

# Incomplete Riemann Solvers Based on Functional Approximations to the Absolute Value Function



José M. Gallardo, Manuel J. Castro, and Antonio Marquina

**Abstract** We give an overview on the work developed in recent years about certain classes of incomplete Riemann solvers for hyperbolic systems. These solvers are based on polynomial or rational approximations to  $|x|$ , and they do not require the knowledge of the complete eigenstructure of the system, but only a bound on the maximum wave speed. Our solvers can be readily applied to nonconservative hyperbolic systems, by following the theory of path-conservative schemes. In particular, this allows for an automatic treatment of source or coupling terms in systems of balance laws. The properties of our schemes have been tested with some challenging numerical experiments involving systems such as the Euler equations, ideal magnetohydrodynamics equations and multilayer shallow water equations.

## 1 Introduction

Since the early work of Godunov [19], Riemann solvers constitute a fundamental ingredient in the design of robust and accurate numerical methods for hyperbolic conservation laws. Usually, Riemann solvers can be classified as *complete* or *incomplete*, depending if all the characteristic waves in the solution of the exact Riemann problem are considered or not. Among the class of complete Riemann solvers, Roe's method [29] is one of the most widely used, as it usually provides the best resolution of the Riemann wave fan. However, when analytic expressions for the eigenstructure of the system are not available or they are difficult to compute, Roe's method may

---

J. M. Gallardo (✉) · M. J. Castro  
Department of Análisis Matemático, Estadística e I. O., y Matemática Aplicada,  
University of Málaga, Málaga, Spain  
e-mail: [jmgallardo@uma.es](mailto:jmgallardo@uma.es)

M. J. Castro  
e-mail: [mjcastro@uma.es](mailto:mjcastro@uma.es)

A. Marquina  
Department of Matemática Aplicada, University of Valencia, Valencia, Spain  
e-mail: [marquina@uv.es](mailto:marquina@uv.es)

be computationally expensive. Therefore, in certain situations it is preferable to consider incomplete Riemann solvers, for which only part of the spectral information is needed. In these cases, an important drawback may be the lack of resolution of internal waves in complex scenarios.

The numerical diffusion of a given numerical flux is determined by its *viscosity matrix*. In the case of Roe's method the viscosity matrix is  $|A|$ , the absolute value of the Roe matrix of the system, which may be difficult to compute as it requires the knowledge of the complete eigenstructure of  $A$ . A number of incomplete Riemann solvers based on appropriate approximations to  $|A|$  have been proposed in the literature. One of the earliest examples is given by the local Lax-Friedrichs (or Rusanov) method, in which  $|A|$  is approximated using only the largest eigenvalue of the system. Another very popular approach is the HLL method [20], where  $|A|$  is approximated by means of a linear polynomial evaluation  $P(A)$ , where  $P(x)$  interpolates  $|x|$  at the smallest and largest eigenvalues of  $A$ . On the other hand, the paper [14] contains the first construction of a simple approximation to  $|A|$  by means of a polynomial that approximates  $|x|$  without interpolating it exactly on the eigenvalues.

The latter approach is the basis of the general framework proposed in [8], where PVM (Polynomial Viscosity Matrix) methods were introduced. The viscosity matrix of a PVM method is built as a polynomial evaluation  $P(A)$  of the Roe matrix or the Jacobian of the flux at some other average value. It is worth noticing that a number of well-known methods in the literature can be viewed as particular cases of PVM schemes: Lax-Friedrichs, Rusanov, HLL, FORCE, Roe, etc. (see also [13, 23, 31]). An additional feature of PVM methods is that they can be defined in the general framework of nonconservative hyperbolic systems, which allows to construct natural extensions of the standard schemes cited before for solving problems in nonconservative form.

To ensure the stability of a PVM method, the graph of the basis polynomial  $P(x)$  must be over the graph of the absolute value function. On the other hand, as  $P(x)$  is closer to  $|x|$  in the uniform norm, the behavior of the associated PVM method will be closer to that of Roe's method. It follows then that it is possible to use accurate approximations to  $|x|$  for building PVM schemes resembling Roe's method, but with a much smaller computational cost. Following this idea, a PVM scheme based on Chebyshev polynomials, which provide optimal uniform approximations to  $|x|$ , was proposed in [10]. This idea was further extended in the same paper to the case of rational functions, which greatly improve the order of approximation to  $|x|$ . The resulting schemes were denoted as RVM (Rational Viscosity Matrix). In fact, RVM schemes based on Newman [24] approximations provided similar performances as Roe's method, but with a much smaller computational cost. As the only difference between PVM and RVM methods rely on the kind of basis function chosen, in this work we will use the term AVM (Approximate Viscosity Matrix) to refer to both of them. We remark that AVM methods constitute a class of general-purpose Riemann solvers, which are constructed using only an estimate of the spectral radius of the Roe matrix or the Jacobian of the system evaluated at an average state. As an additional advantage, unlike Roe's method, no entropy-fix is needed in the presence of sonic points, as long as the basis function does not cross the origin. Recently, a

fully two-dimensional version of AVM schemes has been proposed in [18], where multidimensional effects are taken into account through the approximate solution of two-dimensional Riemann problems.

The Osher-Solomon (OS) scheme [26] is a nonlinear and complete Riemann solver which enjoys a number of interesting features: it is robust, smooth, entropy-satisfying, and good behaved when computing slowly-moving shocks. As a drawback, its practical implementation is complex and computationally expensive, as it requires the computation of a path-dependent integral in phase space (see [30]). For this reason, its practical application has been restricted to certain systems, e.g., the compressible Euler equations. In [15, 16], the authors proposed a variant of the OS method combining linear paths and a Gauss-Legendre quadrature formula. This led to a simplified version of the OS scheme, denoted as DOT (Dumbser-Osher-Toro), which conserves its good properties and it is applicable to general hyperbolic systems. In particular, the viscosity matrix of a DOT solver is defined as a linear combination of the absolute value matrix of the Jacobian of the physical flux evaluated at certain quadrature points. As the practical computation of these matrices can be expensive, they could be approximated in an efficient way following the same technique behind AVM methods. This idea was explored in [11], leading to the class of AVM-DOT solvers. In particular, it was shown that Chebyshev-based AVM-DOT solvers admit a Jacobian-free implementation, in which only evaluations of the physical flux are needed. This kind of methods is particularly interesting when solving systems in which the Jacobian involves complex expressions: see [12].

Both classes of AVM and AVM-DOT solvers can be extended to the case of non-conservative hyperbolic systems, following the theory of path-conservative schemes [28]. In particular, this includes the important case of hyperbolic systems of conservation laws with source terms and nonconservative products. In the conservative case, the proposed schemes have been applied to a number of challenging problems in ideal gas dynamics, magnetohydrodynamics (MHD) and relativistic MHD (RMHD). Multilayer shallow water systems have been considered as a representative example in the nonconservative framework, as they include both source and nonconservative coupling terms [10–12]. In all the cases, the numerical tests indicate that the proposed schemes are robust, running stable and accurate with a satisfactory time step restriction.

## 2 Approximate Viscosity Matrix (AVM) Methods

In this section we give an overview of PVM [8] and related methods developed in recent years [10, 12]. For the sake of clarity, we first focus on the case of a system of conservation laws. Extensions to the nonconservative case will be treated later in Sect. 4.

Let us consider a system of conservation laws

$$\partial_t w + \partial_x F(w) = 0, \tag{1}$$

where  $w(x, t)$  takes values on an open convex set  $\mathcal{O} \subset \mathbb{R}^N$  and  $F: \mathcal{O} \rightarrow \mathbb{R}^N$  is a smooth flux function. The numerical solution of the Cauchy problem for (1) is computed by means of a finite volume method of the form

$$w_i^{n+1} = w_i^n - \frac{\Delta t}{\Delta x} (F_{i+1/2} - F_{i-1/2}), \quad (2)$$

where  $w_i^n$  is an approximation to the average of the exact solution at the cell  $I_i = [x_{i-1/2}, x_{i+1/2}]$  at time  $t^n = n\Delta t$  (the dependence on time will be dropped unless necessary). The numerical flux is assumed to be written as

$$F_{i+1/2} = \frac{F(w_i) + F(w_{i+1})}{2} - \frac{1}{2} Q_{i+1/2} (w_{i+1} - w_i), \quad (3)$$

where the *viscosity matrix*  $Q_{i+1/2}$  controls the numerical diffusion of the scheme.

We will assume that system (1) is *hyperbolic*, i.e., the Jacobian matrix of the flux at each state  $w \in \mathcal{O}$ ,

$$A(w) = \frac{\partial F}{\partial w}(w),$$

can be diagonalized as

$$A = P D P^{-1},$$

where  $D = \text{diag}(\lambda_1, \dots, \lambda_N)$ ,  $\lambda_i$  are the eigenvalues of  $A$ , and the columns of the matrix  $P$  are the associated right eigenvalues of  $A$ . As it is usual, we denote the positive and negative parts of  $A$ , respectively, as

$$A^+ = P D^+ P^{-1}, \quad A^- = P D^- P^{-1},$$

where  $D^\pm = \text{diag}(\lambda_1^\pm, \dots, \lambda_N^\pm)$ , with  $\lambda_i^+ = \max(\lambda_i, 0)$  and  $\lambda_i^- = \min(\lambda_i, 0)$ . It is clear that  $A = A^+ + A^-$ . The absolute value of  $A$  is then defined as

$$|A| = A^+ - A^-.$$

One of the most widely used Riemann solvers for (1) was proposed by Roe in [29]. It usually provides the best resolution of the Riemann wave fan, although for complex systems the method can be computationally expensive. This is due to the fact that Roe's method is a *complete* Riemann solver, in the sense that it uses all the eigenstructure of the system. Therefore, for complex systems, or systems for which the eigenstructure is not known, *incomplete* Riemann solvers may be preferred: they use few characteristic information and are thus easier to implement and computationally efficient.

It is important to note that Roe's method can be written in the form (3) with viscosity matrix  $Q_{i+1/2} = |A_{i+1/2}|$ , where  $A_{i+1/2}$  is a Roe matrix for the system. Several numerical methods have been developed by using approximations to  $|A_{i+1/2}|$  as viscosity matrices; see, e.g., [13, 14, 20, 30, 31] and the references therein.

*PVM Methods*

The original idea of PVM (Polynomial Viscosity Matrix) Riemann solvers [8] was based in approximating  $|A_{i+1/2}|$  using an appropriate polynomial evaluation of  $A_{i+1/2}$ . Assume that  $P(x)$  is a polynomial approximation of  $|x|$  in the interval  $[-1, 1]$ , and let  $\lambda_{i+1/2,\max}$  be the eigenvalue of  $A_{i+1/2}$  with maximum modulus (or an upper bound of it). The numerical flux of the PVM method associated to  $P(x)$  is given by (3) with viscosity matrix

$$Q_{i+1/2} = |\lambda_{i+1/2,\max}| P(|\lambda_{i+1/2,\max}|^{-1} A_{i+1/2}),$$

which provides an approximation to  $|A_{i+1/2}|$ , the viscosity matrix of Roe’s method. Moreover, note that the best  $P(x)$  approaches  $|x|$ , the closer the behavior of the associated PVM scheme will be to that of Roe’s method. It is worth noticing that no spectral decomposition of the matrix  $A_{i+1/2}$  is needed to build a PVM method, but only a bound on its spectral radius. This fact makes PVM methods greatly efficient and applicable to systems in which the eigenstructure is not known or difficult to obtain. In those cases in which a Roe matrix is not available or is difficult to compute,  $A_{i+1/2}$  can be defined as the Jacobian evaluated at some average state.

Several well-known schemes in the literature can be interpreted as PVM methods, for example:

- Lax-Friedrichs:  $P(x) = \frac{\Delta x}{\Delta t}$ .
- HLL:  $P(x) = \alpha_0 + \alpha_1 x$ , where  $P(S_L) = |S_L|$  and  $P(S_R) = |S_R|$ , being  $S_L$  and  $S_R$  approximations to the minimal and maximal speeds of propagation.
- Roe: In this case,  $P(x)$  is the Lagrange polinomial which interpolates the set of points  $(\lambda_{i+1/2}^{(j)}, |\lambda_{i+1/2}^{(j)}|)$ , where  $\lambda_{i+1/2}^{(j)}$  are the eigenvalues of the Roe matrix  $A_{i+1/2}$ .

Other examples include Rusanov, FORCE or Lax-Wendroff methods (see [8]). Another example of PVM method is the one proposed in [14], which constitutes one of the first attempts to construct a simple approximation of  $|A|$  by means of a polynomial that approximates  $|x|$  without interpolating it exactly on the eigenvalues.

The stability of a PVM scheme relies on the properties of the basis polynomial  $P(x)$ . In particular, the following *stability condition* should be verified:

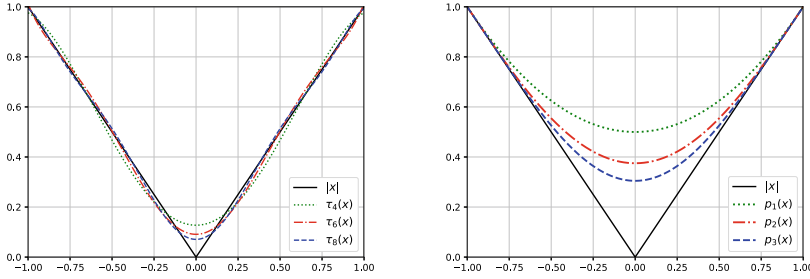
$$|x| \leq P(x) \leq 1, \quad \forall x \in [-1, 1]. \tag{4}$$

It was proven in [8] that condition (4) implies that the associated PVM scheme is linearly  $L^\infty$ -stable under a standard CFL restriction.

A well-known drawback of Roe’s method is the need of an entropy fix to handle sonic flow properly, in order to avoid entropy-violating solutions. In PVM-type schemes there is no need of entropy fix as long as  $P(0) \neq 0$ .

In [10] we proposed a new class of PVM schemes based on *Chebyshev polynomials*, which provide optimal uniform approximations to the absolute value function. The Chebyshev polynomials of even degree  $T_{2k}(x)$  are recursively defined as

$$T_0(x) = 1, \quad T_2(x) = 2x^2 - 1, \quad T_{2k}(x) = 2T_2(x)T_{2k-2}(x) - T_{2k-4}(x).$$



**Fig. 1** Left: Chebyshev approximations  $\tau_{2p}(x)$  for  $p = 2, 3, 4$ . Right: Internal polynomial approximations (6)

Then, for  $p \geq 1$  we consider the polynomial of degree  $2p$  given by (see Fig. 1, left)

$$\tau_{2p}(x) = \frac{2}{\pi} + \sum_{k=1}^p \frac{4}{\pi} \frac{(-1)^{k+1}}{(2k-1)(2k+1)} T_{2k}(x), \quad x \in [-1, 1], \quad (5)$$

which follows after truncation of the series expansion of  $|x|$  in terms of Chebyshev polynomials. It is a classical result [2] that the order of approximation of  $\tau_{2p}(x)$  to  $|x|$  is optimal in the  $L^\infty(-1, 1)$  norm. Moreover, the recursive definition of the polynomials  $T_{2k}(x)$  provides an explicit and efficient way to compute  $\tau_{2p}(x)$  (see the Appendix in [10]).

Notice that  $\tau_{2p}(x)$  does not verify the stability condition (4) strictly: see Fig. 1 (left), where  $\tau_{2p}(x)$  has been drawn for  $p = 2, 3, 4$ . This drawback was partially fixed in [10] in a rough manner, by substituting  $\tau_{2p}(x)$  by  $\tau_{2p}^\varepsilon = \tau_{2p}(x) + \varepsilon$ , where  $\varepsilon$  is chosen as the minimum value such that  $\tau_{2p}^\varepsilon(x)$  fulfills condition (4). However, this could cause incorrect approximations of the external waves.

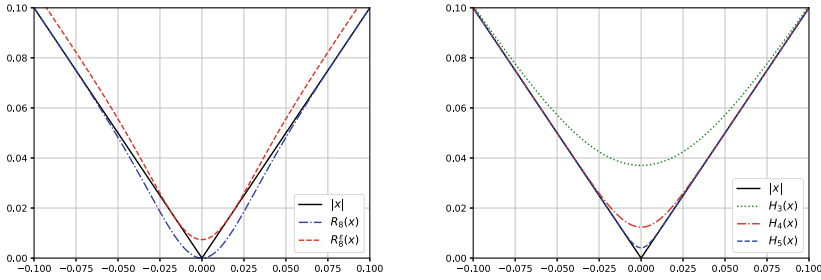
In [12] we proposed another family of polynomials which approximate  $|x|$  in a more elegant way, satisfying the stability condition (4) by construction. This class of *internal* polynomials are iteratively defined as follows (see Fig. 1 (right)):

$$p_0(x) \equiv 1, \quad p_{n+1}(x) = \frac{1}{2}(2p_n(x) - p_n(x)^2 + x^2), \quad n = 0, 1, 2, \dots \quad (6)$$

### RVM Methods

As it is well-known [24], rational functions provide much better approximations to  $|x|$  than polynomials. This fact was used in [10] to create new families of very precise incomplete Riemann solvers, denoted as RVM (Rational Viscosity Matrix), following the same idea of PVM methods but using a rational function as basis instead of a polynomial.

In particular, two families of RVM methods will be considered here. The first one corresponds to *Newman* approximations, which are constructed as follows. For a given  $r \geq 4$ , consider a set of distinct points in  $(0, 1]$ ,  $X = \{0 < x_1 < \dots < x_r \leq 1\}$ , and build the polynomial



**Fig. 2** Left: Comparison between  $R_8(x)$  and  $R_8^\epsilon(x)$  for  $x \in [-0.1, 0.1]$ , with  $\epsilon \approx 7.37e - 3$ . Right: Halley rational approximations  $H_r(x)$  in the interval  $[-0.1, 0.1]$ , for  $r = 3, 4, 5$

$$p(x) = \prod_{k=1}^r (x + x_k).$$

The *Newman rational function* associated to the set  $X$  is then defined by

$$R_r(x) = x \frac{p(x) - p(-x)}{p(x) + p(-x)}.$$

It is easy to see that  $R_r(x)$  interpolates  $|x|$  at the points  $\{-x_r, \dots, -x_1, 0, x_1, \dots, x_r\}$ . Also notice that for even  $r$  both the numerator and denominator of  $R_r(x)$  are of degree  $r$ . The uniform rate of approximation of  $R_r(x)$  to  $|x|$  depends on the choice of the set of nodes  $X$ . Several choices are possible (see [10]); here, we have considered Newman’s original definition, which is given by  $x_k = \xi^k$ , with  $\xi = \exp(-r^{-1/2})$ .

Notice that the stability condition (4) is not fulfilled in any case, so a *modified approximation* of the form  $R_r^\epsilon(x) = R_r(x) + \epsilon$  should be considered, as in the case of Chebyshev polynomials. A comparison between  $R_r(x)$  and  $R_r^\epsilon(x)$  can be seen in Fig. 2 (left). The differences between using  $R_r(x)$  or  $R_r^\epsilon(x)$  are particularly noticeable in the presence of sonic points: in this case,  $R_r^\epsilon(x)$  must be used to avoid entropy-violating solutions.

The second family of rational functions is based on iterative approximations. Note that the absolute value  $|\bar{x}|$  of a given point  $\bar{x} \in [-1, 1]$  can be viewed as the positive root of  $f(x) = x^2 - \bar{x}^2$ . It is then possible to approximate  $|\bar{x}|$  using a root-finding algorithm, such as Newton’s method, or the more precise choice given by the cubic Halley’s method:

$$x_{k+1} = x_k \frac{x_k^2 + 3\bar{x}^2}{3x_k^2 + \bar{x}^2}.$$

Taking  $x_0 = 1$  as initial guess, Halley’s method is well-defined and converges to  $\bar{x}$  (see [5]). The *Halley rational approximations* to  $|x|$  are thus recursively defined as

$$H_0(x) \equiv 1, \quad H_{r+1}(x) = H_r(x) \frac{H_r(x)^2 + 3x^2}{3H_r(x)^2 + x^2}. \tag{7}$$

Notice that the degrees of the numerator and denominator of  $H_r(x)$  are both equal to  $3^r - 1$ . It can be easily verified that  $H_r(x)$  satisfies the stability condition (4) without further modifications. Figure 2 (right) shows the functions  $H_r(x)$  for  $r = 3, 4, 5$ .

As it was commented before, another possibility is to use Newton's method instead of Halley's method. However, numerical experiments show that RVM-Halley methods provide much better resolution of internal waves than RVM-Newton schemes with a comparable computational cost.

#### *AVM Methods*

As it can be seen in the preceding paragraphs, the idea behind PVM and RVM methods is essentially the same, the difference depending only on using polynomials or rational functions to approximate the absolute value function. For this reason, we will encompass both kind of methods under the global name of AVM (Approximate Viscosity Matrix) methods.

Therefore, an AVM method is a finite volume method of the form (2), where the numerical flux is given by (3) with viscosity matrix

$$Q_{i+1/2} = f(A_{i+1/2}),$$

where  $f: \mathbb{R} \rightarrow \mathbb{R}$  is a given function and  $A_{i+1/2}$  is a Roe matrix or the Jacobian of the flux evaluated at some average state. The function  $f$  must verify some conditions:

- $f(x)$  is nonnegative and smooth.
- $f(A_{i+1/2})$  should be *easy* to evaluate; in particular, no spectral decomposition of  $A_{i+1/2}$  should be needed, but only a bound on its spectral radius.
- $L^\infty$ -linear stability, which is accomplished under the condition

$$|x| \leq f(x) \leq \text{CFL} \frac{\Delta x}{\Delta t}, \quad \forall x \in [\lambda_{i+1/2}^{(1)}, \lambda_{i+1/2}^{(N)}],$$

where  $\lambda_{i+1/2}^{(1)} \leq \dots \leq \lambda_{i+1/2}^{(N)}$  are the eigenvalues of  $A_{i+1/2}$ .

- The graph of  $f(x)$  should be as close as possible to the graph of  $|x|$ .

We end this section with some remarks regarding computational efficiency. The implementation of rational-based methods involves the computation of matrix powers and matrix inversions, while Chebyshev-based or internal-based methods can be expressed in such a way that only vector operations are involved. This makes these polynomial methods computationally cheaper than rational ones but, on the other hand, rational methods are far more precise. Thus, a compromise has to be accomplished between accuracy and computational cost. In general, it seems that for problems with solutions containing very complex patterns, rational-based methods perform better than polynomial ones in terms of the ratio between CPU times and computed errors. Otherwise, for mildly complex problems polynomial methods may be the preferred option.

An additional advantage of Chebyshev-based and internal-based methods is that they admit a Jacobian-free implementation (see Sect. 3), which is not possible for rational-based methods. This means that the numerical flux can be constructed using



only evaluations of the physical flux  $F$  at different states, thus avoiding the computation of Jacobian matrices. This point is particularly interesting for systems with complex physical fluxes (as, for example, the equations of RMHD), for which the calculation of the corresponding Jacobian may be a difficult or costly task.

### 3 Approximate DOT Solvers

The so-called *approximate DOT* (Dumbser-Osher-Toro) solvers, introduced in [11], combine the AVM technique with the universal Osher-type solvers proposed in [15]. These methods, that will be denoted in what follows as AVM-DOT, constitute simple and efficient approximations to the classical Osher-Solomon method [26], enjoying most of its interesting features and being applicable to general hyperbolic systems.

The numerical flux of the original Osher-Solomon method is given by

$$F_{i+1/2} = \frac{F(w_i) + F(w_{i+1})}{2} - \frac{1}{2} \int_0^1 |A(\Phi(s))| \Phi'(s) ds, \quad (8)$$

where  $A(w)$  represents the Jacobian of the physical flux  $F$  evaluated at the state  $w$ , and  $\Phi$  is a path in phase-space linking the states  $w_i$  and  $w_{i+1}$ . The path  $\Phi$  for a DOT solver [15] is taken as the segment linking  $w_i$  and  $w_{i+1}$ , and the resulting integral is approximated using a Gauss-Legendre quadrature formula. Thus, the resulting DOT flux adopts the form (3) with viscosity matrix

$$Q_{i+1/2} = \sum_{k=1}^q \omega_k |A(w_i + s_k(w_{i+1} - w_i))|,$$

where  $\omega_k$  and  $s_k$  are the weights and nodes of the quadrature formula. Now, the absolute value of the intermediate matrices could be approximated by using the technique of AVM methods. The resulting AVM-DOT solver associated to a function  $f(x)$  has then the following form:

$$Q_{i+1/2} = \sum_{k=1}^q \omega_k \tilde{F}_{i+1/2}^{(k)},$$

where

$$\tilde{F}_{i+1/2}^{(k)} = |\lambda_{i+1/2, \max}^{(k)}| f(|\lambda_{i+1/2, \max}^{(k)}|^{-1} A_{i+1/2}^{(k)}), \quad (9)$$

for  $k = 1, \dots, q$ . In the above expression,  $\lambda_{i+1/2, \max}^{(k)}$  denotes the eigenvalue of

$$A_{i+1/2}^{(k)} = A(w_i + s_k(w_{i+1} - w_i))$$

with maximum modulus.

Clearly, a kind of AVM methods can be obtained as a particular case of approximate AVM-DOT solvers, simply by taken  $q = 1$  and  $\omega_1 = 1$ .

#### *Jacobian-Free Implementation*

We end this section with some notes on the Jacobian-free implementation of AVM-DOT solvers. As it was already mentioned at the end of Sect. 3, this is only possible for polynomial-based methods. To clarify the process, we will focus on the case of an AVM-DOT solver based on internal polynomial approximations.

As it was indicated in [12], the explicit form of  $p_n(x)$  combined with Horner's method will be considered instead of the recursive form (6). On the other hand, notice that it will not be necessary to compute the viscosity matrix  $Q_{i+1/2}$  explicitly, but only the vector  $Q_{i+1/2}(w_{i+1} - w_i)$  appearing in the numerical flux (3).

To illustrate the procedure, consider the polynomial

$$p_2(x) = \alpha_0 x^4 + \alpha_1 x^2 + \alpha_2 = x^2(\alpha_0 x^2 + \alpha_1) + \alpha_2,$$

where the coefficients are given by  $\alpha_0 = -1/8$ ,  $\alpha_1 = 3/4$  and  $\alpha_2 = 3/8$ . Let  $A \equiv A(w)$  be the Jacobian matrix of  $F$  evaluated at an intermediate state  $w$ , and let  $v$  be an arbitrary state; for simplicity, assume that  $\lambda_{\max} = 1$ . Then the following approximation holds:

$$|A|v \approx p_2(A)v = (A^2(\alpha_0 A^2 + \alpha_1 I) + \alpha_2 I)v.$$

The above expression can be computed using Horner's algorithm:

- Define  $v_0 = v$  and compute  $\tilde{v}_0 = A^2 v_0$ .
- Calculate  $v_1 = \alpha_0 \tilde{v}_0 + \alpha_1 v_0$  and  $\tilde{v}_1 = A^2 v_1$ .
- Compute  $v_2 = \tilde{v}_1 + \alpha_2 v_0$ . Then,  $|A(w)|v \approx p_2(A)v = v_2$ .

The product  $A(w)v$  can be approximated using finite differences:

$$A(w)v \approx \frac{F(w + \varepsilon v) - F(w)}{\varepsilon},$$

which leads to

$$A(w)^2 v \approx \frac{F(w + F(w + \varepsilon v) - F(w)) - F(w)}{\varepsilon} \equiv \Phi_\varepsilon(w; v),$$

where, in practice, the value  $\varepsilon$  should be chosen small relative to the norm of  $w$ . Finally, the vector  $|A(w)|v$  can be approximated using the following steps, in which only vector operations and evaluations of the physical flux  $F$  are needed:

- Define  $v_0 = v$  and compute  $\tilde{v}_0 = \Phi_\varepsilon(w; v_0)$ .
- Calculate  $v_1 = \alpha_0 \tilde{v}_0 + \alpha_1 v_0$  and  $\tilde{v}_1 = \Phi_\varepsilon(w; v_1)$ .
- Compute  $v_2 = \tilde{v}_1 + \alpha_2 v_0$ . Then,  $|A(w)|v \approx v_2$ .

The detailed procedure to build a Jacobian-free AVM-DOT solver with  $p_n(x)$  as basis function is given next. The key point is to approximate the vectors

$$\tilde{P}_{i+1/2}^{(k)} \Delta w = |\lambda_{i+1/2, \max}^{(k)}| p_n(A^{(k)}) \Delta w, \quad k = 1, \dots, q,$$

where  $\Delta w = w_{i+1} - w_i$  and  $A^{(k)} = |\lambda_{i+1/2, \max}^{(k)}|^{-1} A(w_i^{(k)})$ , with  $w_i^{(k)} = w_i + s_k \Delta w$ . Assuming that the coefficients  $\alpha_i$  of the polynomial  $p_n(x)$  have already been computed, the polynomial  $p_n(x)$  can be written as

$$p_n(x) = \alpha_0 x^{2(n+1)} + \alpha_1 x^{2n} + \alpha_2 x^{2(n-1)} + \dots + \alpha_n x^2 + \alpha_{n+1}.$$

Then each term  $\tilde{P}_{i+1/2}^{(k)} \Delta w$  can be approximated using the following algorithm:

- Define  $v_0 = \Delta w$  and compute  $\tilde{v}_0 = |\lambda_{i+1/2, \max}^{(k)}|^{-2} \Phi_\varepsilon(w_i^{(k)}; v_0)$ .
- Calculate  $v_1 = \alpha_0 \tilde{v}_0 + \alpha_1 v_0$ .
- For  $j = 1, \dots, n$ , define  $\tilde{v}_j = |\lambda_{i+1/2, \max}^{(k)}|^{-2} \Phi_\varepsilon(w_i^{(k)}; v_j)$  and compute  $v_{j+1} = \tilde{v}_j + \alpha_{j+1} v_0$ .
- Finally,  $\tilde{P}_{i+1/2}^{(k)} \Delta w \approx |\lambda_{i+1/2, \max}^{(k)}| v_{n+1}$ .

## 4 The Nonconservative Case

The AVM and AVM-DOT solvers introduced in the previous sections can be extended in a natural way to the case of nonconservative hyperbolic systems. We will focus in this section in AVM-DOT solvers, as they are more general. However, all the results can be readily adapted to AVM solvers.

Consider a hyperbolic system in nonconservative form

$$\partial_t W + \mathcal{A}(W) \partial_x W = 0, \tag{10}$$

where the matrix  $\mathcal{A}(W)$  is strictly hyperbolic for each state  $W$  belonging to an open convex subset  $\Omega \subset \mathbb{R}^M$ . The definition of the nonconservative product  $\mathcal{A}(W) \partial_x W$  depends on the choice of a family of paths  $\Phi(s; W_L, W_R)$  joining arbitrary states  $W_L$  and  $W_R$  in the phase space  $\Omega$ : see [22, 28] for details.

The solutions of (10) can be numerically approximated by means of *path-conservative* finite volume schemes of the form [28]:

$$W_i^{n+1} = W_i^n - \frac{\Delta t}{\Delta x} (\mathcal{D}_{i-1/2}^+ + \mathcal{D}_{i+1/2}^-), \tag{11}$$

where  $\mathcal{D}_{i+1/2}^\pm = \mathcal{D}^\pm(W_i^n, W_{i+1}^n)$ . Here  $\mathcal{D}^-$  and  $\mathcal{D}^+$  are two continuous functions from  $\Omega \times \Omega$  to  $\Omega$  satisfying

$$\mathcal{D}^\pm(W, W) = W, \quad \forall W \in \Omega,$$

and

$$\mathcal{D}^-(W_0, W_1) + \mathcal{D}^+(W_0, W_1) = \int_0^1 \mathcal{A}(\Phi(s; W_0, W_1)) \frac{\partial \Phi}{\partial s}(s; W_0, W_1) ds \quad (12)$$

for every  $W_0, W_1 \in \Omega$ , with  $\Phi(0; W_0, W_1) = W_0$  and  $\Phi(1; W_0, W_1) = W_1$ . In particular, the *generalized Roe's scheme* [27] is defined by choosing

$$\mathcal{D}_{i+1/2}^\pm = \frac{1}{2} (\mathcal{A}_\Phi(W_i^n, W_{i+1}^n) \pm |\mathcal{A}_\Phi(W_i^n, W_{i+1}^n)|) (W_{i+1}^n - W_i^n),$$

where  $\mathcal{A}_\Phi$  is a Roe linearization associated to  $\mathcal{A}$  and  $\Phi$ . In this case, the term  $|\mathcal{A}_\Phi(W_i^n, W_{i+1}^n)|$  plays the role of a viscosity matrix. Using Roe's property, it is possible to write the above expression as

$$\begin{aligned} \mathcal{D}_{i+1/2}^\pm &= \frac{1}{2} \int_0^1 \mathcal{A}(\Phi(s; W_i^n, W_{i+1}^n)) \frac{\partial \Phi}{\partial s}(s; W_i^n, W_{i+1}^n) ds \\ &\quad \pm \frac{1}{2} |\mathcal{A}_\Phi(W_i^n, W_{i+1}^n)| (W_{i+1}^n - W_i^n). \end{aligned}$$

Then, it is natural to define the Osher-Solomon scheme for solving the nonconservative system (10) as (11) with

$$\begin{aligned} \mathcal{D}_{i+1/2}^\pm &= \frac{1}{2} \int_0^1 \mathcal{A}(\Phi(s; W_i^n, W_{i+1}^n)) \frac{\partial \Phi}{\partial s}(s; W_i^n, W_{i+1}^n) ds \\ &\quad \pm \frac{1}{2} \int_0^1 |\mathcal{A}(\Phi(s; W_i^n, W_{i+1}^n))| \frac{\partial \Phi}{\partial s}(s; W_i^n, W_{i+1}^n) ds, \quad (13) \end{aligned}$$

or, equivalently,

$$\begin{aligned} \mathcal{D}_{i+1/2}^\pm &= \frac{1}{2} \mathcal{A}_\Phi(W_i^n, W_{i+1}^n) (W_{i+1}^n - W_i^n) \\ &\quad \pm \frac{1}{2} \int_0^1 |\mathcal{A}(\Phi(s; W_i^n, W_{i+1}^n))| \frac{\partial \Phi}{\partial s}(s; W_i^n, W_{i+1}^n) ds. \quad (14) \end{aligned}$$

Notice that (13) is more general than (14), as the latter relies on the existence of a Roe linearization  $\mathcal{A}_\Phi$ . Therefore, (13) could be used in the cases in which a Roe linearization is not known or is difficult to compute.

A good choice of the family of paths  $\Phi$  may be difficult or very costly in practice, and usually relies on the physics of the problem (see [28]). A simple choice, commonly used in the literature, is given by the family of segments:  $\Phi(s; W_L, W_R) = W_L + s(W_R - W_L)$ ; we will consider this choice throughout the rest of the section. Then, denoting  $\mathcal{A}_{i+1/2} = \mathcal{A}_\Phi(W_i^n, W_{i+1}^n)$ , we have:

$$\mathcal{D}_{i+1/2}^{\pm} = \frac{1}{2} \left( \mathcal{A}_{i+1/2} \pm \int_0^1 |\mathcal{A}(W_i^n + s(W_{i+1}^n - W_i^n))| ds \right) (W_{i+1}^n - W_i^n),$$

where the integral  $\int_0^1 |\mathcal{A}(W_i^n + s(W_{i+1}^n - W_i^n))| ds$  can be interpreted as a viscosity term. Next, this integral can be approximated using a Gauss-Legendre quadrature formula, which leads to

$$\mathcal{D}_{i+1/2}^{\pm} = \frac{1}{2} \left( \mathcal{A}_{i+1/2} \pm \sum_{k=1}^q \omega_k |\mathcal{A}_{i+1/2}^{(k)}| \right) (W_{i+1}^n - W_i^n), \quad (15)$$

where  $\mathcal{A}_{i+1/2}^{(k)} = \mathcal{A}(W_i^n + s_k(W_{i+1}^n - W_i^n))$ . Therefore, (15) can be interpreted as a nonconservative extension of the DOT numerical flux. This approach has also been considered in [16].

Once formula (15) has been derived, AVM-DOT schemes for the nonconservative system (10) can be built in a natural way, considering

$$\mathcal{D}_{i+1/2}^{\pm} = \frac{1}{2} \left( \mathcal{A}_{i+1/2} \pm \sum_{k=1}^q \omega_k \tilde{P}_{i+1/2}^{(k)} \right) (W_{i+1}^n - W_i^n),$$

where  $\tilde{P}_{i+1/2}^{(k)}$  is defined in (9).

We will focus now in the particular case of a hyperbolic system of conservation laws with source terms and nonconservative products, that is,

$$\partial_t w + \partial_x F(w) + B(w) \partial_x w = G(w) \partial_x H, \quad (16)$$

where  $w(x, t) \in \mathcal{O}$  (being  $\mathcal{O} \subset \mathbb{R}^N$  open and convex),  $F: \mathcal{O} \rightarrow \mathbb{R}^N$  is a smooth flux function,  $B: \mathcal{O} \rightarrow \mathcal{M}_N(\mathbb{R})$  is a smooth matricial function, and  $G: \mathcal{O} \rightarrow \mathbb{R}^N$  and  $H: \mathbb{R} \rightarrow \mathbb{R}$  are given functions. System (16) can be written in the form (10) adding the trivial equation  $\partial_t H = 0$  and defining

$$W = \begin{pmatrix} w \\ H \end{pmatrix} \in \Omega = \mathcal{O} \times \mathbb{R} \subset \mathbb{R}^{N+1}, \quad \mathcal{A}(W) = \begin{pmatrix} A(w) & -G(w) \\ 0 & 0 \end{pmatrix},$$

where  $A(w) = \frac{\partial F}{\partial w}(w) + B(w)$ . In this case, a Roe linearization  $\mathcal{A}_{i+1/2}$  can be defined as [27]

$$\mathcal{A}_{i+1/2} = \begin{pmatrix} A_{i+1/2} & -G_{i+1/2} \\ 0 & 0 \end{pmatrix},$$

where  $A_{i+1/2} = \mathcal{L}_{i+1/2} + B_{i+1/2}$ ,  $\mathcal{L}_{i+1/2}$  being a Roe matrix for the flux  $F$  in the usual sense, that is,  $\mathcal{L}_{i+1/2}(w_{i+1}^n - w_i^n) = F(w_{i+1}^n) - F(w_i^n)$ ;  $B_{i+1/2}$  is a matrix verifying

$$B_{i+1/2}(w_{i+1}^n - w_i^n) = \left( \int_0^1 B(w_i^n + s(w_{i+1}^n - w_i^n)) ds \right) (w_{i+1}^n - w_i^n),$$

and  $G_{i+1/2}$  is a vector satisfying

$$G_{i+1/2}(H_{i+1} - H_i) = \left( \int_0^1 G(w_i^n + s(w_{i+1}^n - w_i^n)) ds \right) (H_{i+1} - H_i).$$

A simple calculation gives

$$|A(W)| = \begin{pmatrix} |A(w)| - |A(w)|A(w)^{-1}G(w) \\ 0 \quad 0 \end{pmatrix},$$

as long as  $A(w)$  is nonsingular. Substituting in (15), the DOT scheme for solving (16) can be written as

$$w_{i+1}^n = w_i^n - \frac{\Delta t}{\Delta x} (D_{i-1/2}^+ + D_{i+1/2}^-), \quad (17)$$

with

$$D_{i+1/2}^\pm = \frac{1}{2} \left( F(w_{i+1}^n) - F(w_i^n) + B_{i+1/2}(w_{i+1}^n - w_i^n) - G_{i+1/2}(H_{i+1} - H_i) \pm \sum_{k=1}^q \omega_k |A_{i+1/2}^{(k)}| (w_{i+1}^n - w_i^n - (A_{i+1/2}^{(k)})^{-1} G_{i+1/2}^{(k)}(H_{i+1} - H_i)) \right), \quad (18)$$

where  $A_{i+1/2}^{(k)} = A(w_i^n + s_k(w_{i+1}^n - w_i^n))$ , and similarly for  $G_{i+1/2}^{(k)}$ .

Finally, AVM-DOT schemes for (16) are obtained by substituting  $|A_{i+1/2}^{(k)}|$  by  $\tilde{P}_{i+1/2}^{(k)}$  in (18):

$$D_{i+1/2}^\pm = \frac{1}{2} \left( F(w_{i+1}^n) - F(w_i^n) + B_{i+1/2}(w_{i+1}^n - w_i^n) - G_{i+1/2}(H_{i+1} - H_i) \pm \sum_{k=1}^q \omega_k \tilde{P}_{i+1/2}^{(k)} (w_{i+1}^n - w_i^n - (A_{i+1/2}^{(k)})^{-1} G_{i+1/2}^{(k)}(H_{i+1} - H_i)) \right). \quad (19)$$

In the case of a system of conservation laws (that is,  $B = 0$  and  $G = 0$ ), the scheme (17) can be written in the form (2) by simply taking  $F_{i+1/2} = D_{i+1/2}^- + F(w_i^n)$  or, equivalently,  $F_{i+1/2} = -D_{i+1/2}^+ + F(w_{i+1}^n)$ .

## 5 Numerical Experiments

In this section we test the performances of AVM-DOT schemes with some challenging problems related to ideal MHD equations in the conservative case, and to multilayer shallow water equations in the nonconservative case.

Depending on the basis function  $f(x)$ , the AVM-DOT schemes will be denoted:

- DOT-Cheb-2p:  $f(x)$  is the Chebyshev polynomial  $\tau_{2p}(x)$ .
- DOT-Newman-r:  $f(x)$  is taken as the Newman rational function  $R_r(x)$ .
- DOT-Halley-r:  $f(x) = H_r(x)$ , the  $r$ -th Halley rational function.
- DOT: the original DOT method in [15], in which the eigendecomposition is computed numerically.

For higher order schemes, third-order PHM [21] reconstructions have been used, combined with a third-order TVD Runge-Kutta method for time stepping.

## 5.1 Applications to Magnetohydrodynamics

The MHD system of equations is given by [4]

$$\begin{cases} \partial_t \rho = -\nabla \cdot (\rho \mathbf{v}), \\ \partial_t (\rho \mathbf{v}) = -\nabla \cdot \left( \rho \mathbf{v} \mathbf{v}^T + \left( P + \frac{1}{2} \mathbf{B}^2 \right) \mathbf{I} - \mathbf{B} \mathbf{B}^T \right), \\ \partial_t \mathbf{B} = \nabla \times (\mathbf{v} \times \mathbf{B}), \\ \partial_t E = -\nabla \cdot \left( \left( \frac{\gamma}{\gamma - 1} P + \frac{1}{2} \rho q^2 \right) \mathbf{v} - (\mathbf{v} \times \mathbf{B}) \times \mathbf{B} \right), \end{cases} \quad (20)$$

where  $\rho$  represents the mass density,  $\mathbf{v} = (v_x, v_y, v_z)^t$  and  $\mathbf{B} = (B_x, B_y, B_z)^t$  are the velocity and magnetic fields, and  $E$  is the total energy. Denoting by  $q$  and  $B$  the magnitudes of the velocity and magnetic fields, the total energy is

$$E = \frac{1}{2} \rho q^2 + \frac{1}{2} B^2 + \rho \varepsilon,$$

where the specific internal energy  $\varepsilon$  and the hydrostatic pressure  $P$  are related through the equation of state  $P = (\gamma - 1) \rho \varepsilon$ , with  $\gamma$  the adiabatic constant. In addition to the equations, the magnetic field must satisfy the divergence-free condition

$$\nabla \cdot \mathbf{B} = 0. \quad (21)$$

In the numerical experiments, condition (21) has been imposed by means of the projection method in [3]. Notice that for  $\mathbf{B} = \mathbf{0}$ , system (20) reduces to the Euler equations for gases. The eigenstructure of system (20) is completely determined: see, e.g., [4].

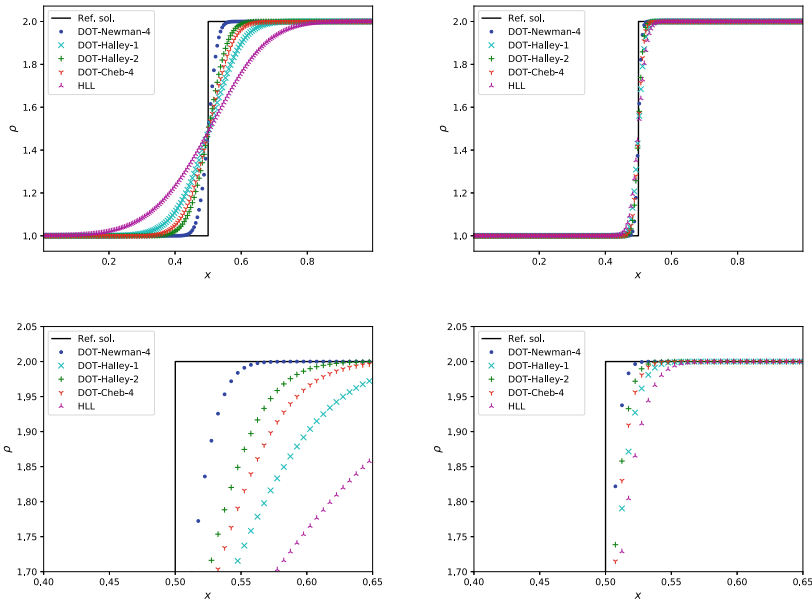
### 5.1.1 Stationary Contact Discontinuity

The purpose of this test, first proposed in [17], is to study the effect of the numerical diffusion in the approximation of a stationary contact discontinuity. This effect, known as *numerical heat conduction*, may cause incorrect heating across the discontinuity. The initial conditions for the Euler equations are given by

$$(\rho, v_x, P) = \begin{cases} (1, 0, 1) & \text{for } x \leq 0.5, \\ (2, 0, 1) & \text{for } x > 0.5, \end{cases}$$

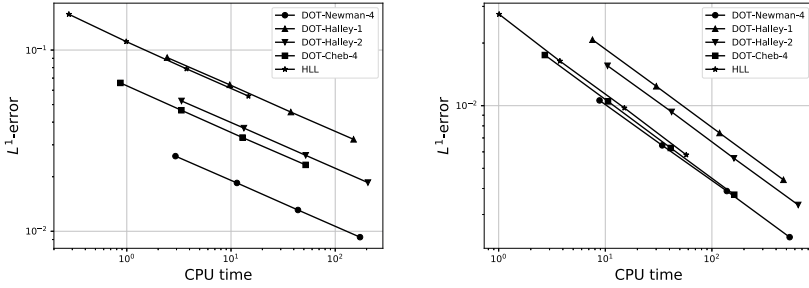
with  $\gamma = 1.4$ . The solution consists in a stationary contact wave located at  $x = 0.5$ . The problem has been solved in the domain  $[0, 1]$  with 200 cells and  $CFL = 0.5$  until a final time  $t = 4$ .

Figure 3 shows the approximations to the density component. Both in the first- and third-order solutions, DOT-Newman-4 gives the best approximation to the solution, followed by DOT-Halley-2, DOT-Cheb-4 and DOT-Halley-1. The HLL scheme gives a very diffusive resolution of the discontinuity. On the other hand, Fig. 4 shows the corresponding efficiency curves which represent, in logarithmic scale, the CPU times versus the  $L^1$  errors with respect to the exact solution for different meshes. In any case, DOT-Newman-4 is the most efficient solver. This test shows that the choice of



**Fig. 3** Test 5.1.1: Left: first order. Right: third order. The solutions obtained with the Roe and DOT schemes coincide with the reference (exact) solution. The lower row shows a zoom near the upper part of the discontinuity





**Fig. 4** Test 5.1.1: Efficiency curves CPU vs.  $L^1$ -error. Left: first order. Right: third order

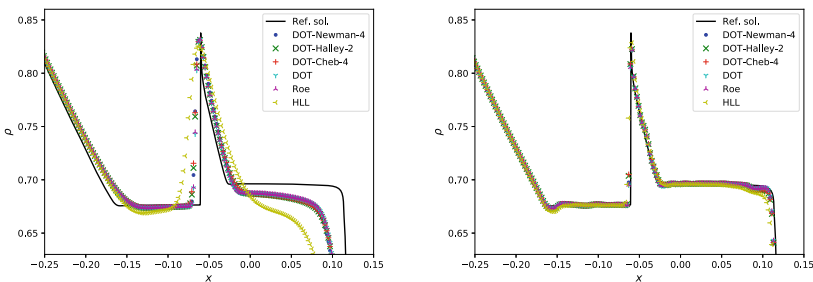
an appropriate first-order solver is important even when it is going to be used as a building block for higher-order schemes.

### 5.1.2 Brio-Wu Shock Tube Problem

This experiment was proposed in [4] to show the formation of a compound wave consisting of a shock followed by a rarefaction wave. The initial conditions are the following:

$$(\rho, v_x, v_y, v_z, B_x, B_y, B_z, P) = \begin{cases} (1, 0, 0, 0, 0.75, 1, 0, 1) & \text{for } x \leq 0, \\ (0.125, 0, 0, 0, 0.75, -1, 0, 0.1) & \text{for } x > 0, \end{cases}$$

with  $\gamma = 2$ . The problem has been solved until time  $t = 0.2$  in the interval  $[-1, 1]$  with a 1000 cell spatial discretization and  $CFL = 0.8$ . The results are shown in Fig. 5: in this case there are no appreciable differences between the solutions computed with Roe, DOT-Newman-4, DOT-Halley-2, DOT-Cheb-4 and DOT. On the other hand, the first order HLL method provides a worse resolution of the compound wave, which is however improved in third order.



**Fig. 5** Test 5.1.2: Zoom of the density compound wave. Left: first order. Right: third order

### 5.1.3 Orszag-Tang Vortex

The Orszag-Tang vortex [25] constitutes a model of transition to supersonic MHD turbulence in which, departing from a smooth state, complex interactions between shock waves are generated as the system evolves.

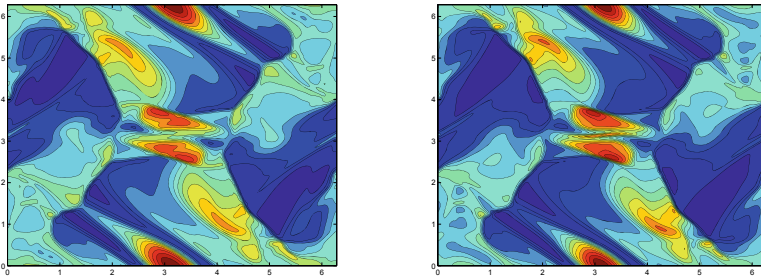
For  $(x, y) \in [0, 2\pi] \times [0, 2\pi]$ , the initial data are given by

$$\begin{aligned} \rho(x, y, 0) &= \gamma^2, & v_x(x, y, 0) &= -\sin(y), & v_y(x, y, 0) &= \sin(x), \\ B_x(x, y, 0) &= -\sin(y), & B_y(x, y, 0) &= \sin(2x), & P(x, y, 0) &= \gamma, \end{aligned}$$

with  $\gamma = 5/3$ . Periodic boundary conditions are imposed in the  $x$ - and  $y$ -directions. The computations have been done using a  $192 \times 192$  uniform mesh and  $\text{CFL}=0.8$ .

Figure 6 shows the results obtained with the third-order DOT-Cheb-4 scheme at time  $t = 3$ ; similar solutions are obtained with the third-order DOT-Newman-4, DOT-Halley-2, and DOT schemes. The results are in very good agreement with those found in the literature, thus showing that our schemes are robust and accurate enough to resolve the complicated structure of this vortex system. Finally, Table 1 shows the relative CPU times with respect to the first-order DOT scheme.

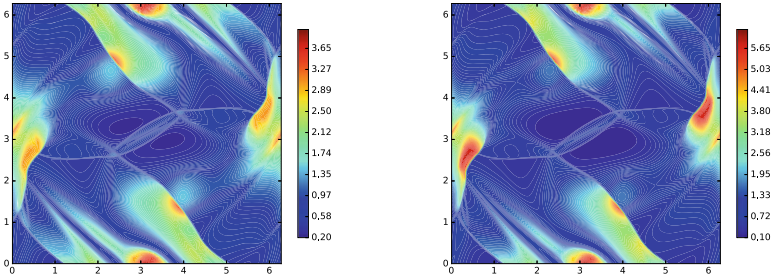
The relativistic version of the Orszag-Tang problem has also been considered. In this case, due to the very complex form of the Jacobians of the system, the Jacobian-free implementation introduced at the end of Sect. 3 has been found to be a very advantageous choice. Figure 7 shows the solution computed at time  $t = 4$  with a



**Fig. 6** Test 5.1.3: Density (left) and pressure (right) computed at time  $t = 3$  with the third-order DOT-Cheb-4 scheme

**Table 1** Test 5.1.3: Relative CPU times with respect to the first-order DOT solver

Method	CPU (first order)	CPU (third order)
DOT	1.00	5.82
DT-Cheb-4	0.16	1.04
DOT-Newman-4	0.38	2.32
DOT-Halley-2	0.50	2.79
HLL	0.05	0.36



**Fig. 7** Relativistic Orszag-Tang vortex: Left: density. Right: pressure

Jacobian-free second-order PVM-int-8 scheme, based on the internal polynomial approximation introduced in Sect. 2. At a qualitative level, our results are in good agreement with those found in, e.g., [32]. For a more detailed discussion about Jacobian-free AVM-DOT solvers applied to relativistic MHD, the reader is referred to [12].

#### 5.1.4 The Rotor Problem

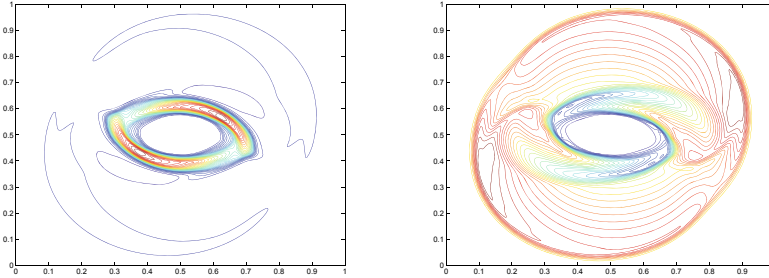
In this section we consider the rotor problem proposed in [1]. At the beginning, a dense disk rotates at the center of the domain, while the ambient fluid remains at rest. These two areas are connected with a taper function, which helps to reduce the initial discontinuity. As time evolves, the rotating dense fluid tends to be confined into an oblate shape, due to the action of the magnetic field.

The computational domain is  $[0, 1] \times [0, 1]$  with periodic boundary conditions. Defining  $r_0 = 0.1$ ,  $r_1 = 0.115$ ,  $f = (r_1 - r)/(r_1 - r_0)$  and  $r = [(x - 0.5)^2 + (y - 0.5)^2]^{1/2}$ , the initial conditions are given by

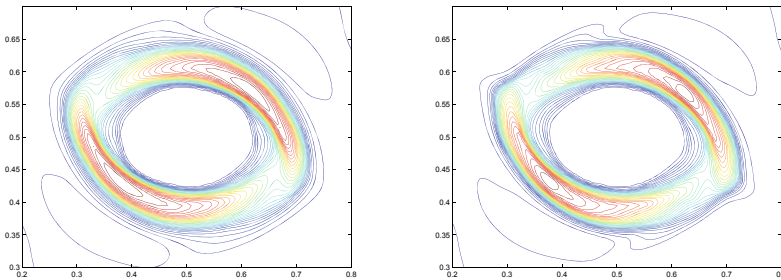
$$(\rho(x, y), v_x(x, y), v_y(x, y)) = \begin{cases} (10, -(y - 0.5)/r_0, (x - 0.5)/r_0) & \text{if } r < r_0, \\ (1 + 9f, -(y - 0.5)f/r, (x - 0.5)f/r) & \text{if } r_0 < r < r_1, \\ (1, 0, 0) & \text{if } r > r_1, \end{cases}$$

with  $B_x = 2.5/\sqrt{4\pi}$ ,  $B_y = 0$  and  $P = 0.5$ . We take  $\gamma = 5/3$ .

Figure 8 shows the solutions obtained with the third-order DOT-Cheb-4 scheme at time  $t = 0.295$  on a  $200 \times 200$  mesh with CFL= 0.8. The results are in good agreement with those presented in [1]. As in the previous test, DOT-Newman-4 and DOT-Halley-2 give similar results as DOT-Cheb-4. On the contrary, the DOT scheme fails for this problem around time  $t \approx 0.187$ . Finally, the third-order HLL and DOT-Cheb-4 methods are compared in Fig. 9. As it can be seen, DOT-Cheb-4 produces more precise results than HLL, which indicates that the choice of a precise first-order solver is important even when designing high-order schemes.



**Fig. 8** Test 5.1.4: Density  $\rho$  (left) and pressure  $P$  (right) computed at time  $t = 0.295$  with the third-order DOT-Cheb-4



**Fig. 9** Test 5.1.4: Comparison between the density solutions obtained with the third-order HLL (left) and DOT-Cheb-4 (right) schemes

## 5.2 Applications to the Two-Layer Shallow Water System

The two-layer shallow water equations constitute a representative model of the non-conservative systems considered in Sect. 4, as they include both source and non-conservative coupling terms (see [7]). The equations governing the one-dimensional flow of two superposed immiscible layers of shallow water fluids can be written in the form (16) by taking

$$w = \begin{pmatrix} h_1 \\ q_1 \\ h_2 \\ q_2 \end{pmatrix}, \quad F(w) = \begin{pmatrix} q_1 \\ \frac{q_1^2}{h_1} + \frac{g}{2}h_1^2 \\ q_2 \\ \frac{q_2^2}{h_2} + \frac{g}{2}h_2^2 \end{pmatrix}, \quad G(w) = \begin{pmatrix} 0 \\ gh_1 \\ 0 \\ gh_2 \end{pmatrix}, \quad B(w) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & gh_1 & 0 \\ 0 & 0 & 0 & 0 \\ rgh_2 & 0 & 0 & 0 \end{pmatrix},$$

where  $h_j$  are the fluid depths,  $q_j = h_j u_j$  represent the discharges,  $u_j$  are the velocities, and  $H(x)$  is the depth function measured from a fixed level of reference;  $g$  is the gravity constant and  $r = \rho_1/\rho_2$  is the ratio of densities. Notice that  $j = 1$  corresponds to the upper layer and  $j = 2$  to the lower one.

To build the AVM-DOT fluxes (19), we define

$$B_{i+1/2} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & gh_{1,i+1/2} & 0 \\ 0 & 0 & 0 & 0 \\ rgh_{2,i+1/2} & 0 & 0 & 0 \end{pmatrix}, \quad G_{i+1/2} = \begin{pmatrix} 0 \\ gh_{1,i+1/2} \\ 0 \\ gh_{2,i+1/2} \end{pmatrix},$$

where

$$h_{k,i+1/2} = \frac{h_{k,i} + h_{k,i+1}}{2}, \quad k = 1, 2.$$

Notice that the exact eigenstructure of the two-layer system is not explicitly known. However, a first order approximation of the maximum wave speed is given by

$$|\lambda_{i+1/2,\max}| \approx |\bar{u}_{i+1/2}| + c_{i+1/2},$$

where

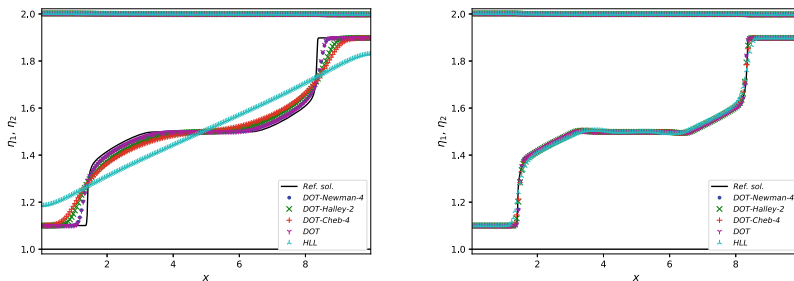
$$\bar{u}_{i+1/2} = \frac{q_{1,i+1/2} + q_{2,i+1/2}}{h_{1,i+1/2} + h_{2,i+1/2}}, \quad c_{i+1/2} = \sqrt{g(h_{1,i+1/2} + h_{2,i+1/2})}.$$

### 5.2.1 Internal Dam-Break

This test was proposed in [9] to simulate a dam-break in a two-layer system. The initial conditions are given by

$$h_1(x, 0) = \begin{cases} 0.9 & \text{if } x < 5, \\ 0.1 & \text{if } x \geq 5, \end{cases} \quad h_2(x, 0) = 1 - h_1(x, 0),$$

and  $q_1(x, 0) = q_2(x, 0) = 0$ , for  $x \in [0, 10]$ . The ratio of densities is taken as  $r = 0.99$ . The problem has been solved using a mesh with 200 grid points until time  $t = 20$ , with CFL number 0.9. Open boundary conditions have been imposed.



**Fig. 10** Test 5.2.1: Free surface and interface. Left: first order. Right: third order

Figure 10 shows the free surface and the interface ( $\eta_j = h_j - H$ ). The best results in first order are obtained with the DOT-Newman-4 and DOT schemes, followed by DOT-Halley-2 and DOT-Cheb-4, while HLL is not able to capture the interface correctly. On the other hand, in third order all the schemes perform equally well, being HLL the one that gives the less precise results.

### 5.2.2 Transcritical Flux with Shock

The initial condition for this test consists in an internal dam-break over a non-flat bottom, which eventually tends towards a stationary transcritical solution with a shock (see [10]). Specifically, the initial conditions are given by  $q_1(x, 0) = q_2(x, 0) = 0$ ,

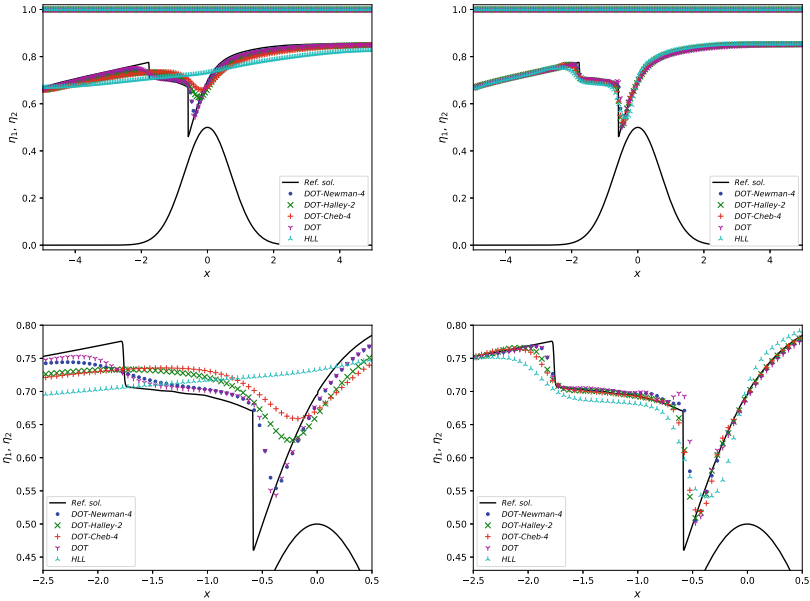
$$h_1(x, 0) = \begin{cases} 0.48 & \text{for } x < 0, \\ 0.02 & \text{for } x \geq 0, \end{cases} \quad h_2(x, 0) = H(x) - h_1(x, 0),$$

and the bottom topography is defined by

$$H(x) = 1 - \frac{1}{2}e^{-x^2}, \quad x \in [-5, 5].$$

Open wall boundary conditions have been imposed, and the ratio of densities has been chosen as  $r = 0.998$ .

The numerical solutions have been computed on a mesh with 200 grid points until final time  $t = 100$ , with CFL number 0.9. The results have been represented in Fig. 11. In first order, the DOT-Newman-4 and DOT schemes provide the best resolution of the interface, followed by DOT-Halley-2 and DOT-Cheb-4; on the other hand, HLL is unable to resolve the complex structure of the interface. The situation improves when going to third order, although again HLL presents a worse resolution near discontinuities. This can be better seen in the bottom row of Fig. 11, where a closer view of the shock has been plotted. Notice also that the DOT scheme presents more pronounced oscillations near the shock than DOT-Newman-4. Finally, the relative CPU times with respect to the first-order DOT scheme are shown in Table 2.



**Fig. 11** Test 5.2.2: Free surface, interface and bottom. Left: first order. Right: third order. The lower row shows a closer view of the shock at the interface

**Table 2** Test 5.2.2: Relative CPU times with respect to the first-order DOT solver

Method	CPU (first order)	CPU (third order)
DOT	1.00	2.98
DOT-Cheb-4	0.15	0.51
DOT-Newman-4	0.37	1.17
DOT-Halley-2	0.44	1.39
HLL	0.06	0.24

**Acknowledgements** This research has been partially supported by the Spanish Government Research project RTI2018-096064-B-C21. The numerical computations have been performed at the Laboratory of Numerical Methods of the University of Málaga.

## References

1. Balsara, D.S., Spicer, D.S.: A staggered mesh algorithm using high order Godunov fluxes to ensure solenoidal magnetic fields in magnetohydrodynamic simulations. *J. Comput. Phys.* **149**, 270–292 (1999)
2. Bernstein, S.: Sur la meilleure approximation de  $|x|$  par des polynômes de degrés donnés. *Acta Math.* **37**, 1–57 (1913)

3. Brackbill, J.U., Barnes, J.C.: The effect of nonzero  $\nabla \cdot B$  on the numerical solution of the magnetohydrodynamic equations. *J. Comput. Phys.* **35**, 426–430 (1980)
4. Brio, M., Wu, C.C.: An upwind differencing scheme for the equations of ideal magnetohydrodynamics. *J. Comput. Phys.* **75**, 400–422 (1988)
5. Candela, V., Marquina, A.: Recurrence relations for rational cubic methods I: the Halley method. *Computing* **44**, 169–184 (1990)
6. Cargo, P., Gallice, G.: Roe matrices for ideal MHD and systematic construction of Roe matrices for systems of conservation laws. *J. Comput. Phys.* **136**, 446–466 (1997)
7. Castro, M.J., Macías, J., Parés, C.: A  $Q$ -scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D shallow water system. *Math. Mod. Num. Anal.* **35**, 107–127 (2001)
8. Castro Díaz, M.J., Fernández-Nieto, E.D.: A class of computationally fast first order finite volume solvers: PVM methods. *SIAM J. Sci. Comput.* **34**, A2173–A2196 (2012)
9. Castro Díaz, M.J., Fernández-Nieto, E.D., Narbona-Reina, G., de la Asunción, M.: A second order PVM flux limiter method. Application to magnetohydrodynamics and shallow stratified flows. *J. Comput. Phys.* **262**, 172–193 (2014)
10. Castro, M.J., Gallardo, J.M., Marquina, A.: A class of incomplete Riemann solvers based on uniform rational approximations to the absolute value function. *J. Sci. Comput.* **60**, 363–389 (2014)
11. Castro, M.J., Gallardo, J.M., Marquina, A.: Approximate Osher-Solomon schemes for hyperbolic systems. *Appl. Math. Comput.* **272**, 347–368 (2016)
12. Castro, M.J., Gallardo, J.M., Marquina, A.: Jacobian-free approximate solvers for hyperbolic systems: Application to relativistic magnetohydrodynamics. *Comput. Phys. Commun.* **219**, 108–120 (2017)
13. Cordier, F., Degond, P., Kumbaro, A.: Phase appearance or disappearance in two-phase flows. *J. Sci. Comput.* **58**, 115–148 (2013)
14. Degond, P., Peyrard, P.F., Russo, G., Villedieu, P.: Polynomial upwind schemes for hyperbolic systems. *C. R. Acad. Sci. Paris Sér. I*(328), 479–483 (1999)
15. Dumbser, M., Toro, E.F.: On universal Osher-type schemes for general nonlinear hyperbolic conservation laws. *Commun. Comput. Phys.* **10**, 635–671 (2011)
16. Dumbser, M., Toro, E.F.: A simple extension of the Osher Riemann solver to non-conservative hyperbolic systems. *J. Sci. Comput.* **48**, 70–88 (2011)
17. Einfeldt, B., Munz, C.D., Roe, P.L., Sjögreen, B.: On Godunov-type methods near low densities. *J. Comput. Phys.* **92**, 273–295 (1991)
18. Gallardo, J.M., Schneider, K.A., Castro, M.J.: On a class of two-dimensional incomplete Riemann solvers. *J. Comput. Phys.* **386**, 541–567 (2019)
19. Godunov, S.K.: Finite difference methods for the computation of discontinuous solutions of the equations of fluid dynamics. *Math. USSR Sbornik* **47**, 271–306 (1959)
20. Harten, A., Lax, P.D., van Leer, B.: On upstream differencing and Godunov type schemes for hyperbolic conservation laws. *SIAM Rev.* **25**, 35–61 (1983)
21. Marquina, A.: Local piecewise hyperbolic reconstructions for nonlinear scalar conservation laws. *SIAM J. Sci. Comput.* **15**, 892–915 (1994)
22. dal Maso, G., LeFloch, P.G., Murat, F.: Definition and weak stability of nonconservative products. *J. Math. Pures Appl.* **74**, 483–548 (1995)
23. Ndjinga, M., Kumbaro, A., de Vuyst, F., Laurent-Gengoux, P.: Numerical simulation of hyperbolic two-phase flow models using a Roe-type solver. *Nucl. Eng. Des.* **238**, 2075–2083 (2008)
24. Newman, D.J.: Rational approximation to  $|x|$ . *Michigan Math. J.* **11**, 11–14 (1964)
25. Orszag, S.A., Tang, C.M.: Small scale structure of two-dimensional magnetohydrodynamic turbulence. *J. Fluid Mech.* **90**, 129–143 (1979)
26. Osher, S., Solomon, F.: Upwind difference schemes for hyperbolic conservation laws. *Math. Comput.* **38**, 339–374 (1982)
27. Parés, C., Castro, M.J.: On the well-balance property of Roe’s method for nonconservative hyperbolic systems. Applications to shallow-water systems. *M2AN* **38**, 821–852 (2004)



28. Parés, C.: Numerical methods for nonconservative hyperbolic systems: a theoretical framework. *SIAM J. Num. Anal.* **44**, 300–321 (2006)
29. Roe, P.L.: Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.* **43**, 357–372 (1981)
30. Toro, E.F.: *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, Heidelberg (2009)
31. Torrilhon, M.: Krylov-Riemann solver for large hyperbolic systems of conservation laws. *SIAM J. Sci. Comput.* **34**, A2072–A2091 (2012)
32. Zanotti, O., Dumbser, M.: A high order special relativistic hydrodynamic and magnetohydrodynamic code with space-time adaptive mesh refinement. *Comput. Phys. Commun.* **188**, 110–127 (2015)