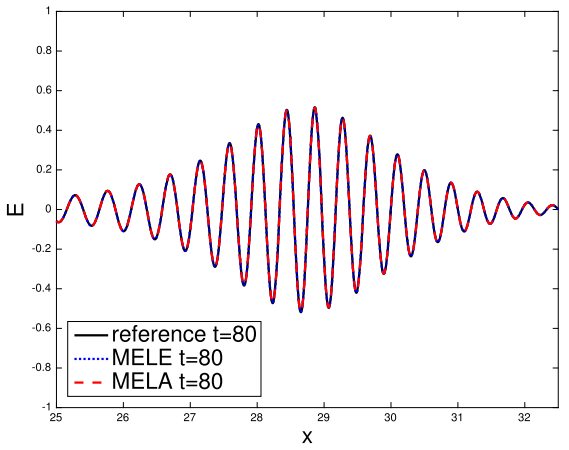
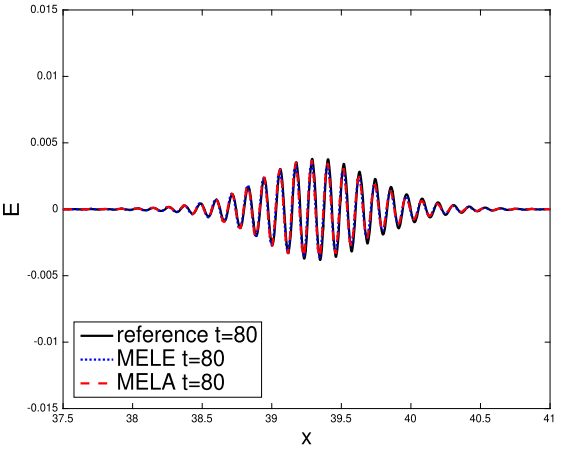


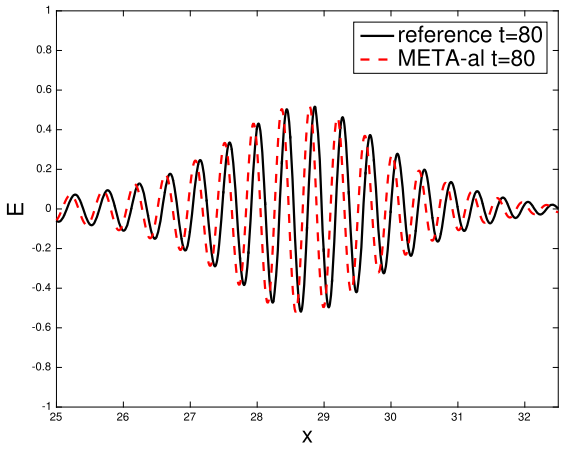

**Fig. 6.7.** Reference solution of the soliton-like propagation in Section 6.2.2, $\varepsilon = 0.1$.



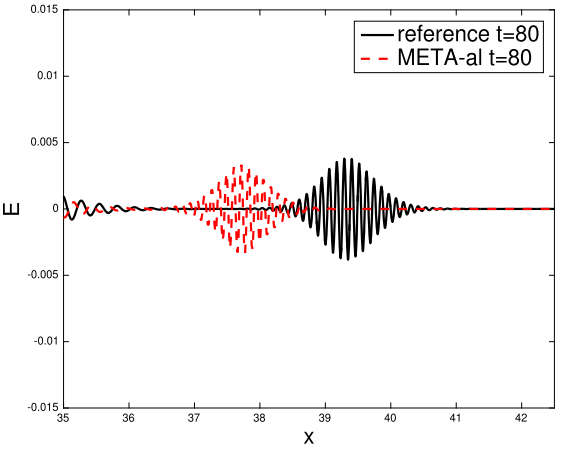

(a) $\varepsilon = 0.1$, MELE (MELA), main pulse



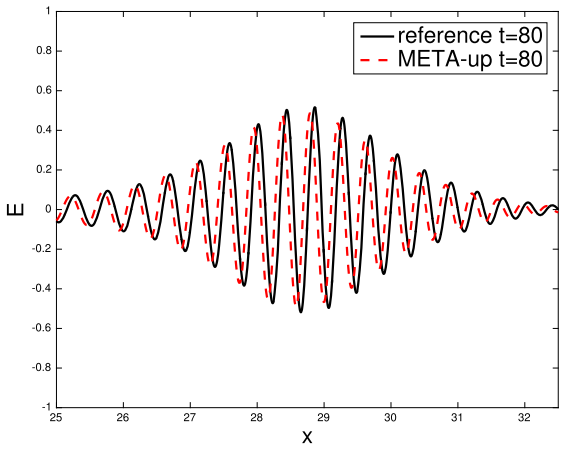

(b) $\varepsilon = 0.1$, MELE (MELA), daughter pulse



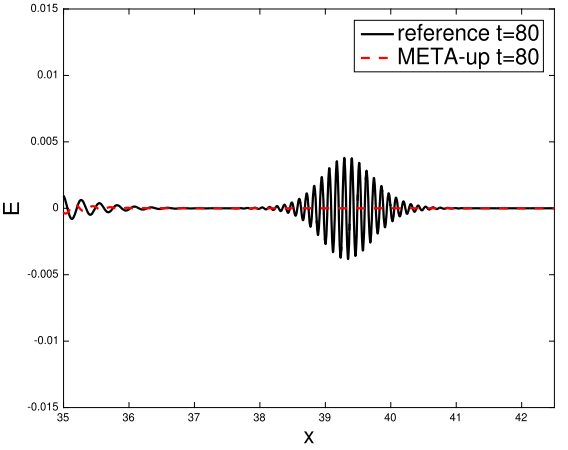

(c) $\varepsilon = 0.1$, META-al, main pulse




(d) $\varepsilon = 0.1$, META-al, daughter pulse

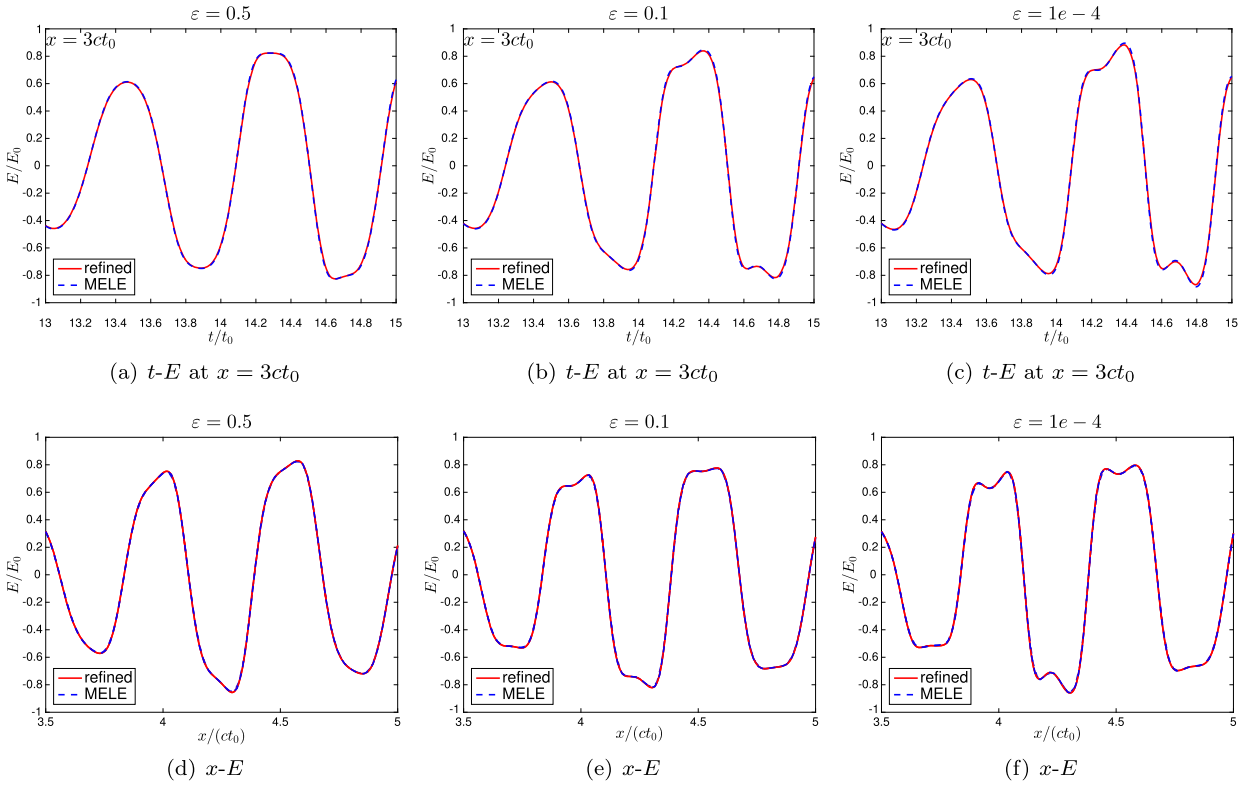



(e) $\varepsilon = 0.1$, META-up, main pulse




(f) $\varepsilon = 0.1$, META-up, daughter pulse

**Fig. 6.8.** The zoomed-in electric field $E$ in the soliton-like propagation at $t = 80$, computed by the second order MELE, MELA and META schemes, $\varepsilon = 0.1$. Left: main pulse; Right: daughter pulse.




(a) $t$-$E$ at $x = 3ct_0$ (b) $t$-$E$ at $x = 3ct_0$ (c) $t$-$E$ at $x = 3ct_0$

(d) $x$-$E$ (e) $x$-$E$ (f) $x$-$E$

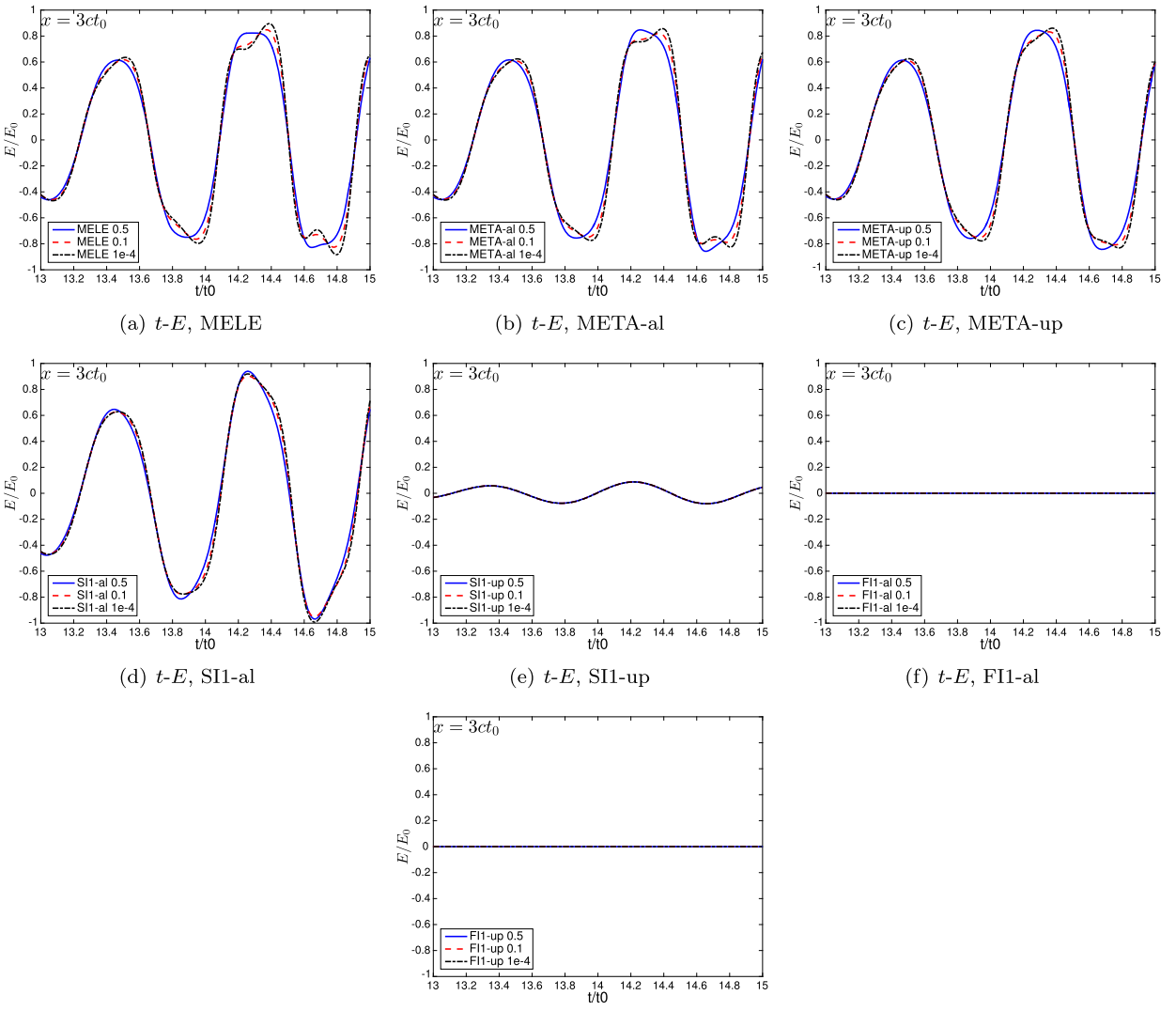
**Fig. 6.9.** The computed electric field $E$ by the MELE scheme for the example in Section 6.2.3, with $\varepsilon = 0.5,\ 0.1,\ 10^{-4}$ (fs). Solid line: finer mesh with $h = \frac{1}{250}$; Dash line: coarser mesh with $h = \frac{1}{250}$. First row: temporal configuration at $x = 3ct_0$; Second row: spatial configuration at the final time $T$.

dominant in the model (see [17]). Using this example, we want to compare different schemes and investigate the influence of the finite relaxation time $\varepsilon$. A reference solution is computed using the second order META-al scheme with $h = \frac{1}{250}$. This mesh is refined enough that all second order schemes produce visually indistinguishable results.

In Fig. 6.9, we present the electric field $E$ computed by the MELE scheme with $h = \frac{1}{50}$, together with the reference solution, for the model with different relaxation times. In the top row, the temporal configuration of $E$ is plotted at $x = 3ct_0$. In the bottom row, the plots are for the spatial configuration of $E$ at the final time. The results of the MELE scheme match the reference solution well. Sharp feature due to the self-steepening effect is observed. The smaller the relaxation time $\varepsilon$ is, the steeper feature the solution has. As observed in [17], the finite relaxation time reduces the self-steepening effect when compared with the Kerr-type model without the relaxation. On such a coarse mesh with $h = \frac{1}{50}$, there is no visual difference in the results from the MELE and the MELA schemes, hence the results of the MELA scheme are not reported here and below.

In Figs. 6.10 and 6.11, we compare the results of all the proposed first and second order schemes on the coarse mesh $h = \frac{1}{50}$. Particularly, in Fig. 6.10, we plot the temporal configuration of $E$ at $x = 3ct_0$, and in Fig. 6.11, we plot the spatial configuration of $E$ at the final time. Second order schemes overall outperform the first order schemes. Compared with the MELE (hence MELA) scheme, the results of the META schemes display less sharp features. The META-up scheme produce relatively more dissipative results. As commented in Remark 6.1, these trapezoidal type methods use larger time step size than the leapfrog type methods yet computational more costly. The resolution of the numerical solution can be improved if the time step is further reduced at the price of an increased computational cost. Among the first order schemes, the semi-implicit SI1-al performs the best. It captures the main structures of the solution. All the other first order schemes, including the SI1-up, FI1-al and FI1-up schemes are too dissipative to produce reasonable approximations on such a coarse mesh. Under the current setting, the MELE and the MELA scheme outperform the other schemes.

Finally, we want to demonstrate the third and fifth harmonic generation due to the cubic nonlinearity of the model. For this, we examine the numerical solutions from the MELA scheme in the frequency domain by applying the discrete Fourier transform to the computed electric field $E$ over 20 periods in time (that is, over $[0, T]$) at $x = 2ct_0$, $x = 3ct_0$ and $x = 4ct_0$, respectively. The results on a mesh with $h = \frac{1}{250}$ with $\varepsilon = 0.5, 0.1, 10^{-4}$ (with unit fs) are presented in Fig. 6.12. One can clearly observe the third and fifth harmonic generation that contributes to the formation of daughter pulses. In addition, the longer the relaxation time $\varepsilon$ is, the smaller magnitude of the daughter waves are. At $x = 4ct_0$, the frequency of the daughter wave is slightly smaller than 3 and 5. Similar observation was made in [17].




(a) $t$-$E$, MELE (b) $t$-$E$, META-al (c) $t$-$E$, META-up

(d) $t$-$E$, SI1-al (e) $t$-$E$, SI1-up (f) $t$-$E$, FI1-al

(g) $t$-$E$, FI1-up

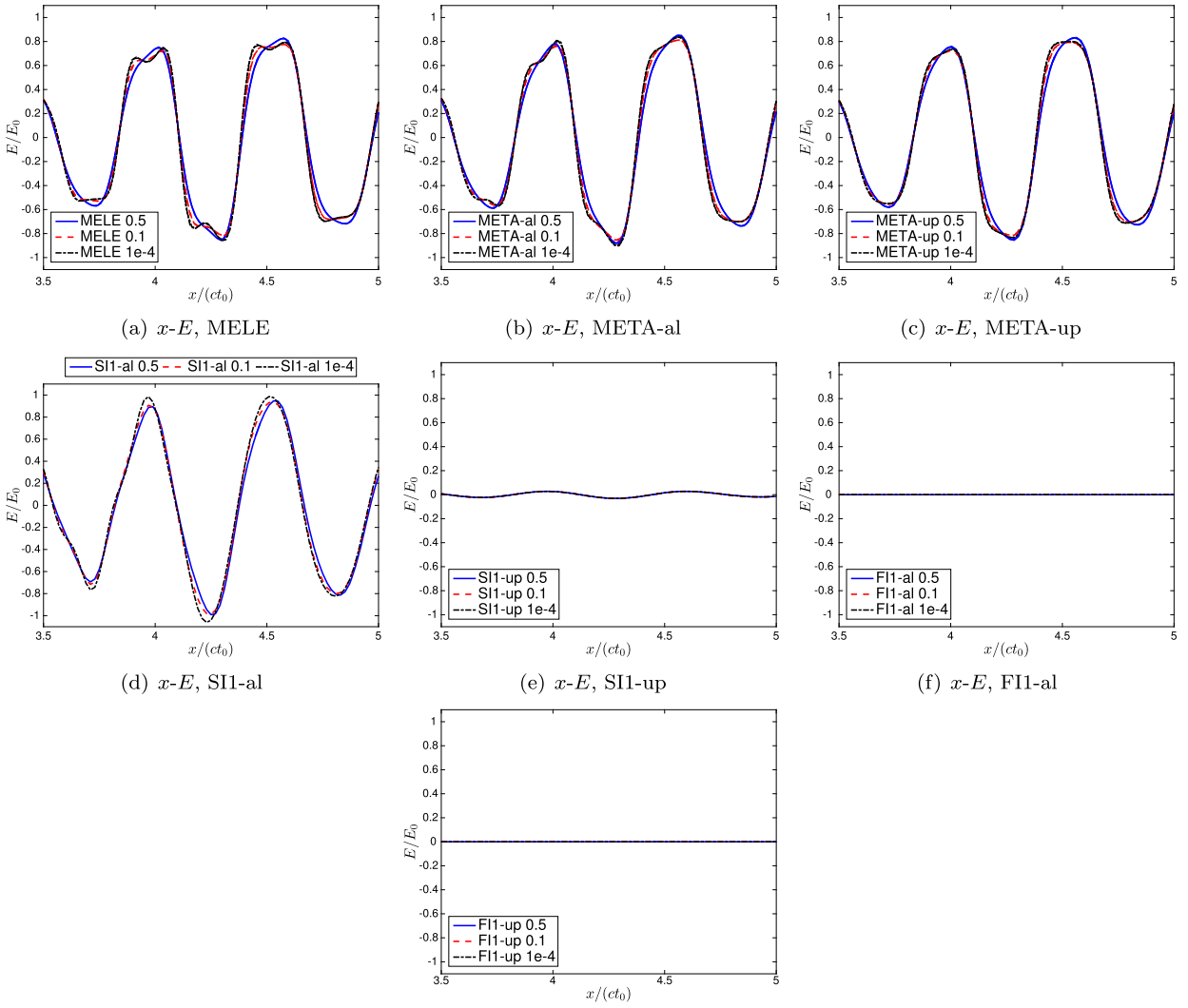
**Fig. 6.10.** The *temporal* configuration of the computed electric field $E$ by the proposed first and second order schemes for the example in Section 6.2.3, with $\varepsilon = 0.5,\ 0.1,\ 10^{-4}$ (fs) at $x = 3ct_0$ on a mesh with $h = \frac{1}{50}$.

In summary, for this example, the second order MELE and the MELA schemes perform the best among all proposed methods. The increase of the relaxation time for the Kerr nonlinearity reduces the self-steepening effect as well as the strength of the harmonic generation. For all the simulations we have conducted, no negative $\chi_{jk}$ is encountered.

## 7. Conclusions and future work

In this paper, we developed AP and PP schemes that are first and second order accurate in time for Maxwell's equations with Kerr-Debye and Lorentz effects in one dimension. For first order in time schemes, the methods are proved to be energy stable, AP and PP, but unsurprisingly, numerical experiments reveal that such methods are deemed too dissipative for wave propagation. For the second order in time methods, the key for the AP and PP properties rely on the numerical discretizations of the ADE, for which we designed an exponential time integrator and the AP and PP properties are preserved simultaneously. Combined with implicit or explicit time discretizations for the Maxwell part, and the algebraic or energy consistent way for the discretization of the constitutive law, we design and numerically compare a set of second order methods, including the META, MELA and MELE schemes as summarized in Table 5.1.

All three second order schemes are AP and PP. The implicit META scheme is unconditionally stable. However, smaller time steps may be preferred in practice for good resolution of numerical solutions. In addition, a global nonlinear system needs to be solved during each time step, and this makes the method computationally expensive. For the MELA and MELE schemes, the nonlinear equations are decoupled, and they are locally defined at each nodal point and can be solved much

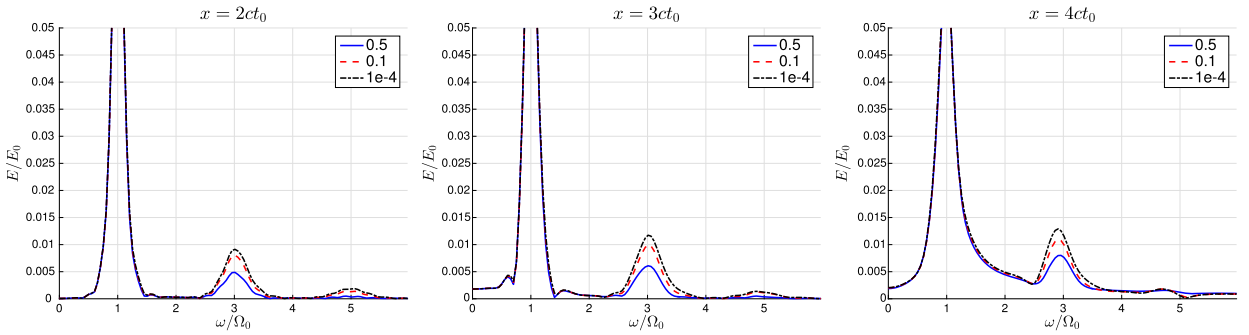



(a) $x$-$E$, MELE (b) $x$-$E$, META-al (c) $x$-$E$, META-up

(d) $x$-$E$, SI1-al (e) $x$-$E$, SI1-up (f) $x$-$E$, FI1-al

(g) $x$-$E$, FI1-up

**Fig. 6.11.** The *spatial* configuration of the computed electric field $E$ by the proposed first and second order schemes for the example in Section 6.2.3, with $\varepsilon = 0.5,\ 0.1,\ 10^{-4}$ (fs) on a mesh with $h = \frac{1}{50}$ at the final time $T$.




(a) $\omega$-$E$ (b) $\omega$-$E$ (c) $\omega$-$E$

**Fig. 6.12.** The computed electric field $E$ by the MELA scheme in the *frequency domain* for the example in Section 6.2.3, with $\varepsilon = 0.5,\ 0.1,\ 10^{-4}$ (fs) and at $x = 2ct_0, 3ct_0, 4ct_0$ on a mesh with $h = \frac{1}{250}$.

faster. The MELE scheme is the only one that has provable energy stability, at a price of possibly requiring special initial guesses for the nonlinear Newton's solver. The numerical performances of the MELE and the MELA scheme are overall comparable. With the consideration of the computational efficiency and the resolution of the numerical solutions, we will recommend the MELA and the MELE schemes over the META (surely also over the first order schemes), with the MELA method not relying on special initial guesses in its Newton's solver.

We want to emphasize that when the linear Lorentz effect is absent, the model considered here becomes the Kerr-Debye system, which is a hyperbolic system with the stiff relaxation. In this regime, as $\varepsilon \to 0$, one gets the limiting Kerr system, a nonlinear hyperbolic system that supports shock structures. The interested readers can refer to [22] for some second order AP and PP discontinuous Galerkin methods for the Kerr-Debye model, and [17] for some numerical studies of the model with or without the linear Lorentz dispersion. Our future endeavors will include extensions to higher dimensions, to other spatial discretizations, to nonreflecting boundary treatments for dispersive nonlinear optical models, and possibly also to higher order temporal discretizations.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. The proof of Theorem 4.2

**Proof.** From Theorem 4.1, and the fact of $\chi$ being non-negative initially, we know $\chi_{jk}^n \geq 0, \forall j,k,n$. This, together with Lemma 3.1, implies

$$\left(\chi_h^n, I_h((E_h^n)^2)\right) = \sum_{j=1}^{N_x}\sum_{k=0}^{K} \frac{\Delta x_j}{2}\widehat{\omega}_k \chi_{jk}^n (E_{jk}^n)^2 \geq 0, \tag{A.66}$$

hence $\mathcal{E}_{h,1}^n \geq 0$.

To perform energy stability analysis, we first consider the semi-implicit scheme (4.14). Take $\phi = H_h^{n+1}$ in (4.14a), $\varphi = E_h^{n+1}$ in (4.14b), and sum up in $j$, we have

$$\frac{\mu_0}{2}\left(||H_h^{n+1}||^2 - ||H_h^n||^2 + ||H_h^{n+1} - H_h^n||^2\right) + \left(D_h^{n+1} - D_h^n, E_h^{n+1}\right) = \Delta t\left(a_h(E_h^n, H_h^{n+1}) + b_h(H_h^n, E_h^{n+1})\right), \tag{A.67}$$

where

$$a_h(E_h,\phi) = \sum_{j=1}^{N_x}\left(-\int_{I_j} E_h \partial_x \phi dx + (\widehat{E}_h)_{j+\frac{1}{2}}\phi_{j+\frac{1}{2}}^- - (\widehat{E}_h)_{j-\frac{1}{2}}\phi_{j-\frac{1}{2}}^+\right), \tag{A.68a}$$

$$b_h(H_h,\varphi) = \sum_{j=1}^{N_x}\left(-\int_{I_j} H_h \partial_x \varphi dx + (\widetilde{H}_h)_{j+\frac{1}{2}}\varphi_{j+\frac{1}{2}}^- - (\widetilde{H}_h)_{j-\frac{1}{2}}\varphi_{j-\frac{1}{2}}^+\right). \tag{A.68b}$$

Based on (4.14c), one obtains

$$\begin{aligned}\left(D_h^{n+1} - D_h^n, E_h^{n+1}\right) = \frac{\epsilon_0\epsilon_\infty}{2}\left(||E_h^{n+1}||^2 - ||E_h^n||^2 + ||E_h^{n+1} - E_h^n||^2\right) \\ + \epsilon_0 a\left(\left(I_h(\chi_h^{n+1}E_h^{n+1}), E_h^{n+1}\right) - \left(I_h(\chi_h^n E_h^n), E_h^{n+1}\right)\right) + \epsilon_0\left(P_h^{n+1} - P_h^n, E_h^{n+1}\right).\end{aligned} \tag{A.69}$$

Multiply $\frac{1}{2}\left(\chi_h^{n+1} - I_h((E_h^{n+1})^2)\right)$ to (4.14d), multiply $P_h^{n+1}$ to (4.14e), and integrate in $x$, we get

$$\frac{1}{4}\left(||\chi_h^{n+1}||^2 - ||\chi_h^n||^2 + ||\chi_h^{n+1} - \chi_h^n||^2\right) - \frac{1}{2}\left(\chi_h^{n+1} - \chi_h^n, I_h\left((E_h^{n+1})^2\right)\right) = -\frac{\Delta t}{2\varepsilon}||\chi_h^{n+1} - I_h((E_h^{n+1})^2)||^2, \tag{A.70}$$

$$\frac{1}{2}\left(||P_h^{n+1}||^2 - ||P_h^n||^2 + ||P_h^{n+1} - P_h^n||^2\right) = \Delta t\left(J_h^{n+1}, P_h^{n+1}\right). \tag{A.71}$$

Multiply $\frac{1}{\omega_p^2}J_h^{n+1}$ to (4.14f) and integrate in $x$, and also use (4.14e), we have

$$\begin{aligned}\frac{1}{2\omega_p^2}\left(||J_h^{n+1}||^2 - ||J_h^n||^2 + ||J_h^{n+1} - J_h^n||^2\right) + \frac{\Delta t}{\omega_p^2\tau}||J_h^{n+1}||^2 + \frac{\omega_0^2}{\omega_p^2}\Delta t\left(J_h^{n+1}, P_h^{n+1}\right) \\ = \Delta t\left(J_h^{n+1}, E_h^{n+1}\right) = \left(P_h^{n+1} - P_h^n, E_h^{n+1}\right).\end{aligned} \tag{A.72}$$

Plug (A.71) into (A.72) to eliminate $\left(J_h^{n+1}, P_h^{n+1}\right)$, and plug (A.72) into (A.69) to eliminate $\left(P_h^{n+1}-P_h^n, E_h^{n+1}\right)$. Combine the resulting (A.69) with (A.67) and (A.70), one gets

$$\begin{aligned}&\widetilde{\mathcal{E}}_h^{n+1}-\widetilde{\mathcal{E}}_h^n+\Lambda_T+\epsilon_0 a\Lambda_\chi\\&=\Delta t\big(a_h(E_h^n,H_h^{n+1})+b_h(H_h^n,E_h^{n+1})\big)-\frac{\epsilon_0 a\Delta t}{2\varepsilon}||\chi_h^{n+1}-I_h\big((E_h^{n+1})^2\big)||^2-\frac{\epsilon_0\Delta t}{\omega_p^2\tau}||J_h^{n+1}||^2,\end{aligned}\tag{A.73}$$

where

$$\widetilde{\mathcal{E}}_h^n=\frac{1}{2}\Big(\mu_0||H_h^n||^2+\epsilon_0\epsilon_\infty||E_h^n||^2+\frac{\epsilon_0 a}{2}||\chi_h^n||^2+\frac{\epsilon_0\omega_0^2}{\omega_p^2}||P_h^n||^2+\frac{\epsilon_0}{\omega_p^2}||J_h^n||^2\Big),\tag{A.74}$$

$$\begin{aligned}\Lambda_T=\frac{1}{2}\Big(&\mu_0||H_h^{n+1}-H_h^n||^2+\epsilon_0\epsilon_\infty||E_h^{n+1}-E_h^n||^2+\frac{\epsilon_0 a}{2}||\chi_h^{n+1}-\chi_h^n||^2\\&+\frac{\epsilon_0\omega_0^2}{\omega_p^2}||P_h^{n+1}-P_h^n||^2+\frac{\epsilon_0}{\omega_p^2}||J_h^{n+1}-J_h^n||^2\Big),\end{aligned}\tag{A.75}$$

$$\Lambda_\chi=\Big(I_h(\chi_h^{n+1}E_h^{n+1}),E_h^{n+1}\Big)-\Big(I_h(\chi_h^nE_h^n),E_h^{n+1}\Big)-\frac{1}{2}\Big(\chi_h^{n+1}-\chi_h^n,I_h\big((E_h^{n+1})^2\big)\Big).\tag{A.76}$$

Based on Lemma 3.2, $\Lambda_\chi$ can be further simplified as

$$\begin{aligned}\Lambda_\chi&=\Big(\chi_h^{n+1},I_h\big((E_h^{n+1})^2\big)\Big)-\Big(\chi_h^n,I_h\big(E_h^nE_h^{n+1}\big)\Big)-\frac{1}{2}\Big(\chi_h^{n+1}-\chi_h^n,I_h\big((E_h^{n+1})^2\big)\Big)\\&=\frac{1}{2}\Big(\big(\chi_h^{n+1},I_h\left((E_h^{n+1})^2\right)\big)-\big(\chi_h^n,I_h\left((E_h^n)^2\right)\big)\Big)+\frac{1}{2}\left(\chi_h^n,I_h\left(\left(E_h^{n+1}-E_h^n\right)^2\right)\right).\end{aligned}\tag{A.77}$$

Substitute (A.77) into (A.73), we obtain

$$\begin{aligned}&\mathcal{E}_{h,1}^{n+1}-\mathcal{E}_{h,1}^n+\Lambda_T+\frac{\epsilon_0 a}{2}\left(\chi_h^n,I_h\left(\left(E_h^{n+1}-E_h^n\right)^2\right)\right)\\&=\Delta t\big(a_h(E_h^n,H_h^{n+1})+b_h(H_h^n,E_h^{n+1})\big)-\frac{\epsilon_0 a\Delta t}{2\varepsilon}||\chi_h^{n+1}-I_h\big((E_h^{n+1})^2\big)||^2-\frac{\epsilon_0\Delta t}{\omega_p^2\tau}||J_h^{n+1}||^2.\end{aligned}\tag{A.78}$$

Similar as for (A.66), we can show $\left(\chi_h^n,I_h\left(\left(E_h^{n+1}-E_h^n\right)^2\right)\right)\geq 0$, and (A.78) will lead to

$$\mathcal{E}_{h,1}^{n+1}-\mathcal{E}_{h,1}^n\leq\Delta t\big(a_h(E_h^n,H_h^{n+1})+b_h(H_h^n,E_h^{n+1})\big)-\Lambda_T.\tag{A.79}$$

What remained is to estimate $a_h(E_h^n,H_h^{n+1})+b_h(H_h^n,E_h^{n+1})$, which depends on the choice of numerical fluxes and can be rewritten as

$$a_h(E_h^n,H_h^{n+1})+b_h(H_h^n,E_h^{n+1})=a_h(E_h^n,H_h^n)+b_h(H_h^n,E_h^n)+a_h(E_h^n,H_h^{n+1}-H_h^n)+b_h(H_h^n,E_h^{n+1}-E_h^n).\tag{A.80}$$

It can be verified that

$$a_h(E_h^n,H_h^n)+b_h(H_h^n,E_h^n)=\begin{cases}0,\quad\text{with alternating flux,}\\-\frac{1}{2}\sum_{j=1}^{N_x}\left(\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_\infty}}[H_h^n]^2+\sqrt{\frac{\epsilon_0\epsilon_\infty}{\mu_0}}[E_h^n]^2\right),\quad\text{with upwind flux.}\end{cases}\tag{A.81}$$

**Case 1: alternating flux.** Using inverse inequalities, Young's inequality and the quasi-uniform assumption for meshes, one gets

$$\begin{aligned}\Delta t|a_h(E_h^n,H_h^{n+1}-H_h^n)|&=\Delta t\Big|\sum_j\int_{I_j}E_h^n\partial_x(H_h^{n+1}-H_h^n)dx-\widehat{(E_h^n)}_{j+\frac{1}{2}}(H_h^{n+1}-H_h^n)^-_{j+\frac{1}{2}}+\widehat{(E_h^n)}_{j-\frac{1}{2}}(H_h^{n+1}-H_h^n)^+_{j-\frac{1}{2}}\Big|\\&\leq\begin{cases}\left(\frac{4C_{inv}^2\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2}\right)\frac{\epsilon_0\epsilon_\infty}{2}||E_h^n||^2+\frac{\mu_0}{2}||H_h^{n+1}-H_h^n||^2,&\text{if}\quad K=0\\\left(\frac{2(\widetilde{C}_{inv}+4C_{inv}^2)\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2}\right)\frac{\epsilon_0\epsilon_\infty}{2}||E_h^n||^2+\frac{\mu_0}{2}||H_h^{n+1}-H_h^n||^2,&\text{if}\quad K\geq 1\end{cases}\qquad(A.82)\\&=\left(\frac{C_K\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2}\right)\frac{\epsilon_0\epsilon_\infty}{2}||E_h^n||^2+\frac{\mu_0}{2}||H_h^{n+1}-H_h^n||^2.\qquad(A.83)\end{aligned}$$

Here $C_K$ is a constant dependent of $K$. Similarly,

$$\Delta t|b_h(H_h^n, E_h^{n+1} - E_h^n)| \leq \Big(\frac{C_K\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2}\Big)\frac{\mu_0}{2}||H_h^n||^2 + \frac{\epsilon_0\epsilon_\infty}{2}||E_h^{n+1} - E_h^n||^2. \tag{A.84}$$

With these estimates and (A.79), we have

$$\begin{aligned}&\Delta t|a_h(E_h^n, H_h^{n+1}) + b_h(H_h^n, E_h^{n+1})| - \Lambda_T\\ &\leq \Big(\frac{C_K\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2}\Big)(\frac{\epsilon_0\epsilon_\infty}{2}||E_h^n||^2 + \frac{\mu_0}{2}||H_h^n||^2) \leq \Big(\frac{C_K\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2}\Big)\mathcal{E}_{h,1}^n,\end{aligned} \tag{A.85}$$

and subsequently

$$\mathcal{E}_{h,1}^{n+1} \leq (1 + \frac{C_K\delta^2\Delta t^2}{\mu_0\epsilon_0\epsilon_\infty h^2})\mathcal{E}_{h,1}^n. \tag{A.86}$$

By requiring $\frac{\Delta t}{h^2}$ to be bounded by any generic constant $C$, we will have $\mathcal{E}_{h,1}^{n+1} \leq (1 + C'\Delta t)\mathcal{E}_{h,1}^n$, with $C'$ depends on $K, \mu_0, \epsilon_0, \epsilon_\infty$ and $\delta$, hence we get the energy stability $\mathcal{E}_{h,1}^n \leq e^{C'T}\mathcal{E}_{h,1}^0, \forall n \leq T/\Delta t$.

**Case 2: upwind flux.** With the upwind flux and apply integration by parts on each $I_j$, we get

$$\begin{aligned}&a_h(E_h^n, H_h^{n+1} - H_h^n) + b_h(H_h^n, E_h^{n+1} - E_h^n)\\ &= \sum_{j=1}^{N_x}\left(\int_{I_j}(\partial_x E_h^n)(H_h^{n+1} - H_h^n)dx + \Big([E_h^n]\{H_h^{n+1} - H_h^n\}\Big)_{j+\frac{1}{2}}\right)\\ &\quad + \sum_{j=1}^{N_x}\left(\int_{I_j}(\partial_x H_h^n)(E_h^{n+1} - E_h^n)dx + \Big([H_h^n]\{E_h^{n+1} - E_h^n\}\Big)_{j+\frac{1}{2}}\right)\\ &\quad + \frac{1}{2}\sum_{j=1}^{N_x}\left(\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_\infty}}[H_h^n][H_h^{n+1} - H_h^n] + \sqrt{\frac{\epsilon_0\epsilon_\infty}{\mu_0}}[E_h^n][E_h^{n+1} - E_h^n]_{j+\frac{1}{2}}\right).\end{aligned} \tag{A.87}$$

When $K = 0$, $\partial_x E_h^n = \partial_x H_h^n = 0$. Using the inverse inequality (3.13a) and Young's inequality, we reach

$$\begin{aligned}&|a_h(E_h^n, H_h^{n+1} - H_h^n) + b_h(H_h^n, E_h^{n+1} - E_h^n)|\\ &\leq \frac{1}{2}\sum_{j=1}^{N_x}\left(\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_\infty}}[H_h^n]^2 + \sqrt{\frac{\epsilon_0\epsilon_\infty}{\mu_0}}[E_h^n]^2\right)_{j+\frac{1}{2}} + \frac{2C_{inv}\delta}{h}\left(\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_\infty}}||H_h^{n+1} - H_h^n||^2 + \sqrt{\frac{\epsilon_0\epsilon_\infty}{\mu_0}}||E_h^{n+1} - E_h^n||^2\right).\end{aligned} \tag{A.88}$$

Combine (A.81), (A.88) and (A.79), we have

$$\begin{aligned}\mathcal{E}_{h,1}^{n+1} - \mathcal{E}_{h,1}^n &\leq \Delta t|a_h(E_h^n, H_h^{n+1}) + b_h(H_h^n, E_h^{n+1})| - \Lambda_T\\ &\leq \frac{2C_{inv}\delta\Delta t}{h}\left(\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_\infty}}||H_h^{n+1} - H_h^n||^2 + \sqrt{\frac{\epsilon_0\epsilon_\infty}{\mu_0}}||E_h^{n+1} - E_h^n||^2\right) - \Lambda_T.\end{aligned} \tag{A.89}$$

We now require the time step $\Delta t$ satisfies $\Delta t \leq \frac{\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{4C_{inv}\delta}h$, to ensure a non-negative right hand side in (A.89), hence the energy stability, $\mathcal{E}_{h,1}^{n+1} \leq \mathcal{E}_{h,1}^n$. Particularly, the time step condition is $\Delta t \leq \frac{\sqrt{\mu_0\epsilon_0\epsilon_\infty}}{4\delta}h$, if one recalls $C_{inv} = 1$ when $K = 0$. In general when $K \geq 1$, using inverse inequalities and Young's inequality, one can establish similar results as in Case 1.

For the fully implicit scheme (4.15), the only difference comes from analyzing the PDE part. More specifically, the equation (A.67) becomes

$$\frac{\mu_0}{2}\Big(||H_h^{n+1}||^2 - ||H_h^n||^2 + ||H_h^{n+1} - H_h^n||^2\Big) + \Big(D_h^{n+1} - D_h^n, E_h^{n+1}\Big) = \Delta t\Big(a_h(E_h^{n+1}, H_h^{n+1}) + b_h(H_h^{n+1}, E_h^{n+1})\Big), \tag{A.90}$$

and (A.79) is replaced by

$$\mathcal{E}_{h,1}^{n+1} - \mathcal{E}_{h,1}^n \leq \Delta t\big(a_h(E_h^{n+1}, H_h^{n+1}) + b_h(H_h^{n+1}, E_h^{n+1})\big) - \Lambda_T \leq 0. \tag{A.91}$$

This energy stability holds for any $\Delta t > 0$, and the method is therefore unconditionally stable. □

## References


[1] Govind P. Agrawal, Nonlinear fiber optics, in: Nonlinear Science at the Dawn of the 21st Century, Springer, 2000, pp. 195–211.
[2] Aregba-Driollet Denise, Godunov scheme for Maxwell's equations with Kerr nonlinearity, Commun. Math. Sci. 13 (8) (2015) 2195–2222.
[3] Nicolaas Bloembergen, Nonlinear Optics, World Scientific, 1996.
[4] Vrushali A. Bokil, Yingda Cheng, Yan Jiang, Fengyan Li, Energy stable discontinuous Galerkin methods for Maxwell's equations in nonlinear optical media, J. Comput. Phys. 350 (2017) 420–452.
[5] Vrushali A. Bokil, Yingda Cheng, Yan Jiang, Fengyan Li, Puttha Sakkaplangkul, High spatial order energy stable FDTD methods for Maxwell's equations in nonlinear optical media in one dimension, J. Sci. Comput. 77 (2018) 330–371.
[6] Robert W. Boyd, Nonlinear Optics, Academic Press, 2003.
[7] Alina Chertock, Shumo Cui, Alexander Kurganov, Tong Wu, Steady state and sign preserving semi-implicit Runge–Kutta methods for ODEs with stiff damping term, SIAM J. Numer. Anal. 53 (4) (2015) 2008–2029.
[8] Bernardo Cockburn, John R. Singler, Yangwen Zhang, Interpolatory HDG method for parabolic semilinear PDEs, J. Sci. Comput. (2018) 1–24.
[9] Anaïs Crestetto, Nicolas Crouseilles, Mohammed Lemou, A particle micro-macro decomposition based numerical scheme for collisional kinetic equations in the diffusion scaling, preprint, arXiv:1701.05069, 2017.
[10] Armel de La Bourdonnaye, High-order scheme for a nonlinear Maxwell system modelling Kerr effect, J. Comput. Phys. 160 (2) (2000) 500–521.
[11] Pierre Degond, Asymptotic-preserving schemes for fluid models of plasmas, preprint, arXiv:1104.1869, 2011.
[12] Pierre Degond, Fabrice Deluzet, Asymptotic-preserving methods and multiscale models for plasma physics, J. Comput. Phys. 336 (2017) 429–457.
[13] Giacomo Dimarco, Lorenzo Pareschi, Exponential Runge–Kutta methods for stiff kinetic equations, SIAM J. Numer. Anal. 49 (5) (2011) 2057–2077.
[14] Jim Douglas, Todd Dupont, The effect of interpolating the coefficients in nonlinear parabolic Galerkin procedures, Math. Comput. 29 (130) (1975) 360–389.
[15] Aaron Fisher, D. White, G. Rodrigue, An efficient vector finite element method for nonlinear electromagnetic modeling, J. Comput. Phys. 225 (2) (2007) 1331–1346.
[16] L. Gilles, S.C. Hagness, L. Vázquez, Comparison between staggered and unstaggered finite-difference time-domain grids for few-cycle temporal optical soliton propagation, J. Comput. Phys. 161 (2) (2000) 379–400.
[17] L. Gilles, J.V. Moloney, L. Vázquez, Electromagnetic shocks on the optical cycle of ultrashort pulses in triple-resonance Lorentz dielectric media with subfemtosecond nonlinear electronic Debye relaxation, Phys. Rev. E 60 (1) (1999) 1051.
[18] Sigal Gottlieb, Chi-Wang Shu, Eitan Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev. 43 (1) (2001) 89–112.
[19] Cheryl V. Hile, William L. Kath, Numerical solutions of Maxwell's equations for nonlinear-optical pulse propagation, J. Opt. Soc. Am. B, Opt. Phys. 13 (6) (1996) 1135–1145.
[20] Jingwei Hu, Ruiwen Shu, A second-order asymptotic-preserving and positivity-preserving exponential Runge-Kutta method for a class of stiff kinetic equations, preprint, arXiv:1807.03728, 2018.
[21] Jingwei Hu, Ruiwen Shu, Xiangxiong Zhang, Asymptotic-preserving and positivity-preserving implicit-explicit schemes for the stiff BGK equation, SIAM J. Numer. Anal. 56 (2) (2018) 942–973.
[22] Juntao Huang, Chi-Wang Shu, A second-order asymptotic-preserving and positivity-preserving discontinuous Galerkin scheme for the Kerr–Debye model, Math. Models Methods Appl. Sci. 27 (03) (2017) 549–579.
[23] Shi Jin, Asymptotic preserving (ap) schemes for multiscale kinetic and hyperbolic equations: a review, in: Lecture Notes for Summer School on Methods and Models of Kinetic Theory, M&MKT, Porto Ercole, Grosseto, Italy, 2010, pp. 177–216.
[24] P. Kinsler, S.B.P. Radnor, J.C.A. Tyrrell, G.H.C. New, Optical carrier wave shocking: detection and dispersion, Phys. Rev. E 75 (6) (2007) 066603.
[25] Dana A. Knoll, David E. Keyes, Jacobian-free Newton–Krylov methods: a survey of approaches and applications, J. Comput. Phys. 193 (2) (2004) 357–397.
[26] Geoffrey New, Introduction to Nonlinear Optics, Cambridge University Press, 2011.
[27] Zhichao Peng, Yingda Cheng, Jing-Mei Qiu, Fengyan Li, Stability-enhanced AP IMEX-LDG schemes for linear kinetic transport equations under a diffusive scaling, 2019.
[28] Mads Peter Sørensen, Garry M. Webb, Moysey Brio, Jerome V. Moloney, Kink shape solutions of the Maxwell-Lorentz system, Phys. Rev. E 71 (3) (2005) 036602.
[29] Allan G. Taylor, et al., User Documentation for KINSOL, a Nonlinear Solver for Sequential and Parallel Computers, Technical report, Lawrence Livermore National Lab., CA, United States, 1998.
[30] Richard W. Ziolkowski, Justin B. Judkins, Full-wave vector Maxwell equation modeling of the self-focusing of ultrashort optical pulses in a nonlinear Kerr medium exhibiting a finite response time, J. Opt. Soc. Am. B, Opt. Phys. 10 (2) (1993) 186–198.