Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 195 (2004) 17–48
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

On the divergence-free condition in Godunov-type schemes for ideal magnetohydrodynamics: the upwind constrained transport method

P. Londrillo [a], L. Del Zanna [b,*]

[a] *INAF – Osservatorio Astronomico di Bologna, via C. Ranzani 1, 40127 Bologna, Italy*
[b] *Università di Firenze, Dipartimento di Astronomia e Scienza dello Spazio, Largo E. Fermi 2, 50125 Firenze, Italy*

Received 13 June 2003; received in revised form 27 August 2003; accepted 30 September 2003

Abstract

We present a general framework to design Godunov-type schemes for multidimensional ideal magnetohydrodynamic (MHD) systems, having the divergence-free relation and the related properties of the magnetic field **B** as built-in conditions. Our approach mostly relies on the *constrained transport* (CT) discretization technique for the magnetic field components, originally developed for the linear induction equation, which assures $[\nabla \cdot \mathbf{B}]_{\text{num}} = 0$ and its preservation in time to within machine accuracy in a finite-volume setting. We show that the CT formalism, when fully exploited, can be used as a general guideline to design the reconstruction procedures of the **B** vector field, to adapt standard upwind procedures for the momentum and energy equations, avoiding the onset of numerical monopoles of O(1) size, and to formulate approximate Riemann solvers for the induction equation. This general framework will be named here *upwind constrained transport* (UCT). To demonstrate the versatility of our method, we apply it to a variety of schemes, which are finally validated numerically and compared: a novel implementation for the MHD case of the second-order Roe-type positive scheme by Liu and Lax [J. Comput. Fluid Dyn. 5 (1996) 133], and both the second- and third-order versions of a central-type MHD scheme presented by Londrillo and Del Zanna [Astrophys. J. 530 (2000) 508], where the basic UCT strategies have been first outlined.

1. Introduction

In extending Godunov-type conservative schemes designed for Euler equations of gas-dynamics to the system of (ideal) magnetohydrodynamics (MHD), in the multidimensional case a main problem arises on how to represent the solenoidal structure of the magnetic field vector **B** and on how to formulate reconstruction procedures and (approximate) Riemann solvers sharing consistency with this property. In the last years a number of works have focused on this specific problem and many different approaches have been

* Corresponding author.
E-mail address: ldz@arcetri.astro.it (L. Del Zanna).

0021-9991/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2003.09.016

proposed. A wide class of (second-order) numerical schemes for regular grids have been analyzed and compared by Toth [1], while contributions covering also higher order schemes, adaptive mesh refinements (AMR) and unstructured grids are in rapid development.

Since we are mainly interested here to analyze *methodological* aspects, we propose a broad classification of the published contributions on this specific topic into *two main groups*:

1. Schemes based on *standard upwind procedures* (henceforth SUP) designed for Euler equations, where also magnetic field components are discretized at cell centers as the other fluid variables. Since in this case the approximated $[\nabla \cdot \mathbf{B}]_{\text{num}}$ based on central derivatives may have a non-vanishing size, different strategies to control or prevent the accumulation in time of related spurious numerical effects (usually referred to as *numerical monopoles*) have been proposed.
   - A first method, suggested by Brackbill and Barnes [2], is to add an elliptic (Poisson) equation to recover the solenoidal property at each time-step. In [1], this procedure has been named *projection scheme* and is currently widely adopted (see [3] for a high order WENO scheme).
   - In the scheme introduced by Powell [4] (see also [5]), the numerical $[\nabla \cdot \mathbf{B}]_{\text{num}}$ quantity is not forced to vanish; the MHD system is reformulated by adding new source terms proportional to this variable in order to recover the original MHD system in non-conservative form. Moreover, the classical seven-mode Riemann wave fans have been enlarged to eight modes. In this modified system, upwinding is applied to all magnetic field components and hence also to the component B_{n} across a discontinuity surface.
   - In a more recent work [6], in order to preserve both the conservative form and the hyperbolic structure of the MHD system, a new time-dependent wave equation is introduced to damp and/or to transport away the non-zero $[\nabla \cdot \mathbf{B}]_{\text{num}}$ contributions.
2. In the second group we include schemes which take advantage of the so-called *constrained transport* (CT) method by Evans and Hawley [7] (originally suggested for the evolution of the induction equation in the linear approximation). It is a main feature of this method to introduce staggered discretizations of magnetic and electric vector fields in the induction equation. In fact, by using these staggered values to approximate the relevant first derivatives, $[\nabla \cdot \mathbf{B}]_{\text{num}} = 0$ in the initial conditions and its exact preservation in time result. The problem here is on how to apply this formalism in a Godunov-type scheme for the full MHD system.
   - Most of the published works combine the above CT discretization with the SUP cell-centered discretization by introducing different empirical recipes (e.g. [8–11]). However, these procedures result in a sort of hybrid schemes and the problem of numerical monopoles is still left open, in our opinion.
   - In our previous work [12] (LD from now on) we have proposed numerical procedures to take advantage of the specific CT discretization benefits, and hence the $[\nabla \cdot \mathbf{B}]_{\text{num}} = 0$ condition, even in the reconstruction steps and in the approximate Riemann solvers. The same method has been then applied to relativistic MHD [13].

The goal of the present paper is twofold. First, by adding analytical arguments to the approach outlined in LD, we propose *a method* to construct and then to characterize *a class* of numerical schemes. Second, we present implementations of a variety of different schemes, to demonstrate the versatility and self-consistency of the method.

Regarding the first goal, our main concern is here to select a set of properties, some of them common to the Euler system and other specific of MHD equations, which in our opinion should be preserved in the numerical discretization. In this way, it is then possible to envisage Godunov-type schemes for MHD having: (a) the divergence-free condition as an *exact built-in property*, (b) reconstruction and upwind procedures consistent with this property. Since the CT formalism comes out to be the necessary starting point to achieve this result, our framework will be named here *upwind constrained transport* (UCT) method.

As a novel numerical application we then propose the UCT implementation of the *positive* scheme by Liu and Lax [14,15], a second-order Roe-type scheme which proves to be accurate and robust. Numerical

validation will be finally presented for several standard two-dimensional test problems, where the results of the new MHD positive scheme are compared with central-type schemes as proposed in LD, extended here to more accurate central-upwind two-speed approximate Riemann solvers, and tested in its second and third-order implementations.

This paper is organized as follows. In the next section, we propose and discuss some general conditions as guidelines for numerical modeling. The main ingredients to formulate general UCT-based Godunov-type schemes for MHD systems, i.e., the discretization form, the proper reconstruction procedures and the approximate Riemann solvers, are presented in Section 2. In Section 3, we specify the method to the positive and central MHD schemes, which will be finally tested and compared in Section 4.

1.1. Conservation laws and consistency demands for numerical MHD

The MHD system has a peculiar form and cannot be simply reduced to a set of conservation laws for scalar variables, as the Euler equations. In fact, if the specific structure of spatial differential operators is taken into account, it is more properly represented by the set of the following two coupled sub-systems:

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot \mathbf{f}(\mathbf{w}) = 0, \tag{1}$$

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E}(\mathbf{w}) = 0, \tag{2}$$

equipped with the non-evolutionary constraint on the $\mathbf{B}$ vector field

$$\nabla \cdot \mathbf{B} = 0, \tag{3}$$

which, once satisfied for initial conditions, is analytically preserved in time by Eq. (2).

The set of Eq. (1) evolves in time the five-component array of scalar functions $\mathbf{u} = [u^l(\mathbf{x},t)]^{\mathrm{T}}$, $l = 1,2,\ldots,5$, while the set (2) evolves the vector field $\mathbf{B} = [B_i(\mathbf{x},t)]^{\mathrm{T}}$, $i = x,y,z$. The overall set of dependent variables are henceforth represented by the eight-component array $\mathbf{w} = [\mathbf{u},\mathbf{B}]^{\mathrm{T}}$. The first array contains the conservative fluid variables $\mathbf{u} = [\rho, q_i, e]^{\mathrm{T}}$, where ρ is the mass density, $q_i = \rho v_i$ are the momentum components, v_i are the fluid velocity components, and $e = p/(\gamma-1) + \rho v^2/2 + B^2/2$ is the total energy density for a perfect gas equation of state, where p is the kinetic pressure and γ is the adiabatic index. The corresponding flux vector components $\mathbf{f}_i = [f_i^l]^{\mathrm{T}}$, $l = 1,2,\ldots,5$ are given by $\mathbf{f}_i = [q_i, M_{i,j}, H_i]^{\mathrm{T}}$, $i,j = x,y,z$, with the momentum flux tensor defined by $M_{i,j} = v_i q_j + \Pi\delta_{i,j} - B_i B_j$ and the energy flux components defined by $H_i = v_i(e+\Pi) - B_i(\mathbf{v}\cdot\mathbf{B})$, where $\Pi = p + B^2/2$ is the total pressure. In sub-system (2), which is the induction equation for the magnetic field vector $\mathbf{B}$, the corresponding flux is simply given by the electric field vector $\mathbf{E} = -\mathbf{v}\times\mathbf{B}$, where the assumption of a perfect conducting plasma (ideal MHD) has been implicitly assumed.

As for the Euler equations, the system (1) and (2) has to be supplied with entropy functions $S = S(\mathbf{w})$ satisfying the condition

$$\frac{\partial S}{\partial t} + \nabla \cdot \mathbf{F}_S(\mathbf{w}) \leqslant 0, \tag{4}$$

which allows to identify, among discontinuous solutions of the MHD system, the (physically) admissible ones. The existence of entropy functions (in fact $S = -\rho s$, where $s \propto \log(p\rho^{-\gamma})$ is the physical entropy per unit mass) is also related to the hyperbolic structure of the MHD equations. For smooth solutions, the system (1) and (2) can be put in the non-conservative (quasi-linear) form

$$\frac{\partial \mathbf{w}}{\partial t} + [\mathbf{J}(\mathbf{w}) \cdot \nabla]\mathbf{w} = 0, \tag{5}$$

where $\mathbf{J} = (J_i)$, $i = x, y, z$, and each J_i is the Jacobian matrix of the eight-component flux array $[f_i^l, E_i]^{\mathrm{T}}$ with respect to the $\mathbf{w}$ variables. It is a well-known property that any linear combination $J(\mathbf{w}, \mathbf{k}) = \sum_i k_i J_i(\mathbf{w})$, for real k_i numbers, and then also each J_i matrix, is hyperbolic at any reference state $\mathbf{w}$. Moreover, as for the Euler equations (see [16]), the (positive) Hessian matrix $S_{\mathbf{w},\mathbf{w}}$ acts as similarity transform to make all J_i symmetrizable.

To underline differences and analogies of the MHD system with respect to the reference Euler system which may have relevance for numerical modeling, some remarks are in order:

- The $\mathbf{u}$ array contains *scalar* variables and the corresponding flux derivatives are expressed by the $\mathrm{div} \equiv [\nabla\cdot]$ conservative operator. Sub-system (1) has then the same formal structure of the Euler system for gas-dynamics. At surface elements where discontinuities take place, this conservation form leads to the usual Rankine–Hugoniot relations. On the other hand, the $\mathbf{B}(\mathbf{x}, t)$ vector is anti-symmetric (an axial vector), components B_i are pseudo-scalars, and the corresponding evolution operator is given by the anti-symmetric $\mathrm{curl} \equiv [\nabla \times \cdot]$ derivative. The conservative form is now expressed by the scalar condition (3) (magnetic flux conservation) and by the $\nabla \times \mathbf{E}$ flux derivatives (conservation along a closed contour). Discontinuous solutions satisfy jump relations just for the tangential components $\mathbf{B}_{\mathrm{t}} = \mathbf{B} \times \mathbf{n}$, where $\mathbf{n}$ indicates the normal direction, whereas the normal field component $B_{\mathrm{n}} = \mathbf{B} \cdot \mathbf{n}$ is continuous. The Rankine–Hugoniot relations, once supplied with an appropriate entropy law, allow to identify the physically correct discontinuous solutions. It is apparent that magnetic discontinuities and the related entropy constraint do not involve the parallel B_{n} component.
- It follows that smoothness properties of MHD variables are also different. Scalar components $u^l(\mathbf{x}, t)$ may develop discontinuous solutions along any space direction and can be then represented on the space of piecewise continuous functions. The vector field $\mathbf{B}(\mathbf{x}, t)$ has more elaborate properties, since the divergence-free condition entails the $\mathbf{B}(\mathbf{x})$ field maps piecewise differentiable (and then continuous) field lines. The conservation law given by Eq. (2) is then essential to preserve in time condition (3) and to assure the smoothness properties of the magnetic field.
- The divergence-free condition enters implicitly in the MHD momentum and energy conservative equations. This can also be expressed by realizing that the Maxwell tensor $\mathbf{T} = \mathbf{I}B^2/2 - \mathbf{BB}$ in the momentum flux has to satisfy

$$\mathbf{B} \cdot (\nabla \cdot \mathbf{T}) = 0, \tag{6}$$

  in order to recover the correct Lorentz force in non-conservative form.
- Finally, the divergence-free condition allows to represent the $\mathbf{B}(\mathbf{x}, t)$ field via a vector potential $\mathbf{A}(\mathbf{x}, t)$, defined by $\mathbf{B} = \nabla \times \mathbf{A}$ and by the gauge condition $\nabla \cdot \mathbf{A} = 0$, which assures the uniqueness of this representation. The new evolution equation is now

$$\frac{\partial \mathbf{A}}{\partial t} + \mathbf{E} = 0. \tag{7}$$

  The above relations and the induction equation (2), together with the condition $\mathbf{E} \cdot \mathbf{B} = 0$ valid for ideal MHD, imply an added conservation law for the magnetic helicity $H = \int (\mathbf{A} \cdot \mathbf{B})\, \mathrm{d}x^3$, carrying informations on the topology of magnetic field lines.

When looking at (finite-dimensional) numerical approximations, a main problem is that no rigorous results on convergence are available, even for the Euler system. In this case, however, by taking advantage of theoretical achievements, like the Lax–Wendroff theorem [17], *heuristic* guidelines are usually adopted in order to:

- retain the conservative form of the original equations in the discretized system;
- assure consistency, in the sense that the approximations of the flux functions and of the differential operators have to recover the exact ones as the spatial and temporal grid sizes go to zero;
- assure non-oscillatory (or even monotonicity preserving) numerical representation of discontinuous data;
- assure consistency with the entropy law, in a way the numerical viscosity induced by the upwind differentiation is compatible with Eq. (4) (see [18]);
- assure stability of the numerical solution.

As already anticipated in Section 1, the main issue addressed here is to select a set of *additional* requirements for the MHD system which should assure that the specific properties of the magnetic field enter as *built-in* conditions of a numerical scheme. We propose the following:

- the discretized first derivatives $\partial_i B_i$ entering the $\nabla \cdot \mathbf{B}$ definition are consistent approximations;
- for initial divergence-free fields the approximated derivatives satisfy $[\nabla \cdot \mathbf{B}]_{\text{num}} = 0$ exactly;
- divergence-free initial conditions are preserved exactly in time by the discretized induction equation.

We then suggest the following definition: *a numerical scheme is consistent with the specific properties of the MHD system* if all above conditions are fulfilled. This definition, together with the guidelines for Euler equations, will enable us to identify and construct a class of Godunov-type schemes for MHD, later referred to as UCT-based schemes.

In this framework, as for the Euler equations, a finite volume setting provides a sufficiently general starting point. Here, we concentrate only on algorithms for regular structured grids, even if the generality of the method allows to extend some basic procedures also to adaptive mesh refinements (AMR, [19]) and to unstructured grids. In particular, De Sterck [20] has developed a general CT formalism for unstructured triangular grids, named MUCT, where rigorous geometrical arguments have been considered to support this approach.

2. The UCT method to design Godunov-type schemes for MHD

2.1. Discretization step: finite-volume formalism

In a finite-volume setting, the 3-D computational domain Ω is first subdivided in Cartesian cells C, with volume V, side sizes h_i, $i = x, y, z$ and faces given by the oriented surface elements $S_i^{\pm}$, $i = x, y, z$, where $\pm$ denotes the sign of face normals. For each face S_i, we then denote as $(L_j^{\pm}, L_k^{\pm})$, $j, k \neq i$ its oriented sides. At this level of the analysis no indexing on a grid is needed, thus allowing to extend the formalism also to a non-uniform partition of the Ω domain, as required in grid refinement techniques. In the following, a *semi-discrete* finite-volume approach will be employed, thus only space averages will be considered and the time dependency will be left for further integration, for example via standard Runge–Kutta algorithms [21].

A conservative discretization of sub-system (1) is accomplished, as usual for Euler system, by integrating each scalar equation on the volume element V of each cell C. By application of the Gauss theorem, one has then

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\mathbf{u}}(t) + \sum_i \frac{1}{h_i}(\overline{\mathbf{f}}_i^{+} - \overline{\mathbf{f}}_i^{-}) = 0, \tag{8}$$

where

$$\overline{\mathbf{u}}(t) = \frac{1}{|V|}\int_V \mathbf{u}(\mathbf{x}, t)\,\mathrm{d}V, \quad \overline{\mathbf{f}}_i^{\pm} = \frac{1}{|S_i^{\pm}|}\int_{S_i^{\pm}} \mathbf{f}_i[\mathbf{w}(\mathbf{x}, t)]\,\mathrm{d}S \tag{9}$$

denote, respectively, volume averages of each scalar component u^l over the cell C, and $\overline{\mathbf{f}}_i^{\pm}$ are flux values averaged on cell faces $S_i^{\pm}$. We note that $\overline{\mathbf{f}}_i$ fluxes are represented as exact *point values* in the (non-averaged) parallel coordinate i, and the corresponding differences in (8) provide the averaged flux derivatives.

In the case of sub-system (2), two different approaches can be pursued. In the SUP approach, magnetic field components are discretized by volume averages $\overline{B}_i$ as other scalar variables u^l, and the electric field components by face averages $\overline{E}_k$, as $\mathbf{f}_i$ fluxes. We have then

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{B}_i(t) + \sum_{j,k} \epsilon_{i,j,k} \frac{1}{h_j}(\overline{\mathbf{E}}_k^+ - \overline{\mathbf{E}}_k^-) = 0, \tag{10}$$

where $\epsilon_{i,j,k}$ is the Levi–Civita symbol and $\pm$ here refers to faces normal to the j-direction.

On the other side, in the CT formalism a discretization preserving the original (vector anti-symmetry) property is accomplished by a *surface* integration on a cell face followed by the application of the Stokes theorem on the line contour of that face. This leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{b}_i(t) + \sum_{j,k} \epsilon_{i,j,k} \frac{1}{h_j}(\overline{\mathbf{E}}_k^+ - \overline{\mathbf{E}}_k^-) = 0, \tag{11}$$

where now

$$\overline{b}_i(t) = \frac{1}{|S_i|}\int_{S_i} B_i(\mathbf{x},t)\,\mathrm{d}S, \quad \overline{E}_k^{\pm} = \frac{1}{|L_k^{\pm}|}\int_{L_k^{\pm}} E_k[\mathbf{w}(\mathbf{x},t)]\,\mathrm{d}L \tag{12}$$

are, respectively, the *staggered* discretized magnetic field variables, defined as integrals over the cell face S_i (we retain the formalism of non-capital b_i components to indicate staggered values to conform with other authors), while $\overline{E}_k^{\pm}$ are now line-averaged electric field components along face edges $L_k^{\pm}$, where the orientation depends on the normal to the face under consideration (see Fig. 1). Here the magnetic field

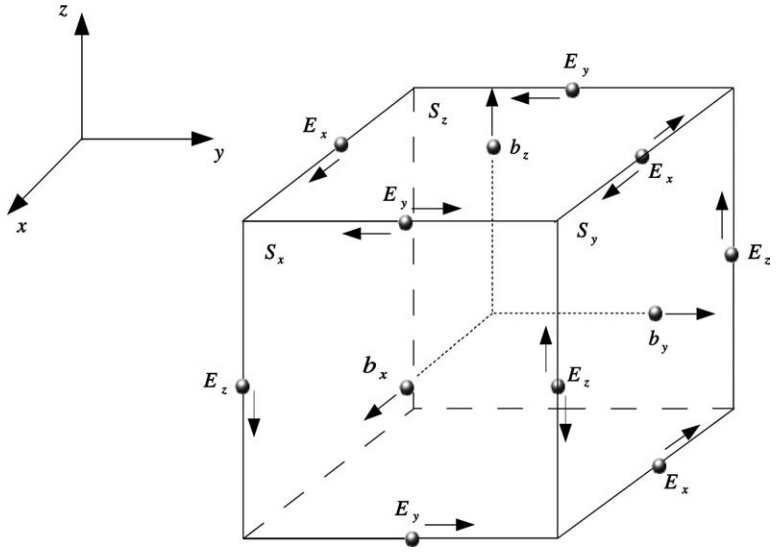


Fig. 1. The staggering of magnetic and electric vector fields in the CT framework. Only the S_x^+, S_y^+ and S_z^+ cell faces are visible. The arrows indicate the respective face normals, placed at intercell centers where b_i magnetic field components are defined, and the relative oriented contours for the application of Stokes' theorem, with arrows placed at edge centers where electric field components are defined.

components are represented as (normalized) *magnetic fluxes*, thus $\overline{b}_i$ are exact point values in the parallel coordinate i (as $\overline{\mathbf{f}}_i$) while $\overline{E}_k$ are point values with respect to their orthogonal coordinates $(i,j) \neq k$.

By using this staggered discretization, which yields a couple of normalized fluxes $\overline{b}_i^{\pm}$ defined at $S_i^{\pm}$ faces for each direction i, it is now possible to represent the volume average of the (parallel) first derivatives $\partial_i B_i(\mathbf{x})$. Therefore, for divergence-free initial conditions (3), we have

$$\sum_i \frac{1}{h_i}(\overline{b}_i^{+} - \overline{b}_i^{-}) = 0, \tag{13}$$

which will be preserved in time algebraically by the induction Eq. (11).

At this general level, anything is exact. Approximations (in space) arise when all MHD variables, starting from the discretized values, have to be reconstructed at the cell faces where fluxes are defined as point values. However, even at this preliminary step, differences in the magnetic field representations of Eq. (11) with respect to Eq. (10) have relevance:

- When primary data for the magnetic field are the staggered $\overline{b}_i^{\pm}$ components, on each cell C one has at disposal *two independent* sets of data. A first consequence is that *no reconstruction* is needed to evaluate (at a second-order approximation) these variables as argument of the corresponding $\overline{\mathbf{f}}_i$ fluxes at a S_i face. Moreover, these staggered data carry informations both on the volume averaged (or centered) values $\overline{B}_i$ and on the first derivative along the parallel coordinate. In fact, $(\overline{b}_i^{+} - \overline{b}_i^{-})/h_i$ provides a second-order (and then consistent) approximation of the $\partial_i B_i$ first derivative in point-wise (or finite difference) sense *inside each cell*.

  Related to the above is the property that each $\overline{b}_i$ component provides a *continuous* sampling across the corresponding S_i cell face. This follows from the definition in Eq. (12) and by taking into account that a divergence-free $\mathbf{B}(\mathbf{x})$ field entails a continuous elemental flux across a discontinuity surface (see [12] for details). In this way, the continuity property of the B_n normal component in the Rankine–Hugoniot jump relations retains a consistent representation in a finite volume setting. Staggered components can also be defined, in a *fully equivalent* way, by using the vector potential $\mathbf{A}$. In fact, by face-averaging the defining condition $\mathbf{B} = \nabla \times \mathbf{A}$, one has

$$\overline{b}_i = \sum_{j,k} \epsilon_{i,j,k} \frac{1}{h_j}(\overline{A}_k^{+} - \overline{A}_k^{-}), \tag{14}$$

  still assuring the divergence-free condition in the form (13). If needed, a time evolution for the numerical (line averaged) $\overline{A}_k$ components, consistent with Eq. (7), can also be formulated:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{A}_k(t) + \overline{E}_k = 0, \tag{15}$$

  preserving in time the representation of Eq. (14). Here, the numerical flux $\overline{E}_k$ are precisely the same as in Eq. (11).
- On the other hand, when primary data are represented by $\overline{B}_i$ volume averages, as in Eq. (10), only *one* numerical value per cell is available. The averaging procedure, which is now applied also along the i coordinate, while still assuring that $\overline{B}_i$ are consistent approximations of B_i point values, it also entails a loss of *direct* information on the point values at the S_i cell faces and on the correspondent parallel derivatives. Moreover, interpolation procedures are not sufficient, for discontinuous data, to recover these informations (related to the divergence-free property) and then some added argument, procedure, or constraint would be necessary, in our opinion, to avoid inconsistent approximations. In the following section we shall provide a more detailed analysis on this problem.

We can then conclude here that a finite volume CT formulation of the induction equation satisfy the general consistency demands presented in Section 1.1, provided the $\partial_i B_i$ are approximated using staggered $\overline{b}_i$ data.

In spite a CT-based discretization for MHD is a longstanding well-known framework for induction equation, its application to the whole MHD system is yet a matter of debate. In fact, it is yet a persistent viewpoint in numerical community that staggered collocation of magnetic field may be useful only to express the divergence-free relation, being otherwise not well suited for upwind formulation in Godunov-type schemes. In contrast with this viewpoint, we propose here to construct and test numerical schemes explicitly based on the CT discretization and on the related properties detailed above.

2.2. Reconstruction step: scalar vs divergence-free fields

In the following, we specialize to a Cartesian partition of the computational domain Ω. For grid indexing $1 \leqslant j \leqslant N_x$, $1 \leqslant k \leqslant N_y$, $1 \leqslant m \leqslant N_z$, a generic cell $C_{j,k,m}$ is defined as

$$C_{j,k,m} \equiv [x_{j-1/2}, x_{j+1/2}] \times [y_{k-1/2}, y_{k+1/2}] \times [z_{m-1/2}, z_{m+1/2}], \tag{16}$$

where each fractional index labels a cell interface (say $x_{j+1/2}$, here with $0 \leqslant j \leqslant N_x$), while a cell center has coordinates (x_j, y_k, z_m), with $x_j = (x_{j-1/2} + x_{j+1/2})/2$, $y_k = (y_{k-1/2} + y_{k+1/2})/2$, and $z_m = (z_{m-1/2} + z_{m+1/2})/2$. For simplicity, now we assume a uniform partition, along all directions, so that ($h_x \equiv x_{j+1/2} - x_{j-1/2}$, $h_y \equiv y_{k+1/2} - y_{k-1/2}$, $h_z \equiv z_{m+1/2} - z_{m-1/2}$) are the constant sizes for all cells. Under these settings, the primary volume-averaged array of fluid variables will be indicated as $\overline{\mathbf{u}}_{j,k,m}$. Surface-averaged staggered magnetic field components, defined at cell interfaces, will be indicated as $(\overline{b}_x)_{j+1/2,k,m}$, $(\overline{b}_y)_{j,k+1/2,m}$, and $(\overline{b}_z)_{j,k,m+1/2}$. The same notation holds then for face-centered flux components $\overline{\mathbf{f}}_i$, $i = x, y, z$, while the edge-centered $\overline{E}_k$ fields have indexing $(\overline{E}_z)_{j+1/2,k+1/2,m}$, and so forth for other components. Finally, the divided differences for functions located at cell interfaces introduced in the previous section, e.g., in Eqs. (8), (11) and (13), will be denoted here as D_i, $i = x, y, z$. For a generic 1-D scalar function f located at inter-cell points $x_{j\pm1/2}$, D_x is then given by

$$[D_x(f)]_j = \frac{1}{h_x}[\Delta_x f]_j; \quad [\Delta_x f]_j = f_{j+1/2} - f_{j-1/2}. \tag{17}$$

In higher order Godunov-type schemes, a one-dimensional scalar variable $u(x)$, represented by cell-centered data $\{\overline{u}_j\}$, is first reconstructed as approximated point values $\tilde{u}(x)$ inside any cell C_j and up to the interior cell faces $x_{j\pm1/2}$, where fluxes $f_x(u)$ have to be evaluated. This is accomplished by using local polynomials $\tilde{u}(x)$: (a) consistent with the cell averages values $\overline{u}_j$ and (b) having monotone or non-oscillatory properties. In a second-order approximation, one has the linear fit

$$\tilde{u}_j(x) = \overline{u}_j + \tilde{D}_x(\overline{u})(x - x_j), \tag{18}$$

where the non-oscillatory derivative $\tilde{D}_x(u)$ is usually constructed using slope limiters. In the simplest case, it is defined as

$$[\tilde{D}_x(\overline{u})]_j = \frac{1}{h_x}\mathrm{mm}([\Delta_x\overline{u}]_{j+1/2}, [\Delta_x\overline{u}]_{j-1/2}), \tag{19}$$

where $[\Delta_x\overline{u}]_{l+1/2} = \overline{u}_{l+1} - \overline{u}_l$ ($l = j, j-1$) and where $\mathrm{mm}(a, b)$ denotes the usual two-point *MinMod* (MM) algorithm. More elaborate limiters can likewise be constructed, using the Van Leer [22] monotonicity constraint or TVD [23] properties. For higher order schemes, ENO-based procedures have been developed (see [26], for a review) assuring a non-oscillatory reconstruction under weaker monotonicity constraints.

The reconstruction in Eq. (18) extends up to the interior face points $x_{j\pm1/2}$ providing a *left approximation* on the S_x^+ face and a *right approximation* on the S_x^- face, along the indicated coordinate, as needed for flux computation. At a given cell interface, say at $x = x_{j+1/2}$, in cases where a jump of size $\Delta_x u = \mathrm{O}(1)$ occurs, the estimated slope coefficients $[\tilde{D}_x]_j$ and $[\tilde{D}_x]_{j+1}$ using minmod limiters both vanish and the reconstructed $\tilde{u}(x)$ variable is represented by piecewise constant (first order) interpolants $\tilde{u}_j = \overline{u}_j$ and $\tilde{u}_{j+1} = \overline{u}_{j+1}$, respectively.

For a multidimensional function $u(\mathbf{x})$ (in the uniform grid defined above), a tensor-product representation with one-dimensional interpolants on each coordinate is usually adopted. For unstructured grids more elaborate procedures are required, but the basic ingredients (consistency with the cell averaged data and non-oscillatory constraints) still hold. In regular grids, the resulting second-order approximation $\tilde{u}(\mathbf{x}) \approx u(\mathbf{x})$ on each cell $C_{j,k,m}$ takes then the form

$$\tilde{u}(\mathbf{x}) = \overline{u} + \tilde{D}_x(\overline{u})(x - x_j) + \tilde{D}_y(\overline{u})(y - y_k) + \tilde{D}_z(\overline{u})(z - z_m), \tag{20}$$

where all quantities are implicitly calculated at cell center and each $\tilde{D}_i(\overline{u})$, is the non-oscillatory 1-D first derivative defined in Eq. (19).

When dealing with the magnetic field $\mathbf{B}(\mathbf{x})$, a different approach is needed, in general, to take into account the vector structure and the specific smoothness properties already quoted in the previous sections. Starting with face-averaged data $\overline{b}_i$ for $i = x, y, z$, satisfying the divergence-free condition (13)

$$D_x(\overline{b}_x) + D_y(\overline{b}_y) + D_z(\overline{b}_z) = 0, \tag{21}$$

the problem of representing each $B_i(\mathbf{x})$ field along the proper parallel coordinate inside a cell is already solved at the linear level, since the slope $D_i(\overline{b}_i)$ is at disposal. Instead, reconstruction is needed along the face (orthogonal) coordinates, where the field is sampled by averaged $[\overline{b}_x(x)]_{k,m}$ data.

For the $B_x(\mathbf{x})$ function one has then

$$\tilde{B}_x(\mathbf{x}) = \tilde{B}_x + D_x(\tilde{b}_x)(x - x_j) + \tilde{D}_y(\tilde{B}_x)(y - y_k) + \tilde{D}_z(\tilde{B}_x)(z - z_m). \tag{22}$$

Here, the cell-centered $\tilde{B}_x$ values result by noticing that $\tilde{B}_x(x, y_k, z_m)$ is the unique linear interpolant of a continuous function with data at the $x_{j\pm1/2}$ points. We have then

$$(\tilde{B}_x)_{j,k,m} = \tfrac{1}{2}[(\overline{b}_x)_{j+1/2} + (\overline{b}_x)_{j-1/2}]_{k,m}, \tag{23}$$

which provides an *approximation* of the cell averages and of the B_x point values:

$$\tilde{B}_x = \overline{B}_x + \mathrm{O}(h_x^2) = B_x + \mathrm{O}(h_x^2). \tag{24}$$

In Eq. (22) it is evident how staggered and cell-centered fields work differently. Along the parallel x coordinate, the $\tilde{B}_x(x)$ function is entirely defined by the $(\overline{b}_x)_{j\pm1/2,k,m}$ data, providing the field values at the cell interfaces, the approximated first derivative $[D_x(\overline{b}_x)]_{j,k,m}$, and the approximated cell-averaged value $(\tilde{B}_x)_{j,k,m}$. Reconstruction is needed instead to evaluate the B_x left–right values at the orthogonal cell faces $(S_y^{\pm}, S_z^{\pm})$, where B_x may have discontinuities and where, correspondingly, $\overline{b}_x$ and cell-centered values $\tilde{B}_x$ behave as the other $\overline{u}$ scalar variables.

A second property related to the divergence-free conditions is that the $\tilde{B}_x(x, \cdot)$ function in Eq. (22) maps a continuous function with first derivative $D_x(\overline{b}_x)$ which may be discontinuous. To evaluate the jump size, one considers the second difference $\Delta_x^2 \overline{b}_x \equiv \Delta_x \Delta_x \overline{b}_x$ centered at the proper interface point $(x_{j+1/2}, y_k, z_k)$. By taking into account the divergence-free relation (21) and the commutativity property of the two-point difference operators, one has

$$\frac{1}{h_x}\Delta_x^2 \overline{b}_x = -\frac{1}{h_y}\Delta_y \Delta_x \overline{b}_y - \frac{1}{h_z}\Delta_z \Delta_x \overline{b}_z, \tag{25}$$

where at least one of the differences $\Delta_x \overline{b}_y$ or $\Delta_x \overline{b}_z$ has size O(1), by definition.

Following the same procedure the (B_y, B_z) components can be represented, on the same cell, by the relations

$$\tilde{B}_y(\mathbf{x}) = \tilde{B}_y + \tilde{D}_x(\tilde{B}_y)(x - x_j) + D_y(\overline{b}_y)(y - y_k) + \tilde{D}_z(\tilde{B}_y)(z - z_m), \tag{26}$$

$$\tilde{B}_z(\mathbf{x}) = \tilde{B}_z + \tilde{D}_x(\tilde{B}_z)(x - x_j) + \tilde{D}_y(\tilde{B}_z)(y - y_k) + D_z(\overline{b}_z)(z - z_m), \tag{27}$$

and remarks made above on different behaviors of staggered $(\overline{b}_y, \overline{b}_z)$ and cell-centered $(\tilde{B}_y, \tilde{B}_z)$ values, respectively, and on the different smoothness properties depending on the involved coordinates, apply.

We remark here that this apparent *duality* in the reconstruction procedure appears to be fully consistent with the physical duality of the MHD Rankine–Hugoniot relations. In fact, using a local characteristic decomposition of the $(\mathbf{u}, \mathbf{B})$ variables, for a specified direction, say x, only the $(\overline{\mathbf{u}}, \tilde{B}_y, \tilde{B}_z)$ variables participate to the Riemann wave fan and may contribute, then, to the discontinuous characteristic modes. On the other hand, the continuous $\overline{b}_x$ variable has no role in the related upwinding procedures.

In upwind schemes, the reconstruction step presented above has relevance not only to recover face centered values, but also to approximate the $\mathbf{B}(\mathbf{x})$ field at any point inside a cell, as it is required in schemes adopting grid refinements (AMR) or multi-grid procedures. In this context, divergence-free interpolants similar to those derived here, even if based on quite different arguments, have been proposed in [24]. A related work [25] has presented a detailed analysis to show that, under linear interpolation based on staggered data, conservative properties and the divergence-free relation can be preserved in cell-subcells refinement procedures.

To summarize, it follows from this analysis that a numerical divergence-free magnetic field can be represented in an unambiguous way by using $\overline{b}_i$ staggered values (or equivalently the related numerical vector potential) as primary data. In this representation, second-order approximated first derivatives are consistent and non-oscillatory (no cell crossing is needed) and the variable $[\nabla \cdot \mathbf{B}]_{\text{num}} \equiv \sum_i [D_i(\overline{b}_i)]_{j,k,m}$ in Eq. (21) results to be exactly zero inside any point of $C_{j,k,m}$. At the same time, the reconstruction procedures at cell interfaces, where variables are discontinuous, provide definite rules on how to formulate upwind differentiation.

2.2.1. Reconstruction and central derivatives in non-CT schemes

A CT-based formalism helps also to analyze the reconstruction problem in the SUP framework, where only cell averaged $\overline{B}_i$ values are at disposal. By restricting to the $B_x(x, \cdot)$ function, a linear interpolant reads

$$\tilde{B}_x(x, \cdot) = [\overline{B}_x]_j + [C_x(\overline{B}_x)]_j(x - x_j) + \cdots, \tag{28}$$

where $C_x(\overline{B}_x)$ denotes a (by now unspecified) *consistent* approximation of the first derivative. For a piecewise differentiable function, at a point $x_{j+1/2}$ this interpolant provides in general *two* values, as left and right approximations, like for all other variables. By imposing there the additional continuity condition, specific to magnetic field components, one has

$$[\overline{B}_x]_j + \tfrac{1}{2}h_x[C_x(\overline{B}_x)]_j = [\overline{B}_x]_{j+1} - \tfrac{1}{2}h_x[C_x(\overline{B}_x)]_{j+1}, \tag{29}$$

and since Taylor expansion for C_x is not applicable across a discontinuous interface, this assures only an *implicit* way to express the numerical derivative in terms of the cell-centered data. This suggests that some added condition, like the non-oscillatory constraint, should be required even for x-wise interpolations, but how to recover the divergence-free condition remains here an open question.We are not aware of any SUP-based scheme where this problem has been properly addressed. It is a common practice, instead, to use for flux computations the mid-point average

$$(\tilde{b}_x)_{j+1/2,k,m} = \tfrac{1}{2}[(\overline{B}_x)_{j,k,m} + (\overline{B}_x)_{j+1,k,m}], \tag{30}$$

resulting in a $O(h_x)$ approximation. In turn, this entails an approximation for the first derivative given by the *central difference*

$$[D_x^{(c)}(\overline{B}_x)]_{j,k,m} = \tfrac{1}{2h_x}[(\overline{B}_x)_{j+1,k,m} - (\overline{B}_x)_{j-1,k,m}]. \tag{31}$$

However, when the relevant (x_{j-1}, x_{j+1}) stencil includes a discontinuity, say at $x_{j+1/2}$, by using the relation (29) with $C_x = D_x(\overline{b}_x)$ we have

$$[D_x^{(c)}(\overline{B}_x)]_{j,k,m} = [D_x(\overline{b}_x)]_{j,k,m} + O(1), \tag{32}$$

the $O(1)$ term resulting from the first derivative jump as estimated in (25). Using the same argument to the other (B_y, B_z) components, a final $[\nabla \cdot \mathbf{B}]_{\text{num}} = O(1)$ results. This is well documented in numerical experience where a central difference is used for discontinuous functions (the Gibbs pathology).

2.2.2. Extension to higher order

The first derivative discontinuity of a divergence-free field has also relevance to extend a CT-based reconstruction to higher orders ($r \geqslant 3$) of spatial accuracy. In Godunov-type schemes, higher order reconstructions for scalar variables are usually provided by ENO-based interpolants, like Weighted-ENO (WENO, [26] and references therein) and Convex-ENO (CENO, [27]). For regular grids, these interpolants, once defined for one-dimensional variables, can be extended to higher dimensions by a tensor-product representation.

For the magnetic field more elaborate procedures are needed, however, by taking into account that non-oscillatory derivatives along different directions are non-commutative. A general strategy we propose is to reconstruct first the vector potential components $\overline{A}_i$ in the usual way, by taking advantage that these are *scalar* variables, and then to define the b_i point values using the basic $\mathbf{B} = \nabla \times \mathbf{A}$ relation. However, using the vector potential alone is not sufficient to guarantee a divergence-free relation. A crucial step to make this procedure effective is to approximate the $\nabla \times \mathbf{A}$ derivatives by consistent two-point, fixed-stencil, high order finite differences.

As an example, in Appendix A we report the third-order implementation of the CENO reconstruction procedures.

2.3. Upwind step: Roe-type approximate Riemann solvers

Using the grid notation of Section 2.2, the MHD equations (8) and (11) take the form

$$\frac{\mathrm{d}}{\mathrm{d}t}[\overline{\mathbf{u}}(t)] + \sum_i D_i(\overline{\mathbf{f}}_i) = 0 \tag{33}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}[\overline{b}_i(t)] + \sum_{j,k} \epsilon_{i,j,k} D_j(\overline{E}_k) = 0, \tag{34}$$

the first set being centered at cell nodes and the second at cell interfaces. The overall system has now to be evaluated using some approximate Riemann solver, the same for all flux functions and for all Cartesian components, at a time. Here we make reference to Roe-type schemes, allowing a full resolution of the characteristic MHD modes, whereas the so-called central and central-upwind schemes, which avoid spectral decomposition, will be briefly treated in Section 3.2.

Let us then first specialize to the 1-D flux differentiation along the x coordinate and denote with $\mathbf{F}_x = [F_x^s]^{\mathrm{T}}$, for $s = 1, 2, \ldots, 7$, the array of all flux components, defined at the $x = x_{j+1/2}$ point for generic (y,z) coordinates of the $S_x = S_x^+$ face. These components are $F_x^{(l)} = f_x^{(l)}$, for $l = 1, 2, \ldots, 5$, whereas the components entering the induction equation are $F_x^{(6)} = -E_z$ and $F_x^{(7)} = E_y$. Correspondingly, we denote with $\mathbf{w}_x = [\rho, q_i, e, B_y, B_z]^{\mathrm{T}}$ the array of variables which need reconstruction as point values at the S_x face and as $[\mathbf{w}_x^{\mathrm{E}}(y,z), \mathbf{w}_x^{\mathrm{W}}(y,z)]$ (East–West, see Fig. 2) the corresponding left–right states. The magnetic field component $b_x(x,\cdot)$ in the parallel direction i satisfies $b_x^{\mathrm{E}} = b_x^{\mathrm{W}}$ by continuity, thus the arguments of flux functions $\mathbf{F}_x(\mathbf{w})$ have to be specified as $(\mathbf{w}_x^a, b_x)$, for $a = \mathrm{E}, \mathrm{W}$. For short, we denote as $\mathbf{F}_x^a = \mathbf{F}_x(\mathbf{w}_x^a, b_x)$.

Approximated Riemann solvers based on local linearization technique (see e.g. [28,29] for the MHD case) rely on the Roe matrix $\hat{A}_x$, defined as numerical Jacobian by

$$\mathbf{F}_x^{\mathrm{W}} - \mathbf{F}_x^{\mathrm{E}} = \hat{A}_x(\tilde{\mathbf{w}}) \cdot (\mathbf{w}_x^{\mathrm{W}} - \mathbf{w}_x^{\mathrm{E}}) \equiv \hat{A}_x \cdot \delta_x \mathbf{w}_x, \tag{35}$$

at any (y,z) point of the indicated S_x cell face. As usual, the $\hat{A}_x$ matrix is evaluated at an appropriate intermediate state $\tilde{\mathbf{w}} = [\tilde{\mathbf{w}}_x(\mathbf{w}_x^{\mathrm{W}}, \mathbf{w}_x^{\mathrm{E}}), b_x]$, and is required to be consistent with the Jacobian matrix J_x presented in Section 1.1. We notice the $\hat{A}_x$ matrix has rank *seven*, as in the pure 1-D case where $B_x = \text{const.}$ holds. In fact, in the multidimensional case the continuity condition $\delta_x b_x = 0$ plays a similar role, implying B_x is locally constant and does not participate to the characteristic wave fan.

To express the local flux variations in Eq. (35) in terms of characteristic modes, let us consider the spectral decomposition $\hat{A}_x = [R\Lambda R^{-1}]_x$, where, if λ_s, $s = 1, 2, \ldots, 7$, are the Roe matrix real eigenvalues, $\Lambda = \mathrm{diag}\{\lambda_s\}$, and columns of the R matrix are the corresponding right eigenvectors. In this representation, let us then split $\hat{A}_x = [\hat{A}_x]^+ + [\hat{A}_x]^-$, where the first term contains components with $\lambda_s > 0$ and the second with $\lambda_s < 0$, respectively. In this form, one has also $|\hat{A}_x| = [\hat{A}_x]^+ - [\hat{A}_x]^-$, where $|\hat{A}_x| = [R|\Lambda|R^{-1}]_x$ and $|\Lambda| = \mathrm{diag}\{|\lambda_s|\}$. By using standard procedures, from the two-point flux function $\mathbf{F}_x(\mathbf{w}_x^{\mathrm{E}}, \mathbf{w}_x^{\mathrm{W}}, b_x)$ a one-valued, continuous and monotone flux can be selected as upwind state by

$$\mathbf{F}_x^{\mathrm{U}} = \mathbf{F}_x^{\mathrm{E}} + [\hat{A}_x]^- \cdot \delta_x \mathbf{w}_x = \mathbf{F}_x^{\mathrm{W}} - [\hat{A}_x]^+ \cdot \delta_x \mathbf{w}_x, \tag{36}$$

so that the usual Roe flux formula comes out as

$$\mathbf{F}_x^{\mathrm{U}}(\mathbf{w}_x^{\mathrm{E}}, \mathbf{w}_x^{\mathrm{W}}, b_x) = \mathbf{F}_x^* - \mathbf{\Phi}_x, \tag{37}$$

where

$$\mathbf{F}_x^* = \tfrac{1}{2}(\mathbf{F}_x^{\mathrm{W}} + \mathbf{F}_x^{\mathrm{E}}); \quad \mathbf{\Phi}_x = \tfrac{1}{2}|\hat{A}_x| \cdot (\mathbf{w}_x^{\mathrm{W}} - \mathbf{w}_x^{\mathrm{E}}), \tag{38}$$

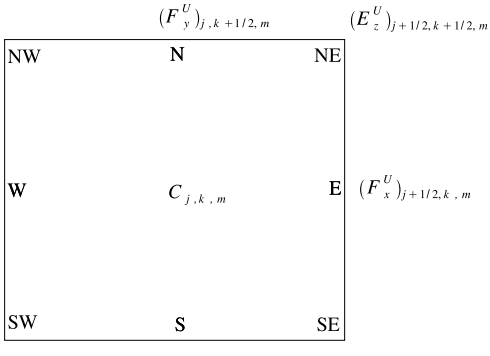


Fig. 2. The notation used for the reconstructed values inside a cell $C_{j,k,m}$ (here a cut through the center, normal to the z-direction, is shown) and the position of upwind fluxes, to be constructed via contributions from neighbouring cells. Fluxes are either defined as two-state functions located at cell interfaces, $(\mathbf{F}_x^{\mathrm{U}})_{j+1/2,k,m}$ and $(\mathbf{F}_y^{\mathrm{U}})_{j,k+1/2,m}$, or as four-state functions located at cell edges, $(\mathbf{E}_z^{\mathrm{U}})_{j+1/2,k+1/2,m}$.

the first term expressing the smooth component leading to a centered two-point formula in flux differentiation and the second the Roe-type component coming from the upwind procedure. At a discontinuity interface, the latter provides numerical dissipation which needs to be consistent with entropy conditions. To that purpose, for shock solutions where $\lambda_s \simeq 0$, a small amount of added dissipation $\lambda_s \to \lambda_s + \eta_s$ has to be introduced as entropy-fix to avoid unphysical behaviors.

With a straightforward extension, for coordinates $i = y, z$, the $\hat{A}_i$ Roe matrices and the upwind fluxes $\mathbf{F}_i^{\mathrm{U}}(\mathbf{w}_i, b_i) = \mathbf{F}_i^* - \mathbf{\Phi}_i$ can likewise be constructed, each function being evaluated at points of the S_i proper orthogonal face.

By taking advantage of the Roe formalism, which is based on *independent* 1-D $\hat{A}_i$ matrices, where characteristic modes are represented locally as planar waves along each direction, it is possible to evaluate, at a time, numerical fluxes collocated at different points. In fact, in Eq. (33) the final five-component numerical fluxes $\overline{\mathbf{f}}_i = \overline{\mathbf{F}}_i^{(1-5)}$ for $i = x, y, z$ are obtained as an average over the proper face S_i. This average involves only interior face points and then only a characteristic wave fan, the one represented by the $\hat{A}_i$ matrix. On the other hand, the remaining $\mathbf{F}_i^{(6-7)}$ flux components appearing in the induction equation (34) are defined as *point values* at the intersections of cell faces, where *different* characteristic wave fans overlap. These flux components can be likewise evaluated by a linear combination of 1-D upwind fluxes along the intersecting direction. It is a main feature of the UCT method that this combination follows a proper upwind selection rule, since a *same flux component* at the *same collocation point* results to have *two independent* representations in terms of characteristic wave fans.

As a prototype, we consider the E_z flux, which is defined at the $(x_{j+1/2}, y_{k+1/2}, z)$ points where faces S_x and S_y intersect. Let then denote as $\mathbf{w}_y = [\rho, q_i, e, B_x, B_z]^{\mathrm{T}}$ the set of variables having a representation in terms of the $\hat{A}_y$ matrix eigenmodes and by $\mathbf{F}_y = \mathbf{F}_y(\mathbf{w}_y^{\mathrm{N}}, \mathbf{w}_y^{\mathrm{S}}, b_y)$, the corresponding flux array, where now $(\mathbf{w}_y^{\mathrm{N}}, \mathbf{w}_y^{\mathrm{S}})$ (North–South) denote the left–right states along the y coordinate. At the indicated intersection points, $E_z = -F_x^{(6)} = F_y^{(6)}$ comes out to be a *four-state* function $E_z(\mathbf{w}^{a,b})$ where $a = \mathrm{N}, \mathrm{S}$, $b = \mathrm{E}, \mathrm{W}$, since the $\mathbf{w}$ argument contains both $\mathbf{w}_x$ and $\mathbf{w}_y$ variables (see again Fig. 2).

The sixth flux component $F_x^{(6)} = -E_z$, defined in Eq. (37) and specialized at the $y = y_{k+1/2}$ point, is represented by two independent contributions coming from states $(\mathbf{w}^{\mathrm{N}}, \mathbf{w}^{\mathrm{S}})$:

$$[F_x^{\mathrm{U}}(\mathbf{w}^{\mathrm{N}})]^{(6)} = f_x^{*\mathrm{N}} - \phi_x^{\mathrm{N}}, \quad [F_x^{\mathrm{U}}(\mathbf{w}^{\mathrm{S}})]^{(6)} = f_x^{*\mathrm{S}} - \phi_x^{\mathrm{S}}, \tag{39}$$

where $f_x^* = F_x^{*(6)}$ and $\phi_x = \Phi_x^{(6)}$. On the other hand, the $F_y^{(6)} = E_z$ flux function, defined for generic x values, is represented by two independent contributions for states $(\mathbf{w}^{\mathrm{E}}, \mathbf{w}^{\mathrm{W}})$ at the $x = x_{j+1/2}$ point:

$$[F_y^{\mathrm{U}}(\mathbf{w}^{\mathrm{E}})]^{(6)} = f_y^{*\mathrm{E}} - \phi_y^{\mathrm{E}}, \quad [F_y^{\mathrm{U}}(\mathbf{w}^{\mathrm{W}})]^{(6)} = f_y^{*\mathrm{W}} - \phi_y^{\mathrm{W}}, \tag{40}$$

where now $f_y^* = F_y^*(6)$ and $\phi_y = \Phi_y^{(6)}$. By taking into account that $F_y^{(6)} = -F_x^{(6)}$ at the same $(x_{j+1/2}, y_{k+1/2})$ point, a one-valued numerical flux function having the continuity and upwind properties along each direction can be then constructed at the linear level by

$$E_z^{\mathrm{U}}(\mathbf{w}) = E_z^* - \phi_y + \phi_x, \tag{41}$$

where

$$E_z^* = \tfrac{1}{4}[E_z^{\mathrm{NE}} + E_z^{\mathrm{SE}} + E_z^{\mathrm{NW}} + E_z^{\mathrm{SW}}]; \quad \phi_y = \tfrac{1}{2}(\phi_y^{\mathrm{E}} + \phi_y^{\mathrm{W}}), \quad \phi_x = \tfrac{1}{2}(\phi_x^{\mathrm{N}} + \phi_x^{\mathrm{S}}), \tag{42}$$

and $E_z^{a,b} = E_z(\mathbf{w}^{a,b})$, with $a = \mathrm{N}, \mathrm{S}$, $b = \mathrm{E}, \mathrm{W}$. The continuity property of this relation implies, in particular, that for any possible orientation, namely along a cell face diagonal or along a cell face side, the corresponding one-dimensional planar mode is taken into account in proper way, to within $\mathrm{O}(|\delta_x \mathbf{w}_x||\delta_y \mathbf{w}_y|)$, by the upwind combination above. The E_z flux, given in Eq. (41) for generic z coordinate of the (S_x, S_y) common side, is finally evaluated by line averaging as $(\overline{E}_z)_{j+1/2,k+1/2,m}$ numerical flux. In a similar way, the

other x and y components can then be constructed by a proper combination of upwind fluxes along the corresponding orthogonal coordinates and line averaging along the parallel coordinate.

This completes the presentation of the UCT formalism. For later reference, we quote here the flux formulas given above, specialized to a first-order approximation, which constitutes the building block of any Godunov-type scheme, needed for higher order extensions. A piecewise constant reconstruction for $\mathbf{w}_x$ variables at the $x = x_{j+1/2}$ point, needed in Eq. (37), gives $\mathbf{w}_x^{\mathrm{E}} = (\mathbf{w}_x)_j$, $\mathbf{w}_x^{\mathrm{W}} = (\mathbf{w}_x)_{j+1}$, at any (y,z) point. Face averaging reduces to the one-point evaluation $\overline{\mathbf{F}}_x(\mathbf{w}_x, b_x) = \mathbf{F}_x(\overline{\mathbf{w}}_x, \overline{b}_x)$ and the first-order numerical flux in Eq. (37), for the five-component $\mathbf{f}$ flux array, reads then (over-bars are henceforth omitted for brevity):

$$(\mathbf{f}_x^{\mathrm{U}})_{j+1/2,k,m} = (\mathbf{f}_x^*)_{j+1/2,k,m} - (\mathbf{\Phi}_x)^{(1-5)}_{j+1/2,k,m}, \tag{43}$$

where, at the indicated (y_k, z_m) points,

$$(\mathbf{f}_x^*)_{j+1/2} = \tfrac{1}{2}[\mathbf{f}_x((\mathbf{w}_x)_{j+1}, (b_x)_{j+1/2}) + \mathbf{f}_x((\mathbf{w}_x)_j, (b_x)_{j+1/2})] \tag{44}$$

and

$$[\mathbf{\Phi}_x]_{j+1/2} = \tfrac{1}{2}|A_x(\tilde{w}_{j+1/2})| \cdot [(\mathbf{w}_x)_{j+1} - (\mathbf{w}_x)_j], \tag{45}$$

and so forth for other directions. To the same first-order approximation, the E_z flux in Eq. (41) can be calculated at $(x_{j+1/2}, y_{k+1/2}, z_m)$ edge points (the z_m centering will be assumed implicitly in the following). By specializing the $E_z = f_y^* = -f_x^* = -(v_x B_y - v_y B_x)$ argument variables, the smooth term results as:

$$(E_z^*)_{j+1/2,k+1/2} = -\tfrac{1}{2}[(\hat{v}_x b_y)_{j+1} + (\hat{v}_x b_y)_j]_{k+1/2} + \tfrac{1}{2}[(\hat{v}_y b_x)_{k+1} + (\hat{v}_y b_x)_k]_{j+1/2}, \tag{46}$$

where $(\hat{v}_x)_{l,k+1/2} = \frac{1}{2}[(v_x)_k + (v_x)_{k+1}]_l$, for $l = j, j+1$, and $(\hat{v}_y)_{j+1/2,l} = \frac{1}{2}[(v_y)_j + (v_x)_{j+1}]_l$, for $l = k, k+1$. It is worth noticing that in Eq. (46), the E_z^* term contains the (b_x, b_y) staggered components and the resulting four-states combination at the $(x_{j+1/2}, y_{k+1/2})$ point *cannot* be reduced simply to an interpolation or averaging form based on the four cell-centered values of the argument. On the other hand, in the dissipative Roe-type fluxes, centering at $y = y_{k+1/2}$ of the ϕ_x term comes out as a two-point average in the orthogonal coordinate and correspondingly for the ϕ_y term, so that

$$(\phi_x)_{j+1/2,k+1/2} = \tfrac{1}{2}[(\phi_x)_k + (\phi_x)_{k+1}]_{j+1/2}, \quad (\phi_y)_{j+1/2,k+1/2} = \tfrac{1}{2}[(\phi_y)_j + (\phi_y)_{j+1}]_{k+1/2}. \tag{47}$$

On the computational side, this form is also economical and of easy implementation, since the four contributes there involved can be derived from the $\mathbf{\Phi}_x$ and $\mathbf{\Phi}_y$ fluxes already worked out for the fluid variables.

2.4. On the problem of numerical monopoles

We consider now some main differences of the present approach with other schemes, by focusing on the problem of numerical monopoles. These unwanted effects may arise when the $\sum_i D_i(\overline{\mathbf{f}}_i)$ term for momentum equations in Eq. (33) fails to recover the proper $[\mathbf{J} \times \mathbf{B}]_{\mathrm{num}}$ Lorentz force in the original non-conservative form, or, equivalently, when the orthogonality condition (6) is not satisfied with sufficiently high accuracy.

This problem has been analyzed in details by [30], in a general discretization setting and with no particular reference to specific upwind differentiations. Here we follow some of his arguments and notations to show the behavior of various classes of MHD schemes in comparison with our UCT method. We specialize to second-order flux differentiation, as it is implemented in most of the schemes in the literature.

By restricting for sake of simplicity to a two-dimensional configuration, the Maxwell stresses $T_{i,j}$, $i, j = x, y$, entering the momentum flux components are given by

$$T_{x,x} = -T_{y,y} = \tfrac{1}{2}(B_y^2 - B_x^2), \quad T_{x,y} = T_{y,x} = -B_x B_y, \tag{48}$$

evaluated at the proper interface points, that is $(x_{j\pm1/2}, y_k)$ for $T_{x,x}$ and $T_{y,x}$, $(x_j, y_{k\pm1/2})$ for $T_{y,y}$ and $T_{x,y}$. We notice that these terms refer specifically to the f_i^* smooth part of the relevant flux components, since Roe-matrix contributions give only a diffusive term of the Lorentz force. For flux differences related to the q_x momentum, one has then the algebraic relations

$$D_x(T_{x,x}) = \mu_x(B_y)D_x(B_y) - \mu_x(B_x)D_x(B_x), \quad D_y(T_{y,x}) = -\mu_y(B_y)D_y(B_x) - \mu_y(B_x)D_y(B_y), \tag{49}$$

all being centered on a (x_j, y_k) point, where D_x and D_y are the usual divided differences of Eq. (17), and where for components $a = B_x, B_y$ we define the averages $\mu_x(a) = (a_{j+1/2,k} + a_{j-1/2,k})/2$ and $\mu_y(a) = (a_{j,k+1/2} + a_{j,k-1/2})/2$. These two-point averaging on the cell center give $\mu_x(a) = a_{j,k} + \mathrm{O}(h_x^2)$ and $\mu_y(a) = a_{j,k} + \mathrm{O}(h_y^2)$, for second-order numerical fluxes. By summing the two differences defined above, one has (let $h = h_x = h_y$)

$$D_x(T_{x,x}) + D_y(T_{y,x}) = -B_x[D_x(B_x) + D_y(B_y)] + B_y[D_x(B_y) - D_y(B_x)] + \mathrm{O}(h^2), \tag{50}$$

where the second term on the right-hand side provides the numerical approximation of the x component of the Lorentz force. By considering then the correspondent q_y momentum flux, the approximation to the orthogonality condition finally becomes:

$$[\mathbf{B}\cdot(\nabla\cdot\mathbf{T})]_{\mathrm{num}} = -(B_x^2 + B_y^2)[\nabla\cdot\mathbf{B}]_{\mathrm{num}} + \mathrm{O}(h^2), \tag{51}$$

which is satisfied to within the truncation error if the numerical divergence-free relation holds (at least) to the same accuracy order. For smooth flows this is clearly true in any discretization of first derivatives. But when discontinuities are present, this term can be even of O(1) size.

In particular:

- In schemes where $(B_x)_{j+1/2,k}$ and $(B_y)_{j,k+1/2}$ are reconstructed as two-point average of the corresponding cell-centered data, as discussed in Section 2.2.1, the resulting $[\nabla\cdot\mathbf{B}]_{\mathrm{num}}$ in Eq. (51) is expressed by the central differences $D_x^{(\mathrm{c})}(B_i)$, giving contributions of $\mathrm{O}(h^2)$ size in smooth regions and of O(1) size near discontinuous interfaces.

  In Powell's eight-wave scheme, the source terms introduced in the momentum equation given by $\mathbf{B}([\nabla\cdot\mathbf{B}]_{\mathrm{num}})$ yield a subtraction of these monopoles, which is another way to express the Lorentz force in its non-conservative form.

  In projection schemes, a $[\nabla\cdot\mathbf{B}]_{\mathrm{num}} = 0$ condition, yet based on cell centered data, is enforced at each time step as an added constraint, so that monopoles, which arise when flux derivatives are computed, are prevented to grow in time.

  In previous CT-based schemes (e.g. [8–10]), where staggered variables are actually at disposal, only cell-centered averaged data $\tilde{B}_i$, derived as in Eq. (23), are then employed in flux computations. In that case the resulting $[\nabla\cdot\mathbf{B}]_{\mathrm{num}}$ term in Eq. (51) is still expressed via derivatives based on central differencing as in Eq. (31), and numerical monopoles are now produced. These can be evaluated now exactly by noticing that

$$D_x^{(\mathrm{c})}(\tilde{B}_x) = D_x(\overline{b}_x) + h_x^2 R_x, \quad R_x = \frac{1}{4h_x^3}\Delta_x^3\overline{b}_x \tag{52}$$

  and by similar expressions for the other (B_y, B_z) components. In this relation the residual $h_x^2 R_x$ is O(1) at points where the $\overline{b}_x$ first derivative is discontinuous, as can be seen from Eq. (25). The final estimate gives $[\nabla\cdot\mathbf{B}]_{\mathrm{num}} = \sum_i h_i^2 R_i = \mathrm{O}(1)$ since the residuals R_i, $i = x, y, z$, being related to the transverse jumps, do not cancel out, in general.

- The class of UCT schemes proposed here prevents the onset of monopoles, since $\overline{b}_i$ staggered data enter directly in the corresponding f_i^* flux components, and $[\nabla \cdot \mathbf{b}]_{\text{num}} = \sum_i [D_i(\overline{b}_i)] = 0$ results in Eq. (51).

We finally notice that, in order to avoid numerical monopoles, flux derivatives have to be computed at the *same* time-stepping level, since time-splitting techniques prevent exact cancellation of $[\nabla \cdot \mathbf{B}]_{\text{num}}$ terms.

3. Examples of application of the UCT method

In the present section, the numerical strategies outlined in Section 2 will be applied to a couple of existing Godunov-type schemes, originally designed for fluid dynamics. We have chosen two completely different schemes, a classical Roe-type scheme based on field-by-field limiting along characteristics, and a simple central-type scheme adopting a two-speed upwind flux with component-wise limiting in the reconstruction algorithms. Both schemes are proposed here in the semi-discrete form, appropriate for our UCT method, and then integrated using TVD Runge–Kutta time discretizations of the appropriate order [21].

3.1. Roe-type: the positive scheme

We present here the MHD implementation of a second order flux-limited scheme proposed by Liu and Lax [14,15]. This approach allows for an easy formulation for multidimensional hyperbolic systems and can then represent a well-behaved alternative to standard TVD-based schemes (see [23]). Moreover, these authors have introduced a new *positivity* principle, which is more appropriate for multidimensional systems, to which TVD do not apply. This stability principle relies, in particular, on the symmetrizable form of the system under investigation and is then well suited also for MHD equations.

Flux-limited schemes are constructed as a proper combination of an accurate *second-order* smooth numerical flux, for example the centered approximation $\mathbf{F}^* = (\mathbf{F}_j + \mathbf{F}_{j+1})/2$ or a Lax–Wendroff term, and a *first-order* diffusive flux of the form $\mathbf{F}^{\text{diss}} = \mathbf{F}^* - \mathbf{\Phi}^{(1)}$, in a way that when the flow is smooth $\mathbf{F}^{\text{U}} = \mathbf{F}^*$ and when discontinuities are present $\mathbf{F}^{\text{U}} = \mathbf{F}^{\text{diss}}$. In a symbolic way, this combination can be expressed as $\mathbf{F}^{\text{U}} = \mathbf{F}^* - (I - L)\mathbf{\Phi}^{(1)}$, where L is a diagonal operator whose entries are *flux limiter* functions $\phi_s(\theta_s)$ acting, in general, on characteristic modes and assuring $\phi_s(\theta_s) = 1$ for smooth modes and $\phi_s(\theta_s) = 0$ otherwise.

In particular, in this *positive* scheme by Liu and Lax, *two* first-order dissipative schemes are combined with different flux limiters: a first (least dissipative) flux of Roe-type and a second (more dissipative) flux of LLF-type, the latter acting also as entropy-fix for the former. In the MHD implementation, the Roe matrices $\hat{A}_i(\tilde{\mathbf{w}})$ are constructed in the space of primitive variables, with a simple two-point average to define the state $\tilde{\mathbf{w}}$ and with the eigenvector formulation given in [31] to remove degeneracies.

The resulting formula, for each Cartesian flux component, reads then:

$$\mathbf{F}^{\text{U}}_{j+1/2} = \tfrac{1}{2}[\mathbf{F}(\mathbf{w}_j) + \mathbf{F}(\mathbf{w}_{j+1})] - \tfrac{1}{2}R[c_{\text{ROE}}|\Lambda|(I - L) + c_{\text{LLF}}\alpha(I - M)]R^{-1}(\mathbf{w}_{j+1} - \mathbf{w}_j), \tag{53}$$

where the two coefficients combining the two dissipative terms are chosen such as $c_{\text{ROE}} + c_{\text{LLF}} = 1$, M is a diagonal term made up of smooth minmod-type limiters that satisfy $0 \leqslant \phi_s(\theta_s) \leqslant 1$, $0 \leqslant \phi_s(\theta_s)/\theta_s \leqslant 1$, L is a diagonal term whose sharper limiters satisfy $0 \leqslant \phi_s(\theta_s) \leqslant 2$, $0 \leqslant \phi_s(\theta_s)/\theta_s \leqslant 2$, for a resulting maximum CFL number of 0.5 (for more technical aspects, the reader is referred to the cited works).

In our experience, when applied to the MHD system by following the UCT strategies described in Section 2, this scheme results to be a robust and accurate flux-limited scheme, with almost no extra computational effort, if compared to standard multidimensional TVD schemes.

3.2. Central-type schemes

In the fluid-dynamics community, schemes adopting simple one or two-speed numerical fluxes with component-wise reconstruction (no characteristic decomposition is thus required), are often referred to as central or central-upwind schemes, since it has been proved these approximate Riemann solvers come out in Roe-type schemes by some form of central averaging over the Riemann characteristic wave fan (see [32,33] for the latest developments). For most applications, even in shock dominated flows, these schemes give yet satisfactory results and are surely more economical than Roe-type schemes, even when third or higher order reconstruction is applied.

In the present paper, two UCT implementations of central-upwind schemes based on the Harten–Lax–Van Leer (HLL) [34] two-speed flux and component-wise reconstruction are considered:

1. the simplest second-order HLL-UCT implementation is performed by reconstructing all (potentially discontinuous) variables by a *monotonized centered* (MC) Van Leer [22] limiter. Flux upwinding is then achieved by a HLL flux formula, as presented below for the MHD system. Time integration is finally given by the second-order Runge–Kutta scheme. This second-order central-type scheme will be here referred to as MC-HLL-UCT;
2. A third-order central scheme, based on the local Lax–Friedrichs (LLF) flux and Liu and Osher Convex-ENO (CENO) scheme [27], has been already presented and tested in our first paper LD. The choice of CENO has been favored by the following interesting features: first, the high order reconstruction algorithm is able to reduce itself to minmod-type limiters (MM or MC) at discontinuities, thus strongly reducing spurious oscillations typical of component-wise reconstruction via ENO interpolants; then, a formulation based on point values, rather than cell averages, allows the use of purely one-dimensional reconstruction routines. The same algorithms are used here in a central-upwind scheme now equipped with an HLL flux formula, and this new third-order central-type scheme will be named CENO-HLL-UCT. For the interested reader, we report in Appendix A the main computational steps used.

In the following, we define the HLL two-speed formulas appropriate for MHD. The HLL upwind flux for the fluid components, say $\mathbf{f}_x$ at $(x_{j+1/2}, y_k, z_m)$, may be written as

$$\mathbf{f}_x^{\mathrm{U}} = \frac{\alpha_x^+ \mathbf{f}_x^{\mathrm{E}} + \alpha_x^- \mathbf{f}_x^{\mathrm{W}} - \alpha_x^+ \alpha_x^- (\mathbf{u}^{\mathrm{W}} - \mathbf{u}^{\mathrm{E}})}{\alpha_x^+ + \alpha_x^-}, \tag{54}$$

where as usual $\mathbf{f}_x^a = \mathbf{f}_x(\mathbf{w}_x^a, b_x)$, $a = \mathrm{E}, \mathrm{W}$, and

$$\alpha_x^{\pm} = \max\{0, \pm\lambda_x^{\pm}(\mathbf{w}_x^{\mathrm{W}}, b_x), \pm\lambda_x^{\pm}(\mathbf{w}_x^{\mathrm{E}}, b_x)\}, \tag{55}$$

and similarly for the other y and z components. Here, to avoid the definition of an intermediate Roe-type state, we have chosen to calculate the dissipative $\alpha_x^{\pm}$ terms by taking the maximum eigenvalues (in MHD systems related to the fast magneto-sonic speeds: $\lambda_x^{\pm} = v_x \pm c_x^{\mathrm{f}}$) between left and right states.

Notice that, by rearranging the terms, it is still possible to rewrite Eq. (54) in the form of Section 2.3 $\mathbf{f}^* - \mathbf{\Phi}$, so that the discussion on $[\nabla \cdot \mathbf{B}]_{\mathrm{num}}$ of Section 2.4 applies unchanged. On the other hand, we can not use the same composition rules outlined in Section 2.3 to derive Roe-type magnetic fluxes. The correct form of the E_z^{U} upwind flux comes out now by averaging over the two overlapping x and y Riemann wave fans at the $(x_{j+1/2}, y_{k+1/2})$ edge, thus

$$E_z^{\mathrm{U}} = \frac{\alpha_x^+ \alpha_y^+ E_z^{\mathrm{NE}} + \alpha_x^+ \alpha_y^- E_z^{\mathrm{SE}} + \alpha_x^- \alpha_y^+ E_z^{\mathrm{NW}} + \alpha_x^- \alpha_y^- E_z^{\mathrm{SW}}}{(\alpha_x^+ + \alpha_x^-)(\alpha_y^+ + \alpha_y^-)} - \frac{\alpha_y^+ \alpha_y^-}{\alpha_y^+ + \alpha_y^-}(b_x^{\mathrm{S}} - b_x^{\mathrm{N}}) + \frac{\alpha_x^+ \alpha_x^-}{\alpha_x^+ + \alpha_x^-}(b_y^{\mathrm{W}} - b_y^{\mathrm{E}}), \tag{56}$$

where the $\alpha_x^{\pm}$ and $\alpha_y^{\pm}$ at edge center points should be calculated by taking the maximum characteristic speeds (in absolute value) among the four reconstructed states, whereas for sake of efficiency we actually consider

the maximum over the two neighboring inter-cell points, where these speeds are already at disposal from fluid fluxes computation.

In spite of a different form, Eq. (56) share with Eq. (41) all the required upwind properties, that is to be a four-state formula and to reduce to the correct 1-D flux for shocks aligned with x, y, or $x-y$ diagonals. The usual settings are instead recovered for LLF fluxes, which are the particular cases of the corresponding HLL ones for $\alpha_x^+ = \alpha_x^- = \alpha_x$ and $\alpha_y^+ = \alpha_y^- = \alpha_y$, where now each single $\alpha_i = \max\{\alpha_i^+, \alpha_i^-\}$ speed defines a symmetrical (central) averaged Riemann fan.

It is interesting to notice that the upwind fluxes of Eq. (54) generalize to MHD the formulas of the [33] central scheme for the Euler system. Moreover, Eq. (56) coincides with that defined in the same paper for Hamilton–Jacobi scalar equations, which is to be expected since each component of the induction equation, for a given velocity field and expressed in terms of the magnetic potential, has exactly the form of such equations.

4. Numerical results on test problems

We present now a standard set of numerical examples to assess accuracy, stability and effective divergence-free properties of our UCT schemes presented above. In the following tables and plots we shall refer to these three schemes as POSITIVE-UCT (the flux-limited Roe-type scheme of Section 3.1), MC-HLL-UCT, and CENO-HLL-UCT (respectively, the second and third-order central-upwind schemes described in Section 3.2). Most of the choices in the following tests have been inspired by those in [1], although a direct comparison of UCT with the other methods for MHD proposed or reported is difficult to achieve, due to the use of different underlying schemes, and it will be not our main concern here. Comparisons will be made instead among our three UCT schemes, and also with respect to their corresponding non-UCT counterparts, i.e., when magnetic field components are discretized and evolved exactly as the other fluid variables (the SUP framework), with the same underlying scheme. We shall refer to these Euler-like versions as *base schemes* (BS). This will enable us to appreciate the effective improvements introduced by the use of our UCT method.

In the following sub-sections we shall provide quantitative measures of the divergence-free properties in our schemes. To distinguish between the two types of numerical representations discussed in Sections 2.2.1 and 2.4, here we shall define, for each cell $C_{j,k,m}$, the two quantities

$$[\nabla\cdot\mathbf{B}]_{\rm num} = \sum_i D_i^{(\rm c)}(\tilde{B}_i), \qquad [\nabla\cdot\mathbf{b}]_{\rm num} = \sum_i D_i(b_i), \tag{57}$$

where $D_i^{(\rm c)}$ are the central derivatives defined in Eq. (31), applied to the cell-centered derived data $\tilde{B}_i$, and D_i are the usual two-point divided differences defined in Eq. (17). As already discussed, while $[\nabla\cdot\mathbf{b}]_{\rm num}$ will be zero to within machine accuracy for second-order UCT schemes, and arbitrarily small (see Appendix A) for CENO-HLL-UCT (it is simply not defined for the BS counterparts), the $[\nabla\cdot\mathbf{B}]_{\rm num}$ variable may have O(1) jumps at discontinuities. Recall that in UCT schemes the onset of monopoles is actually measured by the $[\nabla\cdot\mathbf{b}]_{\rm num}$ variable, rather than by $[\nabla\cdot\mathbf{B}]_{\rm num}$ like in other CT and non-CT methods for MHD where cell-centered fields are employed in fluxes. Notice that, in spite of the use of the staggered magnetic field $\mathbf{b}$ in the initial conditions and in the computations, the output data will be referred to the cell-centered interpolated field $\mathbf{B}$.

To obtain more accurate results in the various test problems, schemes may be tuned by choosing each time the most appropriate slope limiter in the reconstruction routine. In all simulations, a CFL number of 0.5 and an adiabatic coefficient of $\gamma = 5/3$ are used, unless differently specified.

4.1. A convergence test: the oblique CP Alfvén wave

A propagating circularly polarized (CP) Alfvén wave is a well-known exact *non-linear* solution of the multidimensional MHD system, and therefore it is often used to measure accuracy of a numerical

approximation. We notice that this test involves only smooth solutions, thus problems related to the divergence-free condition are not expected to arise here. Following [1], we consider a CP wave propagating on the (x,y) Cartesian plane at an angle $\alpha = 30°$ relative of the x-axis. Periodic boundary conditions can be applied if the computational box is defined by $0 < x < 1/\cos\alpha$, and $0 < y < 1/\sin\alpha$. Let the coordinate along the propagation direction be $\xi = x\cos\alpha + y\cos\alpha$ and $\eta = y\cos\alpha - x\sin\alpha$ be the coordinate along the transverse direction, then the initial values of wave variables may be given as $v_\eta = B_\eta = A\sin(2\pi\xi)$ and $v_z = B_z = A\cos(2\pi\xi)$, where A measures the wave amplitude. The constant parallel components are defined as $v_\xi = 0$ and $B_\xi = 1$, along with a uniform density $\rho = 1$ and pressure $p = 0.1$. The chosen values correspond to a wave period $T = 1$, to a propagation Alfvénic speed $v_A = 1$, and to a sound speed $c_s = \sqrt{\gamma p/\rho} \simeq 0.4$.

To check numerical accuracy, the evolved solution $\mathbf{w}(x,y,t)$ at a time $t_{max} = nT$, for a given number n of periods, is usually compared to the initial condition $\mathbf{w}(x,y,0)$ and the difference $\delta\mathbf{w}$ evaluated in some norm is then tabulated at different grid resolutions. To compare with the Toth results, we also adopt the L_1 norm to evaluate the relative errors and the averaged $\bar{\delta}$ values are measured only by taking into account the transverse vector components v_η, v_z, B_η, and B_z. However, since we have noticed that convergence was not precisely achieved with the values $A = 0.1$ and $n = 5$ suggested by Toth, especially for our most accurate third-order central scheme, we have decided to use here the safer values $A = 0.01$ and $n = 1$. We think that the reason is due to compressible effects: in spite of the fact that a CP Alfvén wave is an exact solution, regardless of its amplitude A, this kind of wave is known to be subject to the so called *parametric decay* instability (see [35] and references therein). This is due to non-linear wave–wave interactions that, via coupling to compressive modes, lead to wave distortion and decay, so that when this happens we are no longer comparing with the true solution.

The results are reported in Table 1 for all our three UCT schemes, with resolutions ranging from 8^2 to 128^2. In our highest accurate scheme CENO-HLL we use the smoother MM limiter, while for the second-order schemes the sharper MC limiter is employed, otherwise results at low resolution are not quite satisfactory, and a small value of $c_{LLF} = 0.01$ is used in POSITIVE. In cases where only smooth fields are involved, UCT schemes may perform worst than their BS counterparts, due to the additional interpolations needed to recover the cell-centered magnetic fields (e.g., in output files). We have verified that in this particular test problem the error for the BS versions (not reported here) are approximately 10% less than for their UCT counterparts.

The results for all UCT schemes are also plotted in Fig. 3, where the convergence rates are apparent. Note that in this kind of problems CENO-HLL is obviously by far the most performing code: the accuracy reached with a resolution of 32^2 is comparable to that obtained by the second-order methods with 128^2.

Concerning divergence-free properties, the particular settings of the problems are such that both $[\nabla\cdot\mathbf{B}]_{num} = 0$ and $[\nabla\cdot\mathbf{b}]_{num} = 0$ to within machine accuracy for the second-order UCT schemes, because the invariance direction η is made to coincide with the cell diagonal, a condition required by the double periodicity (independently on the value of the angle α). For CENO-HLL-UCT, typical values for the maximum values of $[\nabla\cdot\mathbf{B}]_{num}$ and $[\nabla\cdot\mathbf{b}]_{num}$ are 10^{-8} and 10^{-12}, respectively.

Table 1
Averaged errors on the transverse velocity and magnetic field components for the oblique 2-D CP Alfvén wave problem

	$\bar{\delta}_8$	$\bar{\delta}_{16}$	$\bar{\delta}_{32}$	$\bar{\delta}_{64}$	$\bar{\delta}_{128}$
POSITIVE-UCT	0.53097	0.12792	0.04273	0.01254	0.00322
MC-HLL-UCT	0.60488	0.13133	0.04507	0.01392	0.00393
CENO-HLL-UCT	0.31342	0.03759	0.00461	0.00056	0.00012

The errors are measured from the numerical solution at time $t = 1$ (one wave period), compared with the corresponding initial setting, for increasing resolutions and for the various UCT schemes.

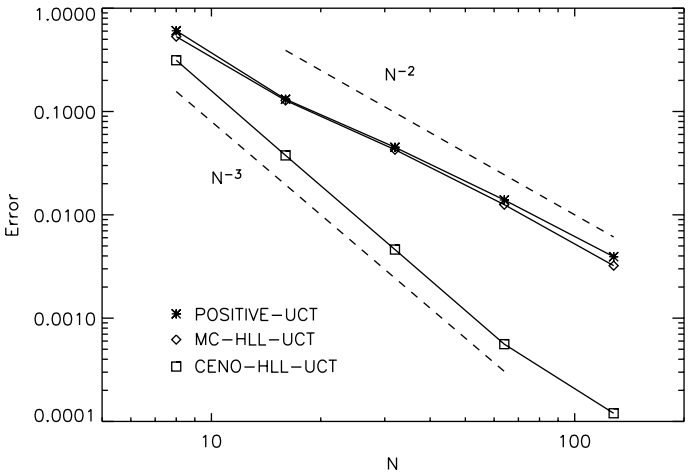


Fig. 3. Convergence test for the oblique 2-D CP Alfvén wave problem. Average L_1 errors on transverse η and z wave components are shown in logarithmic scale for our three UCT schemes. Precise second-order accuracy is achieved only asymptotically for POSITIVE-UCT and MC-HLL-UCT, due to the use of the MC limiter which tends to somehow sharpen wave profiles, while third-order convergence is clearly reached by CENO-HLL-UCT, here employing MM in the reconstruction algorithm, already at very low resolutions.

4.2. Rotated shock-tube problems

These test problems involve the propagation of discontinuities defined by usual 1-D shock-tubes on a 2-D computational plane, and are relevant to some main aspect considered here on the divergence-free condition. However, specific divergence-free properties may be hidden if one (or both) of the following special conditions hold:

1. the initial magnetic field is uniform;
2. the propagation direction lies along cell sides or diagonals.

In the first case, it is clear that any representation of $[\nabla \cdot \mathbf{B}]_{\text{num}}$ will be exactly zero for initial fields. Then its subsequent time evolution will only give a measure of the ability of a numerical scheme to preserve the initial $[\nabla \cdot \mathbf{B}]_{\text{num}}$, even if it is non-vanishing, while a characterizing aspect of any CT-based scheme is precisely the possibility to deal with discontinuous divergence-free fields.

Concerning the second case, the problem is the same as in the previous test, though here involves discontinuities and will be described in details. For initial symmetric conditions where all variables are defined as $w(x,y) = w(\xi)$, where ξ is a coordinate making an angle α with respect to the x-axis, and they do not depend on the transverse η coordinate, it is important to check at later times not only the evolution properties along the ξ coordinate but also the conservation of the transverse invariance. In particular, the $\nabla \cdot \mathbf{B} = 0$ condition expressed in the rotated coordinates is given by

$$\nabla \cdot \mathbf{B} = \partial_\xi B_\xi + \partial_\eta B_\eta = 0, \tag{58}$$

and the equivalent condition $B_\xi(x,y) = \text{const.}$ can be recovered only if B_η (and all other variables, of course) are η independent. However, in numerical applications based on standard Cartesian grids this condition may be achieved only if the ξ and η directions are aligned with the cell diagonals (or the cell sides). In fact, any discontinuity front making a different angle will be discretized with unequal jumps along x and y. Thus, even when the $[\nabla \cdot \mathbf{b}]_{\text{num}} = 0$ condition holds, the $B_\xi = \text{const.}$ relation does so only in approximate sense, with O(1) jumps where discontinuities occur. More strictly related to the errors in the B_ξ variable, which has to be necessarily calculated from the interpolated $\tilde{B}_i$ cell-centered fields, is the $[\nabla \cdot \mathbf{B}]_{\text{num}}$ variable, which in fact shows the highest jumps precisely at discontinuities. We have verified that, when the angle $\alpha = 45°$ is

chosen, both the conditions $B_\xi = \text{const.}$ and $[\nabla \cdot \mathbf{B}]_{\text{num}} = 0$ hold within machine accuracy, as in the previous test.

The numerical domain for the oblique shock tube tests may be reduced, as cleverly suggested by Toth, to just a narrow strip $[0,1] \times [0, 2/N]$, discretized with a $N \times 2$ grid (so that $\mathrm{d}x = \mathrm{d}y$). *Shifted* boundary conditions in the η direction are applied and $\alpha = \tan^{-1} 2 \approx 63.4°$. Each 2-D run, performed with $N = 256$, is compared with the corresponding 1-D test on a 1024 grid, by using the data along the x-axis (the first row) at a final time $t_{\max} \cos\alpha$ (where $t_{\max}$ refers to the 1-D test).

The initial left (L) and right (R) states of the three shock tube problems considered here are reported in Table 2, and the final times are $t_{\max} = 0.08$ for ST-1, $t_{\max} = 0.2$ for ST-2, and $t_{\max} = 0.1$ for ST-3. ST-1 is a coplanar 2-D problem of converging shocks in an initially uniform magnetized background, ST-2 is a non-coplanar case involving Alfvénic discontinuities, and ST-3 is the famous (coplanar) Riemann problem involving the so-called compound (or intermediate) shock. Note that in ST-2 and ST-3 the magnetic field has jumps in the initial data and $\alpha \neq 45°$, so none of the special cases indicated above apply. Other Riemann problems have been checked and the UCT schemes appear to behave well in all cases, including for example non-coplanar tests with compound shocks. The same three problems have been tested in LD for the CENO-LLF-UCT code on a symmetric $N \times N$ grid with $\alpha = 45°$, while Toth shows just the first two tests: ST-1 with precisely the same settings and ST-2, the non-coplanar 2.5-D problem, with $\alpha = 45°$.

Here the most steepening slope limiters and a minimum amount of viscosity are used in our schemes: thus POSITIVE uses the *Superbee* (SB) limiter for all entries of L, $c_{\text{LLF}} = 0.01$, and the limiter in CENO-HLL is MC. In Table 3 we report, for all tests and numerical schemes, the $\overline{\delta}$ average L1 norm over the variables involved in each test, the L1 norm of the errors in B_ξ, the L1 norms of variables $[\nabla \cdot \mathbf{B}]_{\text{num}}$ and $[\nabla \cdot \mathbf{b}]_{\text{num}}$ (respectively, $[\nabla \cdot \mathbf{B}]_{\text{avr}}$ and $[\nabla \cdot \mathbf{b}]_{\text{avr}}$), and their maximum absolute value over the computational domain (respectively, $[\nabla \cdot \mathbf{B}]_{\max}$ and $[\nabla \cdot \mathbf{b}]_{\max}$). The comparison between the x projection of the evolved quantities and the reference 1-D runs for our three UCT schemes are plotted in Figs. 4, 5, and 6, for the ST-1, ST-2, and ST-3 tests, respectively.

For problems involving sharp discontinuities, like shock tubes, schemes based on characteristics are clearly preferable since sharp limiters, which lead to more accurate results, may be often used there without producing spurious oscillations. This can be appreciated from the plots and may be measured more quantitatively from the reported errors in Table 3, which are the lowest for our POSITIVE scheme. If limiters less sharp than SB are used in POSITIVE-UCT the errors obviously increase: with MC we find $\overline{\delta} = 0.0146$ for ST-1, $\overline{\delta} = 0.0153$ for ST-2, and $\overline{\delta} = 0.0203$ for ST-3; with the smoothest MM we find $\overline{\delta} = 0.0201$ for ST-1, $\overline{\delta} = 0.0225$ for ST-2, and $\overline{\delta} = 0.0298$ for ST-3. Concerning the central schemes, the use of the MC limiter in the reconstruction step allows to produce accurate results with a rather low level of oscillations even in the absence of characteristics decomposition. These schemes are just less accurate at contact and Alfvénic discontinuities, since the related characteristic waves do not enter the HLL flux definition.

Table 2
Constant left (L) and right (R) states for the three oblique shock tube problems

	ρ	v_ξ	v_η	v_z	p	B_ξ	B_η	B_z
Test ST-1: L	1	10	0	0	20	$5/\sqrt{4\pi}$	$5/\sqrt{4\pi}$	0
Test ST-1: R	1	−10	0	0	1	$5/\sqrt{4\pi}$	$5/\sqrt{4\pi}$	0
Test ST-2: L	1.08	1.2	0.01	0.5	0.95	$2/\sqrt{4\pi}$	$3.6/\sqrt{4\pi}$	$2/\sqrt{4\pi}$
Test ST-2: R	1	0	0	0	1	$2/\sqrt{4\pi}$	$4/\sqrt{4\pi}$	$2/\sqrt{4\pi}$
Test ST-3: L	1	0	0	0	1	0.75	1	0
Test ST-3: R	0.125	0	0	0	0.1	0.75	−1	0

Table 3
Numerical errors for the three oblique shock tube problems

	$\overline{\delta}$	δB_ξ	$[\nabla \cdot \mathbf{B}]_{\max}$	$[\nabla \cdot \mathbf{B}]_{\text{avr}}$	$[\nabla \cdot \mathbf{b}]_{\max}$	$[\nabla \cdot \mathbf{b}]_{\text{avr}}$
Test ST-1						
POSITIVE-BS	0.0225	0.0051	0.11×10^3	0.30×10^1	–	–
POSITIVE-UCT	0.0126	0.0019	0.94×10^2	0.26×10^1	0.11×10^{-12}	0.67×10^{-15}
MC-HLL-BS	0.0304	0.0090	0.12×10^3	0.32×10^1	–	–
MC-HLL-UCT	0.0227	0.0021	0.10×10^3	0.25×10^1	0.11×10^{-12}	0.44×10^{-15}
CENO-HLL-BS	0.0320	0.0092	0.12×10^3	0.33×10^1	–	–
CENO-HLL-UCT	0.0227	0.0021	0.11×10^3	0.27×10^1	0.22×10^{-6}	0.57×10^{-8}
Test ST-2						
POSITIVE-BS	0.0155	0.0016	0.65×10^1	0.34×10^0	–	–
POSITIVE-UCT	0.0119	0.0006	0.70×10^1	0.29×10^0	0.57×10^{-13}	0.33×10^{-15}
MC-HLL-BS	0.0237	0.0025	0.61×10^1	0.35×10^0	–	–
MC-HLL-UCT	0.0200	0.0005	0.66×10^1	0.21×10^0	0.57×10^{-13}	0.22×10^{-15}
CENO-HLL-BS	0.0201	0.0025	0.71×10^1	0.41×10^0	–	–
CENO-HLL-UCT	0.0154	0.0006	0.63×10^1	0.25×10^0	0.35×10^{-7}	0.14×10^{-8}
Test ST-3						
POSITIVE-BS	0.0365	0.0043	0.16×10^2	0.66×10^0	–	–
POSITIVE-UCT	0.0165	0.0005	0.86×10^1	0.30×10^0	0.14×10^{-13}	0.17×10^{-15}
MC-HLL-BS	0.0586	0.0064	0.22×10^2	0.13×10^1	–	–
MC-HLL-UCT	0.0295	0.0003	0.46×10^1	0.19×10^0	0.28×10^{-13}	0.22×10^{-15}
CENO-HLL-BS	0.0612	0.0079	0.25×10^2	0.17×10^1	–	–
CENO-HLL-UCT	0.0257	0.0004	0.65×10^1	0.23×10^0	0.13×10^{-6}	0.18×10^{-8}

For the various schemes, the errors are calculated from the 2-D 256×2 run with respect to the corresponding 1-D high resolution run with 1024 grid points. The displayed errors are the average $\overline{\delta}$ L1 norm of all the involved variables, the error on the parallel field component B_ξ (that supposed to remain constant), the maximum and averaged values of $[\nabla \cdot \mathbf{B}]_{\text{num}}$, and the maximum and averaged values of $[\nabla \cdot \mathbf{b}]_{\text{num}}$, which is not defined, of course, for non-CT schemes.

Finally, to appreciate the level of damage that monopoles can produce, we plot in Fig. 7 the results of POSITIVE-BS for ST-3, where monopoles are free to arise (even in initial data since here none of the two special conditions apply) and grow. By also looking at the tabulated errors, it is clear that BS results are systematically worst with respect to the correspondent UCT versions (see in particular the errors on B_ξ).

4.3. The Orszag–Tang vortex problem

A well-known model problem to study the transition to MHD turbulence is provided by the so-called Orszag–Tang vortex, which has been later adopted as a standard 2-D test for MHD shock-capturing codes. The initial conditions are here $v_x = -\sin y$, $v_y = \sin x$, $B_x = -\sin y$, $B_y = \sin 2x$, $\rho = \gamma^2$, $p = \gamma$ (so the sound speed and the initial Mach number are both 1), $v_z = B_z = 0$. The computational domain is a square $0 < x, y < 2\pi$, $N \times N$ box with periodic boundary conditions along both directions, while the output time is $t_{\max} = \pi$ (note that in LD the magnetic field was normalized against $B_0 = 1/\sqrt{4\pi}$).

All our UCT schemes have been tested with the MC limiter ($c_{\text{LLF}} = 0.1$ in POSITIVE), at the resolutions of 50^2, 100^2, 200^2, and 400^2. In Table 4, we report the averaged errors (over the six evolving variables). For each scheme, errors are measured with respect to a highest accuracy run at 400^2, obtained as the average of the results from the three UCT schemes. More qualitative comparisons can be appreciated in Fig. 8, where gray-scale images of the temperature for the 200^2 runs, for both the BS and UCT versions, are compared to the respective reference solution (the UCT run with 400^2 grid points). Due to the symmetry of the problem, only the $[0, \pi] \times [0, 2\pi]$ first half of the domain is displayed. The first thing to be noticed is that the BS

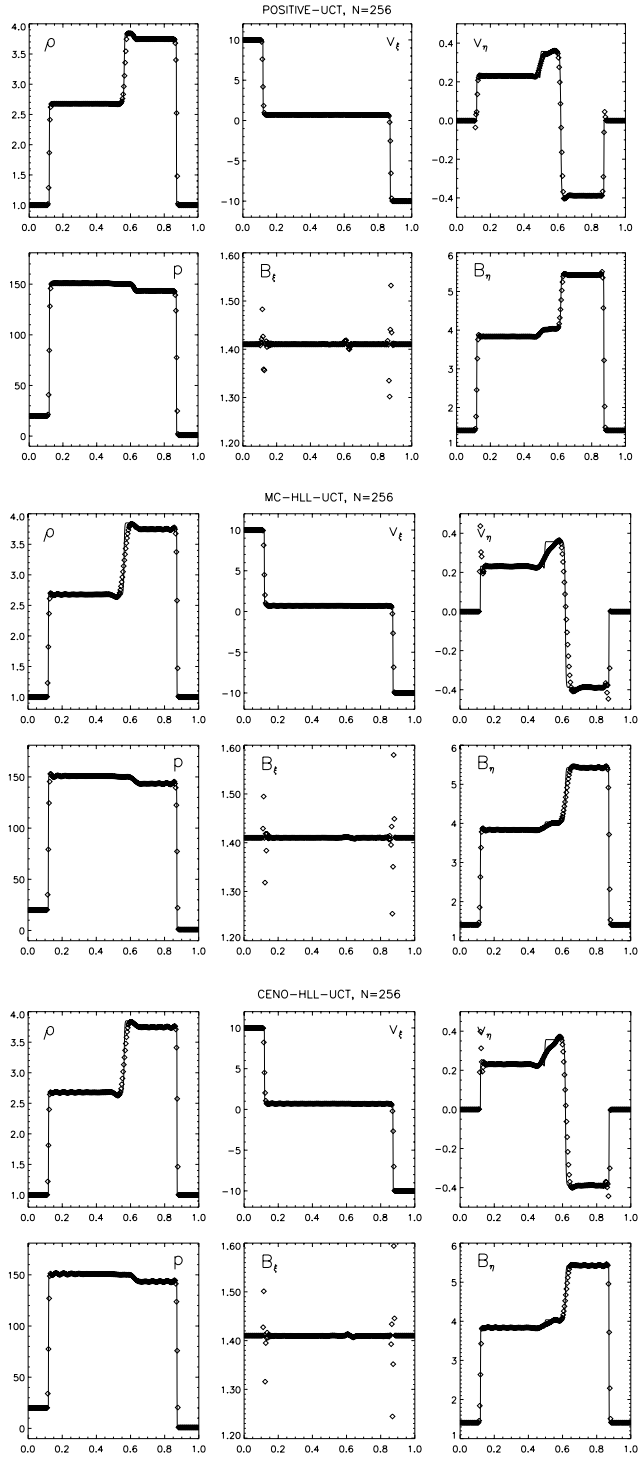


Fig. 4. The oblique ST-1 shock tube problem. The numerical results from our three UCT schemes obtained on a 256×2 grid (symbols) are compared with the 1-D solution on a 1024 grid (solid line).

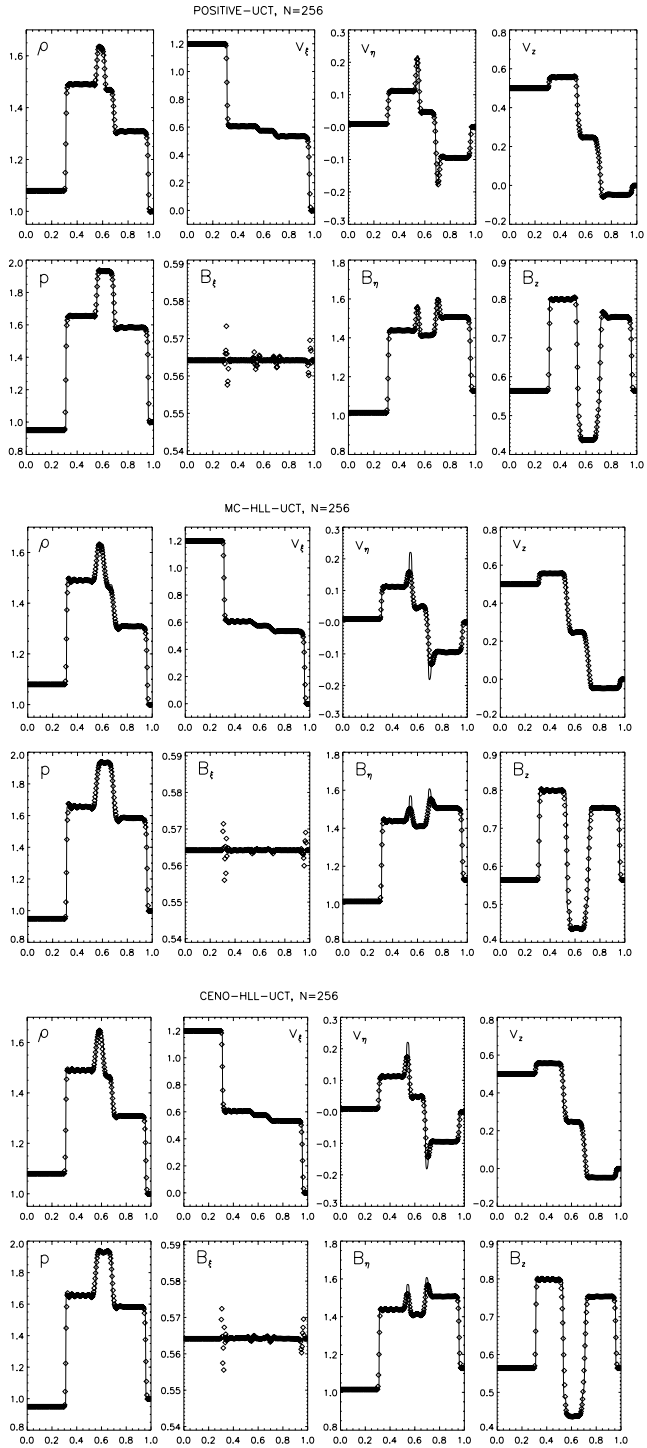


Fig. 5. The oblique ST-2 shock tube problem. The numerical results from our three UCT schemes obtained on a 256×2 grid (symbols) are compared with the 1-D solution on a 1024 grid (solid line).

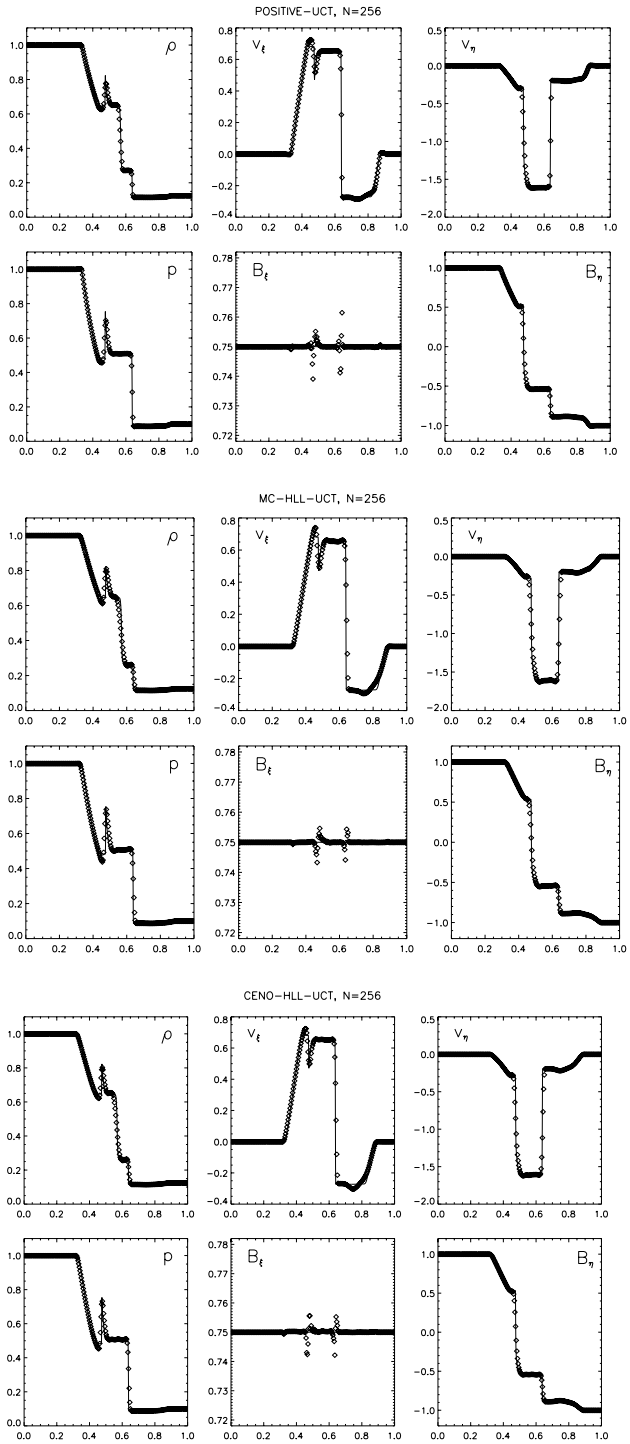


Fig. 6. The oblique ST-3 shock tube problem. The numerical results from our three UCT schemes obtained on a 256×2 grid (symbols) are compared with the 1-D solution on a 1024 grid (solid line).

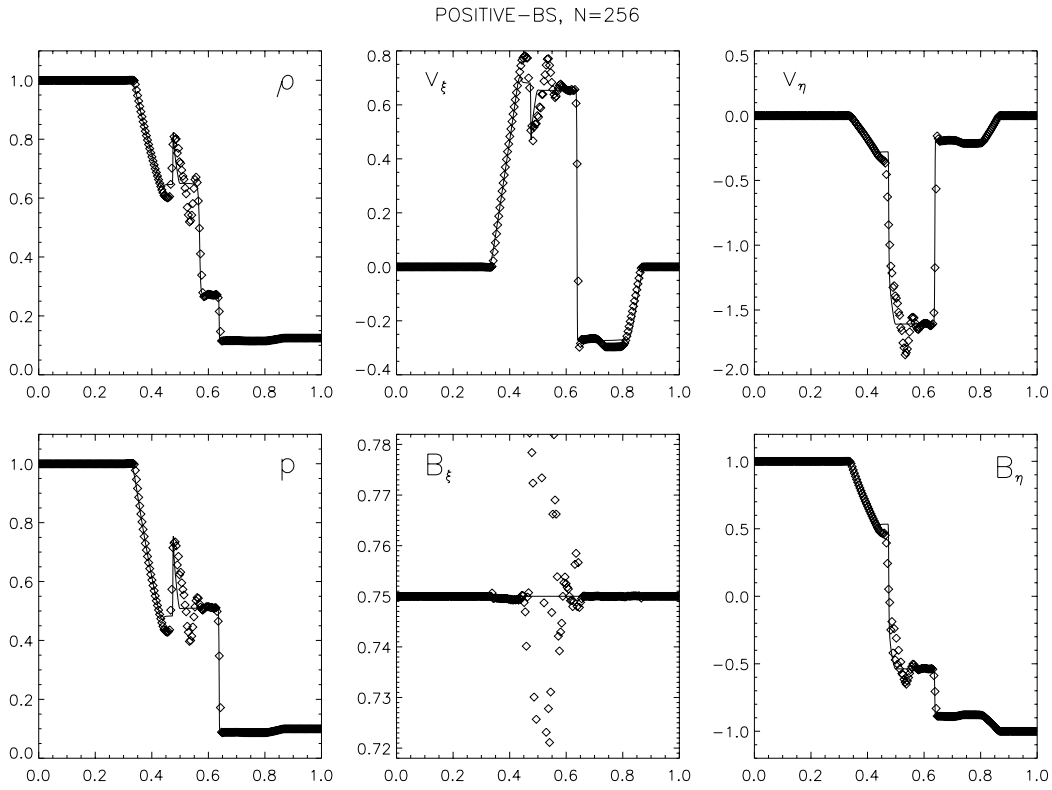


Fig. 7. The oblique ST-3 shock tube problem, this time for the base scheme (BS) version of POSITIVE. Note the presence of errors induced by the numerical monopoles, here free to arise and grow in time.

versions clearly produce incorrect results: the darkest, rather homogeneous features of the reference solutions appear more structured, with whiter filaments, and sometimes spurious oscillations are visible. By also reading Table 4, it is then quite apparent that the three schemes behave similarly in this context. Due to the numerous steepened structures the order of accuracy is close to one in all cases, so that the errors are very similar. CENO-HLL-UCT is the most performing scheme, giving results almost identical to the Roe-type POSITIVE-UCT, while MC-HLL-UCT is of course more dissipative, but its results are still satisfactory.

4.4. The fast rotor problem

In [10], a model problem to study the onset and propagation of strong torsional Alfvén waves was presented and analyzed. A disk of radius $r = 0.1$ made up of dense fluid ($\rho = 10$) rotates with high angular

Table 4
Averaged L1 norms on the involved variables for the OT vortex problem at $t = \pi$

	$\overline{\delta}_{50}$	$\overline{\delta}_{100}$	$\overline{\delta}_{200}$
POSITIVE-UCT	0.1601	0.0780	0.0296
MC-HLL-UCT	0.1898	0.0920	0.0358
CENO-HLL-UCT	0.1477	0.0657	0.0205

Errors are measured against a high resolution (400^2) reference run.

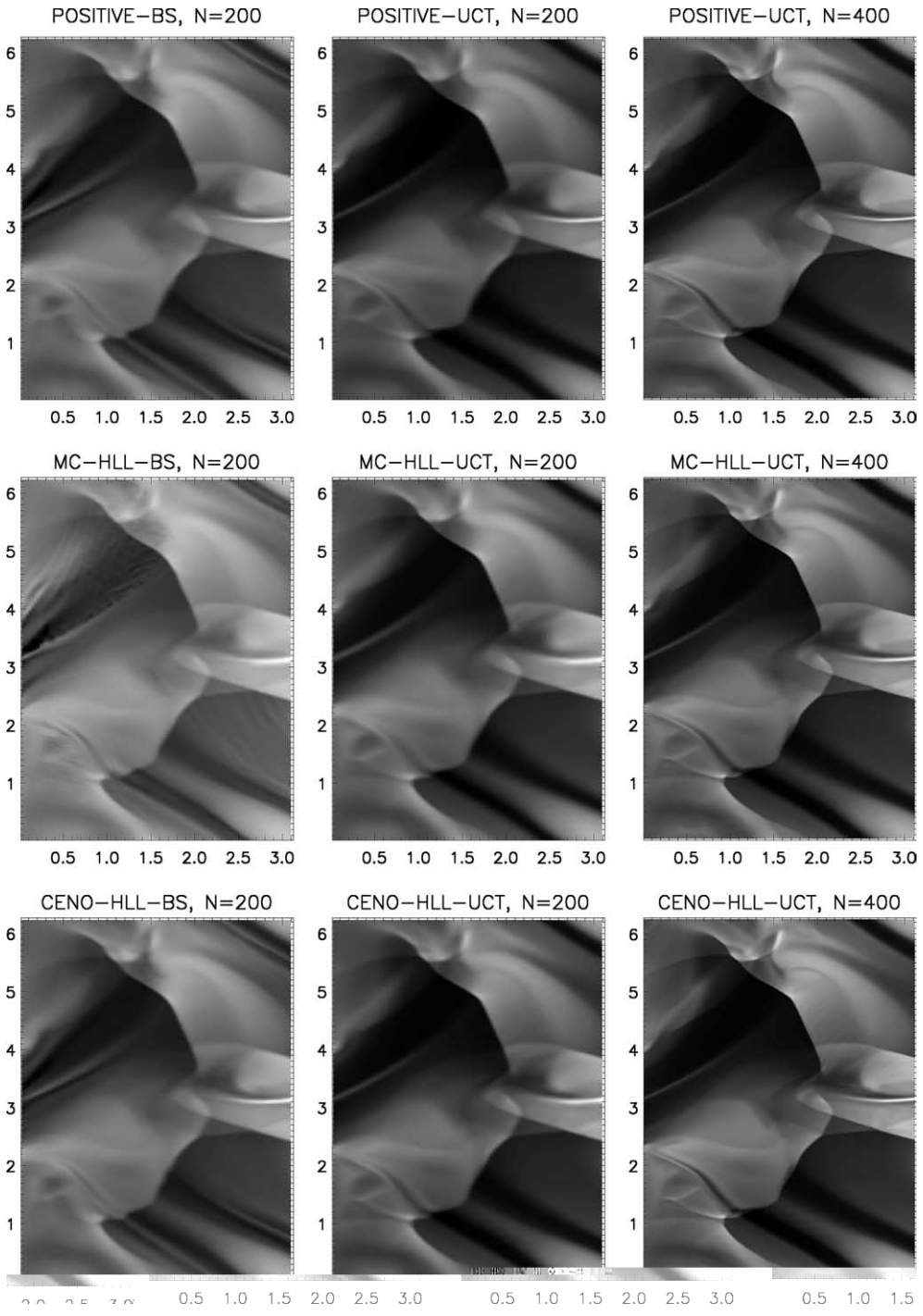
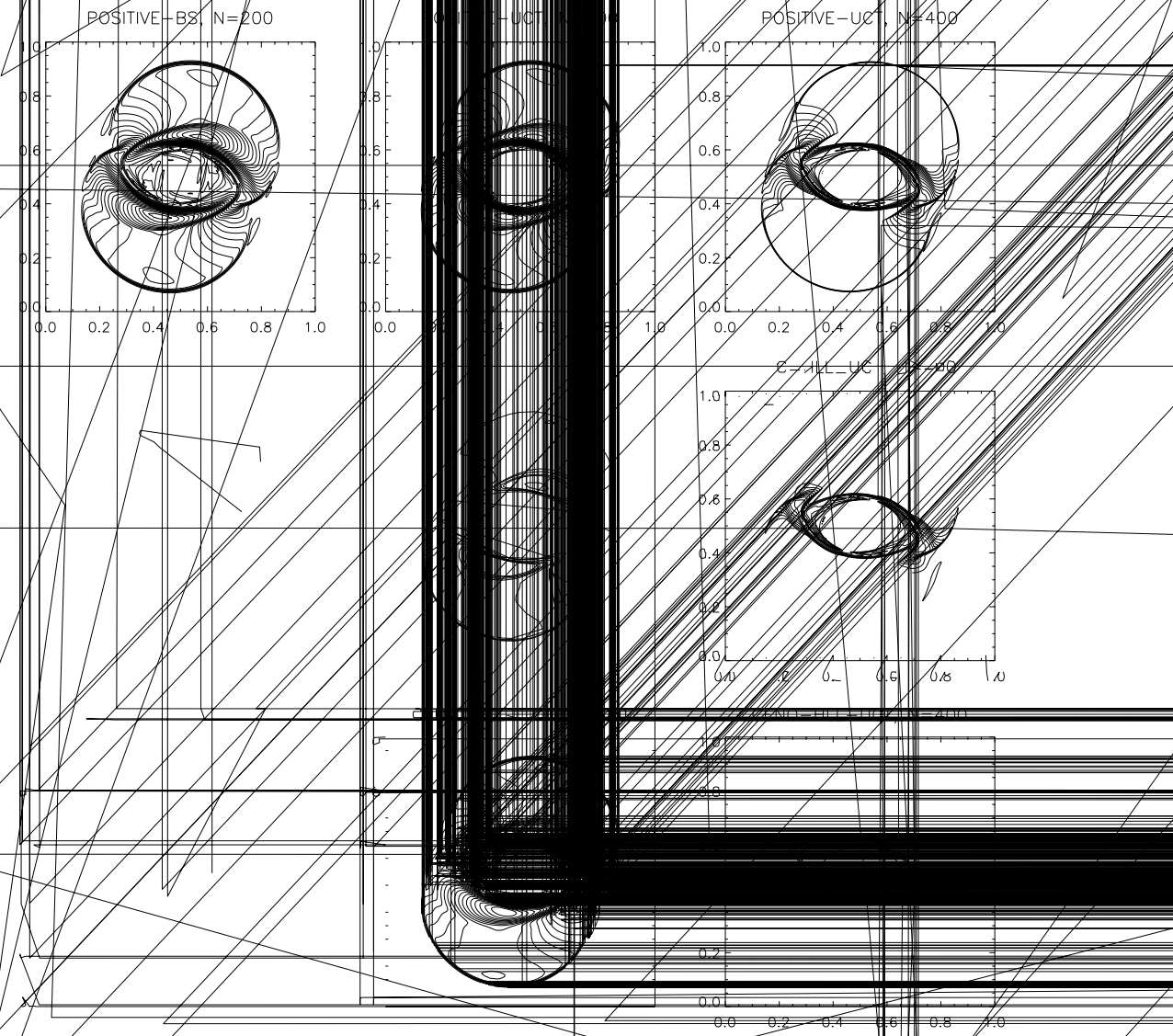


Fig. 8. Gray-scale images of the temperature $T = p/\rho$ distribution in the Orszag–Tang vortex problem. For each scheme, low resolution (200^2) results for the BS and UCT versions are compared with the corresponding high resolution (400^2) UCT reference run.

Table 5
Averaged L1 norms on the involved variables for the fa[illegible] problem at $t = 0.15$

	$\bar{\delta}_{50}$	$\bar{\delta}$	$\bar{\delta}_{200}$
POSITIVE-UCT	0.1751	[illegible]	0.0375
MC-HLL-UCT	0.1770	[illegible]	0.0331
CENO-HLL-UCT	0.1525	[illegible]	0.0239

Errors are measured against a high resolution (400^2 [illegible] reference run.



velocity ($\omega = 20$) in a static, magnetized ($B_x = 5/\sqrt{4\pi}$) background with uniform density and pressure ($\rho = p = 1$). The adiabatic index is $\gamma = 1.4$. To conform with [1], the final time is $t = 0.15$ and the same initial *taper* function is used (note that in LD the same problem was solved by the CENO-LLF-UCT code with $t = 0.18$ and no tapering).

The numerical settings are identical to those employed in the previous test, and also the errors displayed in Table 5 are calculated in the same way. In Fig. 9, the magnetic pressure $p_m = B^2/2$ is shown as iso-contours diagrams for all schemes, first the BS and UCT runs at 200^2 grid points, compared with the corresponding UCT reference run at 400^2, which is also used to calculate the errors for the lower accuracy tests. Here the same remarks made above still apply: the accuracy order is low due to discontinuities, the two second-order schemes behave very similarly (thus MC-HLL-UCT has to be preferred for its efficiency), while CENO-HLL-UCT gives the sharpest profiles, thus in this kind of model problems the accuracy in the reconstruction seems to be more important than the accuracy in resolving the Riemann structures. To conclude, we also notice that the BS schemes appear to behave correctly far from the rotating disk, where the waves are propagating outwards, while a lot of numerical noise is clearly formed inside the disk, where the numerical monopoles have time to accumulate, probably in a way similar to the inclined shock tube problems discussed in Section 4.2.

5. Conclusions

We have presented a method, first outlined in [12], to construct Godunov-type schemes for the MHD system, named *upwind constrained transport* (UCT). The main intent of our work is to assure that specific properties of the magnetic field, related to the basic divergence-free relation, enter as a *built-in* properties also in the approximated systems. To that purpose, by taking advantage of the CT discretization technique, we have presented procedures to define consistent derivative approximations, reconstruction steps and approximate Riemann solvers *all based* on the staggered (or face centered) magnetic field components b_i chosen as primary data. A main advantage in this approach is that no cleaning procedures or ad hoc modifications of the form of the MHD conservative system are required.

The main steps entering a UCT-based schemes are:

1. reconstruction procedures based on the smoothness properties of the divergence-free **B** vector field, as represented in finite volume CT discretization (Section 2.2);
2. the application of standard approximate Riemann solvers for the momentum and energy equations, with the prescription that only variables not related to the divergence-free condition are reconstructed and participate to the upwind differentiation. As a benefit, among others, exact cancellation of numerical monopoles is assured (Sections 2.3 and 2.4);
3. a specific formulation of the approximate Riemann solvers for the induction equation (Section 2.3);
4. a time integration procedure where no time-splitting is adopted.

To demonstrate the validity and flexibility of our UCT method, we have finally applied it to a flux-limited Roe-type scheme (the *positive* scheme by Liu and Lax [14]), which proves to be accurate, robust and well-suited for more demanding applications requiring AMR techniques. This novel scheme has been then tested numerically on a standard set of model problems and compared to central-type second and third-order schemes based on the two-speed HLL solver.

We conclude by remarking that our method, defined here for the classical MHD system in regular structured grids, applies unchanged to the equations of special and general relativistic MHD (see [13]), and many procedures here presented may have a natural generalization for grid refinements and unstructured grids.

Acknowledgements

The authors thank G. Toth and another reviewer for their competent help in improving the manuscript.

Appendix A. Point-value formalism and third-order procedures in the CENO-HLL-UCT scheme

As shown in Section 2.3, to get high $r \geqslant 3$ order schemes in a finite-volume setting, besides a proper r-th order reconstruction of $\overline{\mathbf{w}}$ variables, a final averaging of the $\overline{\mathbf{f}}_i$ flux is needed. In the 3-D case the latter procedure is not cost-effective and more efficient implementations have been proposed for ENO-type schemes by Shu and Osher [21]. In this approach, point values $\mathbf{u}$, instead of cell averages $\overline{\mathbf{u}}$, are advanced in time and flux values at the interfaces are directly reconstructed at the desired order by using flux point values $\mathbf{f}_i(\mathbf{u})$. However, when applied to the MHD system (see [3]), this implementation is not suitable to take properly into account the divergence-free condition, since the flux reconstruction yet involves the B_i cell-centered values. Therefore, as in our previous work (LD), we use more appropriate flux reconstruction techniques. Let then consider the MHD equations, now in the form that comes out by applying the inverse operation of volume averaging to Eqs. (33) and (34):

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{u}(t)] + \sum_i D_i(\hat{\mathbf{f}}_i) = 0, \tag{A.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[\hat{b}_i(t)] + \sum_{j,k} \epsilon_{i,j,k} D_j(\hat{E}_k) = 0. \tag{A.2}$$

Here $\mathbf{u} = \mathbf{u}_{j,k,m}$, now cell-centered point values, constitute the new set of fluid primary data, and correspondingly the conservative flux two-point differences $D_i(\hat{\mathbf{f}}_i)$ are high order approximations of point-value first derivatives. In the induction equation, on the other hand, $\hat{b}_i$ do *not* coincide with point-value representations of the staggered magnetic field components, those named b_i which are needed in flux computations, since their volume averages must now return the *surface*-averaged $\overline{b}_i$ values. Therefore, $\hat{b}_i$ components are actually defined, now as primary data, in the same way $\hat{\mathbf{f}}_i$ fluxes are defined in Eq. (A.1), thus two-point differences $D_i(\hat{b}_i)$ give high order representations of point-value parallel first derivatives. Similarly, $D_j(\hat{E}_k)$ will be here high order representations of the staggered electric fields first derivatives, and their volume average must give back line-averaged electric fields.

It is important to notice that if $\hat{b}_i$ components are evolved as primary data from Eq. (A.7), the point-value version of the divergence-free relation (21), which is written as

$$[\nabla \cdot \mathbf{b}]_{\mathrm{num}} = D_x(\hat{b}_x) + D_y(\hat{b}_y) + D_z(\hat{b}_z) = 0, \tag{A.3}$$

will be preserved in time (if valid at $t = 0$) to within machine accuracy, exactly as in the finite-volume framework of Section 2.1.

Our CENO-HLL-UCT scheme is then based on the following steps:

1. For given $\mathbf{w}_i$ cell-centered point values we reconstruct the face left–right point values using the appropriate CENO algorithm. The upwind fluxes $\mathbf{f}_i^{\mathrm{U}}(\mathbf{w}_i, b_i)$ and $E_k^{\mathrm{U}}(\mathbf{w})$ are then evaluated by using HLL as in Section 2.3.
2. For given values of these $\mathbf{f}_i$ point-value representation of fluxes, the $\hat{\mathbf{f}}_i$ data are defined, for $i = x, y, z$ and at the same points, as

$$\hat{\mathbf{f}}_i = \mathbf{f}_i - \tfrac{1}{24}\tilde{D}_i^{(2)}(\mathbf{f}_i) \tag{A.4}$$

where $\tilde{D}_i^{(2)}$ is a non-oscillatory approximation of the second derivative in the indicated coordinate. In this way, the difference $D_i(\hat{\mathbf{f}}_i)$ provides a high order $r \geqslant 3$ accurate approximation of the flux first derivative. Concerning the magnetic fluxes, given the point values E_k at edge centers, the corresponding $\hat{E}_k$ data must be defined as

$$\hat{E}_z = E_z - \tfrac{1}{24}[\tilde{D}_x^{(2)}(E_z) + \tilde{D}_y^{(2)}(E_z)], \tag{A.5}$$

and similarly for x and y components.

3. Eqs. (A.1) and (A.2) can now be evolved in time by applying the Runge–Kutta algorithm of the appropriate (third) accuracy order.
4. A final computational step is then needed to provide the relation between $\hat{b}_i$ primary data and b_i staggered point-value fields, those used in flux calculations. One has then to solve the *implicit* relations

$$[I - \tfrac{1}{24}\tilde{D}_i^{(2)}](b_i) = \hat{b}_i, \tag{A.6}$$

for $i = x, y, z$, typically by means of iterative methods (see LD). As discussed in details in Section 2.4, the crucial point concerning how to measure magnetic monopoles is that *the same* algorithm to compute first derivatives for fluxes should be actually applied to define the b_i first derivatives in the $[\nabla \cdot \mathbf{b}]_{\text{num}}$ sum of Eq. (A.3), which is exactly preserved *only* for $\hat{b}_i$ primary data. Thus, when derivatives of $[\nabla \cdot \mathbf{b}]_{\text{num}}$ are calculated starting from b_i fields, obtained implicitly from Eq. (A.6), in a similar way as shown in Eq. (A.4) for fluid fluxes (that is the definition actually relevant for numerical monopoles), the $[\nabla \cdot \mathbf{b}]_{\text{num}}$ variable will not be zero to within machine accuracy (see Section 4), though it can be made arbitrarily small depending on the precision of the inversion method employed.

In the actual implementation of the code, if one wants to keep track of magnetic field-lines a slightly different approach, perfectly equivalent to that outlined above, can be followed. The induction Eq. (A.2) may be substituted by

$$\frac{\mathrm{d}}{\mathrm{d}t}[A_k(t)] = E_k, \tag{A.7}$$

where A_k is the point-value representation of the vector potential k component and E_k of the corresponding electric field component (see the line-averaged counterpart, Eq. (15)). The double non-oscillatory derivation in Eq. (A.5) must be now applied to, say, A_z rather than to E_z:

$$\hat{A}_z = A_z - \tfrac{1}{24}[\tilde{D}_x^{(2)}(A_z) + \tilde{D}_y^{(2)}(A_z)], \tag{A.8}$$

and the divergence-free staggered $\hat{b}_i$ fields are now defined (as in Eq. (14), proper for the finite-volume framework) as

$$\hat{b}_i = \sum_{j,k} \epsilon_{i,j,k} D_j(\hat{A}_k), \tag{A.9}$$

to which the above remarks on $[\nabla \cdot \mathbf{b}]_{\text{num}}$ apply unchanged.

References

[1] G. Toth, J. Comput. Phys. 161 (2000) 605.
[2] J.U. Brackbill, D.C. Barnes, J. Comput. Phys. 35 (1980) 426.
[3] G.-S. Jiang, C.-C. Wu, J. Comput. Phys. 150 (1999) 561.
[4] K.G. Powell, ICASE Report no. 94-24, Langley, VA, 1994.
[5] K.G. Powell, P.L. Roe, T.J. Linde, T.J. Gombosi, D.L. De Zeeuw, J. Comput. Phys. 154 (1999) 284.

[6] A. Dedner, F. Kemm, D. Kroner, C.-D. Munz, T. Schitzer, M. Wesemberg, J. Comput. Phys. 175 (2002) 645.
[7] C.R. Evans, J.F. Hawley, Astrophys. J. 332 (1988) 659.
[8] W. Dai, P.R. Woodward, J. Comput. Phys. 142 (1998) 331.
[9] D. Ryu, A. Miniati, T.W. Jones, A. Frank, Astrophys. J. 509 (1998) 244.
[10] D.S. Balsara, D.S. Spicer, J. Comput. Phys. 149 (1999) 270.
[11] S.S. Komissarov, Mon. Not. Roy. Astron. Soc. 343 (1999) 366.
[12] P. Londrillo, L. Del Zanna, Astrophys. J. 530 (2000) 508.
[13] L. Del Zanna, N. Bucciantini, P. Londrillo, Astron. Astroph. 400 (2003) 397.
[14] X.-D. Liu, P.D. Lax, J. Comp. Fluid Dyn. 5 (1996) 133.
[15] P.D. Lax, X.-D. Liu, SIAM J. Sci. Comput. 19 (1998) 319.
[16] A. Harten, P.D. Lax, G.D. Levermore, W.J. Morokoff, SIAM J. Num. Anal. 35 (1998) 2117.
[17] P.D. Lax, B. Wendroff, Comm. Pure Appl. Math. 13 (1960) 217.
[18] E. Tadmor, Math. Comput. 49 (1987) 91.
[19] M.J. Berger, P. Colella, J. Comput. Phys. 82 (1989) 64.
[20] H. De Sterck, Proc. AIAA 15th Comp. Dyn. Conf. 2001-2623 (2001) 1.
[21] C.-W. Shu,, S. Osher, J. Comput. Phys. 83 (1989) 32.
[22] B. Van Leer, J. Comput. Phys. 32 (1979) 101.
[23] A. Harten, SIAM J. Sci. Comput. 21 (1984) 1.
[24] D.S. Balsara, J. Comput. Phys. 174 (2001) 614.
[25] G. Toth, P.L. Roe, J. Comput. Phys. 180 (2002) 736.
[26] C.-W. Shu, NASA ICASE Report, 97-65, 1997, p. 1.
[27] X.-D Liu, S. Osher, J. Comput. Phys. 142 (1998) 304.
[28] P. Cargo, G. Gallice, J. Comput. Phys. 136 (1997) 446.
[29] R.S. Myong, P.L. Roe, J. Comput. Phys. 147 (1998) 545.
[30] G. Toth, J. Comput. Phys. 182 (2002) 346.
[31] P.L. Roe, D.S. Balsara, SIAM J. Num. Anal. 56 (1996) 57.
[32] H. Nessyahu, E. Tadmor, J. Comput. Phys. 87 (1990) 408.
[33] A. Kurganov, S. Noelle, G. Petrova, SIAM J. Sci. Comput. 23 (2001) 707.
[34] A. Harten, P.D. Lax, B. Van Leer, SIAM Rev. 25 (1983) 35.
[35] L. Del Zanna, M. Velli, P. Londrillo, Astron. Astrophys. 307 (2001) 705.