

# Computational electrodynamics in material media with constraint-preservation, multidimensional Riemann solvers and sub-cell resolution – Part II, higher order FVTD schemes



Dinshaw S. Balsara<sup>a,\*</sup>, Sudip Garain<sup>a</sup>, Allen Taflove<sup>b</sup>, Gino Montecinos<sup>c</sup>

<sup>a</sup> University of Notre Dame, United States

<sup>b</sup> Northwestern University, United States

<sup>c</sup> University of Chile, Chile

## ARTICLE INFO

### Article history:

Received 10 July 2017

Received in revised form 9 October 2017

Accepted 10 October 2017

Available online 25 October 2017

### Keywords:

Computational electrodynamics

Involution constraint

Maxwell equations

Godunov schemes

Higher order

Multidimensional Riemann solvers

## ABSTRACT

The Finite Difference Time Domain (FDTD) scheme has served the computational electrodynamics community very well and part of its success stems from its ability to satisfy the constraints in Maxwell's equations. Even so, in the previous paper of this series we were able to present a second order accurate Godunov scheme for computational electrodynamics (CED) which satisfied all the same constraints and simultaneously retained all the traditional advantages of Godunov schemes. In this paper we extend the Finite Volume Time Domain (FVTD) schemes for CED in material media to better than second order of accuracy.

From the FDTD method, we retain a somewhat modified staggering strategy of primal variables which enables a very beneficial constraint-preservation for the electric displacement and magnetic induction vector fields. This is accomplished with constraint-preserving reconstruction methods which are extended in this paper to third and fourth orders of accuracy. The idea of one-dimensional upwinding from Godunov schemes has to be significantly modified to use the multidimensionally upwinded Riemann solvers developed by the first author. In this paper, we show how they can be used within the context of a higher order scheme for CED.

We also report on advances in timestepping. We show how Runge–Kutta IMEX schemes can be adapted to CED even in the presence of stiff source terms brought on by large conductivities as well as strong spatial variations in permittivity and permeability. We also formulate very efficient ADER timestepping strategies to endow our method with sub-cell resolving capabilities. As a result, our method can be stiffly-stable and resolve significant sub-cell variation in the material properties within a zone. Moreover, we present ADER schemes that are applicable to all hyperbolic PDEs with stiff source terms and at all orders of accuracy. Our new ADER formulation offers a treatment of stiff source terms that is much more efficient than previous ADER schemes. The computer algebra system scripts for generating ADER time update schemes for any general PDE with stiff source terms are also given in the electronic supplements to this paper.

Second, third and fourth order accurate schemes for numerically solving Maxwell's equations in material media are presented in this paper. Several stringent tests are also presented to show that the method works and meets its design goals even when material permittivity and permeability vary by an order of magnitude over just a few zones. Furthermore, since the method is unconditionally stable and sub-cell-resolving in the

\* Corresponding author.

E-mail addresses: [dbalsara@nd.edu](mailto:dbalsara@nd.edu) (D.S. Balsara), [sgarain@nd.edu](mailto:sgarain@nd.edu) (S. Garain), [taflove@eecs.northwestern.edu](mailto:taflove@eecs.northwestern.edu) (A. Taflove), [gmontecinos@dim.uchile.cl](mailto:gmontecinos@dim.uchile.cl) (G. Montecinos).

presence of stiff source terms (i.e. for problems involving giant variations in conductivity over just a few zones), it can accurately handle such problems without any reduction in timestep. We also show that increasing the order of accuracy offers distinct advantages for resolving sub-cell variations in material properties. Most importantly, we show that when the accuracy requirements are stringent the higher order schemes offer the shortest time to solution. This makes a compelling case for the use of higher order, sub-cell resolving schemes in CED.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

The numerical solution of Maxwell's equations is crucial to numerous computational electrodynamics (CED) applications in science and engineering. The Finite-difference time-domain (FDTD) method (Yee [52], Taflove and Hagness [43], Taflove, Oskooi and Johnson [44]) has been the method of choice for CED applications for more than fifty years! The reason for the durability of FDTD is that it incorporates many of the features that are intrinsic to the physics of Maxwell's equations. These desirable features include a direct interpretation of the two curl-type equations given by Faraday's and extended Ampere's laws and a natural satisfaction of the constraint equations given by Gauss's law for electric charge and the absence of magnetic monopoles. This is achieved by staggering the electric field and magnetic field components. On a simple Cartesian mesh, every electric field component is surrounded by four circulating magnetic field components and every magnetic field component is surrounded by four circulating electric field components. It is therefore essential that all improved methods for the numerical solution of Maxwell's equations should retain the beneficial aspects of the FDTD method.

A mathematical analysis of Maxwell's equations (Balsara et al. [19]) shows that they consist of an involution-constrained hyperbolic system. Powerful methods have been developed in the last three and a half decades for the high accuracy treatment of hyperbolic systems. Such methods go under the rubric of higher order Godunov methods and they use a zone-centered formulation. Formulations that treat Maxwell's equations with zone-centered variables have been tried (Munz et al. [34], Hesthaven and Warburton [28], Ismagilov [30], Barbas and Velarde [21]) and references therein. The inability to reconcile the different collocations that are used in FDTD methods with higher order Godunov methods has been an impediment to progress. FDTD methods have been unable to benefit from advances in higher order Godunov methodology and higher order Godunov methods have been unable to deliver on some of the nice constraint-preserving properties of the FDTD method. In Balsara et al. [16], [19] steps were taken to broach that divide with a path-breaking synthesis. However, the methods developed in Balsara et al. [19] focused on achieving such a synthesis for second order schemes that solve Maxwell's equations in material media. It is usually challenging for schemes to transition from second order accuracy to accuracies of third and higher orders. In this paper we face up to that challenge by devising constraint-preserving higher order Godunov methods that solve Maxwell's equations in material media with better than second order of accuracy. The methods that we present in this paper synthesize all the strengths of FDTD methods (Taflove [42], Taflove and Hagness [45]) along with all the strengths of higher order Godunov methodology (van Leer [50], [51]) whilst offering the significant advantages of higher order accuracy. These methods are inspired by our prior work on divergence-free numerical MHD (Balsara and Spicer [1], Balsara [2], [3], [4], [7], [11], Balsara and Nkonga [20]). A von Neumann stability analysis of divergence-free MHD schemes is available in Balsara and Käppeli [18] and an analogous von Neumann analysis of FVTD schemes for CED is on the way.

The motivations for going to higher order of accuracy are quite compelling. Higher order schemes provide better phase accuracy and the dissipation is well-controlled and restricted to waves that have wavelength comparable to the mesh size. Retaining accuracy at adaptive mesh interfaces is also easier when the base-level scheme is better than second order. Furthermore, we show that higher order schemes are better at retaining the order of accuracy in regions where the material properties undergo a substantial variation.

The methods presented here are inspired by higher order Godunov schemes in their design philosophy. They use ideas of higher order non-oscillatory reconstruction and upwinding via Riemann solvers that have made the higher order Godunov schemes so popular. However, they also retain a staggering of variables that is inspired by the Yee mesh from FDTD methods. Two special advances go into enabling this synthesis. First, we use a very special form of constraint-preserving reconstruction. Some aspects of this reconstruction were initially developed in the literature for magnetohydrodynamic (MHD) simulations (Balsara [2], [3], [4], Balsara and Dumbser [12], Xu et al. [53], Balsara et al. [16]), but here we import these ideas to computational electrodynamics. The second order extensions to CED were described in Balsara et al. [19]; the higher order extensions were first cataloged in Balsara et al. (2016) so we only provide a brief recapitulation here for the sake of completeness. Second, we draw on the multidimensional Riemann solver technology that has been recently developed (Balsara [7], [8], [11], [14], Balsara, Dumbser and Abgrall [10], Balsara and Dumbser [13], Balsara et al. [15] Balsara and Nkonga [20]). This enables us to get the correct multidimensional upwinding of Maxwell's equations on a Yee-type mesh. The multidimensional Riemann solver for CED was thoroughly described in Section V of Balsara et al. [19] and we do not repeat that description here.

Higher order spatial reconstruction should be matched with higher order temporal evolution in order to have a numerical method that retains overall accuracy. We realize that a computational electrodynamics code has to be responsive to variations in material properties (i.e. the permittivity and permeability) that can change substantially over one or two zones. Furthermore, the source terms (i.e. the conductivity) can be stiff and can also vary in space. Functional Runge–Kutta schemes that can accommodate stiff source terms are only available up to third order of accuracy (Pareschi and Russo [36], Hunsdorfer and Ruuth [29], Kupka et al. [33]) and in Balsara et al. [16] we showed how these can be used very effectively for CED with second and third orders of temporal accuracy. But it is hard to go past third order accuracy with Runge–Kutta time stepping. ADER schemes are not similarly limited, and they extend seamlessly to fourth and higher orders of temporal accuracies. For that reason, we adapt ADER schemes for stiff source terms that were developed by Dumbser et al. [25], Balsara et al. [6], [9] and Dumbser et al. [26]. This method acts on nodal points within a zone. At each of those nodal points it permits variation in the values and the higher moments of the dielectric permittivity and magnetic permeability. Because the values of the material properties and all their higher derivatives are allowed to vary substantially within a zone, we can incorporate sub-cell variations in material properties into our timestepping strategy. The value of the conductivity as well as its higher moments can also vary at the nodal points. The source terms associated with the current are also treated stiffly, permitting an A-stable, sub-cell resolving time update strategy.

Section 2 describes the governing equations and their constraint-preserving numerical discretization. Section 3 briefly recapitulates the reconstruction strategy at third order in three-dimensions. Section 4 presents a very special ADER scheme at high order that is well-suited to numerical treatment of Maxwell's equations. Section 5 presents results. Section 6 analyses accuracy as a function of time to solution, making a very compelling case for the higher order schemes presented here. Section 7 provides conclusions. In the supplement to this paper we give more details associated with the ADER schemes that are described here. In the supplement to this paper we also show how we adapt the temporally second and third order Runge–Kutta methods with treatment of stiff source terms for this application.

## 2. Maxwell's equations and their constraint-preserving numerical discretization

### 2.1. Very brief description of Maxwell's equations

The equations of CED can be written as two evolutionary curl-type equations for the electric displacement (Ampere's law) and the magnetic induction (Faraday's law). The extended Ampere's law for the time evolution of the electric displacement is given by

$$\frac{\partial \mathbf{D}}{\partial t} - \nabla \times \mathbf{H} = -\mathbf{J} \quad (2.1)$$

where  $\mathbf{D}$  is the electric displacement (or electric flux density),  $\mathbf{H}$  is the magnetic field vector and  $\mathbf{J}$  is the electric current density. Faraday's law for the time evolution of the magnetic induction is given by

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = -\mathbf{M} \quad (2.2)$$

where  $\mathbf{B}$  is the magnetic induction (or magnetic flux density),  $\mathbf{E}$  is the electric field and  $\mathbf{M}$  is the magnetic current density. The electric displacement and the magnetic induction vector fields also satisfy the following two non-evolutionary constraint equations given by

$$\nabla \cdot \mathbf{D} = \rho_E \quad (2.3)$$

and

$$\nabla \cdot \mathbf{B} = \rho_M \quad (2.4)$$

where  $\rho_E$  and  $\rho_M$  are the electric and magnetic charge densities. In any physical medium  $\rho_M = 0$ ; however, the imposition of boundary conditions might require the use of non-zero magnetic current densities.

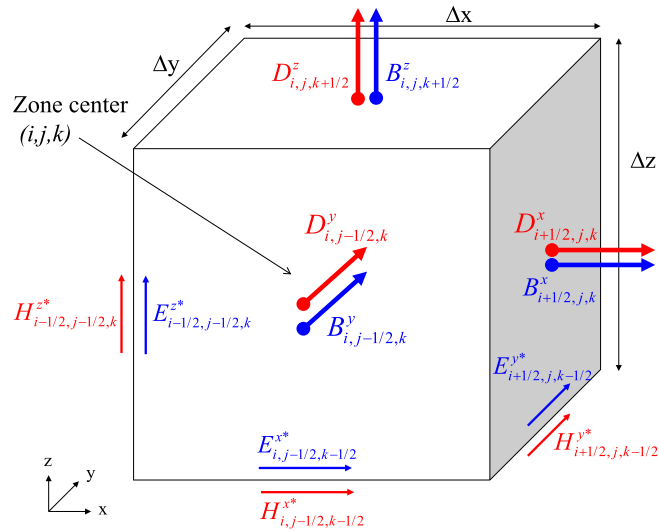
In material media, the electric displacement vector is also related to the electric field vector via a constitutive relation given by

$$\mathbf{D} = \boldsymbol{\epsilon} \mathbf{E} \quad (2.5)$$

where, in general,  $\boldsymbol{\epsilon}$  is a symmetric  $3 \times 3$  permittivity tensor that depends on material properties. Likewise, in material media, the magnetic induction vector is related to the magnetic field vector by

$$\mathbf{B} = \boldsymbol{\mu} \mathbf{H} \quad (2.6)$$

As before,  $\boldsymbol{\mu}$  is a symmetric  $3 \times 3$  magnetic permeability tensor that depends on material properties. We allow the permittivity and magnetic permeability tensors to have a general form. However, the analytical eigenstructure of the hyperbolic system is most easily found by making the simplifying assumption  $\boldsymbol{\epsilon} = \text{diag}\{\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}\}$  and  $\boldsymbol{\mu} = \text{diag}\{\mu_{xx}, \mu_{yy}, \mu_{zz}\}$ . We will also need the inverses of the permittivity and permeability tensors. These  $3 \times 3$  inverse matrices will also be symmetric and we denote them as  $\tilde{\boldsymbol{\epsilon}} = \boldsymbol{\epsilon}^{-1}$  and  $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu}^{-1}$ . The current density is related to the electric field via



**Fig. 1.** Shows us that the primal variables of the scheme, given by the normal components of the magnetic induction and electric field displacement, are facially-located. They undergo an update from Faraday's law and the generalized Ampere's law respectively. The components of the primal magnetic induction vector are shown by the thick blue arrows while the components of the primal electric displacement vector are shown by the thick red arrows. The edge-located electric fields, which are used for updating the facial magnetic induction components, are shown by the thin blue arrows close to the appropriate edge. The edge-located magnetic fields, which are used for updating the facial electric displacement components, are shown by the thin red arrows close to the appropriate edge. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\mathbf{J} = \sigma \mathbf{E} \quad (2.7)$$

where  $\sigma$  is the conductivity. Similarly, the magnetic current density is related to the magnetic field via

$$\mathbf{M} = \sigma^* \mathbf{H} \quad (2.8)$$

where  $\sigma^*$  is the equivalent magnetic loss, which is again zero in physical media, but may be non-zero when imposing boundary conditions in CED.

The CED equations can be written in a flux form, which we formally write as

$$\partial_t \mathbf{U} + \partial_x \mathbf{F} + \partial_y \mathbf{G} + \partial_z \mathbf{K} = \mathbf{S} \quad (2.9)$$

The vector  $\mathbf{U}$  is given by  $\mathbf{U} = (D_x, D_y, D_z, B_x, B_y, B_z)^T$ . The fluxes are explicitly cataloged in eqn. (2.12) of Balsara et al. [19] and the eigenstructure of this hyperbolic system is analyzed in Section II.2 of Balsara et al. [19]. We do not repeat details here.

## 2.2. Narrative description of the FVTD scheme at high order

In Balsara et al. [19], a solution strategy was presented for solving the coupled equations for Faraday's law and Ampere's law on the same control volume. Fig. 1, from Balsara et al. [19], is repeated here just to show the arrangement of primal variables. The electric displacement and magnetic induction vector fields are the primal variables of our scheme. The normal components of these two vector fields are defined at the faces of the control volume shown in Fig. 1. The facial averages of the normal components of the electric displacement vector field are shown by the thick red arrows in Fig. 1 and they constitute one set of primal variables for our scheme. The facial averages of the normal components of the magnetic induction vector field are shown by thick blue arrows and they constitute the other set of primal variables for our scheme. This arrangement of primal variables will also lead to a pair of Yee-type curl equations as long as the magnetic and electric fields can be obtained in some multidimensionally upwinded sense at the edges of the mesh shown in Fig. 1. The edge-averaged components of the magnetic vector field are shown by the thin red arrows in Fig. 1. They can be used for the update of the facially-averaged components of the electric displacement by using a discrete version of Ampere's law in eqn. (2.1). The edge-averaged components of the electric vector field are shown by the thin blue arrows in Fig. 1. They can be used for the update of the facially-averaged components of the magnetic induction by using a discrete version of Faraday's law in eqn. (2.2).

For a multidimensional Godunov scheme, the edge-averaged components of the magnetic and electric fields are obtained with the help of a multidimensional Riemann solver as shown in Section V of Balsara et al. [19]. Further description of multidimensional Riemann solvers is also available (Balsara [7], [8], [11], [14], Balsara, Dumbser and Abgrall [10], Balsara and Dumbser [13], Balsara et al. [15] Balsara and Nkonga [20]). For our purposes, we can treat the multidimensional Riemann solver as an algorithm that lives at the edges of the mesh. This edge-located algorithm takes as its inputs the electric

displacement and magnetic induction vectors from the four zones that abut the edge in question. As its outputs, this algorithm gives us multidimensionally upwinded magnetic and electric vectors at the same edge. Of course, we would like this Riemann solver to give outputs that are as accurate as possible. Happily, it turns out that if the inputs to the Riemann solver are higher order accurate, the outputs from that same Riemann solver will also be higher order accurate. In the next paragraph we describe how these high order inputs are provided to the Riemann solver.

Our strategy for achieving this high accuracy is called the constraint-preserving reconstruction problem. We start with the facial averages shown in Fig. 1. Using electric displacement and magnetic induction variables from neighboring faces, we can obtain higher order moments for the components that reside in the faces of Fig. 1. The constraint-preserving reconstruction can be thought of as yet another algorithm that lives at the center of each zone. It takes as its inputs the facially-averaged electric displacement and magnetic induction components and their higher order moments for any zone. As its output, it produces a constraint-preserving high order polynomial representation of the electric displacement and magnetic induction vector fields that are valid throughout the zone. The constraint-preserving reconstruction is such that it exactly satisfies the non-evolutionary constraints posed by eqns. (2.3) and (2.4). As the accuracy of the facial moments is increased, the algorithm for constraint-preserving reconstruction produces a higher order polynomial that is valid throughout the zone. Section III of Balsara et al. [19] shows how such a constraint-preserving reconstruction operates at second order of accuracy. Section 3 of this paper shows how such a constraint-preserving reconstruction algorithm operates at third order of accuracy. This can be formulated at all orders of accuracy and also on unstructured meshes (Balsara [2], [3], [4], Balsara and Dumbser [12], Xu et al. [53], Balsara et al. [16]). The purpose of the constraint-preserving reconstruction algorithm is to provide high accuracy input values for the multidimensional Riemann solver.

The constraint-preserving reconstruction and the multidimensional Riemann solver can give us a highly accurate set of edge-centered magnetic and electric field components. This enables a spatially accurate discretization of Ampere's and Faraday's laws. However, we desire temporal accuracy that matches the spatial accuracy in our scheme. Up to third order of temporal accuracy can be achieved in two possible ways. We can use Runge–Kutta methods that include the treatment of stiff source terms (Pareschi and Russo [36], Hunsdorfer and Ruuth [29], Kupka et al. [33]); and such a scheme is adapted to CED in the supplement to this paper. Alternatively, we can use ADER methods that give us a full space–time extension of the electric displacement and magnetic induction within a zone by taking the spatial reconstruction as an input. The PDE system in eqn. (2.9) tells us how the temporal evolution takes place for a given spatial reconstruction. Such an ADER scheme also includes treatment of stiff source terms and is described in Section 4 of this paper. These ADER methods are based on our prior papers (Dumbser et al. [25], Balsara et al. [6], [9] and Dumbser et al. [26]). ADER timestepping methods seamlessly extend to all orders, which is why they may be a stronger suit compared to Runge–Kutta methods. The innovations in ADER reported here go beyond the classical ADER methods mentioned previously because we show that there is an unusually efficient way to order the temporal nodes which results in a very efficient treatment of stiff source terms. With either of the methods described in this paragraph we can update the electric displacement using the extended Ampere's law, eqn. (2.1), and the magnetic induction using Faraday's law, eqn. (2.2). This completes our narrative description of the scheme.

In contrast to the methods described here, FDTD methods use a pair of staggered control volumes, making it difficult to devise higher order Godunov schemes for FDTD. While the reconstruction can be carried out as described here even for FDTD methods, the multidimensional Riemann solver requires that all the inputs, which may be discontinuous across the edges, come together at the same point. It is for this fundamental reason that multidimensionally upwinded higher order Godunov schemes cannot be designed for the FDTD staggering of variables.

### 2.3. The update equations

We take integer subscripts as denoting zone-centers and half-integer subscripts as denoting face-centers and edge-centers depending on how they are used; please see Fig. 1. Superscripts denote time-steps so that we describe a one-step update from time  $t^n$  to time  $t^{n+1} = t^n + \Delta t$ . Consistent with Fig. 1 and the narrative from the previous paragraph, we can write a discrete form for the time-update in eqn. (2.1) as

$$\begin{aligned} \bar{D}_{x;i+1/2,j,k}^{n+1} &= \bar{D}_{x;i+1/2,j,k}^n - \bar{J}_{x;i+1/2,j,k} \Delta t \\ &+ \frac{\Delta t}{\Delta y \Delta z} (\Delta x \bar{H}_{z;i+1/2,j+1/2,k} - \Delta x \bar{H}_{z;i+1/2,j-1/2,k} + \Delta y \bar{H}_{y;i+1/2,j,k-1/2} - \Delta y \bar{H}_{y;i+1/2,j,k+1/2}) \end{aligned} \quad (2.10a)$$

$$\begin{aligned} \bar{D}_{y;i,j-1/2,k}^{n+1} &= \bar{D}_{y;i,j-1/2,k}^n - \bar{J}_{y;i,j-1/2,k} \Delta t \\ &+ \frac{\Delta t}{\Delta x \Delta z} (\Delta x \bar{H}_{x;i,j-1/2,k+1/2} - \Delta x \bar{H}_{x;i,j-1/2,k-1/2} + \Delta z \bar{H}_{z;i-1/2,j-1/2,k} - \Delta z \bar{H}_{z;i+1/2,j-1/2,k}) \end{aligned} \quad (2.10b)$$

$$\begin{aligned} \bar{D}_{z;i,j,k+1/2}^{n+1} &= \bar{D}_{z;i,j,k+1/2}^n - \bar{J}_{z;i,j,k+1/2} \Delta t \\ &+ \frac{\Delta t}{\Delta x \Delta y} (\Delta x \bar{H}_{x;i,j-1/2,k+1/2} - \Delta x \bar{H}_{x;i,j+1/2,k+1/2} + \Delta y \bar{H}_{y;i+1/2,j,k+1/2} - \Delta y \bar{H}_{y;i-1/2,j,k+1/2}) \end{aligned}$$

(2.10c)

The current densities,  $\bar{J}_{x;i+1/2,j,k}$ ,  $\bar{J}_{y;i,j-1/2,k}$  and  $\bar{J}_{z;i,j,k+1/2}$  can be evaluated at  $t^n$  or  $t^{n+1}$  resulting in schemes that are explicit or implicit in their treatment of the source terms. In practice, the inclusion of stiff source terms is more intricate and is explained in detail in Sections 4 for ADER schemes. We can write an analogous discrete form for the time-update in eqn. (2.2) as

$$\begin{aligned} \bar{B}_{x;i+1/2,j,k}^{n+1} &= \bar{B}_{x;i+1/2,j,k}^n - \bar{M}_{x;i+1/2,j,k} \Delta t \\ &\quad - \frac{\Delta t}{\Delta y \Delta z} (\Delta z \bar{E}_{z;i+1/2,j+1/2,k} - \Delta z \bar{E}_{z;i+1/2,j-1/2,k} + \Delta y \bar{E}_{y;i+1/2,j,k-1/2} - \Delta y \bar{E}_{y;i+1/2,j,k+1/2}) \end{aligned} \quad (2.11a)$$

$$\begin{aligned} \bar{B}_{y;i,j-1/2,k}^{n+1} &= \bar{B}_{y;i,j-1/2,k}^n - \bar{M}_{y;i,j-1/2,k} \Delta t \\ &\quad - \frac{\Delta t}{\Delta x \Delta z} (\Delta x \bar{E}_{x;i,j-1/2,k+1/2} - \Delta x \bar{E}_{x;i,j-1/2,k-1/2} + \Delta z \bar{E}_{z;i-1/2,j-1/2,k} - \Delta z \bar{E}_{z;i+1/2,j-1/2,k}) \end{aligned} \quad (2.11b)$$

$$\begin{aligned} \bar{B}_{z;i,j,k+1/2}^{n+1} &= \bar{B}_{z;i,j,k+1/2}^n - \bar{M}_{z;i,j,k+1/2} \Delta t \\ &\quad - \frac{\Delta t}{\Delta x \Delta y} (\Delta x \bar{E}_{x;i,j-1/2,k+1/2} - \Delta x \bar{E}_{x;i,j+1/2,k+1/2} + \Delta y \bar{E}_{y;i+1/2,j,k+1/2} - \Delta y \bar{E}_{y;i-1/2,j,k+1/2}) \end{aligned} \quad (2.11c)$$

Similar comments, as above, apply for the treatment of the magnetic current densities.

The computational task consists of specifying the magnetic field components on the right hand side of eqn. (2.10) and the electric field components on the right hand side of eqn. (2.11). The method should also accommodate stiff source terms from the current densities when they are present.

### 3. Constraint-preserving reconstruction at third order

Since this is the first time that this reconstruction strategy is being presented to the CED community, it helps to present all the details in one place. From eqn. (2.3) we see that Gauss' law represents a non-evolutionary constraint on the three-dimensional electric displacement vector field. The components of this vector field are specified at the two-dimensional faces of the zone shown in Fig. 1. Therefore, a strategy needs to be found that starts with the two-dimensional facial components and retrieves the entire three-dimensional vector field in a fashion that is consistent with the constraint imposed by Gauss law. The strategy should be higher order accurate up to the desired order; for the purposes of this Section it should be third order accurate.

Let us place the origin at the center of the zone so that the extent of the zone is given by  $[-\Delta x/2, \Delta x/2] \times [-\Delta y/2, \Delta y/2] \times [-\Delta z/2, \Delta z/2]$ . The time-level is not shown as a superscript for all the equations in this Section because they all apply to the same time level. The zone-averaged components of the electric displacement at the right and left x-faces of Fig. 1 are denoted by  $D_0^{x\pm}$ . The zone-averaged components of the electric displacement at the upper and lower y-faces of Fig. 1 are denoted by  $D_0^{y\pm}$ . The zone-averaged components of the electric displacement at the top and bottom z-faces of Fig. 1 are denoted by  $D_0^{z\pm}$ . These are the face-centered primal variables for the electric displacement. These are the mesh-based variables that we start with when we begin to make a constraint-preserving reconstruction of the vector field.

Let the zone-averaged charge density be denoted by  $q_0$  so that a discrete form of eqn. (2.3) gives

$$q_0 = \frac{D_0^{x+} - D_0^{x-}}{\Delta x} + \frac{D_0^{y+} - D_0^{y-}}{\Delta y} + \frac{D_0^{z+} - D_0^{z-}}{\Delta z} \quad (3.1)$$

This charge density is obtained for all the zones of the mesh. Since the zone-averaged charge density can be defined for all the zones, we can reconstruct it using traditional higher order reconstruction in multiple dimensions; see for example the multidimensional WENO reconstruction in Balsara et al. [6]. (For a description of WENO methods, see also Jiang and Shu [31], Balsara and Shu [5], Dumbser and Käser [24] and Balsara, Garain and Shu [17]). The reconstructed charge density  $\rho_E(x, y, z)$  that is third order accurate is then given by

$$\begin{aligned} \rho_E(x, y, z) &= q_0 + q_x \left( \frac{x}{\Delta x} \right) + q_y \left( \frac{y}{\Delta y} \right) + q_z \left( \frac{z}{\Delta z} \right) + q_{xx} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) + q_{yy} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) \\ &\quad + q_{zz} \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + q_{xy} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) + q_{yz} \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) + q_{xz} \left( \frac{x}{\Delta x} \right) \left( \frac{z}{\Delta z} \right) \end{aligned} \quad (3.2)$$

Eqn. (3.2) will constitute the right hand side of the constraint eqn. (2.3) given by Gauss law.

Notice from Fig. 1 that each face will have adjoining faces. As a result, two-dimensional WENO reconstruction (or any other good, reconstruction strategy) can be used to endow the  $x$ -components of the electric displacement vector in the  $x$ -face of Fig. 1 with variation in the transverse direction. Denoting this by  $D^{x\pm}(y, z)$  in the right and left faces of Fig. 1, we can write

$$D^{x\pm}(y, z) = D_0^{x\pm} + D_y^{x\pm} \left( \frac{y}{\Delta y} \right) + D_z^{x\pm} \left( \frac{z}{\Delta z} \right) + D_{yy}^{x\pm} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) + D_{zz}^{x\pm} \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + D_{yz}^{x\pm} \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) \tag{3.3}$$

The above equation serves to define the moments  $D_y^{x\pm}$ ,  $D_z^{x\pm}$ ,  $D_{yy}^{x\pm}$ ,  $D_{zz}^{x\pm}$  and  $D_{yz}^{x\pm}$ ; and these are the moments that we obtain by applying WENO reconstruction to the  $x$ -component values from neighboring faces in the  $x$ -plane. An analogous procedure can be done in the upper and lower  $y$ -faces of Fig. 1 to give us

$$D^{y\pm}(x, z) = D_0^{y\pm} + D_x^{y\pm} \left( \frac{x}{\Delta x} \right) + D_z^{y\pm} \left( \frac{z}{\Delta z} \right) + D_{xx}^{y\pm} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) + D_{zz}^{y\pm} \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + D_{xz}^{y\pm} \left( \frac{x}{\Delta x} \right) \left( \frac{z}{\Delta z} \right) \tag{3.4}$$

The above equation serves to define the moments  $D_x^{y\pm}$ ,  $D_z^{y\pm}$ ,  $D_{xx}^{y\pm}$ ,  $D_{zz}^{y\pm}$  and  $D_{xz}^{y\pm}$ ; and these are the moments that we obtain by applying WENO reconstruction to the  $y$ -component values from neighboring faces in the  $y$ -plane. An analogous procedure can be done in the top and bottom  $z$ -faces of Fig. 1 to give us

$$D^{z\pm}(x, y) = D_0^{z\pm} + D_x^{z\pm} \left( \frac{x}{\Delta x} \right) + D_y^{z\pm} \left( \frac{y}{\Delta y} \right) + D_{xx}^{z\pm} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) + D_{yy}^{z\pm} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) + D_{xy}^{z\pm} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) \tag{3.5}$$

The above equation serves to define the moments  $D_x^{z\pm}$ ,  $D_y^{z\pm}$ ,  $D_{xx}^{z\pm}$ ,  $D_{yy}^{z\pm}$  and  $D_{xy}^{z\pm}$ ; and these are the moments that we obtain by applying WENO reconstruction to the  $z$ -component values from neighboring faces in the  $z$ -plane. The above three equations show us how the three components of the electric displacement within the six faces of the cube shown in Fig. 1 are endowed with sub-structure.

Despite having substructure within the faces, our narrative so far has not described how the electric displacement vector field is described everywhere within the volume shown in Fig. 1. Such a description is essential if we wish to provide the electric displacement vector as an input to the multidimensional Riemann solver. It is also essential if we wish to develop an ADER predictor step for the PDE system in eqn. (2.9). Let the three components of the electric displacement vector that are valid everywhere inside the zone shown in Fig. 1 be denoted by  $D^x(x, y, z)$ ,  $D^y(x, y, z)$  and  $D^z(x, y, z)$ . At third order, these components should at least have all the modes that are similar to the modes shown in eqn. (3.2). However, we will soon see that further satisfaction of the non-evolutionary constraint imposed by Gauss law, as well as matching the facial variation, requires a few more modes. As a result, for  $D^x(x, y, z)$  we have

$$D^x(x, y, z) = a_0 + a_x \left( \frac{x}{\Delta x} \right) + a_y \left( \frac{y}{\Delta y} \right) + a_z \left( \frac{z}{\Delta z} \right) + a_{xx} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) + a_{yy} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) + a_{zz} \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + a_{xy} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) + a_{yz} \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) + a_{xz} \left( \frac{x}{\Delta x} \right) \left( \frac{z}{\Delta z} \right) + a_{xxx} \left( \left( \frac{x}{\Delta x} \right)^3 - \frac{3}{20} \left( \frac{x}{\Delta x} \right) \right) + a_{xxy} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) \left( \frac{y}{\Delta y} \right) + a_{xxz} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) \left( \frac{z}{\Delta z} \right) + a_{xyy} \left( \frac{x}{\Delta x} \right) \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) + a_{xzz} \left( \frac{x}{\Delta x} \right) \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + a_{xyz} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) \tag{3.6}$$

Likewise, for  $D^y(x, y, z)$  we have

$$D^y(x, y, z) = b_0 + b_x \left( \frac{x}{\Delta x} \right) + b_y \left( \frac{y}{\Delta y} \right) + b_z \left( \frac{z}{\Delta z} \right) + b_{xx} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) + b_{yy} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) + b_{zz} \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + b_{xy} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) + b_{yz} \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) + b_{xz} \left( \frac{x}{\Delta x} \right) \left( \frac{z}{\Delta z} \right) + b_{yyy} \left( \left( \frac{y}{\Delta y} \right)^3 - \frac{3}{20} \left( \frac{y}{\Delta y} \right) \right) + b_{xyy} \left( \frac{x}{\Delta x} \right) \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) + b_{yyz} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) \left( \frac{z}{\Delta z} \right) + b_{xxy} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) \left( \frac{y}{\Delta y} \right) + b_{yzz} \left( \frac{y}{\Delta y} \right) \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + b_{xyz} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) \tag{3.7}$$

Similarly, for  $D^z(x, y, z)$  we have

$$\begin{aligned}
 D^z(x, y, z) = & c_0 + c_x \left( \frac{x}{\Delta x} \right) + c_y \left( \frac{y}{\Delta y} \right) + c_z \left( \frac{z}{\Delta z} \right) + c_{xx} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) + c_{yy} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) \\
 & + c_{zz} \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + c_{xy} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) + c_{yz} \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) + c_{xz} \left( \frac{x}{\Delta x} \right) \left( \frac{z}{\Delta z} \right) \\
 & + c_{zzz} \left( \left( \frac{z}{\Delta z} \right)^3 - \frac{3}{20} \left( \frac{z}{\Delta z} \right) \right) + c_{xzz} \left( \frac{x}{\Delta x} \right) \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) + c_{yzz} \left( \frac{y}{\Delta y} \right) \left( \left( \frac{z}{\Delta z} \right)^2 - \frac{1}{12} \right) \\
 & + c_{xxz} \left( \left( \frac{x}{\Delta x} \right)^2 - \frac{1}{12} \right) \left( \frac{z}{\Delta z} \right) + c_{yyz} \left( \left( \frac{y}{\Delta y} \right)^2 - \frac{1}{12} \right) \left( \frac{z}{\Delta z} \right) + c_{xyz} \left( \frac{x}{\Delta x} \right) \left( \frac{y}{\Delta y} \right) \left( \frac{z}{\Delta z} \right) \quad (3.8)
 \end{aligned}$$

The first two lines in eqns. (3.6), (3.7) and (3.8) are needed for satisfying the order property. The inclusion of the next two lines in eqns. (3.6), (3.7) and (3.8) ensure that the non-evolutionary constraint is satisfied and for ensuring that the facial values are matched. Consequently, the 48 coefficients in the above three equations are required to match the facial values in eqns. (3.3), (3.4) and (3.5) as well as the constraint equation.

Let us now turn to the constraint equation provided by Gauss law. Eqn. (2.3) gives

$$\partial_x D^x(x, y, z) + \partial_y D^y(x, y, z) + \partial_z D^z(x, y, z) = \rho_E(x, y, z) \quad (3.9)$$

Because the left and right sides of eqn. (3.9) are polynomials, satisfying the non-evolutionary constraint in eqn. (3.9) is tantamount to matching the polynomials on either side of eqn. (3.9) at all applicable orders. Since eqn. (3.2) is a quadratic polynomial, and since eqns. (3.6), (3.7) and (3.8) are at most cubic polynomials, we see that we will get ten equations of constraint in three-dimensions. These are explicitly cataloged in Appendix A. In that Appendix, we also explicitly provide a solution for the 48 coefficients in eqns. (3.6), (3.7) and (3.8). These coefficients are chosen so as to explicitly match the quadratic variation of the components of the electric displacement in the six faces that are documented in eqns. (3.3), (3.4) and (3.5). The reason for providing all 48 coefficients explicitly using a leisurely notation in Appendix A is to facilitate easy entry into the field of higher order CED. Appendix B of this paper gives the fourth order case in a format that is optimally adapted for use with fourth order ADER schemes.

In this Section we have documented a constraint-preserving reconstruction strategy for the electric displacement vector field that is consistent with the non-evolutionary constraint in eqn. (2.3). The magnetic induction vector field can also be obtained consistent with the non-evolutionary constraint in eqn. (2.4). The constraint-preserving reconstruction of the magnetic induction vector field follows a strategy that is identical to the one developed in this Section. We therefore do not describe it any further here.

#### 4. ADER timestepping adapted to computational electrodynamics

There are two very important physics issues that are most challenging in CED. The first has to do with the fact that permittivity and permeability are tensors and can undergo up to an order of magnitude change at the interface between dielectrics (Farjadpour et al. [27], Oskooi et al. [35], Taflove, Oskooi, and Johnson [44]). The second has to do with the fact that the conductivity can undergo very substantial variation. This may happen when analyzing skin-depth effects in a metal, but it occurs much more often when dealing with perfectly matched layers (PML) at the boundary of a computational domain (Berenger [22], [23], Katz, Thiel and Taflove [32], Taflove and Hagness [43]).

Section 4.1 provides a quick preview of the ADER predictor step and its utility for CED. Sub-section 4.2 describes the ADER predictor step at third order in some detail. Sub-section 4.3 shows how easy it is to invert the matrices that result from the implicit treatment of stiff source terms. Sub-section 4.4 gives step-wise implementation-related details for an ADER-WENO scheme.

##### 4.1. Preview of ADER predictor step with a focus on its utility for CED

The spatial reconstruction from Section 3 (and Appendices A or B) can provide a three-dimensional, spatially high order reconstruction of the electric displacement and magnetic induction vector fields. Given spatial derivatives, eqn. (2.9) shows that they can be related to the time derivatives for any PDE; Maxwell's equations included. This means that if the spatial reconstruction is known within the zone up to some order of accuracy, we should be able to provide an in-the-small temporal evolution within that zone up to the same order of accuracy. This is what the ADER predictor step provides. In the case of CED, ADER schemes are especially valuable because they allow us to include the sub-cell spatial variation of the permittivity, permeability and conductivities within a zone. Moreover, these material properties can be spatially varying, anisotropic tensors with no loss of generality or accuracy in our algorithm and our code. For dispersive and resistive media, we can even endow these tensors with time-variation, perhaps in a way that is consistent with some governing ordinary differential equation for the evolution of these material properties.



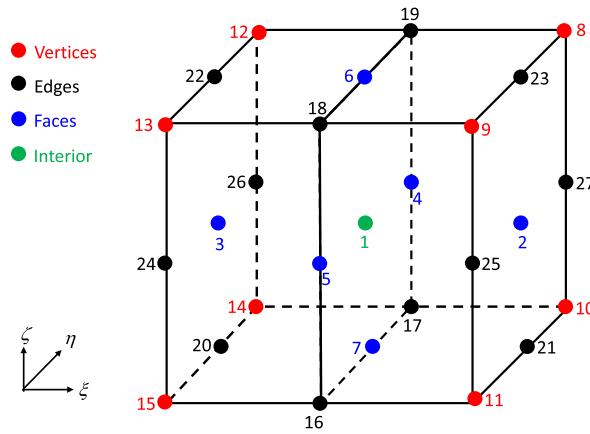


Fig. 2. Shows the nodal points in the reference element that are used for the ADER predictor step at third order. The reference element has coordinates  $(\xi, \eta, \zeta)$ . The nodal points are shown as filled dots, and their color-matched numbering is shown alongside them. Notice that we use a tensor product of the third order accurate Gauss–Lobatto quadrature points. The plane that is shown in the middle of the  $x$ -direction is just intended to guide the eye.

Once we have a higher order space and time representation of the electric displacement and magnetic induction vector fields within all the zones, we can make these variables interact with one another at the zone edges using the multidimensional Riemann solver. This gives us the edge-located electric and magnetic field components shown in Fig. 1. ADER schemes have another valuable feature that is useful for CED. They can treat stiff source terms, which are generated when the conductivities are large. This is done in such a way that the stiff source terms are naturally integrated into the time-evolution of the rest of the PDE. Moreover, sub-cell variations in the conductivities within a zone are also treated in a fully time-implicit fashion. By making an arithmetic average of the space–time integrated stiff source terms at the faces of any two neighboring zones, the facial source terms are also treated in a stable and time-implicit fashion.

#### 4.2. Third order ADER predictor step that is adapted for CED

While the older ADER schemes by Toro, Millington and Nejad [48], Titarev and Toro [46], [47] and Toro and Titarev [49] are not very well-suited to our specific needs in this paper, the newer ADER schemes in predictor–corrector format (Dumbser et al. [25], [26], Balsara et al. [6], [9]) prove to be very valuable. However, even these newer ADER methods are inefficient in their treatment of source terms. The innovations in ADER reported here go beyond the classical ADER methods mentioned previously because we show that there is an unusually efficient way to arrange the ordering of the temporal nodes which results in a very efficient treatment of stiff source terms. The full details about the treatment of source terms will become clear in the next sub-section. To facilitate our narrative, let us make a linear mapping of the space–time extent of each zone in its local coordinates given by  $(x, y, z, t) \in [-\Delta x/2, \Delta x/2] \times [-\Delta y/2, \Delta y/2] \times [-\Delta z/2, \Delta z/2] \times [0, \Delta t]$  to the reference element given by  $(\xi, \eta, \zeta, \tau) \in [-1/2, 1/2] \times [-1/2, 1/2] \times [-1/2, 1/2] \times [0, 1]$ . With this linear mapping, the governing PDE becomes

$$\frac{\partial \mathbf{u}(\xi, \eta, \zeta, \tau)}{\partial \tau} + \frac{\partial \mathbf{f}(\xi, \eta, \zeta, \tau)}{\partial \xi} + \frac{\partial \mathbf{g}(\xi, \eta, \zeta, \tau)}{\partial \eta} + \frac{\partial \mathbf{h}(\xi, \eta, \zeta, \tau)}{\partial \zeta} = \mathbf{s}(\xi, \eta, \zeta, \tau) \tag{4.1}$$

where we have the slight redefinitions (and rescalings) given by  $\mathbf{u}(\xi, \eta, \zeta, \tau) = \mathbf{U}(x, y, z, t)$ ,  $\mathbf{f}(\xi, \eta, \zeta, \tau) = \Delta t \mathbf{F}(x, y, z, t)/\Delta x$ ,  $\mathbf{g}(\xi, \eta, \zeta, \tau) = \Delta t \mathbf{G}(x, y, z, t)/\Delta y$ ,  $\mathbf{h}(\xi, \eta, \zeta, \tau) = \Delta t \mathbf{K}(x, y, z, t)/\Delta z$  and  $\mathbf{s}(\xi, \eta, \zeta, \tau) = \Delta t \mathbf{S}(x, y, z, t)$ . Please note that we are making a small notational realignment in the nomenclature for the  $z$ -flux.

Since we will need to evaluate the time-averaged electric and magnetic fields at the edges of the mesh, see Fig. 1, it is very beneficial if our ADER predictor step gives us the time-evolution of the primal variables at those locations. To obtain up to fourth order of accuracy, we need to have fourth order accurate quadrature at each edge and that is provided by the Simpson quadrature rule. The nodal points that we choose within each reference element are, therefore, a tensor product of the Gauss–Lobatto quadrature points as shown in Fig. 2. In this Sub-section, we describe the ADER predictor step with third order of accuracy. The present ADER scheme is closest to the one in Dumbser et al. [26]. However, it differs in two significant ways. First, we use Gauss–Lobatto-like quadrature points inside the reference element instead of using Gaussian quadrature points. This has the advantage that it directly gives us the primal variables at the edges of the mesh. This is advantageous in overall scheme design because the Riemann solvers are also invoked at the edges of the mesh in a CED code. The overall accuracy of the scheme is not compromised. Second, we use a special set of nodal points in time. This special set of nodal points simplifies the matrix inversion dramatically when stiff source terms are present, as will be shown in the next Sub-section. While we specialize to third order, all the innovations described here extend to all orders and can be applied to any hyperbolic PDE system with stiff source terms.

Fig. 2 shows the nodal points in the reference element that are used for the ADER predictor step at third order. The reference element has coordinates  $(\xi, \eta, \zeta)$ . The nodal points are shown as filled dots. Notice that we use a tensor product

of the third order accurate Gauss–Lobatto quadrature points. The plane that is shown in the middle of the  $x$ -direction is just intended to guide the eye. Label the coordinates in Fig. 2 with an index “ $i$ ”, so that the  $i$ th nodal point is given by  $\mathbf{r}_i = (\xi_i, \eta_i, \zeta_i)$ . We have  $i = 1, \dots, 27$  such nodal points within the reference element at third order as seen from Fig. 2. The specific numbering of nodes that we chose at third order is shown in Fig. 2 and the nodal numbers are color-matched with the nodes they represent. (At fourth order, we would have 64 such nodal points). At third order we consider polynomials that are built from the tensor product of the following polynomial sets:  $\{1, \xi, \xi^2\} \otimes \{1, \eta, \eta^2\} \otimes \{1, \zeta, \zeta^2\}$ . Note that the polynomials that we are considering have as many free coefficients as the number of nodal points in Fig. 2. We now form  $j = 1, \dots, 27$  such polynomials, denoted by  $\phi_j(\mathbf{r})$ , which are obtained by asserting the following rule

$$\phi_j(\mathbf{r}_i) = \delta_{ij} \quad \forall i = 1, \dots, 27 \quad \text{and} \quad j = 1, \dots, 27 \quad (4.2)$$

These  $\phi_j(\mathbf{r})$  form the spatial basis for our scheme. With these spatial basis in hand, we can use the reconstructed vector fields to write the solution vector at the initial time as

$$\mathbf{w}(\xi, \eta, \zeta) = \sum_{j=1}^{27} \mathbf{w}_j \phi_j(\mathbf{r}) \quad (4.3)$$

Here  $\mathbf{w}_j$  is the value of the six-component solution vector from eqn. (2.9) at the nodal point “ $j$ ” within the reference element shown in Fig. 2 at time  $\tau = 0$ .

Since we want the third order accurate temporal evolution of the PDE, we define three temporal nodal points given by the two Gaussian points  $\tau_1 = (1 - 1/\sqrt{3})/2$ ,  $\tau_2 = (1 + 1/\sqrt{3})/2$  and the third temporal node given by  $\tau_3 = 1$ . (For fourth order, it is important to choose the three Gaussian points  $\tau_1 = (1 - \sqrt{3/5})/2$ ,  $\tau_2 = 1/2$ ,  $\tau_3 = (1 + \sqrt{3/5})/2$  and the fourth temporal node given by  $\tau_4 = 1$ ). There is something special about this way of choosing these specific temporal nodal points, and this is an insight that we missed in Dumbser et al. [26]. With this choice of nodal points, the update equation for the solution vector at the last temporal nodal point only depends on the source terms at the previous temporal nodal points at the same spatial location. This can be used to make a dramatic simplification of the matrix inversion when stiff source terms are present – this is a point that we will illustrate in Sub-section 4.3, and the insight carries over to all orders. In simple terms, say we want to treat the stiff source terms in an  $M$ -component hyperbolic system with  $N$ th order of temporal accuracy. The original Dumbser et al. [26] ADER scheme would require us to invert an  $(NM) \times (NM)$  matrix at each spatial nodal point. Our new arrangement of temporal nodes is such that we only have to invert  $N$  matrices, each of size  $M \times M$ , at each spatial nodal point in our new ADER scheme. This is a significant improvement in computational efficiency when dealing with stiff source terms. When the source terms are non-stiff, both ADER methods operate with comparable efficiency and they also have comparable accuracy. In fact, when source terms are non-stiff, SSP-RK methods are also available at all reasonable orders of accuracy (Shu and Osher [39], [40], Shu [41], Spiteri and Ruuth [37], [38]). But please note that even for non-stiff hyperbolic problems SSP-RK methods have been shown to be inefficient relative to the original ADER schemes (see Balsara et al. [6]). However, in CED, it is very common to have very high conductivities, and therefore very stiff source terms. For this reason, we suggest that our newer ADER scheme might be more efficient for hyperbolic problems with stiff source terms than our original ADER scheme in Dumbser et al. [26] and also the RK schemes for treating stiff source terms (Pareschi and Russo [36], Hunsdorfer and Ruuth [29], Kupka et al. [33]).

We define temporal basis functions, which are quadratic polynomials in “ $\tau$ ”. Denote the three such polynomials by  $\theta_k(\tau)$  and we make them satisfy the property

$$\theta_k(\tau_m) = \delta_{km} \quad \forall k = 1, \dots, 3 \quad \text{and} \quad m = 1, \dots, 3 \quad (4.4)$$

We seek a DG-like formulation for ADER where we will use the functions in eqn. (4.4) to define a space–time bases. The assembly of these spatial and temporal basis sets can be automated with a computer algebra system and the polynomials themselves are not needed in the computer code implementation.

We now define the space–time basis functions given by  $\psi_{j,k}(\mathbf{r}, \tau)$  that are defined by

$$\psi_{j,k}(\mathbf{r}, \tau) = \phi_j(\mathbf{r})\theta_k(\tau) \quad \forall j = 1, \dots, 27 \quad \text{and} \quad k = 1, \dots, 3 \quad (4.5)$$

The above equation gives us 81 space–time bases and the full space–time solution vector within the zone is specified by 81 coefficients as follows

$$\mathbf{u}(\xi, \eta, \zeta, \tau) = \sum_{k=1}^3 \sum_{j=1}^{27} \mathbf{u}_{j,k} \psi_{j,k}(\mathbf{r}, \tau) \quad (4.6)$$

Given the definition of our space–time basis functions, we see that the 81 coefficients  $\mathbf{u}_{j,k}$  in the above equation are indeed the nodal values of the solution vector at the 81 space–time nodes. These 81 six-component vectors,  $\{\mathbf{u}_{j,k}; j = 1, \dots, 27 \text{ and } k = 1, \dots, 3\}$ , are the desired output from the ADER predictor step. We want to find them consistent with the initial conditions, given by the set of 27 six-component vectors  $\{\mathbf{w}_j; j = 1, \dots, 27\}$  in eqn. (4.3), and the governing PDE given by eqn. (4.1) in the reference element. We seek a DG-like formulation for ADER where the bases  $\psi_{j,k}(\mathbf{r}, \tau)$  will be used as both the trial and the test functions.

We wish to obtain converged values for  $\{\mathbf{u}_{j,k}; j = 1, \dots, 27 \text{ and } k = 1, \dots, 3\}$  via an iterative process. A very good starting guess with which to start the iteration would be  $\mathbf{u}_{j,k} = \mathbf{w}_j \quad \forall k = 1, \dots, 3$ . Let us say that we have some intermediate guess for  $\{\mathbf{u}_{j,k}; j = 1, \dots, 27 \text{ and } k = 1, \dots, 3\}$ . We can now evaluate the fluxes as  $\mathbf{f}_{j,k} \equiv \mathbf{f}(\mathbf{u}_{j,k})$  and the source terms as  $\mathbf{s}_{j,k} = \mathbf{s}(\mathbf{u}_{j,k})$  at each of those space–time nodal points so that we can write

$$\mathbf{f}(\xi, \eta, \zeta, \tau) = \sum_{k=1}^3 \sum_{j=1}^{27} \mathbf{f}_{j,k} \psi_{j,k}(\mathbf{r}, \tau) \tag{4.7}$$

and

$$\mathbf{s}(\xi, \eta, \zeta, \tau) = \sum_{k=1}^3 \sum_{j=1}^{27} \mathbf{s}_{j,k} \psi_{j,k}(\mathbf{r}, \tau) \tag{4.8}$$

Analogous definitions can be made for the  $y$ - and  $z$ -fluxes. We now explain how we obtain the governing equations with which the iterations can be advanced.

Our solution strategy at third and fourth orders exactly follows the approach laid out in Balsara et al. (2017) at second order. We, nevertheless, give a few brief details for the sake of completeness. The PDE in eqn. (2.9), and analogously eqn. (4.1), controls the dynamics. If we want to iteratively improve the solution vector itself, we should involve the governing PDE in the iteration step. In other words, an improved iterate for eqn. (4.6) can only be obtained if we involve eqn. (4.1). This is most easily done by using our test functions to make a Galerkin projection in space–time of the governing equation. In other words, let  $\phi_{j,k}(\xi, \eta, \zeta, \tau)$  with  $j = 1, \dots, 27$  and  $k = 1, \dots, 3$  be one of the 81 test functions from eqn. (4.6). For  $j = 1, \dots, 27$  and  $k = 1, \dots, 3$  we then demand that

$$\int_0^1 \left\{ \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \left\{ \phi_{j,k}(\xi, \eta, \zeta, \tau) \left[ \frac{\partial \mathbf{u}(\xi, \eta, \zeta, \tau)}{\partial \tau} + \frac{\partial \mathbf{f}(\xi, \eta, \zeta, \tau)}{\partial \xi} + \frac{\partial \mathbf{g}(\xi, \eta, \zeta, \tau)}{\partial \eta} + \frac{\partial \mathbf{h}(\xi, \eta, \zeta, \tau)}{\partial \zeta} - \mathbf{s}(\xi, \eta, \zeta, \tau) \right] \right\} d\xi \right\} d\eta \left\} d\zeta \right\} d\tau = 0 \tag{4.9}$$

Please note that the square bracket in the previous equation is not a matrix. In the above equation, we are simply projecting the governing equation, i.e. eqn. (4.1), into the space of test functions and demanding that the projection is zero. To do this consistent with the initial conditions in eqn. (4.3), we carry out integration by parts for the time-derivative term in eqn. (4.9) to yield

$$\begin{aligned} & \int_{-1/2}^{1/2} \left\{ \int_{-1/2}^{1/2} \left\{ \int_{-1/2}^{1/2} \left\{ \phi_{j,k}(\xi, \eta, \zeta, \tau = 1) \mathbf{u}(\xi, \eta, \zeta, \tau = 1) - \phi_{j,k}(\xi, \eta, \zeta, \tau = 0) \mathbf{w}(\xi, \eta, \zeta) \right\} d\xi \right\} d\eta \right\} d\zeta \\ & - \int_0^1 \left\{ \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \left\{ \int_{-1/2}^{1/2} \left\{ \frac{\partial \phi_{j,k}(\xi, \eta, \zeta, \tau)}{\partial \tau} \mathbf{u}(\xi, \eta, \zeta, \tau) \right\} d\xi \right\} d\eta \right\} d\zeta \right\} d\tau \\ & + \int_0^1 \left\{ \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \left\{ \int_{-1/2}^{1/2} \left\{ \phi_{j,k}(\xi, \eta, \zeta, \tau) \left[ \frac{\partial \mathbf{f}(\xi, \eta, \zeta, \tau)}{\partial \xi} + \frac{\partial \mathbf{g}(\xi, \eta, \zeta, \tau)}{\partial \eta} + \frac{\partial \mathbf{h}(\xi, \eta, \zeta, \tau)}{\partial \zeta} - \mathbf{s}(\xi, \eta, \zeta, \tau) \right] \right\} d\xi \right\} d\eta \right\} d\zeta \right\} d\tau = 0 \end{aligned} \tag{4.10}$$

Operationally, to carry out the Galerkin projection, we simply insert eqns. (4.3), (4.6), (4.7) and (4.8), as well as their analogues, into the square bracket in eqn. (4.10). We then use a computer algebra system to carry out the four-dimensional space–time integrals. The computer algebra system is further asked to simplify the resulting 81 equations that we get from eqn. (4.10). After a little bit of manipulation with the computer algebra system, we get 81 very elegant, very compact, and very simple, evolutionary equations which we will describe briefly in the next paragraph. Computer algebra systems have become so sophisticated that one can just ask them to solve the 81 conditions in eqn. (4.10) for each of the elements of the set  $\{\mathbf{u}_{j,k}; j = 1, \dots, 27 \text{ and } k = 1, \dots, 3\}$  and they will, almost effortlessly, provide a solution.

We write down the three equations that we get from the above procedure at the first nodal point in Fig. 2. I.e., we want to examine the structure of the equations for  $\mathbf{u}_{1,1}$ ,  $\mathbf{u}_{1,2}$  and  $\mathbf{u}_{1,3}$ . The update equation for  $\mathbf{u}_{1,1}$  can be written (with the source terms taken over to the left hand side) as

$$\begin{aligned}
 \mathbf{u}_{1,1} &- \frac{18 - \sqrt{3}}{60} \mathbf{s}(\mathbf{u}_{1,1}) - \frac{12 - 11\sqrt{3}}{60} \mathbf{s}(\mathbf{u}_{1,2}) - \frac{\sqrt{3}}{30} \mathbf{s}(\mathbf{u}_{1,3}) \\
 &= \mathbf{w}_1 + \frac{18 - \sqrt{3}}{60} (\mathbf{f}_{3,1} - \mathbf{f}_{2,1} + \mathbf{g}_{5,1} - \mathbf{g}_{4,1} + \mathbf{h}_{7,1} - \mathbf{h}_{6,1}) \\
 &\quad + \frac{12 - 11\sqrt{3}}{60} (\mathbf{f}_{3,2} - \mathbf{f}_{2,2} + \mathbf{g}_{5,2} - \mathbf{g}_{4,2} + \mathbf{h}_{7,2} - \mathbf{h}_{6,2}) + \frac{\sqrt{3}}{30} (\mathbf{f}_{3,3} - \mathbf{f}_{2,3} + \mathbf{g}_{5,3} - \mathbf{g}_{4,3} + \mathbf{h}_{7,3} - \mathbf{h}_{6,3}) \quad (4.11)
 \end{aligned}$$

The update equation for  $\mathbf{u}_{1,2}$  can be written (with the source terms taken over to the left hand side) as

$$\begin{aligned}
 \mathbf{u}_{1,2} &- \frac{12 + 11\sqrt{3}}{60} \mathbf{s}(\mathbf{u}_{1,1}) - \frac{18 + \sqrt{3}}{60} \mathbf{s}(\mathbf{u}_{1,2}) + \frac{\sqrt{3}}{30} \mathbf{s}(\mathbf{u}_{1,3}) \\
 &= \mathbf{w}_1 + \frac{12 + 11\sqrt{3}}{60} (\mathbf{f}_{3,1} - \mathbf{f}_{2,1} + \mathbf{g}_{5,1} - \mathbf{g}_{4,1} + \mathbf{h}_{7,1} - \mathbf{h}_{6,1}) + \frac{18 + \sqrt{3}}{60} (\mathbf{f}_{3,2} - \mathbf{f}_{2,2} + \mathbf{g}_{5,2} - \mathbf{g}_{4,2} + \mathbf{h}_{7,2} - \mathbf{h}_{6,2}) \\
 &\quad - \frac{\sqrt{3}}{30} (\mathbf{f}_{3,3} - \mathbf{f}_{2,3} + \mathbf{g}_{5,3} - \mathbf{g}_{4,3} + \mathbf{h}_{7,3} - \mathbf{h}_{6,3}) \quad (4.12)
 \end{aligned}$$

The update equation for  $\mathbf{u}_{1,3}$  can be written (with the source terms taken over to the left hand side) as

$$\begin{aligned}
 \mathbf{u}_{1,3} &- \frac{1}{2} \mathbf{s}(\mathbf{u}_{1,1}) - \frac{1}{2} \mathbf{s}(\mathbf{u}_{1,2}) \\
 &= \mathbf{w}_1 + \frac{1}{2} (\mathbf{f}_{3,1} - \mathbf{f}_{2,1} + \mathbf{g}_{5,1} - \mathbf{g}_{4,1} + \mathbf{h}_{7,1} - \mathbf{h}_{6,1}) + \frac{1}{2} (\mathbf{f}_{3,2} - \mathbf{f}_{2,2} + \mathbf{g}_{5,2} - \mathbf{g}_{4,2} + \mathbf{h}_{7,2} - \mathbf{h}_{6,2}) \quad (4.13)
 \end{aligned}$$

It is useful to mention several interesting traits in the above three equations. The right hand sides of eqns. (4.11), (4.12) and (4.13) can be evaluated by using the set of all  $\{\mathbf{u}_{j,k}; j = 1, \dots, 27 \text{ and } k = 1, \dots, 3\}$  from the previous iteration. Notice too that the right hand sides of the above three equations have an elegant finite difference-like quality; this makes computer implementations very simple. This is a trait that we see for the update equations at all 27 nodal points in Fig. 2. (It also extends to fourth and higher orders for this class of ADER schemes.) The left hand sides of eqns. (4.11), (4.12) and (4.13) should be evaluated implicitly because they contain the contributions from stiff source terms. Notice that the source terms  $\mathbf{s}(\mathbf{u}_{1,1})$ ,  $\mathbf{s}(\mathbf{u}_{1,2})$  and  $\mathbf{s}(\mathbf{u}_{1,3})$  are the only source terms that couple to  $\mathbf{u}_{1,1}$ ,  $\mathbf{u}_{1,2}$  and  $\mathbf{u}_{1,3}$ . All these traits are also available in the formulation from Dumbser et al. [26]. However, eqn. (4.13) has a trait that is not available in the formulation from Dumbser et al. [26] and this extra trait opens the door to a very simple solution strategy. Observe from eqn. (4.13) that its left hand side lacks a term with  $\mathbf{s}(\mathbf{u}_{1,3})$ . This will be exploited to produce a very simple solution strategy in the next section. Such a simple solution strategy is not possible for the formulation from Dumbser et al. [26]. It is important to point out that this trait will be observed at all 27 nodal points at third order. Indeed, as long as the temporal basis set from eqn. (4.4) is used, it will be observed at all orders. Physically, eqn. (4.13) expresses the update of  $\mathbf{u}_{1,3}$  at  $\tau_3 = 1$ . It is then evident that the time-integration of the source terms is just given by  $(\mathbf{s}(\mathbf{u}_{1,1}) + \mathbf{s}(\mathbf{u}_{1,2}))/2$ . Consequently,  $\mathbf{s}(\mathbf{u}_{1,3})$  does not appear on the left hand side of eqn. (4.13). In retrospect, this is the physical motivation for choosing the temporal quadrature points that we indeed chose.

The above three equations can also be used to check whether the user’s computer algebra system has produced the correct answers. In the next Sub-section we describe our efficient solution strategy for making a time-implicit solution of eqns. (4.11), (4.12) and (4.13). In the electronic supplements on the JCP website, we provide the computer algebra system codes for ADER schemes at second third and fourth orders. Please cite this paper, and its allied papers (Dumbser et al. [25], [26], Balsara et al. [6], [9]) when the associated ADER scheme is used for any PDE system with stiff source terms.

### 4.3. Efficient treatment of implicit source terms at third order

Let us now describe our efficient solution strategy for making a time-implicit solution of eqns. (4.11), (4.12) and (4.13). Let us begin by defining the following Jacobian matrices for the source terms

$$\mathbf{M}_{1,1} \equiv \frac{\partial \mathbf{s}(\mathbf{u}_{1,1})}{\partial \mathbf{u}_{1,1}}; \quad \mathbf{M}_{1,2} \equiv \frac{\partial \mathbf{s}(\mathbf{u}_{1,2})}{\partial \mathbf{u}_{1,2}}; \quad \mathbf{M}_{1,3} \equiv \frac{\partial \mathbf{s}(\mathbf{u}_{1,3})}{\partial \mathbf{u}_{1,3}} \quad (4.14)$$

Also let  $\mathbf{R}_{1,1}^0$ ,  $\mathbf{R}_{1,2}^0$  and  $\mathbf{R}_{1,3}^0$  be the residuals for eqns. (4.11), (4.12) and (4.13). Let  $\mathbf{I}$  be the  $6 \times 6$  identity matrix. Then the matrix equation that we need to solve for the increments  $\delta \mathbf{u}_{1,1}$ ,  $\delta \mathbf{u}_{1,2}$  and  $\delta \mathbf{u}_{1,3}$  is given by

$$\begin{bmatrix} \mathbf{I} - a_1 \mathbf{M}_{1,1} & -a_2 \mathbf{M}_{1,2} & -a_3 \mathbf{M}_{1,3} \\ -b_1 \mathbf{M}_{1,1} & \mathbf{I} - b_2 \mathbf{M}_{1,2} & -b_3 \mathbf{M}_{1,3} \\ -c_1 \mathbf{M}_{1,1} & -c_2 \mathbf{M}_{1,2} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \delta \mathbf{u}_{1,1} \\ \delta \mathbf{u}_{1,2} \\ \delta \mathbf{u}_{1,3} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{1,1}^0 \\ \mathbf{R}_{1,2}^0 \\ \mathbf{R}_{1,3}^0 \end{bmatrix} \quad (4.15)$$

with

$$\begin{aligned}
 a_1 &= \frac{18 - \sqrt{3}}{60}; & a_2 &= \frac{12 - 11\sqrt{3}}{60}; & a_3 &= \frac{\sqrt{3}}{30}; \\
 b_1 &= \frac{12 + 11\sqrt{3}}{60}; & b_2 &= \frac{18 + \sqrt{3}}{60}; & b_3 &= -\frac{\sqrt{3}}{30}; & c_1 &= \frac{1}{2}; & c_2 &= \frac{1}{2}
 \end{aligned}
 \tag{4.16}$$

For CED, the matrix in eqn. (4.15) is an  $18 \times 18$  matrix and its direct inversion would be rather expensive. This is especially true for an ADER scheme because such  $18 \times 18$  matrices would have to be inverted at all 27 nodal points. It would be very desirable if we could simplify the problem so that we only had to invert two  $6 \times 6$  matrices! Because of the identity matrix in the lower right corner of the matrix in eqn. (4.15) we can perform row operations with the  $6 \times 6$  blocks in order to achieve this goal. We make this strategy explicit in the next paragraph.

Performing row operations on the  $6 \times 6$  blocks, we can write eqn. (4.15) as

$$\begin{bmatrix} \mathbf{I} & \mathbf{N}_{1,2}^1 & \mathbf{N}_{1,3}^1 \\ \mathbf{N}_{2,1}^1 & \mathbf{I} & \mathbf{N}_{2,3}^1 \\ \mathbf{N}_{3,1}^1 & \mathbf{N}_{3,2}^1 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \delta \mathbf{u}_{1,1} \\ \delta \mathbf{u}_{1,2} \\ \delta \mathbf{u}_{1,3} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{1,1}^1 \\ \mathbf{R}_{1,2}^1 \\ \mathbf{R}_{1,3}^1 \end{bmatrix}
 \tag{4.17}$$

with

$$\begin{aligned}
 \mathbf{N}_{1,2}^1 &= \frac{a_1 c_2 - a_2 c_1}{c_1} \mathbf{M}_{1,2}; & \mathbf{N}_{1,3}^1 &= -\frac{a_1}{c_1} \mathbf{I} - a_3 \mathbf{M}_{1,3}; & \mathbf{N}_{2,1}^1 &= \frac{c_1 b_2 - c_2 b_1}{c_2} \mathbf{M}_{1,1}; & \mathbf{N}_{2,3}^1 &= -\frac{b_2}{c_2} \mathbf{I} - b_3 \mathbf{M}_{1,3}; \\
 \mathbf{N}_{3,1}^1 &= -c_1 \mathbf{M}_{1,1}; & \mathbf{N}_{3,2}^1 &= -c_2 \mathbf{M}_{1,2}; & \mathbf{R}_{1,1}^1 &= \mathbf{R}_{1,1}^0 - \frac{a_1}{c_1} \mathbf{R}_{1,3}^0; & \mathbf{R}_{1,2}^1 &= \mathbf{R}_{1,2}^0 - \frac{b_2}{c_2} \mathbf{R}_{1,3}^0; & \mathbf{R}_{1,3}^1 &= \mathbf{R}_{1,3}^0
 \end{aligned}
 \tag{4.18}$$

Performing another round of row operations on the  $6 \times 6$  blocks, we can write eqn. (4.17) as

$$\begin{bmatrix} \mathbf{I} & \mathbf{N}_{1,2}^2 & \mathbf{N}_{1,3}^2 \\ 0 & \mathbf{N}_{2,2}^2 & \mathbf{N}_{2,3}^2 \\ 0 & \mathbf{N}_{3,2}^2 & \mathbf{N}_{3,3}^2 \end{bmatrix} \begin{bmatrix} \delta \mathbf{u}_{1,1} \\ \delta \mathbf{u}_{1,2} \\ \delta \mathbf{u}_{1,3} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{1,1}^2 \\ \mathbf{R}_{1,2}^2 \\ \mathbf{R}_{1,3}^2 \end{bmatrix}
 \tag{4.19}$$

with

$$\begin{aligned}
 \mathbf{N}_{1,2}^2 &= \mathbf{N}_{1,2}^1; & \mathbf{N}_{1,3}^2 &= \mathbf{N}_{1,3}^1; & \mathbf{N}_{2,2}^2 &= \mathbf{I} - \mathbf{N}_{2,1}^1 \mathbf{N}_{1,2}^1; & \mathbf{N}_{2,3}^2 &= \mathbf{N}_{2,3}^1 - \mathbf{N}_{2,1}^1 \mathbf{N}_{1,3}^1; \\
 \mathbf{N}_{3,2}^2 &= \mathbf{N}_{3,2}^1 - \mathbf{N}_{3,1}^1 \mathbf{N}_{1,2}^1; & \mathbf{N}_{3,3}^2 &= \mathbf{I} - \mathbf{N}_{3,1}^1 \mathbf{N}_{1,3}^1; \\
 \mathbf{R}_{1,1}^2 &= \mathbf{R}_{1,1}^1; & \mathbf{R}_{1,2}^2 &= \mathbf{R}_{1,2}^1 - \mathbf{N}_{2,1}^1 \mathbf{R}_{1,1}^1; & \mathbf{R}_{1,3}^2 &= \mathbf{R}_{1,3}^1 - \mathbf{N}_{3,1}^1 \mathbf{R}_{1,1}^1;
 \end{aligned}
 \tag{4.20}$$

Now we evaluate the  $6 \times 6$  matrix  $(\mathbf{N}_{2,2}^2)^{-1}$  and we use it to do another round of row operations on the  $6 \times 6$  blocks. From eqn. (4.19) we then get

$$\begin{bmatrix} \mathbf{I} & \mathbf{N}_{1,2}^2 & \mathbf{N}_{1,3}^2 \\ 0 & \mathbf{I} & (\mathbf{N}_{2,2}^2)^{-1} \mathbf{N}_{2,3}^2 \\ 0 & 0 & \mathbf{N}_{3,3}^2 - \mathbf{N}_{3,2}^2 (\mathbf{N}_{2,2}^2)^{-1} \mathbf{N}_{2,3}^2 \end{bmatrix} \begin{bmatrix} \delta \mathbf{u}_{1,1} \\ \delta \mathbf{u}_{1,2} \\ \delta \mathbf{u}_{1,3} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{1,1}^2 \\ (\mathbf{N}_{2,2}^2)^{-1} \mathbf{R}_{1,2}^2 \\ \mathbf{R}_{1,3}^2 - \mathbf{N}_{3,2}^2 (\mathbf{N}_{2,2}^2)^{-1} \mathbf{R}_{1,2}^2 \end{bmatrix}
 \tag{4.21}$$

A further solution of the  $6 \times 6$  linear system given by the last block of rows in eqn. (4.21) reduces the above system to a solvable upper triangular system with a trivial solution. In other words, solve on the  $6 \times 6$  linear system  $[\mathbf{N}_{3,3}^2 - \mathbf{N}_{3,2}^2 (\mathbf{N}_{2,2}^2)^{-1} \mathbf{N}_{2,3}^2] \delta \mathbf{u}_{1,3} = \mathbf{R}_{1,3}^2 - \mathbf{N}_{3,2}^2 (\mathbf{N}_{2,2}^2)^{-1} \mathbf{R}_{1,2}^2$  first, then the solution for the rest of eqn. (4.21) becomes trivial. Notice that we have been able to invert the system of equations in eqn. (4.15) with just the inversion of one  $6 \times 6$  matrix and the solution of one  $6 \times 6$  linear system. This is a substantial improvement over the inversion of the  $18 \times 18$  matrix in eqn. (4.15)! The method presented here is very general and applies to any PDE system with stiff source terms. The simplified matrix solution strategy also extends to all orders of temporal accuracy when the ADER schemes from this paper are used. When the permittivity and permeability tensors in CED are diagonal, the inversions simply entail inverting a diagonal matrix, which makes the solution even simpler.

#### 4.4. Stepwise implementation-related details for ADER–WDENO schemes for CED

Here we describe the computer implementation of a spatially and temporally third order accurate ADER–WENO scheme for CED in a pointwise form. The corresponding fourth order ADER–WENO scheme can also be constructed once the steps described here are understood.

- 1) At the time of initializing the code, reconstruct the permittivity and permeability tensors, as well as the conductivities, so that they have the appropriate order of reconstruction. For example, at third order, the material properties should also have the same moments as the ones shown in eqn. (3.2).
- 2) We start with the primal variables shown in Fig. 1. In other words, the primal variables for our method are the face-averaged components of the electric displacement and magnetic induction vector fields.

- 3) Using eqn. (3.1) we build the zone-averaged charge density within each zone. Using that charge density and any reasonable WENO scheme applied to the charge density, we build the higher moments of the charge density within each zone; see eqn. (3.2).
- 4) Using any reasonable WENO scheme applied to the primal variables in the *faces* (Balsara and Shu [5], Balsara et al. [6], Balsara, Garain and Shu [17]) we construct the moments of the primal variables within each face, as shown in eqns. (3.3), (3.4) and (3.5).
- 5) Use the constraint-preserving reconstruction strategy from Appendix A to obtain the coefficients of the eqns. (3.6), (3.7) and (3.8). This gives the electric displacement vector field at any point within all the zones of the mesh. Follow an analogous procedure to obtain the magnetic induction vector field within all the zones.
- 6) Using the reconstructed electric displacement and magnetic induction vector fields, evaluate the set of 27 six-component nodal vectors  $\{\mathbf{w}_j; j = 1, \dots, 27\}$  that are needed in eqn. (4.3). Eqns. (3.6), (3.7) and (3.8), which we specified in the previous step, give us the method for evaluating these input values for the ADER predictor step. We will also need the permittivity and permeability tensors and the conductivities at those nodal points. These too have to be evaluated with suitably high order of accuracy. Initialize the iteration with  $\mathbf{u}_{j,k} = \mathbf{w}_j \forall j = 1, \dots, 27$  and  $k = 1, \dots, 3$ . Also evaluate the coefficients in the fluxes and the source terms in eqns. (4.7), (4.8) and their analogues. As the ADER iterations progress, these flux and source coefficients have to be re-evaluated for each new iteration, of course.
- 7) Eqns. (4.11), (4.12) and (4.13) explicitly provide the update equations at the first nodal point shown in Fig. 2. It is expected that the implementer will obtain analogous equations at all the nodes shown in Fig. 2. (Hopefully, the detailed description of ADER schemes in the electronic supplement to this paper will help. We do request that this paper, and its allied papers (Dumbser et al. [25], [26], Balsara et al. [6], [9]), be cited when the associated ADER scheme is used.) Carry out one iteration of these update equations. Use the methods in Sub-section 4.3 to treat implicit source terms, if such stiff source terms are present. This gives us an improved approximation for the set  $\{\mathbf{u}_{j,k}; j = 1, \dots, 27$  and  $k = 1, \dots, 3\}$  that can be used in the next iteration. Our experience has been that at third order, we never need more than three ADER iterations. At fourth order, we need four ADER iterations. We assume that by the end of this ADER iteration step we have the outputs of the ADER iteration, which are the coefficients  $\mathbf{u}_{j,k}$  in eqn. (4.6). This gives us the entire space and time variation of the primal variables within each zone. This variation is also third order accurate in space and time within each zone.
- 8) Using the outputs from step 7, we can build the time-averaged nodal values that are needed at the edge of each zone. These are one set of outputs from the ADER predictor step and will constitute the inputs for the multidimensional Riemann solver.
- 9) Using the outputs from step 7 again, we can build the space and time-averaged stiff source terms at each face of Fig. 2. As a result, each face will have two space–time averaged, and implicitly treated source terms (current densities) from the two zones that abut it. These are the second set of outputs from the ADER predictor step and will constitute the inputs for the facial source terms. Since these source terms have already been implicitly treated within each zone, we just arithmetically average the source terms from either side of each face in order to get the facial current densities that are to be used in the one-step update shown in eqns. (2.10) and (2.11).
- 10) Within each edge, thanks to Fig. 2, we have already identified three nodal points consistent with the Simpson rule. We wish to apply the multidimensional Riemann solver at each of those nodal points. The multidimensional Riemann solver takes as its input the four electric displacement and magnetic induction vectors from the four zones that surround the edge in question. (Unlike the Runge–Kutta step, these are time-averaged values that we have obtained from step 8.) The material properties, evaluated with sufficient order of accuracy, should also be provided as inputs to the multidimensional Riemann solver. For our purposes, the multidimensional Riemann solver returns as an output the edge-aligned components of the electric and magnetic field vectors as shown in Fig. 1. Please see Section V of Balsara et al. [19], and specifically eqns. (4.6) and (4.7), of Balsara et al. [19] for a detailed description of the multidimensional Riemann solver for CED. Because we have three quadrature points, we can use the weighting from the Simpson rule to obtain a higher order line-averaged value for the electric and magnetic field components within each edge. These line-averaged values are also time-averaged. Consequently, they are just the right edge-averaged electric and magnetic field components that are needed in the one-step update shown in eqns. (2.10) and (2.11).
- 11) The previous two steps have shown us how to obtain the space–time averaged current densities at the faces of the mesh as well as the space–time averaged electric and magnetic field components at the edges of the mesh. Assemble these terms using eqns. (2.10) and (2.11) in order to complete one full ADER–WENO step.

This completes our stepwise description of the ADER–WENO scheme for CEM.

## 5. Results

We first provide several tests that demonstrate that our methods meet their designed order of accuracy. Subsequently, we report on several interesting and stringent test problems. Uniform meshes are used for all of these tests.

**Table 1**

Shows the accuracy analysis for the second order ADER–WENO scheme for the propagation of an electromagnetic wave in vacuum. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{\text{inf}}$ error	$D_y L_{\text{inf}}$ accuracy
$16^2$	7.8656E–05		1.2448E–04	
$32^2$	1.5608E–05	2.33	2.4474E–05	2.35
$64^2$	3.7288E–06	2.07	5.8575E–06	2.06
$128^2$	9.2983E–07	2.00	1.4606E–06	2.00
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{\text{inf}}$ error	$B_z L_{\text{inf}}$ accuracy
$16^2$	4.0557E–02		6.2464E–02	
$32^2$	8.2597E–03	2.30	1.3014E–02	2.26
$64^2$	1.9705E–03	2.07	3.0920E–03	2.07
$128^2$	4.8785E–04	2.01	7.6614E–04	2.01

**Table 2**

Shows the accuracy analysis for the third order ADER–WENO scheme for the propagation of an electromagnetic wave in vacuum. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{\text{inf}}$ error	$D_y L_{\text{inf}}$ accuracy
$16^2$	3.8899E–05		5.7293E–05	
$32^2$	4.5684E–06	3.09	7.0625E–06	3.02
$64^2$	5.4770E–07	3.06	8.5467E–07	3.05
$128^2$	6.7628E–08	3.02	1.0571E–07	3.02
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{\text{inf}}$ error	$B_z L_{\text{inf}}$ accuracy
$16^2$	2.0039E–02		3.0544E–02	
$32^2$	2.4112E–03	3.06	3.7362E–03	3.03
$64^2$	2.9210E–04	3.05	4.5495E–04	3.04
$128^2$	3.6006E–05	3.02	5.6308E–05	3.01

**Table 3**

Shows the accuracy analysis for the fourth order ADER–WENO scheme for the propagation of an electromagnetic wave in vacuum. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{\text{inf}}$ error	$D_y L_{\text{inf}}$ accuracy
$16^2$	1.8613E–06		3.7388E–06	
$32^2$	9.2805E–08	4.33	1.3348E–07	4.81
$64^2$	5.6665E–09	4.03	8.9423E–09	3.90
$128^2$	3.6462E–10	3.96	5.6873E–10	3.97
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{\text{inf}}$ error	$B_z L_{\text{inf}}$ accuracy
$16^2$	1.0631E–03		1.8582E–03	
$32^2$	5.2974E–05	4.33	7.8582E–05	4.56
$64^2$	3.1636E–06	4.07	4.9861E–06	3.98
$128^2$	2.0062E–07	3.98	3.1241E–07	4.00

### 5.1. Propagation of a plane electromagnetic wave in two dimensions

The three-dimensional version of this test has already been reported in Balsara et al. [16] and the two-dimensional version of this test has been reported in Balsara et al. [19]. We do not repeat the problem set-up here. This test problem consists of a plane polarized electromagnetic wave propagating in a vacuum along the north-east diagonal of a two dimensional Cartesian mesh spanning  $[-0.5, 0.5] \times [-0.5, 0.5]$ . Periodic boundary conditions were enforced. The problem was run with a CFL of 0.45 to a final time of unity on the computational mesh.

Table 1 shows the accuracy analysis for the second order ADER–WENO scheme. We see that the scheme meets its designed second order accuracy. Table 2 shows the accuracy analysis for the third order ADER–WENO scheme. We see that the scheme is indeed third order accurate. Table 3 shows the accuracy analysis for the fourth order ADER–WENO scheme. We see that the scheme is indeed fourth order accurate. It is noteworthy that our fourth order scheme, acting on a  $16^2$  mesh can produce an accuracy that is comparable to a second order scheme, acting on a  $128^2$  mesh. It shows the value of the higher order formulations for computational electromagnetism developed in this paper.

**Table 4**

Shows the accuracy analysis for the second order ADER–WENO scheme for the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
$120^2$	4.4861E–05		1.7293E–02	
$240^2$	2.6292E–05	0.77	1.7450E–02	–0.01
$480^2$	8.5367E–06	1.62	5.4298E–03	1.68
$960^2$	1.9295E–06	2.15	1.2868E–03	2.08
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
$120^2$	1.8617E–02		2.7253E+00	
$240^2$	9.3940E–03	0.99	2.9933E+00	–0.14
$480^2$	3.0817E–03	1.61	9.7009E–01	1.63
$960^2$	6.9444E–04	2.15	2.0541E–01	2.24

**Table 5**

Shows the accuracy analysis for the third order ADER–WENO scheme for the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
$120^2$	3.9216E–05		1.7314E–02	
$240^2$	2.3791E–05	0.72	1.7018E–02	0.02
$480^2$	5.6346E–06	2.08	4.0016E–03	2.09
$960^2$	7.8591E–07	2.84	5.4984E–04	2.86
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
$120^2$	1.5523E–02		2.7303E+00	
$240^2$	8.3494E–03	0.89	2.9504E+00	–0.11
$480^2$	2.0493E–03	2.03	7.5348E–01	1.97
$960^2$	2.8917E–04	2.83	1.0480E–01	2.85

**Table 6**

Shows the accuracy analysis for the fourth order ADER–WENO scheme for the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
$120^2$	3.0490E–05		1.7238E–02	
$240^2$	1.4925E–05	1.03	1.0730E–02	0.68
$480^2$	6.2729E–07	4.57	4.3484E–04	4.63
$960^2$	2.5970E–08	4.59	1.6122E–05	4.75
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
$120^2$	1.0819E–02		2.7229E+00	
$240^2$	5.3574E–03	1.01	1.9133E+00	0.51
$480^2$	2.3015E–04	4.54	8.0452E–02	4.57
$960^2$	1.0808E–05	4.41	2.6899E–03	4.90

## 5.2. Compact Gaussian electromagnetic pulse incident on a refractive disk

This problem was drawn from Zakharian et al. (2006) and further refined in Balsara et al. [19]. The physical domain spans  $[-7, 7] \times [-7, 7]$ . It shows the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. The waves inside the Gaussian pulse have a wavelength of 1.5 m. The disk, with a radius of 0.75 m, has a refractive index of 3 relative to the ambient, which has a refractive index of unity. This means that the permittivity in the disk is almost an order of magnitude larger than the permittivity in the ambient. A taper of 0.08 m was applied to the relative permittivity of the disk so that it smoothly goes from a value of 9 inside the disk to a value of 1 outside it. In Balsara et al. [19] we show further figures that show the refraction of the wave by the disk. The figures, as well as the problem set-up are not repeated here.

Tables 4, 5 and 6 show the accuracy analysis for second, third and fourth order accurate ADER–WENO schemes for this problem. Tables 7, 8 and 9 show the corresponding results for second, third and fourth order accurate RK–WENO schemes for this problem. This enables us to compare and contrast between the schemes. The fourth order RK–WENO scheme used fourth order accurate WENO reconstruction in space, but it only used a third order accurate RK scheme for



**Table 7**

Shows the accuracy analysis for the second order RK–WENO scheme for the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
$120^2$	7.2342E–05		1.7361E–02	
$240^2$	4.9696E–05	0.54	2.0295E–02	–0.23
$480^2$	2.2189E–05	1.16	1.1110E–02	0.87
$960^2$	6.1890E–06	1.84	3.5042E–03	1.66
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
$120^2$	3.3603E–02		2.7402E+00	
$240^2$	2.0922E–02	0.68	3.3818E+00	–0.30
$480^2$	8.5658E–03	1.29	1.7471E+00	0.95
$960^2$	2.3723E–03	1.85	5.6219E–01	1.64

**Table 8**

Shows the accuracy analysis for the third order RK–WENO scheme for the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
$120^2$	4.6479E–05		1.7353E–02	
$240^2$	2.6840E–05	0.79	1.8232E–02	–0.07
$480^2$	6.4831E–06	2.05	4.4831E–03	2.02
$960^2$	9.1133E–07	2.83	6.1661E–04	2.86
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
$120^2$	1.9435E–02		2.7331E+00	
$240^2$	9.6392E–03	1.01	3.1132E+00	–0.19
$480^2$	2.3993E–03	2.01	8.4065E–01	1.89
$960^2$	3.4171E–04	2.81	1.1744E–01	2.84

**Table 9**

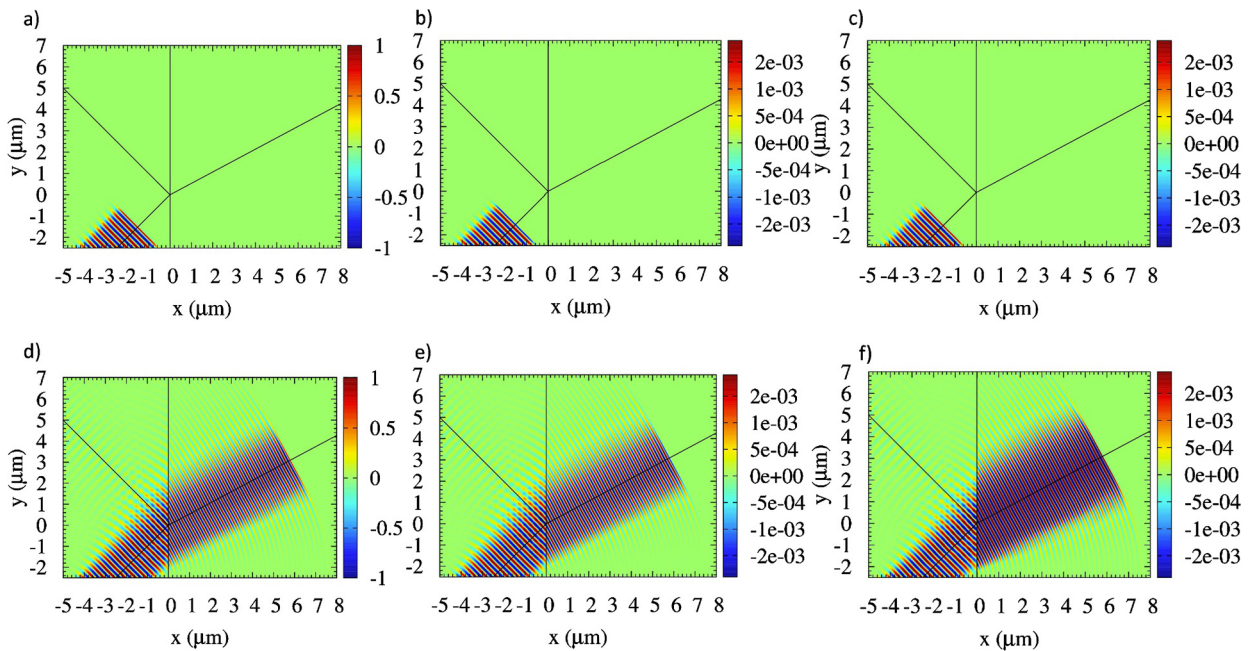
Shows the accuracy analysis for the fourth order RK–WENO scheme for the propagation of a Gaussian pulse of radiation that is incident on a refractive disk. The timestep was decreased with the decreasing mesh size so that  $\Delta t \sim \Delta x^{4/3}$ . The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
$120^2$	3.2473E–05		1.7344E–02	
$240^2$	1.6214E–05	1.00	1.1735E–02	0.56
$480^2$	7.4650E–07	4.44	5.1434E–04	4.51
$960^2$	3.9596E–08	4.24	2.1422E–05	4.59
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
$120^2$	1.1771E–02		2.7269E+00	
$240^2$	5.8246E–03	1.01	2.0886E+00	0.38
$480^2$	2.7220E–04	4.42	9.4284E–02	4.47
$960^2$	1.6876E–05	4.01	3.5698E–03	4.72

the time-integration. As a result, the timestep was reduced with decreasing mesh size using the variation  $\Delta t \sim \Delta x^{4/3}$ . This ensures that the temporal accuracy matches the fourth order spatial accuracy of the RK–WENO scheme. For practical applications, when the timestep is not reduced in this fashion, the spatially fourth order RK–WENO scheme described here will be effectively third order accurate due to its lower order RK algorithm.

From Tables 4 and 5 we see that the second and third order ADER–WENO schemes only reach their design accuracy on a  $960^2$  zone mesh. However, from Table 6 we see that the fourth order ADER–WENO scheme has already reached its design accuracy on a  $480^2$  zone mesh. Furthermore, the fourth order scheme on a  $480^2$  zone mesh is already as accurate as a third order scheme on a  $960^2$  zone mesh. On a  $480^2$  zone mesh, the taper corresponds to  $\sim 3$  grid points. It shows that the fourth order scheme indeed detects the sub-cell variation in the dielectric properties and responds to them. Also notice from Table 5 that the third order scheme is already second order accurate on a  $480^2$  zone mesh. So, the third order scheme is also quite effective at detecting the spatial variation in permittivity and incorporating that variation into the solution.

We can also compare Tables 4, 5 and 6 for ADER–WENO schemes to the corresponding Tables 7, 8 and 9 for RK–WENO schemes. Comparing Table 4 to Table 7 we see that the second order ADER–WENO scheme offers a significant accuracy advantage compared to the second order RK–WENO scheme. Comparing Table 5 to Table 8 we see that the third order



**Fig. 3.** Figs. 3a, 3b, and 3c visualize  $B_z$ ,  $D_x$ , and  $D_y$  at the initial time. Figs. 3d, 3e, and 3f show the same at a final time of  $4 \times 10^{-14}$  s. The surface of the dielectric slab is identified by the vertical black line. The oblique black lines demarcate the angle of incidence, the angle of refraction and the angle of reflection. (For interpretation of the colors in this figure, the reader is referred to the web version of this article.)

ADER–WENO scheme has accuracy that is comparable to the third order RK–WENO scheme. If we nominally compare Table 6 to Table 9 we would see that the same trend is repeated at fourth order. However, recall that the fourth order RK–WENO code had to use a progressively reduced CFL as the mesh was refined, and this is a significant disadvantage for practical computation.

### 5.3. Refraction of a compact electromagnetic beam by a dielectric slab

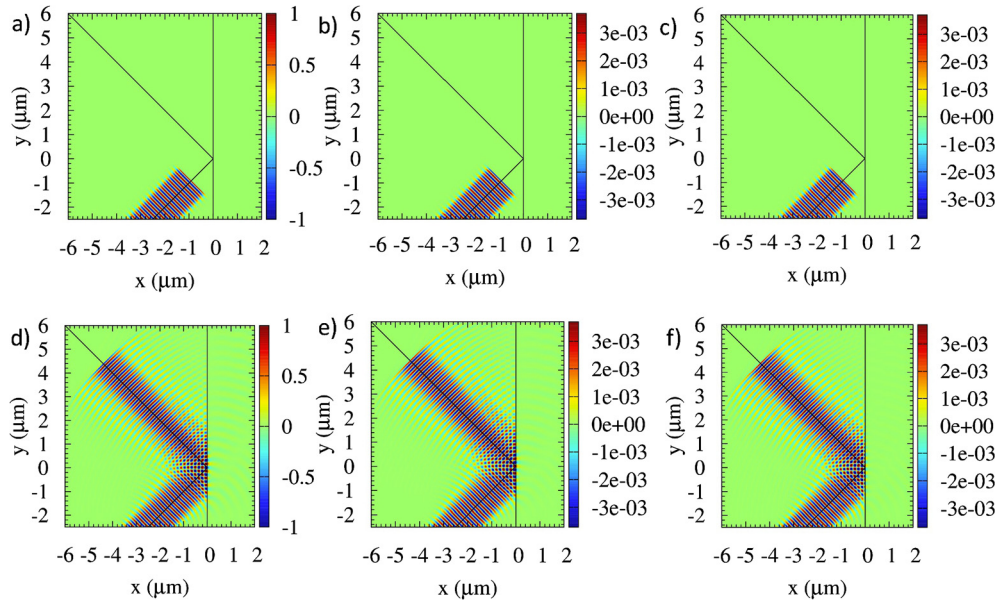
This problem was described in detail in Balsara et al. [19]), so we do not repeat the description here. It is set up on a rectangular  $xy$ -domain that spans  $[-5, 8] \times [-2.5, 7]$   $\mu\text{m}$ . For the simulation shown, we use a fourth-order-accurate scheme using a  $1300 \times 950$  zone mesh. The permittivity increases in a tapered fashion from  $\epsilon_0$  for  $x < 0$  to  $2.25\epsilon_0$  for  $x > 0$ . A compact electromagnetic beam is incident on the slab at an angle of incidence given by  $45^\circ$ .

Figs. 3a, 3b, and 3c show, respectively,  $B_z$ ,  $D_x$ , and  $D_y$  at the initial time. Figs. 3d, 3e, and 3f show the same at the final time of  $4.0 \times 10^{-14}$  s. The surface of the dielectric slab is shown by a vertical black line. The inclined black lines show the angles of incidence, refraction and reflection and these black lines are over-plotted on the field components to guide the eye. According to Snell's law, the angle of refraction should be  $28.12^\circ$  since the refractive index of the dielectric slab is 1.5. We see that the code obtains the correct angle of refraction. We also observe that some of the radiation is reflected from the surface of the slab. The presence of a reflected wave is consistent with the Fresnel conditions for transmission and reflection of radiation at dielectric interfaces.

### 5.4. Total internal reflection of a compact electromagnetic beam by a dielectric slab

This problem was also described in detail in Balsara et al. [19], so we do not repeat the description here. This problem is set up on a rectangular  $xy$ -domain that spans  $[-6, 2] \times [-2.5, 6]$   $\mu\text{m}$ . For the simulation shown, we use a third-order-accurate scheme using a  $800 \times 850$  zone mesh. Here,  $\epsilon$  is chosen such that it has a value of  $4.0\epsilon_0$  for  $x \leq 0$  and tapers rapidly to the ambient value of  $\epsilon_0$  for  $x > 0$ . This value of permittivity for  $x < 0$  implies a refractive index of 2 for the dielectric slab. For this mesh, the taper width that is applied to the variation in the permittivity is 0.25 times a zone width. The wavelength in the dielectric medium corresponds to about 30 zones.

For such a slab, the critical angle for total internal reflection is  $30^\circ$ . As in the previous Sub-section, the angle of incidence of the incident radiation is  $45^\circ$ , with the result that the incident radiation will undergo total internal reflection. Figs. 4a, 4b, and 4c show, respectively,  $B_z$ ,  $D_x$ , and  $D_y$  at the initial time. Figs. 4d, 4e, and 4f show the same at the final time of  $5.0 \times 10^{-14}$  s. The surface of the dielectric slab is shown by a vertical black line. The inclined black lines for the incident and reflected rays are over-plotted on the field components to guide our eye. We see from the figures that the radiation undergoes total internal reflection, as expected.



**Fig. 4.** Figs. 4a, 4b, and 4c visualize  $B_z$ ,  $D_x$ , and  $D_y$  at the initial time. Figs. 4d, 4e, and 4f show the same at final time of  $5 \times 10^{-14}$  s. The surface of the dielectric slab is identified by the vertical black line. The oblique black lines demarcate the angle of incidence and the angle of total internal reflection. (For interpretation of the colors in this figure, the reader is referred to the web version of this article.)

It is also possible to go further in one's study of total internal reflection. This interest is warranted because total internal reflection plays an important role in the design of optical fibers for carrying digital data and also in the design of nanophotonic devices. Let  $\theta_{inc}$  be the angle of incidence for the total internally refracting ray in Fig. 4. Let the optically dense material to the left of the black lines in Fig. 4 have a refractive index of  $n_1$  and let the relatively less dense material to the right of the black lines in Fig. 4 have a refractive index of  $n_2$ . While the radiation undergoes total internal reflection in the sense of geometrical optics, there exists an evanescent wave in the less optically dense medium "2". In this region, the electromagnetic field exhibits two properties. First, the  $B_z$  (and also the  $D_y$ ) components of the radiation in medium "2" should die off exponentially with an exponential decay length given by

$$\delta = \frac{\lambda_0}{2\pi \sqrt{n_1^2 \sin^2(\theta_{inc}) - n_2^2}}$$

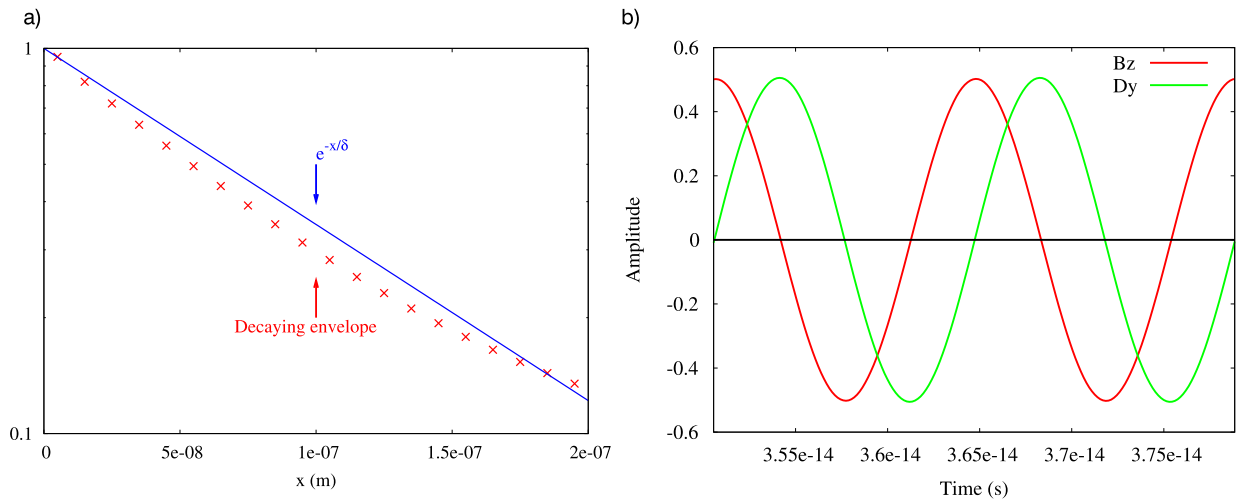
In the above equation,  $\lambda_0$  is the free space wavelength. For the values chosen in this problem, we have an exponential decay length given by  $\delta = 0.95 \times 10^{-7}$  m. Second, realize that the flux of radiant energy into medium "2" should average to zero. This is tantamount to requiring that the variation of  $B_z$  and  $D_y$  in medium "2" should be  $90^\circ$  out of phase. That is the only way that the time-average of the Poynting flux can be zero. These two pieces of physics are important markers that demonstrate the quality of a numerical scheme for this technologically important application.

Fig. 5a shows the envelope of the amplitude of  $B_z$  in air as a function of distance from the dielectric interface. The amplitude is expected to decay exponentially which is why we use a log-linear plot. The red crosses show the data while the blue solid line shows the theoretical exponential decay. A distance of two exponential decay lengths from the interface is shown. Fig. 5b shows the variation of  $B_z$  and  $D_y$  as a function of time at a station in the air. For there to be no energy transfer into the air from the total internal reflection, the two field variables should be  $90^\circ$  out of phase; a result which is confirmed by Fig. 5b.

### 5.5. Compact electromagnetic beam impinging on a conducting slab

This test problem was also described in Balsara et al. [19] and is intended to demonstrate the ability of our method to handle gigantic variations in conductivity over very short distances. The problem is set up on a rectangular  $xy$ -domain that spans  $[-25, 5] \times [-40, 40]$  m. Please see the cited paper for further details. For the simulations shown, we use a fourth-order-accurate scheme using a  $750 \times 2000$  zone mesh. A copper slab is located in the domain  $x \geq 0$  and the conductivity goes from zero for  $x < 0$  to a conductivity of  $\sigma_0 = 5.9 \times 10^7$  S/m for  $x > 0$  with a taper of one zone! This represents a gigantic change in the stiffness of the equations. We realize that the skin depth is unresolved in our simulations and we also realize that a perfect electrical conductor (PEC) surface in CED would be treated differently. Of course, we expect the radiation to be reflected at the copper surface.

Figs. 6a, 6b, and 6c show, respectively,  $B_z$ ,  $D_x$ , and  $D_y$  at the initial time. Figs. 6d, 6e, and 6f show the same at the final time of  $1.83 \times 10^{-7}$  s. The surface of the conducting slab is shown by a black line. It is interesting to observe that with very



**Fig. 5.** Fig. 5a shows the envelope of the amplitude of  $B_z$  in air as a function of distance from the dielectric interface. The amplitude is expected to decay exponentially which is why we use a log-linear plot. The red crosses show the data while the blue solid line shows the theoretical exponential decay. Fig. 5b shows the variation of  $B_z$  and  $D_y$  as a function of time at a station in the air. For there to be no energy transfer into the air from the total internal reflection, the two field variables should be 90 degrees out of phase; a result which is confirmed by Fig. 5b. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reasonable resolutions, most of the incident intensity in the radiation is reflected back at the copper surface. Fig. 7 shows the interference pattern that develops between the incident and reflected radiation.

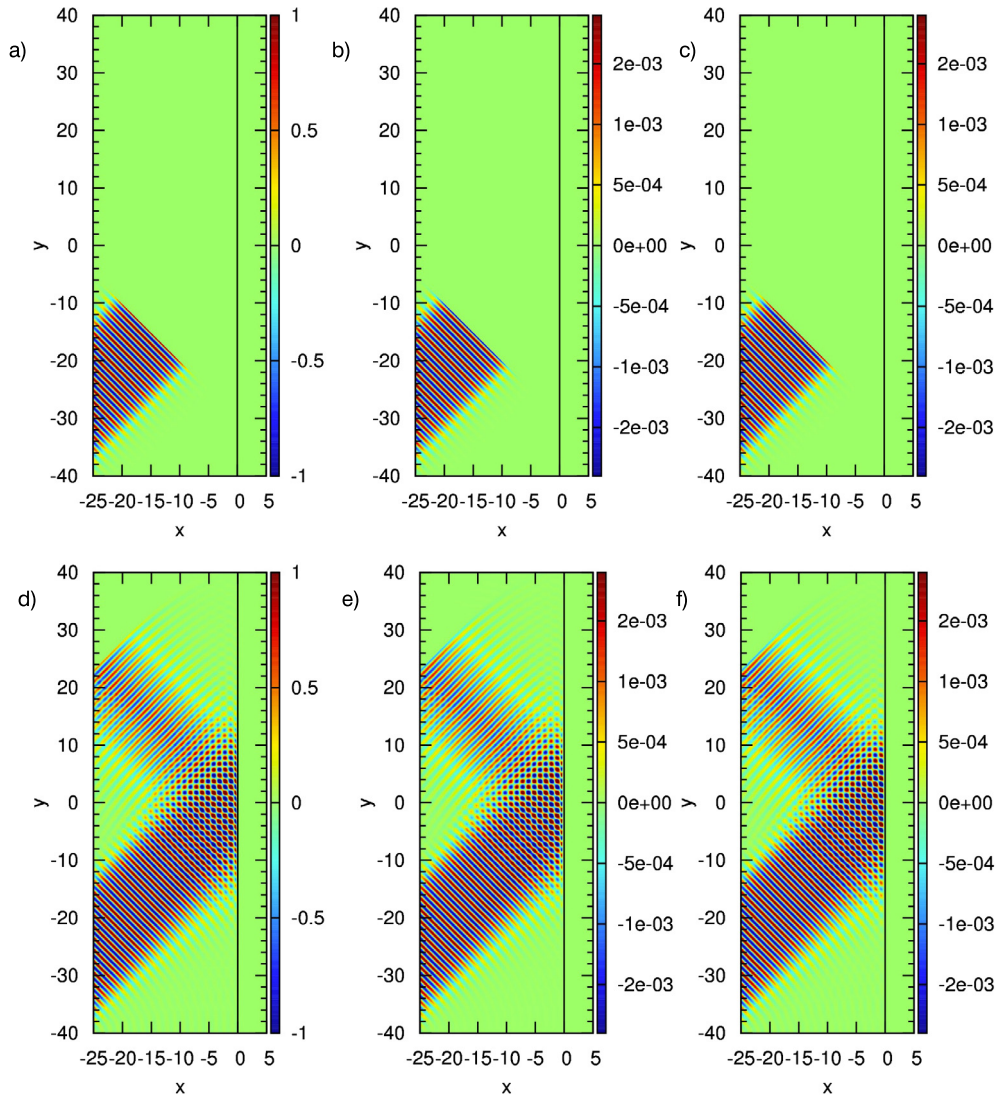
### 5.6. Decay of a sinusoidal wave inside a conductor

The previous problem did not resolve the skin depth of the conductor. In this one-dimensional problem we resolve the skin depth of the conductor. This test problem was described in Balsara et al. [19] and we do not repeat the description here. The problem is set up on a 1D domain that spans  $[0, 10\delta]$  where  $\delta$  is the skin depth of the conductor. For the simulations shown, we use a second-order-accurate scheme using a 100 zone mesh. Thus, we use 10 zones per skin depth for this simulation. Two simulations were carried out. In our first simulation we used a conductivity of  $\sigma = 2.0 \times 10^3$  S/m which corresponds to the conductivity of amorphous carbon. In our second simulation we used  $\sigma = 5.9 \times 10^7$  S/m which corresponds to the conductivity of copper. For carbon, we choose  $\nu = 1.679 \times 10^{13}$  Hz which gives  $\delta = 3.44 \times 10^{-6}$  m and a wavelength of the plane wave as  $\lambda = 1.38 \times 10^{-5}$  m inside this material. We stop this simulation at a time of  $2.38 \times 10^{-13}$  s. For copper, we choose  $\nu = 1.0 \times 10^{13}$  Hz which gives  $\delta = 2.06 \times 10^{-8}$  m and a wavelength of the plane wave as  $\lambda = 1.3 \times 10^{-7}$  m inside this material. We stop this simulation at a time of  $2.0 \times 10^{-13}$  s.

We will now show that the numerical results accurately match the analytic results even when the skin depth is resolved with no more than ten zones. A fourth-order-accurate ADER-WENO scheme was used. Figs. 8a and 8b show the variations of  $B_z$  with radial distance inside carbon and copper, respectively. The exponentially decaying envelopes are also over-plotted. In Figs. 8c and 8d we present the structure of the numerical envelopes (red crosses) and the analytical envelopes (blue curve) on a semi-log scale for carbon and copper, respectively. The numerically-obtained envelopes match extraordinarily well with the analytical envelopes. Fig. 10 of Balsara et al. [19] shows the same plot at the same resolution for a second order FVDT scheme for CED. We see a dramatic improvement in our fourth order data when it is compared to the second order data for this same problem.

### 5.7. Accuracy analysis for the decay of a sinusoidal wave in a conductor – demonstrating accuracy in the presence of extremely stiff source terms

Sometimes, methods that are designed to retain accuracy for hyperbolic problems without source terms do not retain the same accuracy when stiff source terms are present. We demonstrate that this is not the case for our methods. We analyze the accuracy of wave propagation in a copper conductor. We focus on just the portion of the data that is shown in Fig. 8d and we compare the numerically-derived field components with the analytical field components in order to extract an order of accuracy. Please note, however, that the accuracy analysis was done somewhat differently. When performing an accuracy analysis, we provide the solution as a function of time on the ghost zones with  $x < 0$  and also on the ghost zones with  $x > 1.0 \times 10^{-7}$  m. For the active computational domain with  $0 \leq x \leq 1.0 \times 10^{-7}$  m we only initialize the solution at time  $t = 0$  and we measure the error within the dynamically active part of the mesh at a time of  $t = 4.0 \times 10^{-13}$  s. This corresponds to a situation where the wave propagates through the computational domain about five times.

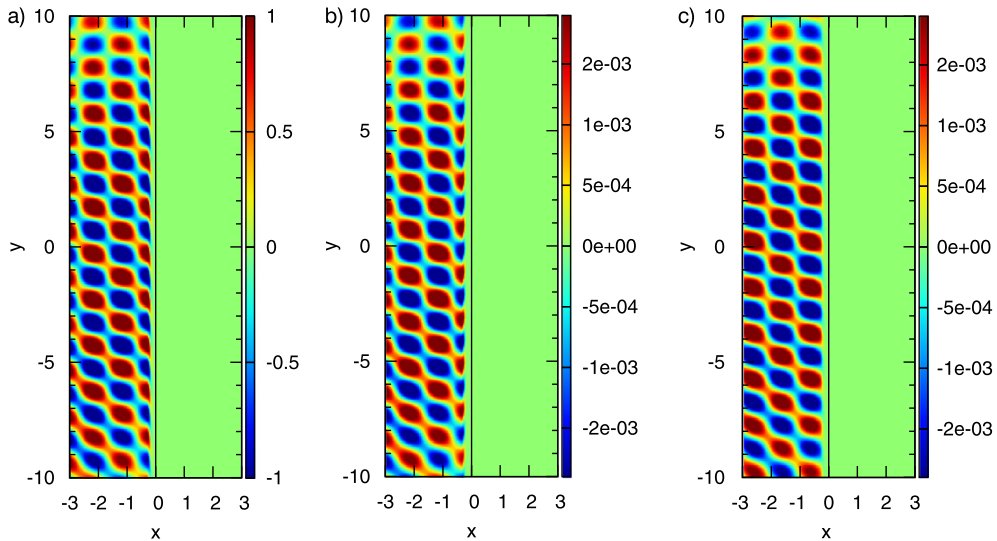


**Fig. 6.** Figs. 6a, 6b and 6c show the initial configuration of  $B_z$ ,  $D_x$  and  $D_y$  for the test problem where a beam of radiation is incident upon a conductor at  $45^\circ$ . Figs. 6d, 6e and 6f show the same variables at the final time of  $1.83 \times 10^{-7}$  s when the beam has reflected off the surface of the conductor. The conductor is located at  $x=0$  in the figure and is shown by a black line. (For interpretation of the colors in this figure, the reader is referred to the web version of this article.)

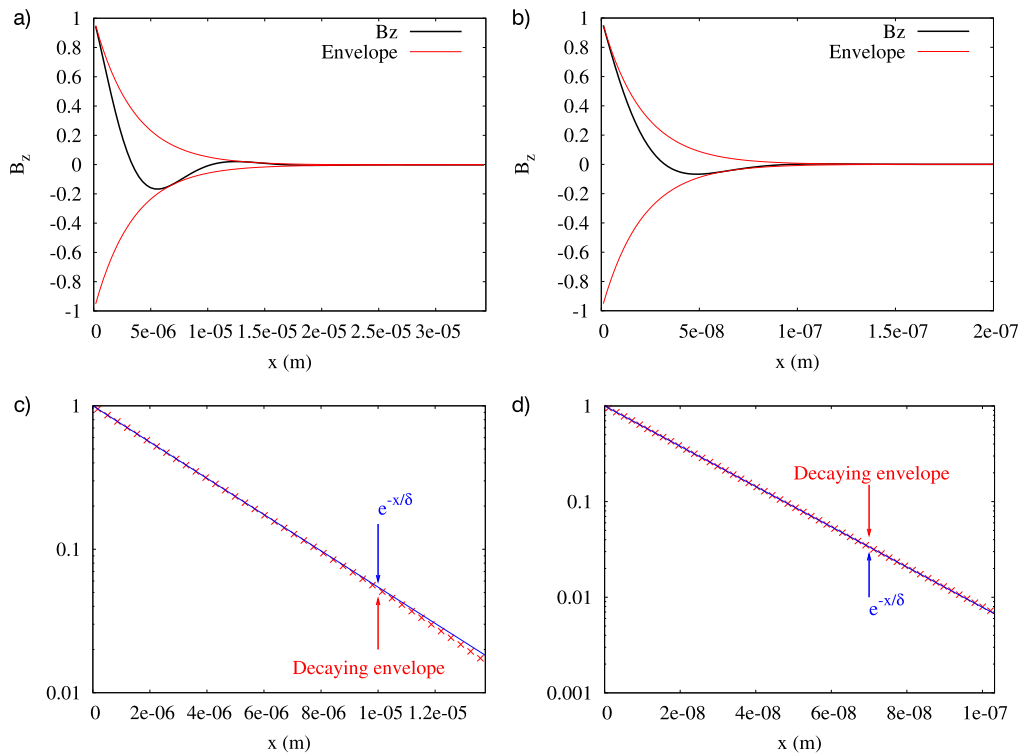
Tables 10, 11 and 12 show the accuracy, measured in the  $L_1$  and  $L_\infty$  norms, for second, third and fourth orders respectively. The ADER–WENO schemes were used for all the reported accuracy tests. The second and fourth order schemes show a serendipitous increase in accuracy. However, as in Sub-section 5.2, we see that all the ADER–WENO schemes achieve their design accuracies; and this accuracy requirement is met even when extremely stiff source terms are present. We also see that the higher order schemes reach their design accuracies on a smaller number of mesh points than the lower order schemes. This reinforces our conclusion that higher order constraint-preserving FVTD schemes are very versatile and valuable schemes for CED.

### 5.8. Long-distance or long-time propagation of electromagnetic radiation

Several problems in CED tend to involve scattering. In such problems, radiation has to propagate with minimal amplitude or phase change over many mesh points before it interacts with a scattering structure. As a result, it is interesting to demonstrate that a method for CED can propagate electromagnetic radiation over many thousands of mesh points with minimal dissipative or dispersive errors. Propagation along mesh lines, or at  $45^\circ$  to the mesh, is quite easy for most schemes. For that reason, this problem is designed to propagate radiation at an angle that is different from the above-mentioned favorable angles. Instead of propagating the radiation over a large mesh, we propagate it over a square mesh with  $120 \times 120$  zones and periodic boundaries. The radiation is propagated over twenty orbits, so that this is equivalent to making it



**Fig. 7.** Figs. 7a, 7b and 7c show a zoom-in of Figs. 6d, 6e and 6f and these figures are centered on the reflection region, spanning the sub-domain  $[-3, 3] \times [-10, 10]$  m. The location of the midpoint in the taper for the conductivity is again shown by a black line. (For interpretation of the colors in this figure, the reader is referred to the web version of this article.)



**Fig. 8.** Figs. 8a and 8b show the radial variations of  $B_z$  (black lines) and the decaying envelopes (red lines) inside carbon and copper, respectively. Figs. 8c and 8d present the structure of the numerically-obtained envelopes (red crosses) and the analytical envelopes (blue curve) on a semi-log scale for carbon and copper, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

propagate over thousands of zones. Consequently, long time propagation over a periodic mesh is equivalent to long distance propagation over a non-periodic mesh.

This problem is closely analogous to a test problem proposed in Balsara [3] to demonstrate the long-time propagation of torsional Alfvén waves in MHD. We use a  $120 \times 120$  zone mesh with uniform zones. The computational domain spans  $[-r/2, r/2] \times [-r/2, r/2]$  in the  $xy$ -plane with  $r = 6$ . The wave is made to propagate at an angle given by  $\tan^{-1}(1/r) =$

**Table 10**

Shows the accuracy analysis for the second order ADER–WENO scheme for the decay of a sinusoidal wave inside a copper conductor. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones per skin depth	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
5	5.1861E–07		1.0174E–06	
10	5.8877E–08	3.14	1.2488E–07	3.03
20	7.0982E–09	3.05	1.8261E–08	2.77
40	1.0610E–09	2.74	2.5650E–09	2.83
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
5	4.9259E–02		1.0510E–01	
10	7.0452E–03	2.81	1.7416E–02	2.59
20	9.3963E–04	2.91	2.3337E–03	2.90
40	1.1686E–04	3.01	2.8815E–04	3.02

**Table 11**

Shows the accuracy analysis for the third order ADER–WENO scheme for the decay of a sinusoidal wave inside a copper conductor. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones per skin depth	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
5	3.7217E–07		8.5076E–07	
10	4.2162E–08	3.14	9.1763E–08	3.21
20	4.8410E–09	3.12	1.2439E–08	2.88
40	7.6775E–10	2.66	1.8599E–09	2.74
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
5	3.3262E–02		7.2717E–02	
10	4.7679E–03	2.80	1.1811E–02	2.62
20	6.3506E–04	2.91	1.5716E–03	2.91
40	8.0823E–05	2.97	1.9900E–04	2.98

**Table 12**

Shows the accuracy analysis for the fourth order ADER–WENO scheme for the decay of a sinusoidal wave inside a copper conductor. A CFL of 0.45 was used. The errors and accuracy in the  $y$ -component of the electric displacement vector and  $z$ -component of the magnetic induction are shown.

Zones per skin depth	$D_y L_1$ error	$D_y L_1$ accuracy	$D_y L_{inf}$ error	$D_y L_{inf}$ accuracy
5	1.0115E–08		2.2503E–08	
10	2.4695E–10	5.36	6.2465E–10	5.17
20	7.1326E–12	5.11	1.6365E–11	5.25
40	2.9685E–13	4.59	6.9852E–13	4.55
Zones	$B_z L_1$ error	$B_z L_1$ accuracy	$B_z L_{inf}$ error	$B_z L_{inf}$ accuracy
5	9.1428E–04		2.5348E–03	
10	3.1764E–05	4.85	8.5105E–05	4.90
20	1.0120E–06	4.97	2.6776E–06	4.99
40	3.1178E–08	5.02	8.2333E–08	5.02

$\tan^{-1}(1/6) = 9.462^0$  with respect to the  $y$ -direction of the mesh. The direction of wave propagation is along the unit vector given by

$$\hat{\mathbf{n}} = n_x \hat{\mathbf{x}} + n_y \hat{\mathbf{y}} = \frac{1}{\sqrt{r^2 + 1}} \hat{\mathbf{x}} + \frac{r}{\sqrt{r^2 + 1}} \hat{\mathbf{y}}$$

The phase of the wave is taken to be

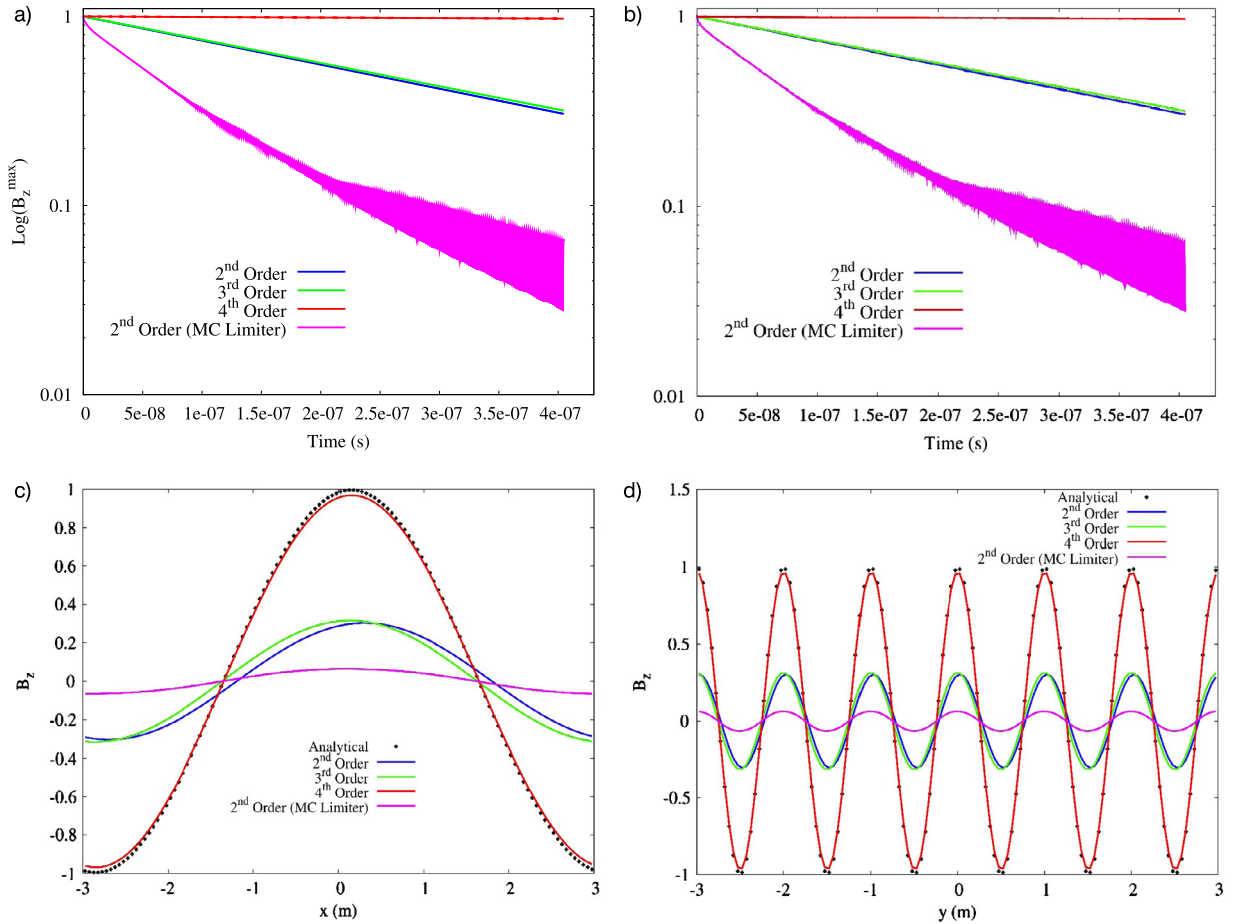
$$\phi = \frac{2\pi}{n_y} (n_x x + n_y y - ct)$$

where “ $c$ ” is the speed of light. The vector potentials for the electric displacement and the magnetic induction are given by

$$\mathbf{C}(x, y, z, t) = -\frac{n_y}{2\pi} \sin(\phi) \hat{\mathbf{z}}$$

and

$$\mathbf{A}(x, y, z, t) = \frac{n_y}{2\pi n_x} \sin(\phi) \hat{\mathbf{y}}$$



**Fig. 9.** Figs. 9a and 9b show the maximum of the Log (base 10) of the amplitude of  $B_z$  and  $D_y$  as a function of time. Fig. 9c shows the variation in  $B_z$  as a function of  $x$  along the line  $y=0$  at the final time. Fig. 9d shows the variation in  $B_z$  as a function of  $y$  along the line  $x=0$  at the final time. In Figs. 9c and 9d, the black crosses show the exact solution. The blue curve shows the second order result with our ADER–WENO scheme; the green curve shows the third order result with our ADER–WENO scheme; the red curve shows the fourth order result with our ADER–WENO scheme; the magenta curve shows the result from a scheme that used the MC limiter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The components of the electric displacement vector field are obtained at the zone faces by using the relationship  $\mathbf{D} = c\epsilon_0(\nabla \times \mathbf{C})$ . The components of the magnetic induction vector field are obtained at the zone faces by using the relationship  $\mathbf{B} = \nabla \times \mathbf{A}$ . The exact expressions for the electric displacement vector field and the magnetic induction vector field are given by

$$\mathbf{D} = c\epsilon_0(-n_y \cos(\phi)\hat{x} + n_x \cos(\phi)\hat{y})$$

and

$$\mathbf{B} = \cos(\phi)\hat{z}$$

In the  $y$ -direction, each sinusoidal variation occupies 20 zones, which makes this a moderate resolution problem. The problem was run to a final time of  $4.05 \times 10^{-7}$  s with a CFL of 0.45. This final time corresponds to 20 orbits on the periodic mesh, which is equivalent to propagating the electromagnetic wave over 2400 zones of a uniform mesh.

Figs. 9a and 9b show the maximum of the Log (base 10) of the amplitude of  $B_z$  and  $D_y$  as a function of time. We see that the scheme that uses second order accurate MC-limiting, and is shown by the magenta curve, has very poor wave propagation characteristics and much of the power of the radiation is dissipated by the end of the simulation. Now please focus on the blue and green curves in Figs. 9a and 9b. Our second order WENO reconstruction uses the linear part of the  $r=3$  WENO, which means that it uses many elements of a third order WENO reconstruction. For that reason, we see that the second and third order ADER–WENO schemes in Figs. 9a and 9b show comparable decay. The red curve in Figs. 9a and 9b shows the results from the fourth order ADER–WENO scheme. We see that the waves have been propagated by the fourth order ADER–WENO scheme with almost no reduction in amplitude, indicating that they have been propagated with minimal dissipation. Fig. 9c shows the variation in  $B_z$  as a function of  $x$  along the line  $y=0$  at the final time. Fig. 9d shows the



variation in  $B_z$  as a function of  $y$  along the line  $x = 0$  at the final time. In Figs. 9c and 9d, the black crosses show the exact analytical solution. The blue curve shows the second order result with our ADER–WENO scheme; the green curve shows the third order result with our ADER–WENO scheme; the red curve shows the fourth order result with our ADER–WENO scheme; the magenta curve shows the result from a scheme that used the MC limiter. Of course, the MC-based scheme is clearly inferior. We also see that the results from the second order ADER–WENO scheme have a small phase shift relative to the results from the third order ADER–WENO scheme. This shows that while the two schemes have comparable dissipation, the second order ADER–WENO scheme has higher dispersion error than its higher order counterparts. We also see that the results from the fourth order ADER–WENO scheme have minimal dispersion and minimal dissipation because they match up almost perfectly with the analytical results shown by the black crosses in Figs. 9c and 9d.

We have also conducted a similar long-distance wave-propagation test using a standard, second-order accurate, open-region FDTD grid employing total-field/scattered-field grid zoning and a perfectly matched layer (PML) absorbing outer boundary. We found that standard FDTD propagates a sinusoidal wave without observable dissipation. However, using the parameters specified for the wave-propagation test results shown in Fig. 9, the FDTD-computed spatial oscillation period was 0.997 times the exact value, and the corresponding propagating wave exhibited a significant phase lag relative to the exact solution. If directly compared to the results shown in Fig. 9c, the FDTD-computed result would be represented by a unit-amplitude sinusoid shifted to the left relative to the analytical result by 2.1 m; and in Fig. 9d, the sinusoids would be shifted to the left by 0.35 m. These results are consistent with previous Fourier-based analyses of FDTD numerical wave propagation (for example, see Chapter 4 of Taflov and Hagness [43]).

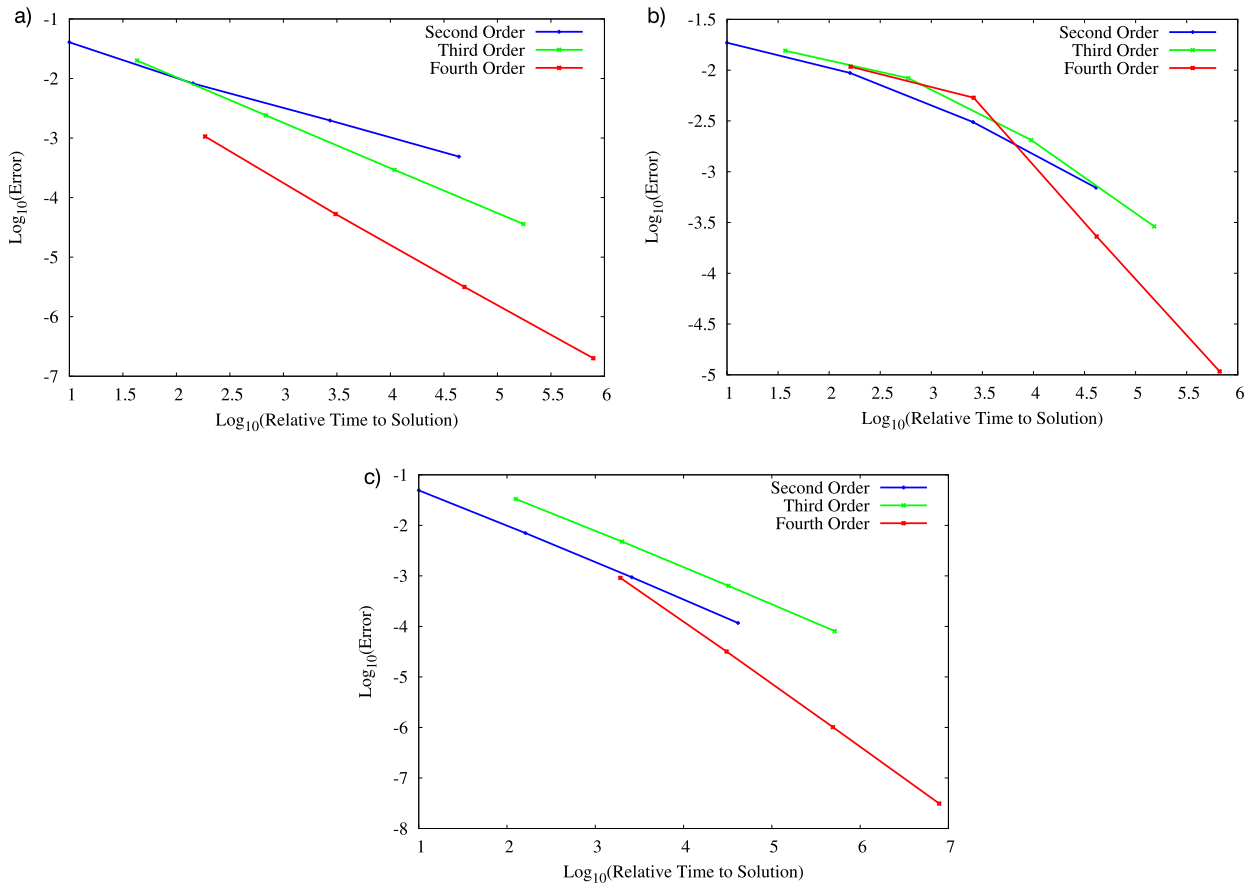
## 6. The value of higher order schemes in CED – error versus time to solution

The third order ADER–WENO scheme has a computational complexity that is 3.7 times that of a second order ADER–WENO scheme on the same mesh. Similarly, the fourth order ADER–WENO scheme has a computational complexity that is 4.3 times that of a third order ADER–WENO scheme on the same mesh. A naïve assessment would, therefore, indicate that it is more advantageous to use a second order scheme in all circumstances; but such an assessment is entirely specious. What matters in a practical scientific or engineering calculation is one's ability to obtain an answer up to a specified accuracy in the shortest amount of time. It is also true that one may want to carry out several survey simulations with reduced accuracy. However, once the problem has been optimized with inexpensive survey simulations, one would want a few simulations at extremely high accuracy. What matters in practical computation is the ability to meet a specified error tolerance with a simulation that minimizes the time to solution. Please also recall that our second order WENO reconstruction uses the linear part of the  $r = 3$  WENO, which means that it uses many elements of a third order WENO reconstruction. As a result, we expect that the second order ADER–WENO scheme will deliver a result that is moderately comparable to the third order ADER–WENO scheme.

Figs. 10a, 10b and 10c show the Log (base 10) of the error versus the Log (base 10) of the relative time to solution for the second, third and fourth order ADER–WENO schemes. We apply the schemes to several of the test problems documented here. Fig. 10a catalogues the error versus the time to solution for the propagation of a plane electromagnetic wave in vacuum from Subsection 5.1 Fig. 10b documents analogous data for the compact Gaussian pulse that is incident on a refractive disk from Subsection 5.2. Fig. 10c shows similar data for the decay of a sinusoidal wave in a copper conductor from Subsection 5.7. The time to solution does depend on the dimensionality of the problem. In practice one wants to do large three-dimensional simulations where a doubling of the resolution increases the computational cost by a factor of 16. While the test problems are two-dimensional, we have used this three-dimensional scaling of computational cost because that is more meaningful in a modern context where large three-dimensional computations are the norm. In all instances, we see that if the accuracy requirements are not stringent, then the second order scheme obtains the solution in the shortest amount of time. However, Figs. 10a and 10b at least show us that at some point, the third order scheme reaches more stringent accuracies with a shorter time to solution. If one only had access to a second order and a third order scheme then at some stringent accuracy requirement the third order scheme would become the scheme of choice. Figs. 10a, 10b and 10c all show us that when the accuracy requirements become very stringent, the fourth order scheme provides the shortest time to solution for a specified accuracy. Thus when all three schemes are available, the fourth order scheme will become the scheme of choice when the scientific or technological problem requires high accuracy. Because of the serendipitous improvement in accuracy in Subsection 5.7, the third order scheme is never an optimal scheme in Fig. 10c. However, we hope that Fig. 10 makes a compelling case for higher order schemes in CED.

## 7. Conclusions

FDTD and FVTD schemes for CED have evolved along disjoint pathways for a while. FDTD does not use reconstruction strategies or Riemann solvers. However, via an exact representation of the curl-type equations on a Yee mesh, it preserves the divergence constraints. FVTD methods do use reconstruction strategies as well as Riemann solvers. However, at least in the past, they have used zone-centered collocations which preclude a preservation of the divergence constraints. In Paper I of this series a synthesis of the two approaches was provided at second order of accuracy. The FVTD schemes that we presented use a staggered mesh formulation, resulting in an exact preservation of the constraints, just like FDTD schemes. However, FDTD schemes have also been available at fourth order of accuracy. Unfortunately, the fourth order variants of



**Fig. 10.** Figs. 10a, 10b and 10c show the Log (base 10) of the error versus the Log (base 10) of the relative time to solution for the second, third and fourth order ADER-WENO schemes. We apply the schemes to several of the test problems documented here. Fig. 10a catalogues the error versus the time to solution for the propagation of a plane electromagnetic wave in vacuum from sub-Section 5.1 Fig. 10b documents analogous data for the compact Gaussian pulse that is incident on a refractive disk from Sub-section 5.2. Fig. 10c shows similar data for the decay of a sinusoidal wave in a copper conductor from Sub-section 5.7.

FDTD schemes for CED have not seen much practical use because they do not preserve accuracy at boundaries where the dielectric properties undergo significant variation. In contrast, the third and fourth order FVTD schemes for CED that we present here can function robustly and meet their design accuracies even when the permittivity and permeability undergo significant variation over a very small number of zones. Furthermore, the methods reported here extend to all orders of accuracy, resulting in a significant expansion of capabilities in CED computations. The methods also take well to problems with large variations in conductivity as well as substantial variations in permittivity and permeability, with the consequence that CED simulations with high order of accuracy involving realistic materials are now within range. This focus on high accuracy, in situations involving realistic materials, is one of the major advances reported for the first time in this paper.

The higher order FVTD methods we devise here are completely general. They can accommodate to any tensorial structure for the permittivity and permeability tensors; though a detailed demonstration of this property will be presented in a subsequent paper. In the current paper we have also shown that our higher order FVTD methods can function robustly even when there are gigantic variations in the conductivity spread over only a few zones. This demonstration of a high order FVTD method that can resolve sub-cell variation in the material properties while retaining its designed order of accuracy is an extremely important, useful and novel contribution arising from this paper. All the higher order FVTD methods presented here accurately preserve the same constraints that are preserved by FDTD schemes.

Three fundamental advances make the present innovations possible. First, we have designed constraint-preserving reconstruction strategies for CED that have been extended to third and fourth orders of accuracy in Section 3 and Appendices A and B. Second, in Balsara et al. [19] we have already described multidimensional Riemann solvers for CED and in this paper we show how they are used within the context of high accuracy schemes. The high order of spatial accuracy has to be matched with high order of temporal accuracy and our third advance consists of offering temporally higher order strategies for advancing the CED equations. This third innovation is presented here in two variations. In the first variation, which we report in the electronic supplement to this paper, we adapt Runge–Kutta IMEX schemes to CED. However, Runge–Kutta IMEX schemes restrict us to second and third orders of accuracy in time. In the second variation, which we report in Section 4,

we innovate ADER schemes for CED that can provide a treatment of stiff source terms at any order. These advances in ADER schemes should also be very useful for treating any general hyperbolic system with stiff source terms. The stiff source terms have to be treated implicitly, however, their presence does not cause any degradation in the CFL that can be used.

We also present a suite of stringent tests to show that our constraint-preserving FVTD methods work and meet their design goals even when material permittivity and permeability vary by an order of magnitude over just a few zones. Furthermore, since our methods are unconditionally stable and sub-cell-resolving in the presence of stiff source terms (i.e. in problems involving giant variations in conductivity over just a few zones), they can accurately handle such problems without any reduction in timestep. We also show that increasing the order of accuracy offers distinct advantages for resolving sub-cell variations in material properties. Most importantly, we show that when the accuracy requirements are stringent the higher order schemes offer the shortest time to solution. This makes a compelling case for the use of higher order, sub-cell resolving schemes in CED.

## Acknowledgements

DSB acknowledges support via NSF grants NSF-DMS-1361197, NSF-ACI-1533850, NSF-DMS-1622457, NSF-ACI-1713765. Several simulations were performed on a cluster at UND that is run by the Center for Research Computing. Computer support on NSF's XSEDE and Blue Waters computing resources is also acknowledged.

## Appendix A

The constraint equation given by eqn. (3.9) gives us six constraints at second order. They are given by

$$\begin{aligned} 3a_{xxx}/\Delta x + b_{xxy}/\Delta y + c_{xxz}/\Delta z &= q_{xx}; & a_{xyy}/\Delta x + 3b_{yyy}/\Delta y + c_{yyz}/\Delta z &= q_{yy}; \\ a_{xzz}/\Delta x + b_{yzz}/\Delta y + 3c_{zzz}/\Delta z &= q_{zz}; & 2a_{xxy}/\Delta x + 2b_{xyy}/\Delta y + c_{xyz}/\Delta z &= q_{xy}; \\ a_{xyz}/\Delta x + 2b_{yyz}/\Delta y + 2c_{yzz}/\Delta z &= q_{yz}; & 2a_{xxz}/\Delta x + b_{xyy}/\Delta y + 2c_{xzz}/\Delta z &= q_{xz} \end{aligned} \quad (\text{A.1})$$

Minimizing the energy of the electric displacement, which depends quadratically on the coefficients, enables us to obtain a further simplification given by the three equations:  $a_{xxy} = b_{xyy}$ ,  $b_{yyz} = c_{yzz}$  and  $a_{xxz} = c_{xzz}$ . The three constraints at first order are given by

$$\begin{aligned} 2a_{xx}/\Delta x + b_{xy}/\Delta y + c_{xz}/\Delta z &= q_x; & a_{xy}/\Delta x + 2b_{yy}/\Delta y + c_{yz}/\Delta z &= q_y; \\ a_{xz}/\Delta x + b_{yz}/\Delta y + 2c_{zz}/\Delta z &= q_z \end{aligned} \quad (\text{A.2})$$

The constraint at zeroth order will eventually turn out to be equivalent to the definition of the mean charge density in eqn. (3.1) and it is given by

$$\begin{aligned} (a_x - 3a_{xxx}/20 - a_{xxy}/12 - a_{xzz}/12)/\Delta x + (b_y - 3b_{yyy}/20 - b_{xxy}/12 - b_{yzz}/12)/\Delta y \\ + (c_z - 3c_{zzz}/20 - c_{xxz}/12 - c_{yyz}/12)/\Delta z = q_0 - (q_{xx} + q_{yy} + q_{zz})/12 \end{aligned} \quad (\text{A.3})$$

Eqn. (A.3) is slightly different from eqns. (A.1) and (A.2) in that it does not help us to obtain the coefficients. Instead it helps us to validate that the coefficients are consistent with the discrete constraint in eqn. (3.1).

We again focus on Fig. 1. We wish to relate the polynomial coefficients that have been defined in the interior of the cube (i.e. eqns. (3.6), (3.7) and (3.8)) to the variation that has been posited in the faces of the cube (i.e. eqns. (3.3), (3.4) and (3.5)). Matching the quadratic variation in the right and left  $x$ -faces gives

$$\begin{aligned} a_{yy} = (D_{yy}^{x+} + D_{yy}^{x-})/2; & a_{xyy} = D_{yy}^{x+} - D_{yy}^{x-}; & a_{zz} = (D_{zz}^{x+} + D_{zz}^{x-})/2; \\ a_{xzz} = D_{zz}^{x+} - D_{zz}^{x-}; & a_{yzz} = (D_{yz}^{x+} + D_{yz}^{x-})/2; & a_{xyz} = D_{yz}^{x+} - D_{yz}^{x-} \end{aligned} \quad (\text{A.4})$$

Matching the quadratic variation in the upper and lower  $y$ -faces gives

$$\begin{aligned} b_{xx} = (D_{xx}^{y+} + D_{xx}^{y-})/2; & b_{xxy} = D_{xx}^{y+} - D_{xx}^{y-}; & b_{zz} = (D_{zz}^{y+} + D_{zz}^{y-})/2; \\ b_{yzz} = D_{zz}^{y+} - D_{zz}^{y-}; & b_{xz} = (D_{xz}^{y+} + D_{xz}^{y-})/2; & b_{xyz} = D_{xz}^{y+} - D_{xz}^{y-} \end{aligned} \quad (\text{A.5})$$

Matching the quadratic variation in the top and bottom  $z$ -faces gives

$$\begin{aligned} c_{xx} = (D_{xx}^{z+} + D_{xx}^{z-})/2; & c_{xxz} = D_{xx}^{z+} - D_{xx}^{z-}; & c_{yy} = (D_{yy}^{z+} + D_{yy}^{z-})/2; \\ c_{yyz} = D_{yy}^{z+} - D_{yy}^{z-}; & c_{xy} = (D_{xy}^{z+} + D_{xy}^{z-})/2; & c_{xyz} = D_{xy}^{z+} - D_{xy}^{z-} \end{aligned} \quad (\text{A.6})$$

Applying the second order constraints from eqn. (A.1) gives

$$\begin{aligned}
a_{xxx} &= \frac{\Delta x}{3} \left( q_{xx} - \frac{b_{xxy}}{\Delta y} - \frac{c_{xxz}}{\Delta z} \right); & b_{yyy} &= \frac{\Delta y}{3} \left( q_{yy} - \frac{a_{xyy}}{\Delta x} - \frac{c_{yyz}}{\Delta z} \right); & c_{zzz} &= \frac{\Delta z}{3} \left( q_{zz} - \frac{a_{xzz}}{\Delta x} - \frac{b_{yzz}}{\Delta y} \right); \\
a_{xxy} = b_{xyy} &= \frac{\Delta x \Delta y}{2(\Delta x + \Delta y)} \left( q_{xy} - \frac{c_{xyz}}{\Delta z} \right); & b_{yyz} = c_{yzz} &= \frac{\Delta y \Delta z}{2(\Delta y + \Delta z)} \left( q_{yz} - \frac{a_{xyz}}{\Delta x} \right); \\
c_{xzz} = a_{xxz} &= \frac{\Delta x \Delta z}{2(\Delta x + \Delta z)} \left( q_{xz} - \frac{b_{xyz}}{\Delta y} \right)
\end{aligned} \tag{A.7}$$

We now turn our attention to the linear variation.

Matching the linear variation in the right and left  $x$ -faces gives

$$\begin{aligned}
a_y &= (D_y^{x+} + D_y^{x-})/2 - a_{xy}/6; & a_{xy} &= D_y^{x+} - D_y^{x-}; \\
a_z &= (D_z^{x+} + D_z^{x-})/2 - a_{xz}/6; & a_{xz} &= D_z^{x+} - D_z^{x-}
\end{aligned} \tag{A.8}$$

Matching the linear variation in the upper and lower  $y$ -faces gives

$$\begin{aligned}
b_x &= (D_x^{y+} + D_x^{y-})/2 - b_{xy}/6; & b_{xy} &= D_x^{y+} - D_x^{y-}; \\
b_z &= (D_z^{y+} + D_z^{y-})/2 - b_{yz}/6; & b_{yz} &= D_z^{y+} - D_z^{y-}
\end{aligned} \tag{A.9}$$

Matching the linear variation in the top and bottom  $z$ -faces gives

$$\begin{aligned}
c_x &= (D_x^{z+} + D_x^{z-})/2 - c_{xz}/6; & c_{xz} &= D_x^{z+} - D_x^{z-}; \\
c_y &= (D_y^{z+} + D_y^{z-})/2 - c_{yz}/6; & c_{yz} &= D_y^{z+} - D_y^{z-}
\end{aligned} \tag{A.10}$$

Applying the first order constraints from eqn. (A.2) gives

$$a_{xx} = \frac{\Delta x}{2} \left( q_x - \frac{b_{xy}}{\Delta y} - \frac{c_{xz}}{\Delta z} \right); \quad b_{yy} = \frac{\Delta y}{2} \left( q_y - \frac{a_{xy}}{\Delta x} - \frac{c_{yz}}{\Delta z} \right); \quad c_{zz} = \frac{\Delta z}{2} \left( q_z - \frac{a_{xz}}{\Delta x} - \frac{b_{yz}}{\Delta y} \right) \tag{A.11}$$

We now turn our attention to the constant terms in the faces.

Matching the constant terms in the right and left  $x$ -faces gives

$$a_0 = (D_0^{x+} + D_0^{x-})/2 - a_{xx}/6; \quad a_x = (D_0^{x+} - D_0^{x-}) - a_{xxx}/10 \tag{A.12}$$

Matching the constant terms in the upper and lower  $y$ -faces gives

$$b_0 = (D_0^{y+} + D_0^{y-})/2 - b_{yy}/6; \quad b_y = (D_0^{y+} - D_0^{y-}) - b_{yyy}/10 \tag{A.13}$$

Matching the constant terms in the top and bottom  $z$ -faces gives

$$c_0 = (D_0^{z+} + D_0^{z-})/2 - c_{zz}/6; \quad c_z = (D_0^{z+} - D_0^{z-}) - c_{zzz}/10 \tag{A.14}$$

When the relevant equations from eqn. (A.4) to (A.14) are inserted in eqn. (A.3) we retrieve eqn. (3.1). This emphasizes the role of eqn. (A.3) as a consistency condition. This completes our detailed description of third order constraint-preserving reconstruction of the electric displacement vector field in three-dimensions. When making a computer implementation it is most efficient to implement the equations sequentially from eqn. (A.4) to (A.14). In other words, the computer implementation follows the text of this paper. Constraint-preserving reconstruction of the magnetic induction vector field works similarly.

## Appendix B

We now present the basics of constraint-preserving reconstruction at fourth order. Let the zone size be  $\Delta x$ ,  $\Delta y$  and  $\Delta z$  in the  $x$ -,  $y$ - and  $z$ -directions. While the third order case was presented in coordinate space, it is more economical to present the results for the fourth order case in the space of a reference element spanning  $[-1/2, 1/2]^3$ . The present formulation expands on the fourth order constraint-preserving reconstruction from Appendix B of Balsara et al. [16]. That formulation is only suitable for Runge–Kutta timestepping because it provides an order-preserving reconstruction at zone-boundaries. However, Appendix B of Balsara et al. [16] ignores some important terms that only operate in the interior of a zone and are needed in ADER timestepping. At the right and left  $x$ -faces of the reference element, the fourth order accurate facial reconstruction of the  $x$ -component of the electric displacement is given by

$$\begin{aligned}
D^{x\pm}(y, z) &= D_0^{x\pm} + D_y^{x\pm} y + D_z^{x\pm} z + D_{yy}^{x\pm} (y^2 - 1/12) + D_{zz}^{x\pm} (z^2 - 1/12) + D_{yz}^{x\pm} yz \\
&\quad + D_{yyy}^{x\pm} (y^3 - 3y/20) + D_{zzz}^{x\pm} (z^3 - 3z/20) + D_{yyz}^{x\pm} (y^2 - 1/12)z + D_{yzz}^{x\pm} y(z^2 - 1/12)
\end{aligned} \tag{B.1}$$

A WENO strategy for obtaining the moments of the above equation is described in (Balsara and Shu [5], Balsara et al. [6]). The fourth order accurate facial reconstruction for the  $y$ -component of the electric field at the upper and lower  $y$ -faces are given by

$$D^{y\pm}(x, z) = D_0^{y\pm} + D_x^{y\pm}x + D_z^{y\pm}z + D_{xx}^{y\pm}(x^2 - 1/12) + D_{zz}^{y\pm}(z^2 - 1/12) + D_{xz}^{y\pm}xz + D_{xxx}^{y\pm}(x^3 - 3x/20) + D_{zzz}^{y\pm}(z^3 - 3z/20) + D_{xxz}^{y\pm}(x^2 - 1/12)z + D_{xzz}^{y\pm}x(z^2 - 1/12) \tag{B.2}$$

The analogous facial reconstruction for the z-component of the electric field at the top and bottom z-faces are given by

$$D^{z\pm}(x, y) = D_0^{z\pm} + D_x^{z\pm}x + D_y^{z\pm}y + D_{xx}^{z\pm}(x^2 - 1/12) + D_{yy}^{z\pm}(y^2 - 1/12) + D_{xy}^{z\pm}xy + D_{xxx}^{z\pm}(x^3 - 3x/20) + D_{yyy}^{z\pm}(y^3 - 3y/20) + D_{xxz}^{z\pm}(x^2 - 1/12)y + D_{xyy}^{z\pm}x(y^2 - 1/12) \tag{B.3}$$

We wish to find a solution in the interior of the zone in question, consistent with the facial variations in eqns. (B.1), (B.2) and (B.3).

We also want the solution for the electric displacement in the interior of the zone in question to be consistent with the constraint in eqn. (2.3). The most economical way to initialize such a computation is to transcribe  $D_0^{x\pm} \rightarrow D_0^{x\pm}/\Delta x$  and similarly for all the other coefficients in eqn. (B.1). For all the coefficients in eqn. (B.2) we make transcriptions that are analogous to  $D_0^{y\pm} \rightarrow D_0^{y\pm}/\Delta y$ . Likewise, for all the coefficients in eqn. (B.3) we make transcriptions that are analogous to  $D_0^{z\pm} \rightarrow D_0^{z\pm}/\Delta z$ . The analogue of eqn. (3.1) then becomes

$$q_0 = (D_0^{x+} - D_0^{x-}) + (D_0^{y+} - D_0^{y-}) + (D_0^{z+} - D_0^{z-}) \tag{B.4}$$

Once such a zone-averaged charge density is obtained in all the zones, we can apply zone-centered WENO reconstruction to the zone-averaged charge density to get

$$\rho_E(x, y, z) = q_0 + q_x x + q_y y + q_z z + q_{xx}(x^2 - 1/12) + q_{yy}(y^2 - 1/12) + q_{zz}(z^2 - 1/12) + q_{xy}xy + q_{yz}yz + q_{xz}xz + q_{xxx}(x^3 - 3x/20) + q_{yyy}(y^3 - 3y/20) + q_{zzz}(z^3 - 3z/20) + q_{xxy}(x^2 - 1/12)y + q_{xxz}(x^2 - 1/12)z + q_{xyy}x(y^2 - 1/12) + q_{yyz}(y^2 - 1/12)z + q_{xzz}x(z^2 - 1/12) + q_{yzz}y(z^2 - 1/12) + q_{xyz}xyz \tag{B.5}$$

Eqn. (B.5) gives us the right-hand side for eqn. (2.3).

Let the x-component of the electric displacement within the unit cube be described by the following polynomial

$$D^x(x, y, z) = a_0 + a_x x + a_y y + a_z z + a_{xx}(x^2 - 1/12) + a_{yy}(y^2 - 1/12) + a_{zz}(z^2 - 1/12) + a_{xy}xy + a_{yz}yz + a_{xz}xz + a_{xxx}(x^3 - 3x/20) + a_{yyy}(y^3 - 3y/20) + a_{zzz}(z^3 - 3z/20) + a_{xxy}(x^2 - 1/12)y + a_{xxz}(x^2 - 1/12)z + a_{xyy}x(y^2 - 1/12) + a_{yzz}y(z^2 - 1/12) + a_{xyz}xyz + a_{xxxx}(x^4 - 3x^2/14 + 3/560) + a_{xyyy}x(y^3 - 3y/20) + a_{xzzz}x(z^3 - 3z/20) + a_{xxxxy}(x^3 - 3x/20)y + a_{xxxz}(x^3 - 3x/20)z + a_{xxyy}(x^2 - 1/12)(y^2 - 1/12) + a_{xyyz}x(y^2 - 1/12)z + a_{xxzz}(x^2 - 1/12)(z^2 - 1/12) + a_{xyzz}xy(z^2 - 1/12) + a_{xxyyz}(x^2 - 1/12)yz + a_{xxxxy}(x^4 - 3x^2/14 + 3/560)y + a_{xxyyy}(x^2 - 1/12)(y^3 - 3y/20) + a_{xxxxz}(x^4 - 3x^2/14 + 3/560)z + a_{xxzzz}(x^2 - 1/12)(z^3 - 3z/20) \tag{B.6}$$

Notice that eqn. (B.6) has all the terms that are needed for up to fourth order accurate reconstruction. But it also has extra terms for ensuring that the facial moments can be matched and that the constraints can be preserved. The zone-centered terms  $a_{xxy}$  and  $a_{xxz}$  in the above equation cannot be satisfied by just examining the values in the faces of that zone and we will specify a WENO method that looks at the x-face values of the adjoining zones in the x-direction. Let the y-component of the electric displacement within the unit cube be described by the following polynomial

$$D^y(x, y, z) = b_0 + b_x x + b_y y + b_z z + b_{xx}(x^2 - 1/12) + b_{yy}(y^2 - 1/12) + b_{zz}(z^2 - 1/12) + b_{xy}xy + b_{yz}yz + b_{xz}xz + b_{xxx}(x^3 - 3x/20) + b_{yyy}(y^3 - 3y/20) + b_{zzz}(z^3 - 3z/20) + b_{xxy}(x^2 - 1/12)y + b_{xxz}(x^2 - 1/12)z + b_{xyy}x(y^2 - 1/12) + b_{yyz}(y^2 - 1/12)z + b_{xzz}x(z^2 - 1/12) + b_{yzz}y(z^2 - 1/12) + b_{xyz}xyz + b_{yyyy}(y^4 - 3y^2/14 + 3/560) + b_{xxxxy}(x^3 - 3x/20)y + b_{yzzz}y(z^3 - 3z/20) + b_{xxyy}(x^2 - 1/12)(y^2 - 1/12) + b_{xxyz}(x^2 - 1/12)yz + b_{xyyy}x(y^3 - 3y/20) + b_{yyyz}(y^3 - 3y/20)z + b_{xyzz}xy(z^2 - 1/12) + b_{yyzz}(y^2 - 1/12)(z^2 - 1/12) + b_{xyyz}x(y^2 - 1/12)z + b_{xyyyy}x(y^4 - 3y^2/14 + 3/560) + b_{xxxxy}(x^3 - 3x/20)(y^2 - 1/12) + b_{yyyyz}(y^4 - 3y^2/14 + 3/560)z + b_{yyzzz}(y^2 - 1/12)(z^3 - 3z/20) \tag{B.7}$$

The zone-centered terms  $b_{xyy}$  and  $b_{yyz}$  in the above equation cannot be satisfied by just examining the values in the faces of that zone and we will need a WENO method that looks at the  $y$ -face values of the adjoining zones in the  $y$ -direction. Let the  $z$ -component of the electric displacement within the unit cube be described by the following polynomial

$$\begin{aligned}
D^z(x, y, z) = & c_0 + c_x x + c_y y + c_z z + c_{xx}(x^2 - 1/12) + c_{yy}(y^2 - 1/12) \\
& + c_{zz}(z^2 - 1/12) + c_{xy}xy + c_{yz}yz + c_{xz}xz + c_{xxx}(x^3 - 3x/20) + c_{yyy}(y^3 - 3y/20) \\
& + c_{zzz}(z^3 - 3z/20) + c_{xxy}(x^2 - 1/12)y + c_{xxz}(x^2 - 1/12)z \\
& + c_{xyy}x(y^2 - 1/12) + c_{yyz}(y^2 - 1/12)z + c_{xzz}x(z^2 - 1/12) + c_{yzz}y(z^2 - 1/12) + c_{xyz}xyz \\
& + c_{zzzz}(z^4 - 3z^2/14 + 3/560) + c_{xxxz}(x^3 - 3x/20)z + c_{yyyz}(y^3 - 3y/20)z \\
& + c_{xxyz}(x^2 - 1/12)yz + c_{xxzz}(x^2 - 1/12)(z^2 - 1/12) + c_{xyyz}x(y^2 - 1/12)z \\
& + c_{yyzz}(y^2 - 1/12)(z^2 - 1/12) + c_{xzzz}x(z^3 - 3z/20) + c_{yzzz}y(z^3 - 3z/20) \\
& + c_{xyzz}xy(z^2 - 1/12) + c_{xzzzz}x(z^4 - 3z^2/14 + 3/560) + c_{xxxz}(x^3 - 3x/20)(z^2 - 1/12) \\
& + c_{yzzzz}y(z^4 - 3z^2/14 + 3/560) + c_{yyzzz}(y^3 - 3y/20)(z^2 - 1/12)
\end{aligned} \tag{B.8}$$

As before, the zone-centered terms  $c_{xzz}$  and  $c_{yzz}$  in the above equation cannot be satisfied by just examining the values in the faces of that zone and we will need a WENO method that looks at the  $z$ -face values of the adjoining zones in the  $z$ -direction. Once the coefficients in eqns. (B.6), (B.7) and (B.8) are defined, we can require that they satisfy the divergence constraint  $\partial_x D^x(x, y, z) + \partial_y D^y(x, y, z) + \partial_z D^z(x, y, z) = \rho_E(x, y, z)$  at all orders in the polynomial expansion. The need to balance all the terms in the constraint equation in a nice and symmetrical way accounts for many of the fourth and fifth order polynomial terms in eqns. (B.6), (B.7) and (B.8).

We now present a strategy for satisfying all the coefficients in all of the above equations in a format that also satisfies all the constraints. We do this in such a way that the person seeking to make a computer implementation simply has to sequentially follow all the steps that are given below. Let us denote  $D_y^{x++}$  and  $D_z^{x++}$  to be the facial values of the  $y$ - and  $z$ -gradients in the right  $x$ -face of the right neighbor to the zone that we are considering. Similarly, let  $D_y^{x--}$  and  $D_z^{x--}$  to be the facial values of the  $y$ - and  $z$ -gradients in the left  $x$ -face of the left neighbor to the zone that we are considering. We can define three possible values of  $a_{xxy}$ , denote them by  $a_{xxy}^1$  and  $a_{xxy}^2$  for each of the two smaller stencils and  $a_{xxy}^3$  for the larger stencil. We can write

$$a_{xxy}^1 = (D_y^{x++} - 2D_y^{x-} + D_y^{x-})/2; \quad a_{xxy}^2 = (D_y^{x+} - 2D_y^{x-} + D_y^{x--})/2; \quad a_{xxy}^3 = (a_{xxy}^1 + a_{xxy}^2)/2 \tag{B.9}$$

We can achieve a CWENO-style non-linear hybridization of the three stencils as follows

$$\begin{aligned}
w^1 &= 1/((a_{xxy}^1)^2 + \varepsilon); & w^2 &= 1/((a_{xxy}^2)^2 + \varepsilon); & w^3 &= 100/((a_{xxy}^3)^2 + \varepsilon); \\
\bar{w}^1 &= w^1/(w^1 + w^2 + w^3); & \bar{w}^2 &= w^2/(w^1 + w^2 + w^3); & \bar{w}^3 &= w^3/(w^1 + w^2 + w^3); \\
a_{xxy} &= \bar{w}^1 a_{xxy}^1 + \bar{w}^2 a_{xxy}^2 + \bar{w}^3 a_{xxy}^3
\end{aligned} \tag{B.10}$$

where  $\varepsilon$  is any very tiny positive number. A similar strategy can be applied to obtain  $a_{xxz}$  by using  $D_z^{x++}$ ,  $D_z^{x+}$ ,  $D_z^{x-}$  and  $D_z^{x--}$ . We can also use a similar line of reasoning to obtain  $b_{xyy}$  and  $b_{yyz}$ ; an analogous line of reasoning gives us  $c_{xzz}$  and  $c_{yzz}$ . By matching the quadratic cross terms in the faces we also obtain

$$a_{xyz} = D_{yz}^{x+} - D_{yz}^{x-}; \quad b_{xyz} = D_{xz}^{y+} - D_{xz}^{y-}; \quad c_{xyz} = D_{xy}^{z+} - D_{xy}^{z-} \tag{B.11}$$

We now wish to satisfy all the fourth order constraints and some of the second order constraints. This is accomplished by first defining three auxiliary variables,  $\alpha$ ,  $\beta$  and  $\gamma$ . We obtain

$$\begin{aligned}
\alpha &= 35(q_{xy} - 2a_{xxy} - 2b_{xyy} - c_{xyz})/12; & \beta &= 35(q_{yz} - a_{xyz} - 2b_{yyz} - 2c_{yzz})/12; \\
\gamma &= 35(q_{xz} - 2a_{xxz} - b_{xyz} - 2c_{xzz})/12
\end{aligned} \tag{B.12}$$

The above equation also ensures that the second order constraints involving the cross terms are satisfied. The six fourth order constraints are also satisfied in an energy-minimizing fashion by making the following choices

$$\begin{aligned}
a_{xxxx} &= b_{xyyy} = \alpha; & b_{xxyy} &= a_{xyyy} = -2\alpha; & b_{yyyy} &= c_{yzzz} = \beta; & b_{yyzz} &= c_{yyyz} = -2\beta; \\
a_{xxxx} &= c_{zzzz} = \gamma; & a_{xxzz} &= c_{xxzz} = -2\gamma
\end{aligned} \tag{B.13}$$

By matching the third order facial values in the right and left  $x$ -faces we obtain

$$\begin{aligned}
a_{yyy} &= (D_{yyy}^{x+} + D_{yyy}^{x-})/2 - a_{xyyy}/6; & a_{xyyy} &= D_{yyy}^{x+} - D_{yyy}^{x-}; \\
a_{zzz} &= (D_{zzz}^{x+} + D_{zzz}^{x-})/2 - a_{xxzz}/6; & a_{xxzz} &= D_{zzz}^{x+} - D_{zzz}^{x-}; \\
a_{yyz} &= (D_{yyz}^{x+} + D_{yyz}^{x-})/2; & a_{xyyz} &= D_{yyz}^{x+} - D_{yyz}^{x-}; & a_{yzz} &= (D_{yzz}^{x+} + D_{yzz}^{x-})/2; & a_{xyzz} &= D_{yzz}^{x+} - D_{yzz}^{x-}
\end{aligned} \tag{B.14}$$

Similarly, matching the third order facial values in the upper and lower y-faces gives

$$\begin{aligned} b_{xxx} &= (D_{xxx}^{y+} + D_{xxx}^{y-})/2 - b_{xxxxy}/6; & b_{xxxxy} &= D_{xxx}^{y+} - D_{xxx}^{y-}; \\ b_{zzz} &= (D_{zzz}^{y+} + D_{zzz}^{y-})/2 - b_{yyzzz}/6; & b_{yzzz} &= D_{zzz}^{y+} - D_{zzz}^{y-}; \\ b_{xxz} &= (D_{xxz}^{y+} + D_{xxz}^{y-})/2; & b_{xxyy} &= D_{xxz}^{y+} - D_{xxz}^{y-}; & b_{xzz} &= (D_{xzz}^{y+} + D_{xzz}^{y-})/2; & b_{xyzz} &= D_{xzz}^{y+} - D_{xzz}^{y-} \end{aligned} \tag{B.15}$$

Likewise, matching the third order facial values in the top and bottom z-faces gives

$$\begin{aligned} c_{xxx} &= (D_{xxx}^{z+} + D_{xxx}^{z-})/2 - c_{xxxzz}/6; & c_{xxxz} &= D_{xxx}^{z+} - D_{xxx}^{z-}; \\ c_{yyy} &= (D_{yyy}^{z+} + D_{yyy}^{z-})/2 - c_{yyyz}/6; & c_{yyyz} &= D_{yyy}^{z+} - D_{yyy}^{z-}; \\ c_{xxy} &= (D_{xxy}^{z+} + D_{xxy}^{z-})/2; & c_{xxyz} &= D_{xxy}^{z+} - D_{xxy}^{z-}; & c_{xyy} &= (D_{xyy}^{z+} + D_{xyy}^{z-})/2; & c_{xyyz} &= D_{xyy}^{z+} - D_{xyy}^{z-} \end{aligned} \tag{B.16}$$

Applying the third order constraints while minimizing the energy in the field gives

$$\begin{aligned} a_{xxx} &= (q_{xxx} - b_{xxxxy} - c_{xxxz})/4; & b_{yyy} &= (q_{yyy} - a_{xyyy} - c_{yyyz})/4; & c_{zzz} &= (q_{zzz} - a_{xzzz} - b_{yzzz})/4; \\ b_{xxxxy} &= 3(q_{xxy} - c_{xxyz})/20; & a_{xxxz} &= 14b_{xxxxy}/9; & c_{xxz} &= 3(q_{xxz} - b_{xxyy})/20; & a_{xxxz} &= 14c_{xxxz}/9; \\ a_{xxyy} &= 3(q_{xxy} - c_{xyyz})/20; & b_{xyyy} &= 14a_{xxyy}/9; & c_{yyz} &= 3(q_{yyz} - a_{xyyz})/20; & b_{yyyz} &= 14c_{yyyz}/9; \\ a_{xzz} &= 3(q_{xzz} - b_{xyyz})/20; & c_{xzzz} &= 14a_{xzzz}/9; & b_{yzzz} &= 3(q_{yzz} - a_{xyyz})/20; & c_{yzzz} &= 14b_{yzzz}/9; \\ a_{xxyz} &= b_{xyyz} = c_{xyzz} = q_{xyz}/6 \end{aligned} \tag{B.17}$$

Matching the second order facial values in the right and left x-faces gives

$$\begin{aligned} a_{yy} &= (D_{yy}^{x+} + D_{yy}^{x-})/2 - a_{xxyy}/6; & a_{xyy} &= D_{yy}^{x+} - D_{yy}^{x-}; \\ a_{zz} &= (D_{zz}^{x+} + D_{zz}^{x-})/2 - a_{xxzz}/6; & a_{xzz} &= D_{zz}^{x+} - D_{zz}^{x-}; & a_{yz} &= (D_{yz}^{x+} + D_{yz}^{x-})/2 - a_{xxyz}/6 \end{aligned} \tag{B.18}$$

Similarly, matching the second order facial values in the upper and lower y-faces gives

$$\begin{aligned} b_{xx} &= (D_{xx}^{y+} + D_{xx}^{y-})/2 - b_{xxyy}/6; & b_{xxy} &= D_{xx}^{y+} - D_{xx}^{y-}; \\ b_{zz} &= (D_{zz}^{y+} + D_{zz}^{y-})/2 - b_{yyzz}/6; & b_{yzz} &= D_{zz}^{y+} - D_{zz}^{y-}; & b_{xz} &= (D_{xz}^{y+} + D_{xz}^{y-})/2 - b_{xyyz}/6 \end{aligned} \tag{B.19}$$

Likewise, matching the second order facial values in the top and bottom z-faces gives

$$\begin{aligned} c_{xx} &= (D_{xx}^{z+} + D_{xx}^{z-})/2 - c_{xxzz}/6; & c_{xxz} &= D_{xx}^{z+} - D_{xx}^{z-}; \\ c_{yy} &= (D_{yy}^{z+} + D_{yy}^{z-})/2 - c_{yyyz}/6; & c_{yyz} &= D_{yy}^{z+} - D_{yy}^{z-}; & c_{xy} &= (D_{xy}^{z+} + D_{xy}^{z-})/2 - c_{xyyz}/6; \end{aligned} \tag{B.20}$$

Some of the second order constraints were resolved in eqn. (B.12). But we still have three more second order constraints which can now be resolved as

$$a_{xxx} = (q_{xx} - b_{xxy} - c_{xxz})/3; \quad b_{yyy} = (q_{yy} - a_{xyy} - c_{yyz})/3; \quad c_{zzz} = (q_{zz} - a_{xzz} - b_{yzz})/3 \tag{B.21}$$

Matching the first order facial values at the right and left x-faces gives

$$\begin{aligned} a_y &= (D_y^{x+} + D_y^{x-})/2 - a_{xxy}/6 - a_{xxxxy}/70; & a_{xy} &= (D_y^{x+} - D_y^{x-}) - a_{xxxxy}/10; \\ a_z &= (D_z^{x+} + D_z^{x-})/2 - a_{xzz}/6 - a_{xxxz}/70; & a_{xz} &= (D_z^{x+} - D_z^{x-}) - a_{xxxz}/10 \end{aligned} \tag{B.22}$$

Similarly, matching the first order facial values at the upper and lower y-faces gives

$$\begin{aligned} b_x &= (D_x^{y+} + D_x^{y-})/2 - b_{xyy}/6 - b_{xyyyy}/70; & b_{xy} &= (D_x^{y+} - D_x^{y-}) - b_{xyyy}/10; \\ b_z &= (D_z^{y+} + D_z^{y-})/2 - b_{yyz}/6 - b_{yyyz}/70; & b_{yz} &= (D_z^{y+} - D_z^{y-}) - b_{yyyz}/10 \end{aligned} \tag{B.23}$$

Likewise, matching the first order facial values at the top and bottom z-faces gives

$$\begin{aligned} c_x &= (D_x^{z+} + D_x^{z-})/2 - c_{xzz}/6 - c_{xzzzz}/70; & c_{xz} &= (D_x^{z+} - D_x^{z-}) - c_{xzzz}/10; \\ c_y &= (D_y^{z+} + D_y^{z-})/2 - c_{yzz}/6 - c_{yzzzz}/70; & c_{yz} &= (D_y^{z+} - D_y^{z-}) - c_{yzzz}/10 \end{aligned} \tag{B.24}$$

The three first order constraints then give us

$$\begin{aligned} a_{xx} &= (q_x - b_{xy} - c_{xz})/2 - 3q_{xxx}/40 - (q_{xyy} + q_{xzz})/24 + 3a_{xxx}/14 + (a_{xxyy} + a_{xxzz})/12 \\ &\quad + (b_{xyzz} + c_{xyyz})/24 + 3(b_{xxxxy} + b_{xyyy} + c_{xxxz} + c_{xzzz})/40; \\ b_{yy} &= (q_y - a_{xy} - c_{yz})/2 - 3q_{yyy}/40 - (q_{xxy} + q_{yzz})/24 + 3b_{yyy}/14 + (b_{xxyy} + b_{yyzz})/12 \\ &\quad + (a_{xyzz} + c_{xxyy})/24 + 3(a_{xxxxy} + a_{xyyy} + c_{yyyz} + c_{yzzz})/40; \\ c_{zz} &= (q_z - a_{xz} - b_{yz})/2 - 3q_{zzz}/40 - (q_{xxz} + q_{yyz})/24 + 3c_{zzz}/14 + (c_{xxzz} + c_{yzzz})/12 \\ &\quad + (a_{xyyz} + b_{xxyy})/24 + 3(a_{xxxz} + a_{xzzz} + b_{yyyz} + b_{yzzz})/40 \end{aligned} \tag{B.25}$$

Matching the zeroth order facial values at the right and left  $x$ -faces give

$$a_0 = (D_0^{x+} + D_0^{x-})/2 - a_{xx}/6 - a_{xxxx}/70; \quad a_x = (D_0^{x+} - D_0^{x-}) - a_{xxx}/10 \quad (\text{B.26})$$

Similarly, matching the zeroth order facial values at the upper and lower  $y$ -faces gives

$$b_0 = (D_0^{y+} + D_0^{y-})/2 - b_{yy}/6 - b_{yyyy}/70; \quad b_y = (D_0^{y+} - D_0^{y-}) - b_{yyy}/10 \quad (\text{B.27})$$

Likewise, matching the zeroth order facial values at the top and bottom  $z$ -faces gives

$$c_0 = (D_0^{z+} + D_0^{z-})/2 - c_{zz}/6 - c_{zzzz}/70; \quad c_z = (D_0^{z+} - D_0^{z-}) - c_{zzz}/10 \quad (\text{B.28})$$

The constant part of the constraint is actually a consistency condition and gives us consistency with the equation given in eqn. (B.4). It is given by

$$\begin{aligned} & a_x + b_y + c_z - 3(a_{xxx} + b_{yyy} + c_{zzz})/20 - (a_{xyy} + a_{xzz} + b_{xxy} + b_{yzz} + c_{xxz} + c_{yyz})/12 \\ & = q_0 - (q_{xx} + q_{yy} + q_{zz})/12 \end{aligned} \quad (\text{B.29})$$

Recall that we divided all the coefficients of eqns. (B.1), (B.2) and (B.3) by  $\Delta x$ ,  $\Delta y$  and  $\Delta z$  respectively before embarking on the constraint-preserving reconstruction described in the above paragraphs. We now undo that process so that all the coefficients in eqn. (B.6) are multiplied by  $\Delta x$ ; so that we have  $a_0 \rightarrow a_0 \Delta x$  for example. Similarly, all the coefficients of eqn. (B.7) are multiplied by  $\Delta y$ ; so that we have  $b_0 \rightarrow b_0 \Delta y$  for example. Likewise, all the coefficients of eqn. (B.8) are multiplied by  $\Delta z$ ; so that we have  $c_0 \rightarrow c_0 \Delta z$  for example.

## Appendix C. Supplementary material

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jcp.2017.10.013>.

## References

- [1] D.S. Balsara, D.S. Spicer, A staggered mesh algorithm using high order Godunov fluxes to ensure solenoidal magnetic fields in magnetohydrodynamic simulations, *J. Comput. Phys.* 149 (1999) 270–292.
- [2] D.S. Balsara, Divergence-free adaptive mesh refinement for magnetohydrodynamics, *J. Comput. Phys.* 174 (2001) 614–648.
- [3] D.S. Balsara, Second-order-accurate schemes for magnetohydrodynamics with divergence-free reconstruction, *Astrophys. J. Suppl. Ser.* 151 (2004) 149–184.
- [4] D.S. Balsara, Divergence-free reconstruction of magnetic fields and WENO schemes for magnetohydrodynamics, *J. Comput. Phys.* 228 (2009) 5040–5056.
- [5] D.S. Balsara, C.-W. Shu, Monotonicity preserving weighted non-oscillatory schemes with increasingly high order of accuracy, *J. Comput. Phys.* 160 (2000) 405–452.
- [6] D.S. Balsara, T. Rumpf, M. Dumbser, C.-D. Munz, Efficient, high accuracy ADER–WENO schemes for hydrodynamics and divergence-free magnetohydrodynamics, *J. Comput. Phys.* 228 (2009) 2480–2516.
- [7] D.S. Balsara, Multidimensional HLLC Riemann solver; application to Euler and magnetohydrodynamic flows, *J. Comput. Phys.* 229 (2010) 1970–1993.
- [8] D.S. Balsara, A two-dimensional HLLC Riemann solver for conservation laws: application to Euler and magnetohydrodynamic flows, *J. Comput. Phys.* 231 (2012) 7476–7503.
- [9] D.S. Balsara, M. Dumbser, C. Meyer, H. Du, Z. Xu, Efficient implementation of ADER schemes for Euler and magnetohydrodynamic flow on structured meshes – comparison with Runge–Kutta methods, *J. Comput. Phys.* 235 (2013) 934–969.
- [10] D.S. Balsara, M. Dumbser, R. Abgrall, Multidimensional HLL and HLLC Riemann solvers for unstructured meshes – with application to Euler and MHD flows, *J. Comput. Phys.* 261 (2014) 172–208.
- [11] D.S. Balsara, Multidimensional Riemann problem with self-similar internal structure – Part I – application to hyperbolic conservation laws on structured meshes, *J. Comput. Phys.* 277 (2014) 163–200.
- [12] D.S. Balsara, M. Dumbser, Divergence-free MHD on unstructured meshes using high order finite volume schemes based on multidimensional Riemann solvers, *J. Comput. Phys.* 299 (2015) 687–715.
- [13] D.S. Balsara, M. Dumbser, Multidimensional Riemann problem with self-similar internal structure – Part II – application to hyperbolic conservation laws on unstructured meshes, *J. Comput. Phys.* 287 (2015) 269–292.
- [14] D.S. Balsara, Three dimensional HLL Riemann solver for structured meshes; application to Euler and MHD flow, *J. Comput. Phys.* 295 (2015) 1–23.
- [15] D.S. Balsara, J. Vides, K. Gurski, B. Nkonga, M. Dumbser, S. Garain, E. Audit, A two-dimensional Riemann solver with self-similar sub-structure – alternative formulation based on least squares projection, *J. Comput. Phys.* 304 (2016) 138–161.
- [16] D.S. Balsara, T. Amano, S. Garain, J. Kim, High order accuracy divergence-free scheme for the electrodynamics of relativistic plasmas with multidimensional Riemann solvers, *J. Comput. Phys.* 318 (2016) 169–200.
- [17] D.S. Balsara, S. Garain, C.-W. Shu, An efficient class of WENO schemes with adaptive order, *J. Comput. Phys.* 326 (2016) 780–804.
- [18] D.S. Balsara, R. Käppeli, von Neumann stability analysis of globally divergence-free RKDG and PNPM schemes for the induction equation using multidimensional Riemann solvers, *J. Comput. Phys.* 336 (2017) 104–127.
- [19] D.S. Balsara, A. Taflove, S. Garain, G. Montecinos, Computational electrodynamics in material media with constraint-preservation, multidimensional Riemann solvers and sub-cell resolution – Part I, second-order FVTD schemes, submitted for publication, *J. Comput. Phys.* 349 (2017) 604–635.
- [20] D.S. Balsara, B. Nkonga, Formulating multidimensional Riemann solvers in similarity variables – Part III: a multidimensional analogue of the HLLI Riemann solver for conservative hyperbolic systems, *J. Comput. Phys.* 346 (2017) 25–48.
- [21] A. Barbas, P. Velarde, Development of a Godunov method for Maxwell’s equations with adaptive mesh refinement, *J. Comput. Phys.* 300 (2016) 186–201.
- [22] J.-P. Berenger, A perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.* 114 (2) (1994) 185–200.
- [23] J.-P. Berenger, Three-dimensional perfectly matched layer for the absorption of electromagnetic waves, *J. Comput. Phys.* 127 (2) (1996) 363–379.
- [24] M. Dumbser, M. Käser, Arbitrary high order non-oscillatory finite volume schemes on unstructured meshes for linear hyperbolic systems, *J. Comput. Phys.* 221 (2007) 693–723.
- [25] M. Dumbser, D. Balsara, E.F. Toro, C.D. Munz, A unified framework for the construction of one-step finite volume and discontinuous Galerkin schemes on unstructured meshes, *J. Comput. Phys.* 227 (2008) 8209–8253.



- [26] M. Dumbser, O. Zanotti, A. Hidalgo, D.S. Balsara, ADER–WENO finite volume schemes with space–time adaptive mesh refinement, *J. Comput. Phys.* 248 (2013) 257–286.
- [27] A. Farjadpour, D. Roundy, A. Rodriguez, M. Ibanescu, P. Bernel, J. Joannopoulos, S. Johnson, P. Burr, *Opt. Lett.* 31 (2006) 2972–2974.
- [28] J. Hesthaven, T. Warburton, Nodal high-order methods on unstructured grids: I. Time-domain solution of Maxwell's equations, *J. Comput. Phys.* 181 (2002) 186–221.
- [29] W. Hunsdorfer, S.J. Ruuth, IMEX extensions of linear multistep methods with general monotonicity and boundedness properties, *J. Comput. Phys.* 225 (2007) 2016–2042.
- [30] T.Z. Ismagilov, Second order finite volume scheme for Maxwell's equations with discontinuous electromagnetic properties on unstructured meshes, *J. Comput. Phys.* 282 (2015) 33–42.
- [31] G.-S. Jiang, C.-W. Shu, Efficient implementation of weighted ENO schemes, *J. Comput. Phys.* 126 (1996) 202–228.
- [32] D.S. Katz, E.T. Thiel, A. Taflove, Validation and extension to three dimensions of the Berenger PML absorbing boundary condition for FD-TD meshes, *IEEE Microw. Guided Wave Lett.* 4 (8) (1994) 268.
- [33] F. Kupka, N. Huppenhofer, I. Higuera, O. Koch, Total-variation-diminishing implicit–explicit Runge–Kutta methods for the simulation of double-diffusive convection in astrophysics, *J. Comput. Phys.* 231 (2012) 3561–3586.
- [34] C.-D. Munz, P. Omnes, R. Schneider, E. Sonnendrücker, U. Voß, Divergence correction techniques for Maxwell solvers based on a hyperbolic model, *J. Comput. Phys.* 161 (2000) 484.
- [35] A. Oskooi, C. Kottke, S. Johnson, *Opt. Lett.* 34 (2009) 2778–2780.
- [36] L. Pareschi, G. Russo, Implicit–explicit Runge–Kutta schemes and applications to hyperbolic systems with relaxation, *J. Sci. Comput.* 25 (2005) 129.
- [37] R.J. Spiteri, S.J. Ruuth, A new class of optimal high-order strong-stability-preserving time-stepping schemes, *SIAM J. Numer. Anal.* 40 (2002) 469–491.
- [38] R.J. Spiteri, S.J. Ruuth, Non-linear evolution using optimal fourth-order strong-stability-preserving Runge–Kutta methods, *Math. Comput. Simul.* 62 (2003) 125–135.
- [39] C.-W. Shu, S.J. Osher, Efficient implementation of essentially non-oscillatory shock capturing schemes, *J. Comput. Phys.* 77 (1988) 439–471.
- [40] C.-W. Shu, S.J. Osher, Efficient implementation of essentially non-oscillatory shock capturing schemes II, *J. Comput. Phys.* 83 (1989) 32–78.
- [41] C.-W. Shu, Total variation-diminishing time discretizations, *SIAM J. Sci. Stat. Comput.* 9 (1988) 1073–1084.
- [42] A. Taflove, Review of the formulation and applications of the finite-difference time-domain method for numerical modeling of electromagnetic wave interactions with arbitrary structures, *Wave Motion* 10 (1988) 547–582.
- [43] A. Taflove, S. Hagness, *Computational Electrodynamics*, third edition, Artech House, 2005.
- [44] A. Taflove, A. Oskooi, S. Johnson, *Advances in FDTD Computational Electrodynamics – Photonics and Nanotechnology*, Artech House, 2013.
- [45] A. Taflove, S. Hagness, Finite difference time domain solution of Maxwell's equations, *WEEEEE Rev.* (2016) 8303.
- [46] V.A. Titarev, E.F. Toro, ADER: arbitrary high order Godunov approach, *J. Sci. Comput.* 17 (1–4) (2002) 609–618.
- [47] V.A. Titarev, E.F. Toro, ADER schemes for three-dimensional nonlinear hyperbolic systems, *J. Comput. Phys.* 204 (2005) 715–736.
- [48] E.F. Toro, R.C. Millington, R.C.M. Nejad, Towards very high-order Godunov schemes, in: E.F. Toro (Ed.), *Godunov Methods: Theory and Applications* (edited review), Kluwer Academic/Plenum Publishers, 2001, pp. 905–937.
- [49] E.F. Toro, V.A. Titarev, Solution of the generalized Riemann problem for advection reaction equations, *Proc. R. Soc. Lond. Ser. A* 458 (2002) 271–281.
- [50] B. van Leer, Towards the ultimate conservative difference scheme, IV: a new approach to numerical convection, *J. Comput. Phys.* 23 (1977) 276–299.
- [51] B. van Leer, Towards the ultimate conservative difference scheme, V: a second order sequel to Godunov's method, *J. Comput. Phys.* 32 (1979) 101–136.
- [52] K.S. Yee, Numerical solution of initial boundary value problems involving Maxwell equation in an isotropic media, *IEEE Trans. Antennas Propag.* 14 (1966) 302.
- [53] Z. Xu, D.S. Balsara, H. Du, Divergence-free WENO reconstruction-based finite volume scheme for ideal MHD equations on triangular meshes, *Commun. Comput. Phys.* 19 (04) (2016) 841–880.