



# A New ADER Method Inspired by the Active Flux Method

Christiane Helzel<sup>1</sup> · David Kerkmann<sup>1</sup> · Leonardo Scandurra<sup>1</sup>

Received: 16 April 2018 / Revised: 3 May 2019 / Accepted: 11 June 2019 / Published online: 20 June 2019  
© The Author(s) 2019

## Abstract

We study numerical methods that are inspired by the active flux method of Eymann and Roe and present several new results for one and two-dimensional hyperbolic problems. For one-dimensional linear problems we show that the unlimited active flux method can be interpreted as an ADER method. This interpretation motivates the construction of new third order accurate methods for nonlinear hyperbolic conservation laws. In the two-dimensional case, equivalent methods are only obtained for scalar linear problems. For two-dimensional linear systems the methods are no longer equivalent. For the two-dimensional acoustic equations we compare the accuracy of the two resulting approaches. While commonly used methods for hyperbolic problems are based on discontinuous reconstructions, the active flux method uses a continuous, piecewise quadratic reconstruction. For nonlinear problems we identify a situation in which the continuous reconstruction leads to an unstable approximation. We propose a limiting strategy which overcomes this problem. Our limited version of the active flux method uses the same local stencil as the original method.

**Keywords** Finite volume methods · Hyperbolic conservation laws · Active flux method · ADER method · High-order methods

**Mathematics Subject Classification** 65M08 · 65M25

## 1 Introduction

Recently, Eymann, Roe and coauthors [2–4,12,14] introduced a new numerical method for hyperbolic conservation laws, which they called the active flux method. For sufficiently smooth linear problems the method is third order accurate [4]. In [12], third order accurate

---

This work was supported by the DFG through HE 4858/4-1.

---

✉ Christiane Helzel  
Christiane.Helzel@hhu.de

David Kerkmann  
David.Kerkmann@hhu.de

Leonardo Scandurra  
scandurr@uni-duesseldorf.de

<sup>1</sup> Institute of Mathematics, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

approximations were also obtained for Burgers' equation. The method has also been used to discretise nonlinear systems of conservation laws, in particular the Euler equations of gas dynamics, by employing a splitting approach [4,14].

Eymann and Roe [4] proposed to reconsider all main components of today's CFD codes. Their goal is to construct methods which are truly multidimensional, more than second order accurate and which have a compact stencil.

The degrees of freedom of the active flux method, which consist of cell average values as well as interface values of the conserved quantities, are chosen in such a way that the method leads to third order accurate approximations using only the most local stencil. Note that interface values are point values of the conserved quantities located at the interfaces. Since the interface values are used by two or more grid cells, this choice of the degrees of freedom reduces memory requirements and computational expense. In the recent paper [14], Roe, Maeng and Fan showed that the active flux method produces accurate results at lower computational costs than other high-order accurate methods. In particular they studied the advection equation, the two-dimensional acoustic equations as well as the pressureless Euler equations.

The choice of the degrees of freedom suggests a continuous reconstruction of the conserved quantities. This continuous reconstruction distinguishes the active flux method from commonly used methods for hyperbolic problems such as Godunov type schemes [8], discontinuous Galerkin schemes [5] or WENO methods [15].

The time evolution of the interface values is performed before the time evolution of the cell average values. Interface value updates should respect the multidimensional physics and the accuracy requirements. A finite volume method is used to evolve average values of the conserved quantities. This guarantees that the method is conservative. The numerical fluxes use the point values at the interfaces at different time levels. The name of the scheme refers to the fact that the interface values are updated independently of the conserved quantities, see [4].

Here we consider methods which use the same degrees of freedom, the same piecewise quadratic reconstruction (whenever possible as explained later) as well as the same finite volume update for the cell average values as suggested in the original active flux method by Eymann and Roe. We show that for linear one-dimensional problems the active flux method can be interpreted as a new version of an ADER scheme, as introduced by Titarev and Toro [17]. We are not aware of an ADER method which uses these same degrees of freedom. In regions where we use the continuous reconstruction, our ADER version does not require the solution of Riemann problems for the conserved quantities but only for certain derivative values of the conserved quantities. The ADER interpretation of the active flux method allows us to construct new third order accurate methods for nonlinear systems of conservation laws, which share many if not all the attractive features of the active flux method. Furthermore, this interpretation provides a strategy for the construction of active flux methods for nonlinear problems which does not rely on a splitting approach. For the Euler equations of gas dynamics, a splitting method might be the best approach for the evolution of the interface values of the active flux method, see Roe [12]. Other authors have also explored splitting methods for the Euler equations, in particular in order to obtain better methods for low Mach number flows, see for example [6,18]. However, in order to construct finite volume methods for general hyperbolic problems, that use the same degrees of freedom as the active flux method, we believe that it is useful to have a general approach for the evolution of the interface values.

Numerical methods for hyperbolic problems use some form of limiting in order to avoid unphysical oscillations near shock waves or steep gradients of the solution profile. In [4], Eymann and Roe suggested a limiting based on neighbouring points of the solution in time as

well as space. In [4,12] they point out that the development of appropriate limiting strategies is still ongoing work. In Sect. 5 we present different possibilities for the use of limiters in the active flux method. For linear problems we suggest new continuous reconstructions which use the same stencil as the original active flux method but avoid the occurrence of new minima or maxima in the reconstructed function. These new continuous reconstructions can be used in the original active flux method after some changes, as discussed in Sect. 5.1. They can not be used in the ADER version of the method. For Burgers’ equation we identify a situation in which a continuous reconstruction without further modification leads to an unstable approximation. For this situation we propose a discontinuous reconstruction which cures this failure. This limiting approach can be used both in the original active flux method as well as in the ADER version.

It is interesting to note, that 1977 van Leer [20] mentioned a method (scheme V) for the one-dimensional advection equation, which agrees with the active flux method. Although he pointed out advantages in terms of accuracy and efficiency, he wrote: “its value for the ideal compressible flow equations at present seems doubtful”. A related but second order accurate finite volume method for advective transport problems, which uses the same compact stencil and distinguishes between the interface update and the update of the cell average values was much more recently proposed by Karabasov and Goloviznin [7]. Furthermore, Zeng [21] proposed related, up to fifth order accurate, hybrid finite difference–finite volume methods.

Our paper is organised as follows: In Sect. 2 we briefly introduce the original form of the active flux method for the advection equation as well as the acoustic equations. In Sect. 3 we show that in the linear case the active flux method can be interpreted as an ADER method. In Sect. 4 we use this interpretation to derive new methods for nonlinear hyperbolic problems and in particular for the Euler equations. In Sect. 5 we present different concepts to introduce limiters for our new one-dimensional schemes. In Sect. 6 we explore the extension of our one-dimensional results to the two-dimensional case. Throughout the paper we show numerous numerical results which illustrate our findings.

## 2 Brief Review of the Active Flux Method

We study one-dimensional hyperbolic problems of the form

$$\begin{aligned} q_t + f(q)_x &= 0 & -\infty < x < \infty, t > 0 \\ q(x, 0) &= q_0(x) & -\infty < x < \infty, \end{aligned} \tag{1}$$

where  $q : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}^m$  is a vector of conserved quantities and  $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is a vector valued flux function. We assume that the system (1) is hyperbolic, i.e. we assume that the flux Jacobian matrix is diagonalisable with real eigenvalues for all  $q$  of interest.

The degrees of freedom of the active flux method are cell averaged values of the conserved quantities as well as point values of the conserved quantities at the grid cell interfaces. We denote these quantities by  $Q_i^n$  and  $Q_{i+\frac{1}{2}}^n$ , i.e.

$$\begin{aligned} Q_i^n &\approx \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x, t_n) dx \\ Q_{i+\frac{1}{2}}^n &\approx q(x_{i+\frac{1}{2}}, t_n). \end{aligned}$$

Since each interface value belongs to two grid cells but is only counted once, the one-dimensional active flux method has two degrees of freedom per grid cell. This is the same

freedom employed by the discontinuous Galerkin method with linear reconstruction, which can be shown to be the semidiscrete limit of scheme V and active flux respectively, see Roe [11].

In order to describe one time step of the active flux method, we assume that we have third order accurate approximations  $Q_{i+\frac{1}{2}}^n$  and  $Q_i^n$  at the previous time level  $t_n$ . The cell averaged values of the conserved quantities at the new time level are obtained via a finite volume approach. On an equidistant grid the method is written in the form

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

where  $F_{i+\frac{1}{2}}$  denotes the numerical flux across the grid cell interface located at  $x_{i+\frac{1}{2}}$ . The numerical flux  $F_{i+\frac{1}{2}}$  is an approximation to the time averaged flux, i.e.

$$F_{i+\frac{1}{2}} \approx \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f(q(x_{i+\frac{1}{2}}, t)) dt. \tag{3}$$

In the active flux method the integral on the right hand side of (3) is approximated by using Simpson’s formula

$$\begin{aligned} \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} f(q(x_{i+\frac{1}{2}}, t)) dt \approx & \frac{1}{6} (f(q(x_{i+\frac{1}{2}}, t_n)) + 4f(q(x_{i+\frac{1}{2}}, t_{n+\frac{1}{2}})) \\ & + f(q(x_{i+\frac{1}{2}}, t_{n+1}))). \end{aligned}$$

We recall that the degree of accuracy of Simpson’s formula is  $p = 3$ , since  $p = 3$  is the largest integer such that the quadrature formula is exact for  $x^k$  for all  $k = 0, \dots, p$ . This means that we could obtain fourth order accurate approximations of the cell average values if the point values  $q(x_{i+\frac{1}{2}}, t_n + \tau)$  for  $\tau \in \{0, \Delta t/2, \Delta t\}$  would be fourth order accurate. Here the point values are obtained from a piecewise quadratic reconstruction and are thus only third order accurate. This limits the overall accuracy to third order.

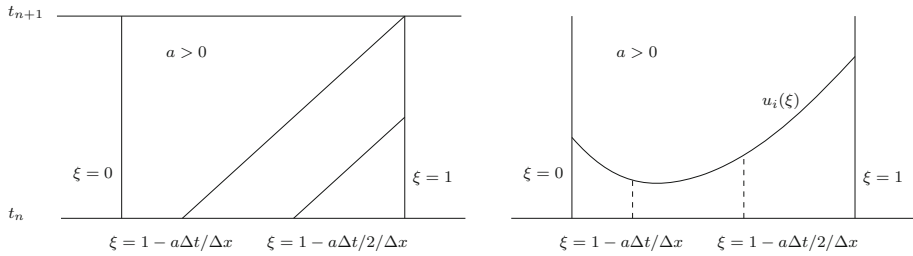
Note that the methods considered in this paper only differ in the way how the point values of the conserved quantities are computed. We denote the approximations of the conserved quantities at the grid cell interface by  $Q_{i+\frac{1}{2}}^n$ ,  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i+\frac{1}{2}}^{n+1}$ . The numerical flux of the finite volume method has the form

$$F_{i+\frac{1}{2}} = \frac{1}{6} \left( f(Q_{i+\frac{1}{2}}^n) + 4f(Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}) + f(Q_{i+\frac{1}{2}}^{n+1}) \right). \tag{4}$$

From the cell averaged values  $Q_i^n$  and the interface values  $Q_{i+\frac{1}{2}}^n$  of the conserved quantities, we can construct a continuous, piecewise quadratic function  $u(x)$  for  $-\infty < x < \infty$ . In grid cell  $i$ , the reconstructed function has the form

$$\begin{aligned} u_i(\xi) = & Q_{i-\frac{1}{2}}^n (2\xi - 1)(\xi - 1) + (6Q_i^n - Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n) \xi(1 - \xi) + Q_{i+\frac{1}{2}}^n \xi(2\xi - 1) \\ = & Q_{i-\frac{1}{2}}^n (3\xi^2 - 4\xi + 1) + Q_i^n (6\xi - 6\xi^2) + Q_{i+\frac{1}{2}}^n (3\xi^2 - 2\xi), \end{aligned} \tag{5}$$

with  $\xi := (x - x_{i-\frac{1}{2}}) / \Delta x$  and  $0 \leq \xi \leq 1$ . Note that it is the piecewise quadratic reconstruction which limits the method in its current form to third order accuracy. Eymann and Roe [3,4] used characteristic theory in order to compute  $q(x_{i+\frac{1}{2}}, t_{n+\frac{1}{2}})$  and  $q(x_{i+\frac{1}{2}}, t_{n+1})$ . For linear hyperbolic problems, they computed exact solutions at  $(x_{i+\frac{1}{2}}, t_{n+\frac{1}{2}})$  and  $(x_{i+\frac{1}{2}}, t_{n+1})$  of the



**Fig. 1** Illustration of the computation of interface values for the advection equation with positive advection speed. Left: space time diagram and characteristics; Right: reconstruction in grid cell  $i$  with reference to the values which contribute to the flux at the interface  $i + \frac{1}{2}$ ;

initial value problem with continuous, piecewise quadratic data  $u(x)$  at time  $t_n$ . We will now illustrate the method for two linear examples.

### 2.1 Advection

For the advection equation

$$q_t + aq_x = 0, \tag{6}$$

with  $q : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$  and  $a \in \mathbb{R}$ , the conserved quantity at the interface is approximated by

$$q(x_{i+\frac{1}{2}}, t_n + \tau) \approx \begin{cases} u_i(1 - a \tau/\Delta x) : & a > 0 \\ u_{i+1}(-a \tau/\Delta x) : & a < 0. \end{cases}$$

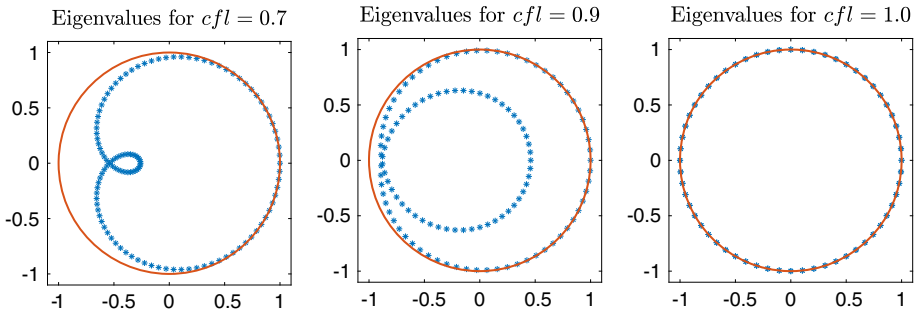
For the advection equation with positive advection speed  $a > 0$ , the averaged flux at  $x_{i+\frac{1}{2}}$  satisfies

$$\begin{aligned} \frac{a}{\Delta t} \int_{t_n}^{t_{n+1}} q(x_{i+\frac{1}{2}}, t) dt &= \frac{a}{\Delta t} \int_{t_n}^{t_{n+1}} q(x_{i+\frac{1}{2}} - a(t - t_n), t_n) dt \\ &= \frac{1}{\Delta t} \int_{x_{i+\frac{1}{2}} - a\Delta t}^{x_{i+\frac{1}{2}}} q(x, t_n) dx \\ &\approx \frac{\Delta x}{\Delta t} \int_{1 - \frac{a\Delta t}{\Delta x}}^1 u_i(\xi) d\xi \\ &= \frac{a}{6} \left( u_i \left( 1 - \frac{a\Delta t}{\Delta x} \right) + 4u_i \left( 1 - \frac{a\Delta t}{2\Delta x} \right) + u_i(1) \right) \end{aligned} \tag{7}$$

An approximation is only introduced when we replace  $q(x, t_n)$  by the piecewise quadratic function  $u$ . The spatial integration of the quadratic reconstruction using Simpson’s rule provides the exact integral value. Figure 1 illustrates the flux computation.

The active flux method is stable provided that the standard CFL condition  $a\Delta t/\Delta x \leq 1$  is satisfied. In [4], Eymann and Roe explained why this time step restriction is necessary. We are not aware of a formal proof that the method is in fact stable for all according time steps. For the advection equation we can write the method in the form

$$Q^{n+1} = A(\Delta t)Q^n,$$



**Fig. 2** Eigenvalues of the matrix  $A(\Delta t)$ , which describes the one-dimensional active flux method for advective transport with periodic boundary conditions

where  $Q^n$  is a vector which contains all the degrees of freedom at time  $t_n$ . If we impose periodic boundary conditions, then eigenvalues of the matrix  $A(\Delta t)$  can be studied in order to understand the stability. While we do not have an analytical formula for the eigenvalues, we plot them for a fixed grid in Fig. 2. All the eigenvalues are inside the unit circle, which experimentally confirms the linear stability of the method. We performed this test for many more CFL numbers in the range  $0 < CFL \leq 1$  and show the results of three examples.

### 2.2 Acoustics

In our second example we consider a linear hyperbolic system of the general form

$$q_t + Aq_x = 0.$$

As an example for such a linear system we consider the acoustic equations [8], i.e. we set

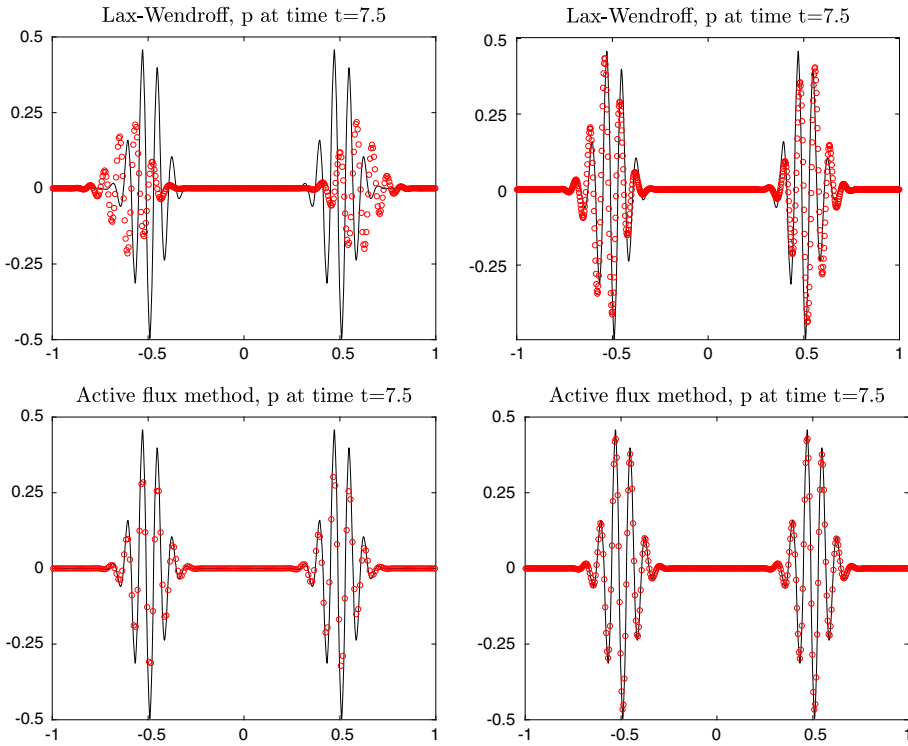
$$q = \begin{pmatrix} p \\ v \end{pmatrix}, \quad A = \begin{pmatrix} 0 & K_0 \\ 1/\rho_0 & 0 \end{pmatrix}. \tag{8}$$

Here  $p$  and  $v$  denote the pressure and the velocity component in the  $x$ -direction,  $\rho_0$  is a constant density,  $K_0$  is called the bulk modulus of compressibility,  $Z_0 = \rho_0 c_0$  is called the impedance and  $c_0 = \sqrt{K_0/\rho_0}$  is the speed of sound.

Using (5), the piecewise quadratic reconstruction for the acoustic equations is a vector valued function. We use an upper index to denote the two components of  $u$ . With the active flux method, the solution of the acoustic equations at  $(x_{i+\frac{1}{2}}, t_n + \tau)$  is approximated by

$$q \left( x_{i+\frac{1}{2}}, t_n + \tau \right) \approx \frac{1}{2} \left( u_i^1 \left( 1 - \frac{c_0 \tau}{\Delta x} \right) + u_{i+1}^1 \left( \frac{c_0 \tau}{\Delta x} \right) + Z_0 \left( u_i^2 \left( 1 - \frac{c_0 \tau}{\Delta x} \right) - u_{i+1}^2 \left( \frac{c_0 \tau}{\Delta x} \right) \right) \right. \\ \left. + \frac{1}{Z_0} \left( u_i^1 \left( 1 - \frac{c_0 \tau}{\Delta x} \right) - u_{i+1}^1 \left( \frac{c_0 \tau}{\Delta x} \right) \right) + u_i^2 \left( 1 - \frac{c_0 \tau}{\Delta x} \right) + u_{i+1}^2 \left( \frac{c_0 \tau}{\Delta x} \right) \right). \tag{9}$$

Formula (9) computes the exact solution of a modified problem, obtained by replacing the exact solution at time  $t_n$  by the piecewise quadratic reconstruction. For  $\tau = \Delta t/2$  and  $\tau = \Delta t$  we obtain the values of the conserved quantities at the half time step, i.e.  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$ , as well as at the full time step, i.e.  $Q_{i+\frac{1}{2}}^{n+1}$ . Those values are used in Simpson’s rule to obtain the numerical flux of the finite volume method. By applying this numerical flux in (2) we compute the cell averaged values at the new time level. This completes the description of one time step.



**Fig. 3** Comparison of the Lax–Wendroff method (first row) and the active flux method (second row) for acoustics. Results of  $p$  at time  $t = 7.5$  with 400 (left column) and 800 (right column) degrees of freedom are shown. The solid line is the exact solution, the red circles indicate the cell average values of  $p$  as computed by the two different methods (Color figure online)

In Fig. 3 we show numerical results of an accuracy study for the acoustic equations. For this computation the parameters in (8) are set to  $\rho_0 = K_0 = 1.4$ , which leads to  $c_0 = 1$ . The initial values are given by

$$\begin{aligned}
 p(x, 0) &= \exp(-100x^2) \sin(80(x - 0.5)) \\
 u(x, 0) &= 0.
 \end{aligned}$$

We impose the periodicity condition  $q(-1, t) = q(1, t)$ . The initial wave package is split into two waves, one moving to the left and one moving to the right with the speed of sound  $c_0 = 1$ . Due to the periodicity condition, the exact solution agrees with the initial condition at times  $t = 2n, n \in \mathbb{N}$ . We show numerical results at time  $t = 7.5$ , e.g. shortly before the waves overlap in the center of the domain for the fourth time. With the active flux method, the wave package is located at the correct position. This is not the case if we use the Lax–Wendroff method. Furthermore, we note that the numerical results of the active flux method with 200 and 400 grid cells are more accurate than the results of the Lax–Wendroff method with 400 and 800 grid cells, respectively. Recall that the active flux method with 200 grid cells has 400 degrees of freedom. This example clearly shows the advantages of third order accurate methods, in particular if one is interested in the resolution of small scale features.

The active flux method in the form presented so far relies on exact or third order accurate evolution formulas to compute  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i+\frac{1}{2}}^{n+1}$  from the piecewise quadratic function  $u$ . For one-dimensional linear hyperbolic problems, exact evolution formulas can easily be derived. For the nonlinear Euler equations, Eymann and Roe used a splitting approach, which decomposes the Euler flux in a part that describes acoustic wave propagation and a part which describes advective transport. This allowed them to use evolution formulas for the simpler subproblems. We will consider an alternative unsplit approach in Sect. 4.

### 3 Alternative Derivation of the Active Flux Method for Linear Problems

In this section we derive a numerical method which shares basic properties with the active flux method. In particular, the method uses the same degrees of freedom and the same piecewise quadratic spatial reconstruction.

While the original active flux method uses exact evolution formulas to approximate  $q(x_{i+\frac{1}{2}}, t_{n+\frac{1}{2}})$  and  $q(x_{i+\frac{1}{2}}, t_{n+1})$ , we now instead perform a Taylor series expansion with respect to the time variable

$$q(x_{i+\frac{1}{2}}, t_n + \tau) = q(x_{i+\frac{1}{2}}, t_n) + \tau q_t(x_{i+\frac{1}{2}}, t_n) + \frac{1}{2} \tau^2 q_{tt}(x_{i+\frac{1}{2}}, t_n) + \mathcal{O}(\tau^3) \tag{10}$$

and replace time derivatives on the right hand side of (10) by spatial derivatives.

#### 3.1 Advection

For the advection equation (6) the time derivatives are replaced by

$$q_t(x_{i+\frac{1}{2}}, t_n) = -aq_x(x_{i+\frac{1}{2}}, t_n), \quad q_{tt}(x_{i+\frac{1}{2}}, t_n) = a^2 q_{xx}(x_{i+\frac{1}{2}}, t_n). \tag{11}$$

Now we need approximations for  $q_x$  and  $q_{xx}$  at the interface, which we denote by  $Q_x(x_{i+\frac{1}{2}}, t_n)$  and  $Q_{xx}(x_{i+\frac{1}{2}}, t_n)$ . Using the reconstruction (5) we compute

$$\begin{aligned} Q_x(x_{i+\frac{1}{2}}, t_n) &= \begin{cases} \frac{1}{\Delta x} u'_i(1) : & a > 0 \\ \frac{1}{\Delta x} u'_{i+1}(0) : & a < 0, \end{cases} \\ Q_{xx}(x_{i+\frac{1}{2}}, t_n) &= \begin{cases} \frac{1}{\Delta x^2} u''_i(1) : & a > 0 \\ \frac{1}{\Delta x^2} u''_{i+1}(0) : & a < 0, \end{cases} \end{aligned} \tag{12}$$

with

$$\begin{aligned} u'_i(\xi) &= Q_{i-\frac{1}{2}}^n(4\xi - 3) + \left(6Q_i^n - Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n\right)(1 - 2\xi) + Q_{i+\frac{1}{2}}^n(4\xi - 1) \\ u''_i(\xi) &= 6Q_{i-\frac{1}{2}}^n - 12Q_i^n + 6Q_{i+\frac{1}{2}}^n. \end{aligned} \tag{13}$$

Thus, the conserved quantity at the grid cell interface can be approximated by

$$q(x_{i+\frac{1}{2}}, t_n + \tau) \approx Q_{i+\frac{1}{2}}^n - \tau a Q_x(x_{i+\frac{1}{2}}, t_n) + \frac{1}{2} \tau^2 a^2 Q_{xx}(x_{i+\frac{1}{2}}, t_n). \tag{14}$$



For  $\tau = \Delta t/2$  and  $\tau = \Delta t$ , we insert these approximations into Simpson’s formula and obtain the numerical flux

$$F_{i+\frac{1}{2}} = aQ_{i+\frac{1}{2}}^n - \frac{1}{2}\Delta ta^2 Q_x(x_{i+\frac{1}{2}}, t_n) + \frac{1}{6}\Delta t^2 a^3 Q_{xx}(x_{i+\frac{1}{2}}, t_n).$$

**Theorem 3.1** *The two formulations of the active flux method for the advection equation described in Sects. 2.1 and 3.1 are equivalent.*

**Proof** It is enough to show that both approaches lead to the same approximation of  $q(x_{i+\frac{1}{2}}, t_n + \tau)$ . We consider the case  $a > 0$  and assume that  $a\tau/\Delta x \leq 1$ . The method from Sect. 2.1 leads to

$$\begin{aligned} q(x_{i+\frac{1}{2}}, t_n + \tau) &\approx u_i(1 - a\tau/\Delta x) \\ &= Q_{i-\frac{1}{2}}^n \left( 3 \left( 1 - \frac{a\tau}{\Delta x} \right)^2 - 4 \left( 1 - \frac{a\Delta t}{\Delta x} \right) + 1 \right) \\ &\quad + Q_i^n \left( 6 \left( 1 - \frac{a\tau}{\Delta x} \right) - 6 \left( 1 - \frac{a\tau}{\Delta x} \right)^2 \right) \\ &\quad + Q_{i+\frac{1}{2}}^n \left( 3 \left( 1 - \frac{a\tau}{\Delta x} \right)^2 - 2 \left( 1 - \frac{a\tau}{\Delta x} \right) \right) \\ &= Q_{i-\frac{1}{2}}^n \left( -\frac{2a\tau}{\Delta x} + \frac{3a^2\tau^2}{\Delta x^2} \right) \\ &\quad + Q_i^n \left( \frac{6a\tau}{\Delta x} - \frac{6a^2\tau^2}{\Delta x^2} \right) + Q_{i+\frac{1}{2}}^n \left( 1 - \frac{4a\tau}{\Delta x} + \frac{3a^2\tau^2}{\Delta x^2} \right) \\ &= Q_{i+\frac{1}{2}}^n - a\frac{\tau}{\Delta x} \left( 2Q_{i-\frac{1}{2}}^n - 6Q_i^n + 4Q_{i+\frac{1}{2}}^n \right) \\ &\quad + \frac{1}{2}a^2\frac{\tau^2}{\Delta x^2} \left( 6Q_{i-\frac{1}{2}}^n - 12Q_i^n + 6Q_{i+\frac{1}{2}}^n \right) \end{aligned}$$

Using (14) and (12), the method from the current section leads to

$$\begin{aligned} q(x_{i+\frac{1}{2}}, t_n + \tau) &\approx Q_{i+\frac{1}{2}}^n - a\frac{\tau}{\Delta x}u'_i(1) + \frac{1}{2}a^2\frac{\tau^2}{\Delta x^2}u''_i(1) \\ &= Q_{i+\frac{1}{2}}^n - a\frac{\tau}{\Delta x} \left( Q_{i-\frac{1}{2}}^n - (6Q_i^n - Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n) + 3Q_{i+\frac{1}{2}}^n \right) \\ &\quad + \frac{1}{2}a^2\frac{\tau^2}{\Delta x^2} \left( 6Q_{i-\frac{1}{2}}^n - 12Q_i^n + 6Q_{i+\frac{1}{2}}^n \right) \end{aligned}$$

Thus, both approaches are equivalent. The case  $a < 0$  can be considered analogously.

□

### 3.2 Acoustics

For linear hyperbolic systems of the general form

$$q_t + Aq_x = 0, \tag{15}$$

we use the relations

$$q_t(x_{i+\frac{1}{2}}, t_n) = -Aq_x(x_{i+\frac{1}{2}}, t_n), \quad q_{tt}(x_{i+\frac{1}{2}}, t_n) = A^2q_{xx}(x_{i+\frac{1}{2}}, t_n)$$

to replace the time derivatives in (10) by terms involving spatial derivatives. The vector valued conserved quantities at the grid cell interface are now approximated using the relation

$$q \left( x_{i+\frac{1}{2}}, t_n + \tau \right) \approx Q_{i+\frac{1}{2}}^n - \tau A Q_x \left( x_{i+\frac{1}{2}}, t_n \right) + \frac{1}{2} \tau^2 A^2 Q_{xx} \left( x_{i+\frac{1}{2}}, t_n \right). \tag{16}$$

For linear systems, the interface values of the derivatives are obtained by solving Riemann problems of the form

$$\begin{aligned} (Q_x)_t + A(Q_x)_x &= 0 \\ Q_x(x, t_n) &= \begin{cases} \frac{1}{\Delta x} u'_i(1) : & x < x_{i+\frac{1}{2}} \\ \frac{1}{\Delta x} u'_{i+1}(0) : & x > x_{i+\frac{1}{2}} \end{cases} \end{aligned} \tag{17}$$

as well as

$$\begin{aligned} (Q_{xx})_t + A(Q_{xx})_x &= 0 \\ Q_{xx}(x, t_n) &= \begin{cases} \frac{1}{\Delta x^2} u''_i(1) : & x < x_{i+\frac{1}{2}} \\ \frac{1}{\Delta x^2} u''_{i+1}(0) : & x > x_{i+\frac{1}{2}} \end{cases} \end{aligned} \tag{18}$$

For the acoustic equations (8), the solution of the Riemann problem with data of the general form

$$q(x_{i+\frac{1}{2}}, t_n) = \begin{cases} (p_L, v_L) : & x_{i+\frac{1}{2}} < x \\ (p_R, v_R) : & x_{i+\frac{1}{2}} > x \end{cases}$$

consists of three constants states, which are separated by two contact discontinuities moving with speed  $-c_0$  and  $c_0$ . The interface value is given by

$$q(x_{i+\frac{1}{2}}, t_n + \tau) = \frac{1}{2} \begin{pmatrix} (p_L + p_R) - Z_0(v_R - v_L) \\ (v_L + v_R) - (p_R - p_L)/Z_0 \end{pmatrix}.$$

This formula can also be used to compute the interface values for the Riemann problems (17) and (18).

This alternative derivation of the active flux method shares basic properties with both the original active flux method of Eymann and Roe as well as the ADER method of Titarev and Toro [17]. It uses the same degrees of freedom and the same piecewise quadratic reconstruction as well as Simpson’s rule for the flux computation as proposed by Eymann and Roe. However, the quantities  $q(x_{i+\frac{1}{2}}, t_{n+\frac{1}{2}})$  and  $q(x_{i+\frac{1}{2}}, t_{n+1})$  are computed in analogy to the ADER approach. It is straightforward to show that both formulations lead to the same numerical results.

**Theorem 3.2** *For one-dimensional linear hyperbolic systems the ADER version of the active flux method is equivalent to the active flux method that uses the exact evolution formula.*

**Proof** By transforming the system to characteristic variables and using the results for advection, it is straight forward to show that both approaches lead to the same values  $q(x_{i+\frac{1}{2}}, t_n + \tau)$ . □

In particular the two formulations of the active flux method for the acoustic equations described in Sects. 2.2 and 3.2 are equivalent. In the next section we will see that the ADER interpretation of the active flux method can be extended to nonlinear problems. The resulting method is no longer equivalent with existing active flux methods.

### 4 New Active Flux Methods for Nonlinear Problems

In this section we present new active flux methods for nonlinear problems, which use an ADER approach in order to compute the conserved quantities at the nodes of Simpson’s rule.

#### 4.1 Burgers’ Equation

We start our consideration of nonlinear problems with the Burgers equation, i.e. we consider

$$q_t + \left(\frac{1}{2}q^2\right)_x = 0. \tag{19}$$

Our method is again based on the Taylor series expansion of the exact solution (10). For sufficiently smooth solutions, we can replace time derivatives of the conserved quantity by terms involving spatial derivatives, using the relations

$$\begin{aligned} q_t(x_{i+\frac{1}{2}}, t_n) &= -q(x_{i+\frac{1}{2}}, t_n) q_x(x_{i+\frac{1}{2}}, t_n), \\ q_{tt}(x_{i+\frac{1}{2}}, t_n) &= 2q(x_{i+\frac{1}{2}}, t_n) (q_x(x_{i+\frac{1}{2}}, t_n))^2 + (q(x_{i+\frac{1}{2}}, t_n))^2 q_{xx}(x_{i+\frac{1}{2}}, t_n). \end{aligned}$$

The conserved quantities at the grid cell interfaces are approximated in the form

$$\begin{aligned} q(x_{i+\frac{1}{2}}, t_n + \tau) &\approx Q_{i+\frac{1}{2}}^n - \tau Q_{i+\frac{1}{2}}^n Q_x(x_{i+\frac{1}{2}}, t_n) \\ &\quad + \frac{1}{2} \tau^2 \left( 2Q_{i+\frac{1}{2}}^n (Q_x(x_{i+\frac{1}{2}}, t_n))^2 + (Q_{i+\frac{1}{2}}^n)^2 Q_{xx}(x_{i+\frac{1}{2}}, t_n) \right). \end{aligned} \tag{20}$$

The derivatives of  $q$  at the grid cell interfaces are again obtained by solving generalised Riemann problems. We get

$$\begin{aligned} Q_x(x_{i+\frac{1}{2}}, t_n) &= \begin{cases} \frac{1}{\Delta x} u'_i(1) : & Q_{i+\frac{1}{2}}^n > 0 \\ \frac{1}{\Delta x} u'_{i+1}(0) : & Q_{i+\frac{1}{2}}^n < 0, \end{cases} \\ Q_{xx}(x_{i+\frac{1}{2}}, t_n) &= \begin{cases} \frac{1}{\Delta x^2} u''_i(1) : & Q_{i+\frac{1}{2}}^n > 0 \\ \frac{1}{\Delta x^2} u''_{i+1}(0) : & Q_{i+\frac{1}{2}}^n < 0, \end{cases} \end{aligned} \tag{21}$$

with  $u'_i(\xi)$  and  $u''_i(\xi)$  as defined in (13).

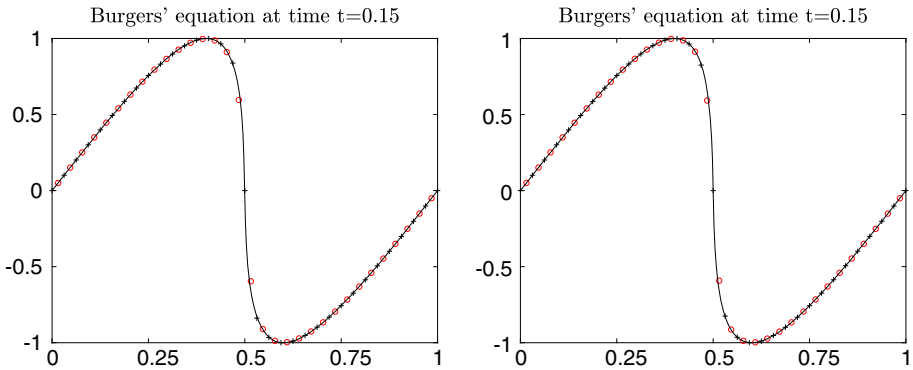
Evaluations of the right hand side of (20) for  $\tau = \Delta t/2$  and  $\tau = \Delta t$  provide  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i+\frac{1}{2}}^{n+1}$ . Now the numerical flux of the resulting finite volume method has the form

$$F_{i+\frac{1}{2}} = \frac{1}{6} \left( f(Q_{i+\frac{1}{2}}^n) + 4f(Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}) + f(Q_{i+\frac{1}{2}}^{n+1}) \right), \tag{22}$$

with  $f(q) = \frac{1}{2}q^2$ .

Most recently, Roe [12] discussed active flux methods for the one-dimensional Burgers equation. He computes a wave speed at each grid cell interface, using the relation

$$s_{i+\frac{1}{2}}(\tau) = \begin{cases} Q_{i+\frac{1}{2}}^n / (1 + \tau u'_i(1)) : & Q_{i+\frac{1}{2}}^n > 0 \\ Q_{i+\frac{1}{2}}^n / (1 + \tau u'_{i+1}(0)) : & Q_{i+\frac{1}{2}}^n < 0. \end{cases} \tag{23}$$



**Fig. 4** Approximation of Burgers' equation with initial data (25) at time  $t = 0.15$  using the ADER version (left) as well as Roe' version of the active flux method (right). The solid line is a highly resolved reference solution. The circles indicate cell averaged values and the crosses indicate interface values of the conserved quantities

These wave speeds are used to calculate the conserved quantity at the grid cell interface:

$$q(x_{i+\frac{1}{2}}, t_n + \tau) \approx \begin{cases} u_i(1 - s_{i+\frac{1}{2}}(\tau) \tau / \Delta x) : & s_{i+\frac{1}{2}}(\tau) > 0 \\ u_{i+1}(-s_{i+\frac{1}{2}}(\tau) \tau / \Delta x) : & s_{i+\frac{1}{2}}(\tau) < 0 \end{cases} \quad (24)$$

He compares the approach with a method which instead sets the local wave speed equal to  $Q^n_{i+\frac{1}{2}}$ . Here we only consider the method with wave speed correction according to (23), since the simpler method without wave speed correction is only second order accurate.

We now present numerical results for Burgers' equation with initial data of the form

$$q(x, 0) = \sin(2\pi x). \quad (25)$$

In Fig. 4, we show numerical results at time  $t = 0.15$  for the ADER version of the active flux as well as Roe's version. For this computation we used a grid with 32 grid cells on the interval  $[0, 1]$ . The solid line is a highly resolved reference solution obtained by using  $2^{15}$  grid cells. The time steps were chosen according to the CFL condition  $CFL \leq 0.9$ , with  $CFL := \max_i |Q^n_{i+\frac{1}{2}}| \Delta t / \Delta x$ .

In Table 1, we show results of a convergence study using Roe's method with wave speed correction as well as our ADER version of the active flux method. The active flux method with wave-speed correction leads to more accurate results on coarse grids. However, the ADER approach also performs very well and on well resolved grids the accuracy is almost the same.

### 4.2 Euler Equations

In this section we consider approximations of the one-dimensional Euler equations. The Euler equations can be written in the general form of a conservation law (1) with

$$q = \begin{pmatrix} \rho \\ \rho v \\ E \end{pmatrix}, \quad f(q) = \begin{pmatrix} \rho v \\ \rho v^2 + p \\ v(E + p) \end{pmatrix}. \quad (26)$$

**Table 1** Convergence study for the Burgers equation (19) with data (25) at time  $t = 0.15$ , using Roe’ method with wave-speed correction as well as our ADER version of active flux method

Grid	Roe’s method with wave-speed correction		ADER version of the AF method	
	$L_1$ -error	EOC	$L_1$ -error	EOC
64	5.59063d-5		8.14128d-5	
128	8.40714d-6	2.73	4.09186d-5	0.99
256	1.20434d-6	2.80	1.12111d-5	1.86
512	6.08588d-7	0.98	1.47012d-6	2.92
1024	1.47617d-7	2.04	2.04482d-7	2.84
2048	2.27928d-8	2.70	2.61081d-8	2.96
4096	3.11783d-9	2.86	3.19291d-9	3.03

Here  $\rho, v, p$  and  $E$  denote the density, the velocity component in the  $x$  direction, the pressure and the energy, respectively. We use the ideal gas equation of state

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho v^2,$$

with adiabatic exponent  $\gamma = 1.4$ .

We explain our approach for the computation of  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i+\frac{1}{2}}^{n+1}$ , since this is the only part of the method which needs to be adapted to the specific equations. We again want to use the Taylor series expansion (10) in order to compute these interface values of the conserved quantities. For nonlinear hyperbolic systems of the general form (1), we express the time derivatives of the conserved quantities using the relations

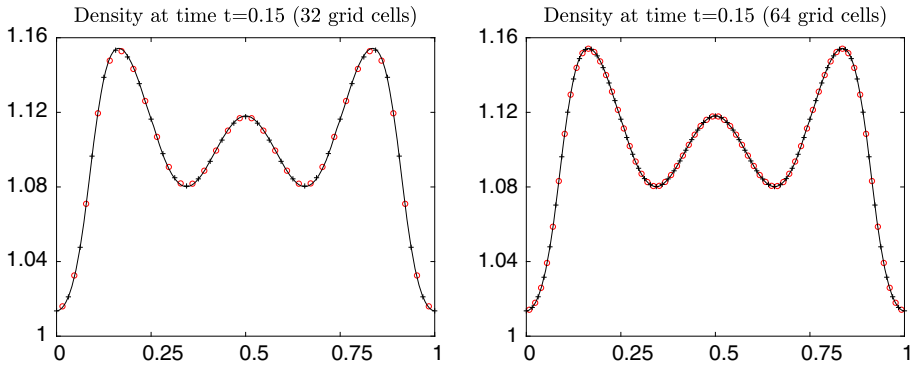
$$\begin{aligned} q_t &= -f(q)_x = -f'(q)q_x \\ q_{tx} &= q_{xt} = -(f'(q)q_x)_x = -(f''(q)q_xq_x + f'(q)q_{xx}) \\ q_{tt} &= -(f'(q)q_x)_t = -(f''(q)q_tq_x + f'(q)q_{xt}). \end{aligned}$$

Thus, the conservative quantities at the grid cell interface can be expressed in the form

$$\begin{aligned} q(x_{i+\frac{1}{2}}, t_n + \tau) &\approx Q_{i+\frac{1}{2}}^n - \tau f'(Q_{i+\frac{1}{2}}) Q_x(x_{i+\frac{1}{2}}, t_n) \\ &+ \frac{1}{2}\tau^2 \left( f''(Q_{i+\frac{1}{2}}) f'(Q_{i+\frac{1}{2}}) Q_x(x_{i+\frac{1}{2}}, t_n) Q_x(x_{i+\frac{1}{2}}, t_n) \right. \\ &+ f'(Q_{i+\frac{1}{2}}) f''(Q_{i+\frac{1}{2}}) Q_x(x_{i+\frac{1}{2}}, t_n) Q_x(x_{i+\frac{1}{2}}, t_n) \\ &\left. + (f'(Q_{i+\frac{1}{2}}))^2 Q_{xx}(x_{i+\frac{1}{2}}, t_n) \right). \end{aligned} \tag{27}$$

The exact form of  $f'(q)$  and  $f''(q)$  can be found in ‘‘Appendix A’’. The values  $Q_x(x_{i+\frac{1}{2}}, t_n)$  and  $Q_{xx}(x_{i+\frac{1}{2}}, t_n)$  are obtained by solving Riemann problems of the form

$$\begin{aligned} (Q_x)_t + f'(Q_{i+\frac{1}{2}}^n)(Q_x)_x &= 0 \\ Q_x(x, t_n) &= \begin{cases} \frac{1}{\Delta x} u'_i(1) : & x < x_{i+\frac{1}{2}} \\ \frac{1}{\Delta x} u'_{i+1}(0) : & x \geq x_{i+\frac{1}{2}} \end{cases} \end{aligned} \tag{28}$$



**Fig. 5** Approximation of the Euler equations (26) with initial data (30) at time  $t = 0.25$  using the ADER version of the active flux method. The solid line is a highly resolved reference solution. The circles indicate cell averaged values and the crosses indicate interface values of the conserved quantities. We show results for density using 32 (left) and 64 (right) grid cells

**Table 2** Convergence study for Euler equations (26) with data (30) at time  $t = 0.25$

Grid	$L_1$ -error in density	EOC
32	2.22499d-4	
64	2.76332d-5	3.01
128	3.55307d-6	2.96
256	4.57967d-7	2.96

as well as

$$(Q_{xx})_t + f'(Q^n_{i+\frac{1}{2}})(Q_{xx})_x = 0$$

$$Q_{xx}(x, t_n) = \begin{cases} \frac{1}{\Delta x^2} u''_i(1) : & x < x_{i+\frac{1}{2}} \\ \frac{1}{\Delta x^2} u''_{i+1}(0) : & x \geq x_{i+\frac{1}{2}}. \end{cases} \tag{29}$$

To study the accuracy of the resulting method for smooth solutions of the Euler equations, we consider initial values of the form

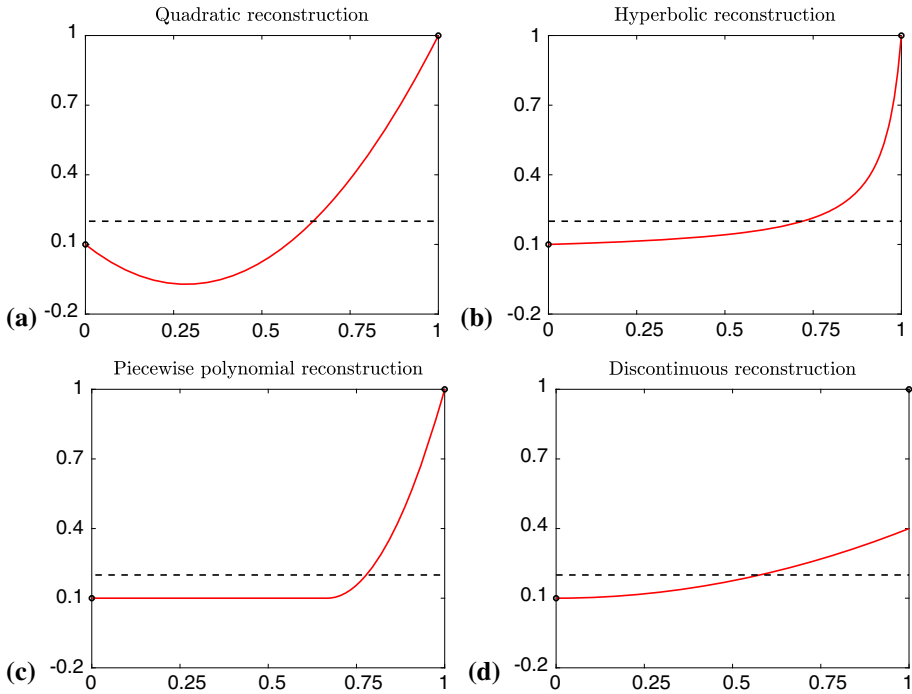
$$\rho(x, 0) = p(x, 0) = 1 + \frac{1}{2} \exp\left(-80\left(x - \frac{1}{2}\right)^2\right)$$

$$v(x, 0) = 0 \tag{30}$$

on the interval  $[0, 1]$  with periodicity condition. Our time steps satisfy  $CFL \leq 0.9$ . In Fig. 5 we show numerical results for the density at time  $t = 0.25$  on two-different grids together with a highly resolved reference solution. In Table 2 we show the results of our convergence study, which confirm the third order accuracy of the method.

### 5 Limiters for the Active Flux Method

Using modified equation analysis, as explained in [19, Chapter 2], it can be shown that methods for the advection equation with an even order of accuracy behave differently at discontinuities or steep gradients than odd-ordered methods. While the leading error term in



**Fig. 6** Different reconstructions of the active flux method. The dashed line indicates the cell average value, the ‘o’ symbols at the boundaries indicate the edge values and the red solid line indicates the reconstruction. For all of these plots we used  $Q_{i-\frac{1}{2}}^n = 0.1$ ,  $Q_{i+\frac{1}{2}}^n = 1$  and  $Q_i^n = 0.2$  (Color figure online)

the modified equation analysis of a second order accurate method creates dispersive waves, the leading error term of a third order accurate method provides damping of spurious oscillations. Thus, in a third order accurate method much less limiting is needed in order to eliminate oscillations. Nevertheless, it is a challenge to find the appropriate amount of limiting. In this section we will discuss different limiting strategies for the active flux method.

The piecewise quadratic reconstruction (5), which is used in the unlimited active flux method, can create new minima or maxima. New minima or maxima will occur if  $u'(0)u'(1) < 0$ . Problems need to be expected if there is a relatively large difference between  $Q_{i-\frac{1}{2}}^n$  and  $Q_{i+\frac{1}{2}}^n$ , while the cell average value  $Q_i^n$  is relatively close to one of these interface values. Such a situation is shown in Fig. 6a.

In order to avoid unphysical oscillations, we want to replace the piecewise quadratic reconstruction, described in Eq. (5), if needed. We will now discuss different reconstructions which are shown in Fig. 6b–d and discuss their use in active flux methods.

### 5.1 Continuous, Piecewise Monotone Reconstructions

In Fig. 6b, c we show reconstructions which are monotone within one grid cell. In Fig. 6b, we show a hyperbola with edge values  $Q_{i-\frac{1}{2}}^n$  and  $Q_{i+\frac{1}{2}}^n$  and average value  $Q_i^n$ . This reconstruction is motivated by Marquina’s work [10], who used a hyperbolic reconstruction

in numerical methods for conservation laws. Here, the degrees of freedom of the active flux method are used in the definition of the hyperbola. The reconstruction has the general form

$$u_i(\xi) = a + \frac{b}{\xi - \frac{1}{2} + c}. \tag{31}$$

The coefficients  $a, b, c$  are obtained by solving a nonlinear system of equations, which is obtained from this ansatz by imposing

$$u_i(0) = Q_{i-\frac{1}{2}}^n, u_i(1) = Q_{i+\frac{1}{2}}^n \text{ and } \int_0^1 u_i(\xi)d\xi = Q_i^n.$$

We solved the resulting nonlinear equations using an iterative Newton-type method.

In Fig. 6c we show a continuous, piecewise polynomial reconstruction, which (in each grid cell) consists of a constant part followed by a parabola. This is inspired by a limiter proposed in Roe et al. [13]. The cell average is maintained and the interpolation condition at the two boundary points of the domain is satisfied. This reconstruction is much easier to compute than the hyperbolic reconstruction and can be expressed in the form

$$\begin{aligned} &\text{If } \left( Q_i^n - \frac{1}{2} \left( Q_{i-\frac{1}{2}}^n + Q_{i+\frac{1}{2}}^n \right) \right) \left( Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n \right) > 0, \text{ then} \\ &u_i(\xi) = \begin{cases} Q_{i-\frac{1}{2}}^n : & \xi < \xi^* \\ Q_{i-\frac{1}{2}}^n + \left( Q_{i+\frac{1}{2}}^n - Q_{i-\frac{1}{2}}^n \right) (\xi - \xi^*)^2 / (1 - \xi^*)^2 : & \xi > \xi^* \end{cases} \\ &\text{with } \xi^* = \left( 2Q_{i-\frac{1}{2}}^n + Q_{i+\frac{1}{2}}^n - 3Q_i^n \right) / \left( Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n \right) \end{aligned}$$

otherwise

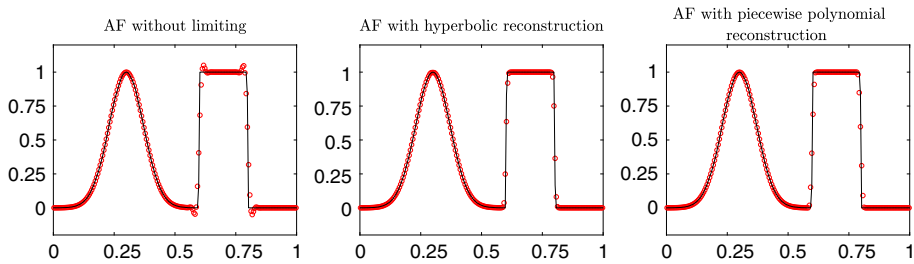
$$\begin{aligned} &u_i(\xi) = \begin{cases} Q_{i+\frac{1}{2}}^n : & \xi < \xi^* \\ Q_{i+\frac{1}{2}}^n + \left( Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n \right) (\xi - \xi^*) / \xi^* : & \xi > \xi^* \end{cases} \\ &\text{with } \xi^* = 3 \left( Q_i^n - Q_{i+\frac{1}{2}}^n \right) / \left( Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n \right). \end{aligned}$$

We now discuss the use of these reconstructions in an active flux method for the advection equation. If the reconstruction does not produce new minima or maxima, the same is true for the exact solution of the advection equation with such data. However, by using the active flux method as described in Sect. 2, the method would in general produce unphysical oscillations. The reason for these oscillations is that Simpson’s rule does not integrate the hyperbola or the piecewise polynomial function exactly. The last equality in (7) does not hold if  $u_i$  is a hyperbola or a piecewise defined function.

For the advection equation we can easily replace Simpson’s rule by an exact integration formula, using the relations described in the first two lines of Eq. (7). For the piecewise polynomial reconstruction we can alternatively use a combined Simpson’s rule for the two intervals which are divided by  $\xi^*$ . The new value  $Q_{i+\frac{1}{2}}^{n+1}$  is still computed by using the characteristics but based on the new reconstructions.

These new reconstructions produce very good results as indicated in Fig. 7 (middle) and (right) for a standard advection test problem. In Fig. 7 (left) we show the results of the unlimited active flux method. For this test we used initial data





**Fig. 7** Advection test computed with different versions of the active flux method using 200 grid cells. The solution is shown at time  $t = 1$ , i.e. after one rotation. The solution using the unlimited active flux method is shown in the left plot. In the middle and the right plot we show results for the hyperbolic and piecewise polynomial reconstruction, respectively

$$q(x, 0) = \begin{cases} 1 + \exp(-100(x - 0.3)^2) : & x \in [0.6, 0.8] \\ \exp(-100(x - 0.3)^2) : & x \in [0, 1] \setminus [0.6, 0.8]. \end{cases}$$

The advection speed has been set to one. We used a grid with 200 grid cells and time steps corresponding to  $CFL \leq 0.9$ . The periodicity condition  $q(0, t) = q(1, t)$  is imposed. Thus, at time  $t = 1$  the exact solution agrees with the initial values.

Unfortunately, these new reconstructions are not well suited if we use the ADER interpretation of the active flux method. Note that the original form of the active flux method and its ADER interpretation are only equivalent if we use the piecewise parabolic reconstruction (5). For the hyperbolic reconstructions, the derivatives  $u_i^{(k)}(\xi)$  do not decrease fast enough for  $k > 2$  and the Taylor series expansion does not converge. Therefore, the ADER approach does not converge. Active flux with exact evolution can still be used. However, the exact flux can no longer be approximated by Simpson’s rule, since this would be equivalent to an approach that replaces the hyperbola or the piecewise polynomial reconstruction by a quadratic function which might introduce new minima or maxima.

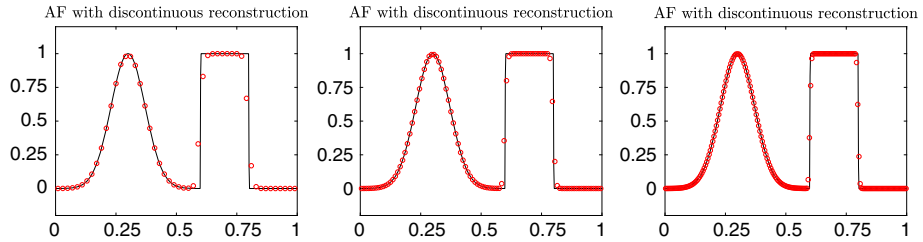
For the piecewise polynomial reconstruction, the use of the ADER interpretation is only valid up to a time which depends on  $\xi^*$ . This leads to a reduced time step restriction, with time steps that might become arbitrarily small since  $\xi^*$  can be located anywhere in the interval  $(0, 1)$ . The reason for this time step restriction is that the reconstruction is not twice continuously differentiable in each grid cell. Therefore, the Taylor series expansion (10) with (11) is only valid for small time steps.

### 5.2 Discontinuous Quadratic Reconstruction

In Fig. 6d we show a quadratic reconstruction, which has the correct cell average but interpolates only one interface value. We require that the reconstruction uses the interface value which is closer to the cell average value. Furthermore, we require that the first derivative of the reconstruction is equal to zero at that boundary point. This implies that the second boundary point is as close as possible to the original boundary point. The reconstruction can be expressed in the form

$$\text{If } \left( Q_i^n - \frac{1}{2} \left( Q_{i-\frac{1}{2}}^n + Q_{i+\frac{1}{2}}^n \right) \right) \left( Q_{i-\frac{1}{2}}^n - Q_{i+\frac{1}{2}}^n \right) > 0, \text{ then}$$

$$u_i(\xi) = Q_{i-\frac{1}{2}}^n + 3 \left( Q_i^n - Q_{i-\frac{1}{2}}^n \right) \xi^2$$



**Fig. 8** Advection test computed with the limited discontinuous reconstruction. The solution with (left) 50, (middle) 100 and (right) 200 grid cells is shown at time  $t = 1$ , i.e. after one rotation

otherwise

$$u_i(\xi) = Q_{i+\frac{1}{2}}^n + 3 \left( Q_i^n - Q_{i+\frac{1}{2}}^n \right) (\xi - 1)^2.$$

The discontinuous quadratic reconstruction can be used both for the original version of the active flux method as well as for the ADER version. At interface points where the reconstruction is discontinuous, i.e. at points  $x_{i+\frac{1}{2}}$  with  $u_i(1) \neq u_{i+1}(0)$ , we need to recompute the interface value  $Q_{i+\frac{1}{2}}^n$ , since this value is used in Simpson’s formula for the flux. This is done by solving the Riemann problem with data

$$Q(x, t_n) = \begin{cases} u_i(1) & : x < x_{i+\frac{1}{2}} \\ u_{i+1}(0) & : x > x_{i+\frac{1}{2}} \end{cases}$$

The values  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i+\frac{1}{2}}^{n+1}$  can be computed as before using the recomputed value  $Q_{i+\frac{1}{2}}^n$ .

In Fig. 8 we show numerical results for the advection problem from the previous section on three different grids.

### 5.3 Criteria for the Use of a Limited Reconstruction for Linear Problems

For the advection problem we used one of the new reconstructions in grid cell  $i$ , if the reconstruction from Eq. (5) produces a new minima or maxima within the grid cell. This is the case, if

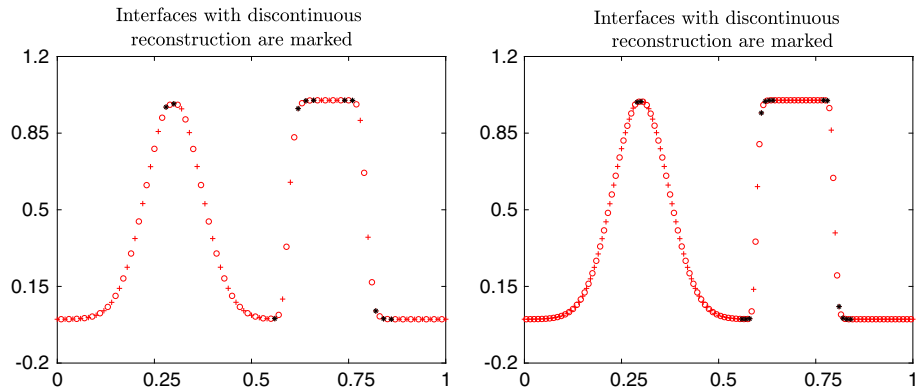
$$\begin{aligned} & u'_i(0)u'_i(1) < 0 \\ \Leftrightarrow & \left( Q_{i-\frac{1}{2}}^n - 3Q_i^n + 2Q_{i+\frac{1}{2}}^n \right) \left( -2Q_{i-\frac{1}{2}}^n + 3Q_i^n - Q_{i+\frac{1}{2}}^n \right) < 0. \end{aligned} \tag{32}$$

The only exception in which we continue to use the original piecewise quadratic reconstruction is, if

$$Q_i^n > \max \left( Q_{i-\frac{1}{2}}^n, Q_{i+\frac{1}{2}}^n \right) \quad \text{or} \quad Q_i^n < \min \left( Q_{i-\frac{1}{2}}^n, Q_{i+\frac{1}{2}}^n \right), \tag{33}$$

In this situation it is not possible to construct a monotone function with the correct cell average and interface values. Such a situation arises near local extrema of the solution.

For a fixed grid, all of our attempts to reduce the number of cells marked for limiting lead to oscillations. In Fig. 9, we show again the results of the advection test with limited,



**Fig. 9** Cell average ‘o’ and point values ‘+’ for the last time step of the advection test. The grid cell interfaces with a discontinuous reconstruction are marked with a black ‘\*’. The left plot shows a simulation with 50 grid cells, the right plot shows a simulation with 100 grid cells

discontinuous reconstruction. Now the grid cell interfaces with discontinuous reconstruction are marked.

### 5.4 Limiting Strategies for Burgers’ Equation

For the Burgers equation (19), the wave speed at each grid cell interface is determined by the conserved quantity  $q$ . This requires a change of the limiting strategy compared to the advection equation. The unlimited active flux method with continuous, piecewise quadratic reconstruction is unstable in situation where  $q$  changes sign from positive to negative, as explained in Sect. 5.5.

If  $u'_i(0)u'_i(1) > 0$  with  $u_i$  described in (5), we use the standard piecewise quadratic reconstruction (5).

If  $u'_i(0)u'_i(1) < 0$ , the piecewise quadratic reconstruction would lead to new minima or maxima. In this situation, we use a discontinuous reconstruction. Near shock waves, i.e. if  $Q^n_{i-\frac{1}{2}} > Q^n_{i+\frac{1}{2}}$ , we use a piecewise linear reconstruction of the form

$$u_i(\xi) = Q^n_i + \sigma_i \left( \xi - \frac{1}{2} \right), \tag{34}$$

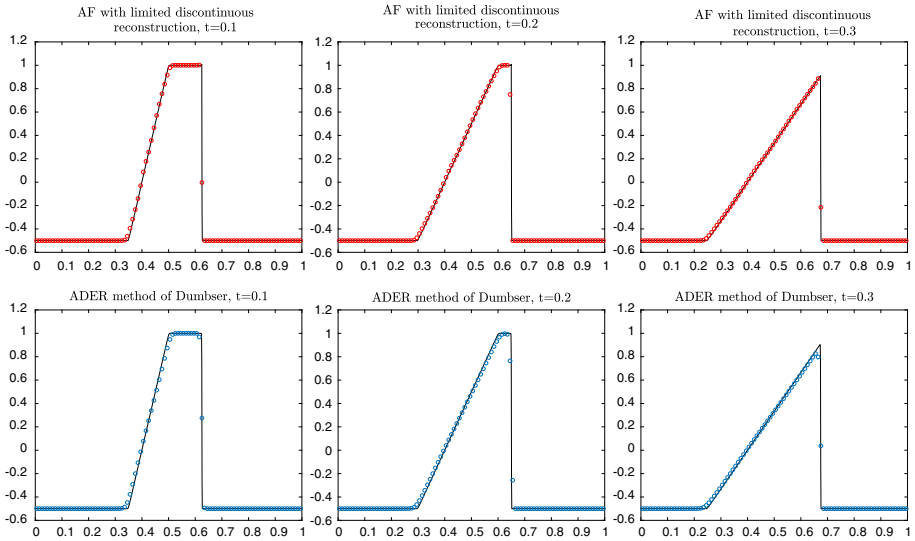
with

$$\sigma_i = \text{minmod} \left( 2 \left( Q^n_i - Q^n_{i-\frac{1}{2}} \right), 2 \left( Q^n_{i+\frac{1}{2}} - Q^n_i \right) \right) \tag{35}$$

and the standard minmod function

$$\text{minmod}(a, b) = \begin{cases} a & : ab > 0 \text{ and } |a| \leq |b| \\ b & : ab > 0 \text{ and } |b| \leq |a| \\ 0 & : ab \leq 0. \end{cases}$$

This also implies, that in the situations described by (33), a piecewise constant reconstruction is used. Near rarefaction waves, i.e. if  $Q^n_{i-\frac{1}{2}} < Q^n_{i+\frac{1}{2}}$ , we continue to use the discontinuous, piecewise quadratic reconstruction described in Sect. 5.2.



**Fig. 10** Test computation for Burgers’ equation. Top line: using a discontinuous, limited, piecewise quadratic reconstruction together with the ADER interpretation of the active flux method. Bottom line: using the third order accurate ADER-DG method of Dumbser. The solid line is a reference solution, computed using 1000 grid cells. The ‘o’ symbols denote cell average values of the conserved quantity for a computation with 100 grid cells

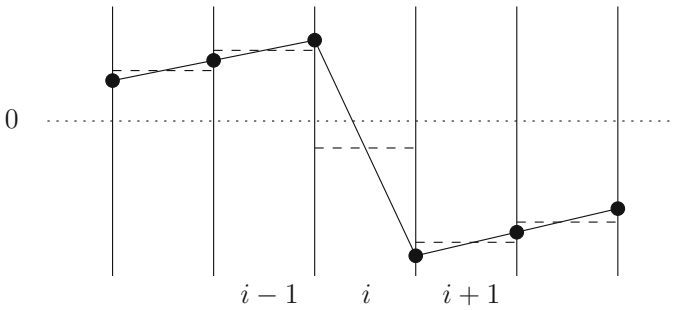
For our test computation we consider Burgers’ equation with initial data

$$q(x, 0) = \begin{cases} 1 & 0.4 \leq x \leq 0.6 \\ -0.5 & \text{otherwise} \end{cases}$$

In Fig. 10, we show numerical results using the limited version of the active flux method based on the ADER approach. We computed the solution at different times using time steps corresponding to  $CFL \leq 0.8$ . We compare the numerical solution with results of the 3rd order accurate version of a one-step ADER finite volume method with space–time DG predictor that was kindly provided to us by Michael Dumbser.

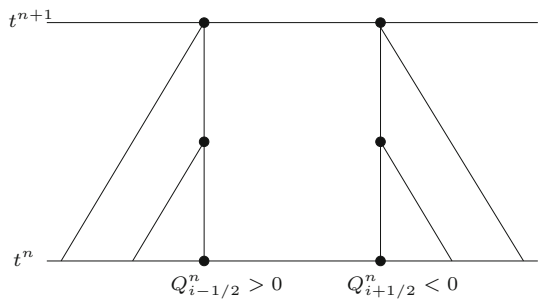
### 5.5 Problems of Continuous Reconstructions

We will now describe a situation in which the active flux method with continuous, piecewise quadratic reconstruction fails to approximate the correct solution structure. We consider Burgers’ equation and data of the form illustrated in Fig. 11. In this situation, we have a grid cell  $i$  with  $Q_{i-\frac{1}{2}}^n > 0$  and  $Q_{i+\frac{1}{2}}^n < 0$ . In the standard form of the active flux method for Burgers’ equation [i.e. (23) and (24)], the numerical fluxes  $F_{i-\frac{1}{2}}$  and  $F_{i+\frac{1}{2}}$  are not influenced by the cell average value  $Q_i^n$ . This is indicated in Fig. 12. By interpreting the interface value as local propagation speed, the approximation of  $q(x_{i\pm\frac{1}{2}}, t_n + \tau)$  is based on a positive propagation speed at interface  $i - 1/2$  and a negative speed at interface  $i + 1/2$ . The values  $Q_{i-\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i-\frac{1}{2}}^{n+1}$  are computed using the reconstruction  $u_{i-1}$ , while the values  $Q_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and  $Q_{i+\frac{1}{2}}^{n+1}$  are computed using  $u_{i+1}$ . Therefore, the reconstruction from cell  $i$  is never used in the



**Fig. 11** Illustration of initial data for Burgers’ equation, which require a discontinuous reconstruction

**Fig. 12** Schematic description of the active flux method for Burgers’ equation for a grid cell with  $Q^n_{i-\frac{1}{2}} > 0$  and  $Q^n_{i+\frac{1}{2}} < 0$



update. In the special situation where  $Q^n_i = 0$  and  $F_{i-\frac{1}{2}} = F_{i+\frac{1}{2}}$ , the method approximates the correct solution structure. In general, the fluxes  $F_{i-\frac{1}{2}}$  and  $F_{i+\frac{1}{2}}$  will differ and the cell value  $Q_i$  will increase or decrease indefinitely during the next time steps. Our discontinuous reconstruction offers one possibility to cure this failure. In a personal conversation, Roe pointed out that there might be other possibilities to fix this problem, which still use a continuous reconstruction.

### 5.6 Euler Computations

Finally, we show numerical results for the Euler equations, which require the use of limiters. Again we compare our numerical results with results of the 3rd order accurate version of Dumbser’s ADER-DG method.

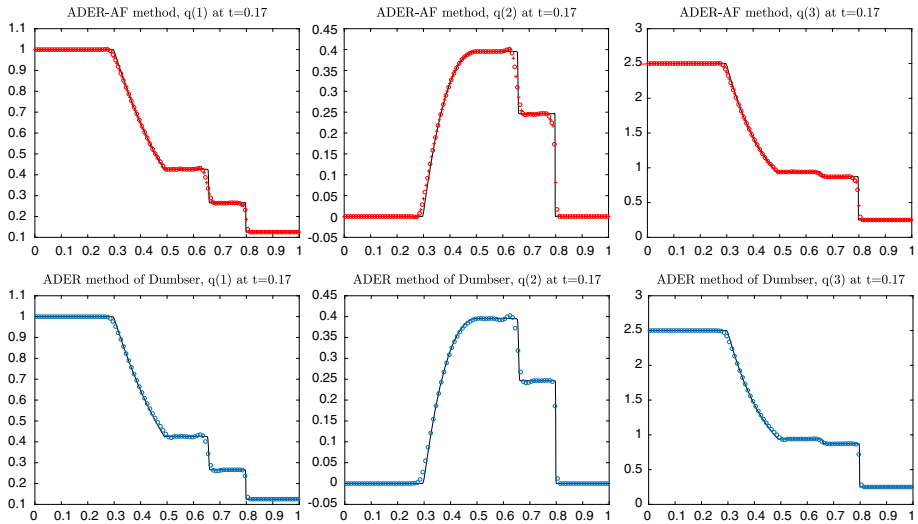
Our first test problem is Sod’s shock tube problem, i.e. we compute solutions of the Euler equations (26) with initial values of the form

$$(\rho, v, p)(x, 0) = \begin{cases} (1, 0, 1) : & x < \frac{1}{2} \\ (0.125, 0, 0.1) : & x \geq \frac{1}{2}. \end{cases}$$

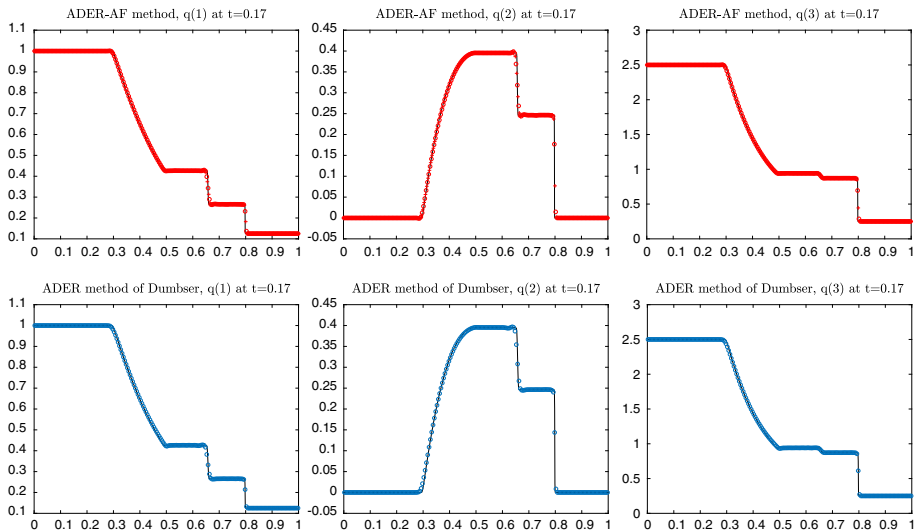
Numerical results at time  $t = 0.17$  computed with 100 and 200 grid cells as well as a highly resolved reference solution are shown in Figs. 13 and 14.

Our second test problem is the well known Shu–Osher test problem [16], i.e. (26) with initial data

$$(\rho, v, p)(x, 0) = \begin{cases} (3.857143, 2.629369, 10.3333) : & x < -4 \\ (1 + 0.2 \sin(5x), 0, 1) : & x \geq -4. \end{cases}$$



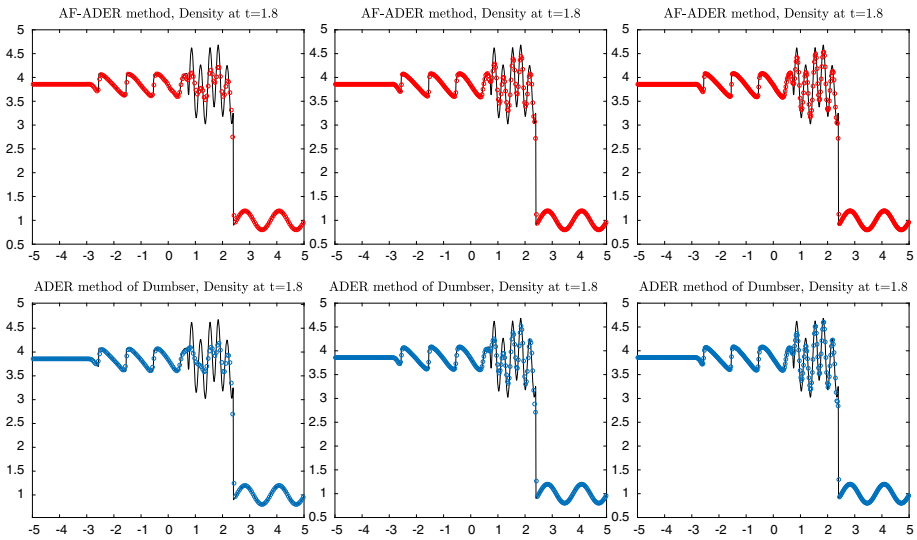
**Fig. 13** Test computation for Euler equations using a discontinuous, limited, piecewise quadratic reconstruction together with the ADER interpretation of the active flux method. The solid line is a reference solution, computed using 2000 grid cells. The ‘o’ symbols denote cell average and the ‘+’ symbols denote the interface values computation with 100 grid cells. Time steps correspond to  $CFL \leq 0.9$ . In the first row we show the results of the new ADER version of the active flux method, in the second row we show results obtained with the 3rd order accurate ADER-DG method of Dumbser



**Fig. 14** Same as in Fig. 13, but with 200 grid cells

In Fig. 15, we show numerical results for the density at time 1.8, computed on grids with 200, 300 and 400 cells. A highly resolved solution is shown as solid line.

Our goal is to use the unlimited ADER version of the active flux method as described in Sect. 4.2 in as many grid cells as possible. For all components of the conserved quantities we use the same type of reconstruction, i.e. either the standard continuous, piecewise quadratic



**Fig. 15** First row: Results for the Shu–Osher test problem using the active flux method with 200, 300 and 400 grid cells. The solid line is a reference solution, which was obtained using 2000 grid cells. In the second row we show the results using the third order accurate ADER-DG method of Dumbser

reconstruction, the discontinuous, piecewise quadratic reconstruction of Sect. 5.2 or the discontinuous, piecewise linear reconstruction (34), (35). For the Euler equations, most of the structure is seen in density. Therefore, a special consideration is given to this quantity in our limiting strategy. If in a grid cell  $i$  the standard piecewise quadratic reconstruction (5) in density does not produce a new minima or maxima, i.e. if  $u'_i(0)u'_i(1) > 0$ , then we use the active flux method in the unlimited form as described in Sect. 4.2. If instead  $u'_i(0)u'_i(1) < 0$ , then we make further tests in order to decide on the appropriate reconstruction. Along shock waves, i.e. if velocity and pressure decreases, we use the piecewise linear reconstruction (34), (35). In smoothly varying regions near local extrema of the flow, we continue to use the standard reconstruction. Otherwise, we use the discontinuous, piecewise quadratic reconstruction. The discontinuous, quadratic reconstruction will typically be used at the start or the end of a rarefaction wave.

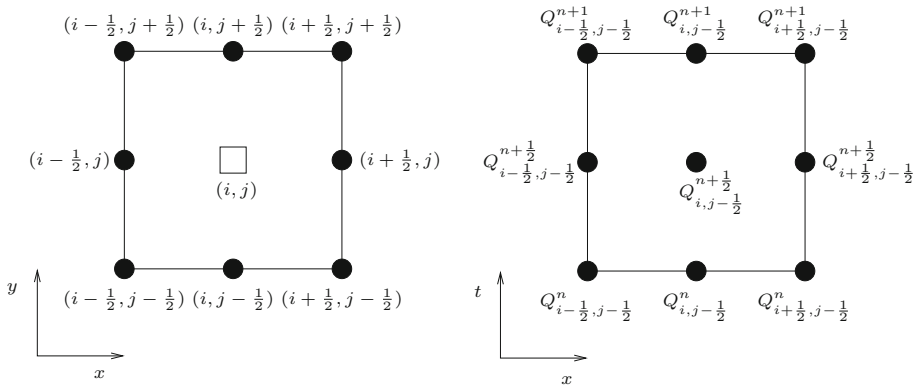
The Shu–Osher test confirms that our limited version of the active flux method can capture both the small-scale smooth flow features as well as the shock wave. Furthermore, it compares well with the ADER-DG method of Dumbser. The unlimited version of the active flux method is unstable for the two test problems considered in this section, while the limited version is stable for time steps which satisfy the inequality  $CFL \leq 1$ .

### 6 Multidimensional Problems

In this section we discuss possible extensions of our scheme to the two-dimensional case. The equations of interest have the general form

$$\begin{aligned}
 q_t + f(q)_x + g(q)_y &= 0 & -\infty < x < \infty, -\infty < y < \infty, t > 0 \\
 q(x, y, 0) &= q_0(x, y) & -\infty < x < \infty, -\infty < y < \infty,
 \end{aligned}
 \tag{36}$$

where  $q : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}^m$  is a vector of conserved quantities and  $f, g : \mathbb{R}^m \rightarrow \mathbb{R}^m$  are vector valued flux functions. Again, we assume hyperbolicity of the system.



**Fig. 16** Left: Configuration in cell  $(i, j)$ . Right: Area of integration in space–time

In this section, we will first describe the setup of our grid, the degrees of freedom of the methods as well as the reconstruction which has been used. We will then see that the equivalence to the active flux method that was derived in the one-dimensional case carries over to the two-dimensional linear advection equation if we solve one-dimensional Riemann problems in a certain way described below. For two-dimensional linear systems we did not find an ADER version which is equivalent with the method proposed by Roe and coauthors. Nevertheless, we will present a third order accurate ADER type evolution procedure which leads to numerical results that compare well with the original method.

Here we will restrict our considerations to two-dimensional Cartesian grids, although other grids such as triangular grids can certainly be used. The original active flux method by Eymann and Roe uses unstructured triangular grids [4,12]. Those grids allow the reconstruction of piecewise quadratic functions with fewer degrees of freedom. Our restriction to Cartesian grids simplifies the comparison of the ADER evolution of interface values with the exact multidimensional evolution that was used by Roe and coauthors. The degrees of freedom that contribute to any given cell will therefore be the four corners and the midpoints of the four edges as well as the cell average. As the edge points are used by two cells and the points in each corner are used by four cells, each grid cell requires the update of four degrees of freedom.

To simplify the calculations, all cells will be mapped to a reference cell  $[-1, 1]^2$ . The cell averages  $Q^n_{i,j}$ , cell corners  $Q^n_{i+\frac{1}{2},j+\frac{1}{2}}$  and cell edges  $Q^n_{i+\frac{1}{2},j}$ ,  $Q^n_{i,j+\frac{1}{2}}$  will be defined in the usual way. Figure 16 (left) shows the arrangement in one cell. The reconstruction of the conserved quantities will be performed with the contributing nine degrees of freedom to determine in each grid cell a polynomial  $\hat{u}$  of the form

$$\hat{u}(\xi, \eta) = c_{00} + c_{10}\xi + c_{01}\eta + c_{20}\xi^2 + c_{11}\xi\eta + c_{02}\eta^2 + c_{21}\xi^2\eta + c_{12}\xi\eta^2 + c_{22}\xi^2\eta^2,$$

with  $c_{ij} \in \mathbb{R}$ ,  $i, j \in \{0, 1, 2\}$ . This polynomial will more conveniently be expressed in the form

$$\hat{u}(\xi, \eta) = \sum_{i=1}^9 c_i N_i(\xi, \eta). \tag{37}$$

The basis functions  $N_i$  and coefficients  $c_i$ ,  $i = 1, \dots, 9$  can be found in Table 3.



**Table 3** Basis functions and coefficients for the two-dimensional reconstruction (37)

$i$	$c_i$	$N_i$
1	$Q_{i-\frac{1}{2},j-\frac{1}{2}}$	$\frac{1}{4}(\xi^2 - \xi)(\eta^2 - \eta)$
2	$Q_{i,j-\frac{1}{2}}$	$\frac{1}{2}(1 - \xi^2)(\eta^2 - \eta)$
3	$Q_{i+\frac{1}{2},j-\frac{1}{2}}$	$\frac{1}{4}(\xi^2 + \xi)(\eta^2 - \eta)$
4	$Q_{i+\frac{1}{2},j}$	$\frac{1}{2}(\xi^2 + \xi)(1 - \eta^2)$
5	$Q_{i+\frac{1}{2},j+\frac{1}{2}}$	$\frac{1}{4}(\xi^2 + \xi)(\eta^2 + \eta)$
6	$Q_{i,j+\frac{1}{2}}$	$\frac{1}{2}(\eta^2 + \eta)(1 - \xi^2)$
7	$Q_{i-\frac{1}{2},j+\frac{1}{2}}$	$\frac{1}{4}(\xi^2 - \xi)(\eta^2 + \eta)$
8	$Q_{i-\frac{1}{2},j}$	$\frac{1}{2}(\xi^2 - \xi)(1 - \eta^2)$
9	$\frac{1}{16} \left( 36Q_{i,j} - (Q_{i-\frac{1}{2},j-\frac{1}{2}} + Q_{i+\frac{1}{2},j-\frac{1}{2}} + Q_{i+\frac{1}{2},j+\frac{1}{2}} + Q_{i-\frac{1}{2},j+\frac{1}{2}}) \right. \\ \left. - 4(Q_{i,j-\frac{1}{2}} + Q_{i+\frac{1}{2},j} + Q_{i,j+\frac{1}{2}} + Q_{i-\frac{1}{2},j}) \right)$	$(1 - \xi^2)(1 - \eta^2)$

The resulting piecewise polynomial reconstruction on the whole grid will be denoted by  $u(x, y)$ . This function is not only continuous in the corner and edge points, but on every point on the edge, as the three points on this edge determine the one-dimensional parabola uniquely. Therefore, we will have continuity of all partial derivatives with respect to  $x$  on the horizontal interfaces as well as continuity of all partial derivatives with respect to  $y$  on the vertical interfaces. We will use this fact to our advantage later.

In Fig. 16 (right) we show a space time interface for a horizontal grid cell interface. The dots in this figure indicate the nodes of the two-dimensional Simpson formula. The update of the cell average values is described by a finite volume method of the form

$$Q_{i,j}^{n+1} = Q_{i,j}^n - \frac{\Delta t}{\Delta x} \left( F_{i+\frac{1}{2},j} - F_{i-\frac{1}{2},j} \right) - \frac{\Delta t}{\Delta y} \left( G_{i,j+\frac{1}{2}} - G_{i,j-\frac{1}{2}} \right), \tag{38}$$

where  $F_{i\pm\frac{1}{2},j}$  and  $G_{i,j\pm\frac{1}{2}}$  are approximations of space–time integrals. At a horizontal grid cell interface, for example, we have

$$G_{i,j-\frac{1}{2}} \approx \frac{1}{\Delta t \Delta x} \int_{t_n}^{t_{n+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} g(q(x, y_{j-\frac{1}{2}}, t)) dx dt.$$

These integrals are approximated by a two-dimensional version of Simpson’s rule, i.e.

$$\begin{aligned} G_{i,j-\frac{1}{2}} := & \frac{1}{36} \left( g \left( Q_{i-\frac{1}{2},j-\frac{1}{2}}^n \right) + 4g \left( Q_{i,j-\frac{1}{2}}^n \right) + g \left( Q_{i+\frac{1}{2},j-\frac{1}{2}}^n \right) \right) \\ & + 4g \left( Q_{i-\frac{1}{2},j-\frac{1}{2}}^{n+\frac{1}{2}} \right) + 16g \left( Q_{i,j-\frac{1}{2}}^{n+\frac{1}{2}} \right) + 4g \left( Q_{i+\frac{1}{2},j-\frac{1}{2}}^{n+\frac{1}{2}} \right) \\ & + g \left( Q_{i-\frac{1}{2},j-\frac{1}{2}}^{n+1} \right) + 4g \left( Q_{i,j-\frac{1}{2}}^{n+1} \right) + g \left( Q_{i+\frac{1}{2},j-\frac{1}{2}}^{n+1} \right) \end{aligned} \tag{39}$$

and analogously for  $F_{i\pm\frac{1}{2},j}$ . Note that the same flux approximation is used by the active flux method on triangular shaped grids, see for example [4].

For the advection equation as well as the acoustic equations, exact evolution formulas can be used to update the interface values. Such an approach was used in the original active flux method and will also be used below for comparison.

For general hyperbolic problems exact evolution formulas are not available and we are therefore interested in exploring an ADER type approach. Thus, the update of the cell interface values at the intermediate and new time level is obtained by a Taylor series expansion of the exact solution evaluated at the interface points. Time derivatives are replaced by spatial derivatives using the form of the considered partial differential equation. The spatial derivatives of the conserved quantities at interface points are approximated by solving one-dimensional Riemann problems as discussed in more detail below. Especially at the corner points different approximations seem to be possible.

## 6.1 Advection

We start with the advection equation

$$q_t + aq_x + bq_y = 0.$$

For this simple equation the exact evolution formula, i.e.

$$q(x, y, t_n + \tau) = q(x - a\tau, y - b\tau, t_n),$$

can be used in order to approximate the interface values of the conserved quantity at the intermediate and new time level. During each time step, we use all the interface values and all the cell average values at the previous time level, i.e. the level  $t_n$ , to reconstruct the piecewise polynomial function  $u(x, y)$ . Then we can approximate the interface values of the conservative variable at time  $t_{n+\frac{1}{2}}$  and  $t_{n+1}$  by using the approximation

$$q(x, y, t_n + \tau) \approx u(x - a\tau, y - b\tau).$$

Once all the interface values for  $\tau = \Delta t/2$  and  $\tau = \Delta t$  are computed, we use the finite volume method (38) with (39) to compute the cell average values at the new time level  $t_{n+1}$ .

If an exact and easy to evaluate evolution formula is available, then it is of course best to simply use this formula. For general hyperbolic problems such a formula is not available and it is therefore valuable to explore an alternative approach for the approximation of interface values. For the advection equation, Taylor series expansion provides an update formula of the form

$$\begin{aligned} q(x, y, t_n + \tau) &= q(x, y, t_n) + \tau(-aq_x - bq_y) \\ &\quad + \frac{\tau^2}{2}(a^2q_{xx} + abq_{xy} + abq_{yx} + b^2q_{yy}) + \mathcal{O}(\tau^3), \end{aligned} \quad (40)$$

where the dependencies on the variables in the derivatives were dropped for simplicity.

Across a vertical grid cell interface, the reconstructed quantities  $q$ ,  $q_y$  and  $q_{yy}$  are continuous. Thus, those interface values can directly be used in the Taylor series expansion in order to approximate the interface values at the midpoint  $(i - \frac{1}{2}, j)$  of this vertical interface. All the other spatial derivatives of the conserved variable are obtained by solving a one-dimensional Riemann problem in the  $x$  direction, for example an approximation of  $q_x(x_{i-\frac{1}{2}}, y_j, t_n + 0)$  is obtained by solving the Riemann problem

$$q_t + aq_x = 0$$

$$q(x, t_n) = \begin{cases} u_x(x_{i-\frac{1}{2}} - 0, y_j) & x < x_{i-\frac{1}{2}} \\ u_x(x_{i-\frac{1}{2}} + 0, y_j) & x > x_{i-\frac{1}{2}} \end{cases}$$

We obtain

$$q_x(x_{i-\frac{1}{2}}, y_j, t_n + 0) \approx \begin{cases} u_x(x_{i-\frac{1}{2}} - 0, y_j) : & a > 0 \\ u_x(x_{i-\frac{1}{2}} + 0, y_j) : & a < 0 \end{cases}$$

Analogously, at a horizontal interface the values  $q$ ,  $q_x$  and  $q_{xx}$  can be obtained from the continuous reconstruction and all the other spatial derivatives can be computed by solving a one-dimensional Riemann problem.

At a corner  $(i - \frac{1}{2}, j + \frac{1}{2})$ , for each non-mixed derivative, the four initial values of the two-dimensional Riemann problem will collapse into two values due to the same continuity argument. We will therefore approximate this two-dimensional Riemann problem by the resulting one-dimensional Riemann problem as it is done in the edges. We solve two Riemann problem in the  $x$ -direction in order to compute  $q_x$  and  $q_{xx}$ . Furthermore, we solve two Riemann problems in the  $y$ -direction to compute  $q_y$  and  $q_{yy}$ . The only troublesome terms are the mixed derivatives  $q_{xy}$  and  $q_{yx}$ . For the advection equation, there is no difference in how these two are treated because of the scalar coefficients. Thus, we only consider  $q_{xy}$ . Since we cannot use any continuity of this derivative, we do in fact have four distinct initial values in the general case. We denote these four values by  $q_{xy}^{BL} := u_{xy}(x_{i-\frac{1}{2}} - 0, y_{j+\frac{1}{2}} - 0)$ ,  $q_{xy}^{BR} := u_{xy}(x_{i-\frac{1}{2}} + 0, y_{j+\frac{1}{2}} - 0)$ ,  $q_{xy}^{TL} := u_{xy}(x_{i-\frac{1}{2}} - 0, y_{j+\frac{1}{2}} + 0)$  and  $q_{xy}^{TR} := u_{xy}(x_{i-\frac{1}{2}} + 0, y_{j+\frac{1}{2}} + 0)$ , where the abbreviations L, R, T, B stand for left, right, top, bottom, respectively. First we solve two Riemann problems in the  $x$  direction, i.e. we solve

$$q_t + aq_x = 0$$

$$q(x, 0) = \begin{cases} q_{xy}^{*L} : & x \leq x_{i-\frac{1}{2}} \\ q_{xy}^{*R} : & x > x_{i-\frac{1}{2}} \end{cases},$$

with  $* \in \{B, T\}$ . Again, we are interested in the solution of these Riemann problems at the interface, i.e. at  $x_{i-\frac{1}{2}}$ . Thus, for the advection equation these solutions are given by

$$q_{xy}^* = \begin{cases} q_{xy}^{*L} : & a > 0 \\ q_{xy}^{*R} : & a < 0 \end{cases}$$

Now we use the solutions of these two Riemann problems in order to define initial values for a Riemann problem in the  $y$ -direction, i.e. we solve

$$q_t + bq_y = 0$$

$$q(y, 0) = \begin{cases} q_{xy}^B : & y \leq y_{j+\frac{1}{2}} \\ q_{xy}^T : & y > y_{j+\frac{1}{2}} \end{cases}.$$

Finally, we obtain

$$q_{xy}(x_{i-\frac{1}{2}}, y_{j+\frac{1}{2}}, t_n + 0) \approx \begin{cases} q_{xy}^B : & b > 0 \\ q_{xy}^T : & b < 0 \end{cases}$$

For the advection equation, the same solution would be obtained by first solving two Riemann problems in the  $y$ -direction, followed by a Riemann problem in the  $x$ -direction. We recover the correct solution of the two-dimensional Riemann problem.

Notice that, as we want to construct a third order method, we typically cut the Taylor series expansion after the second derivative. For the one-dimensional method, all derivatives of the reconstruction past  $u''$  vanish, i.e.  $u^{(n)} = 0, n > 2$ . This is not the case for our two-dimensional reconstruction. To obtain a slightly more accurate solution, all the non zero spatial derivative terms that arise when we replace  $q_{ttt}$  and  $q_{ttt}$  by spacial derivatives can be approximated analogously. For our reconstruction on a Cartesian grid, these are the terms  $q_{xxy}, q_{xyy}$  and  $q_{xxyy}$ . All of these terms can be handled analogously to the terms described above. If all of these terms are included in the Taylor series expansion then the resulting ADER approach is again equivalent to the active flux method with exact evolution formula for the update of the interface values. This is due to the fact that Taylor series expansion recovers the correct characteristic path for our reconstructed values.

In our accuracy test, we solve the advection equation with  $a = 1$  and  $b = 0.7$  on the domain  $[0, 1] \times [0, 1]$  with periodic boundary conditions. The initial data are given by

$$q(x, y, 0) = \sin(4\pi(x + y)).$$

At time  $t = 1$  the numerical solution is compared with the exact solution and the error as a function of the mesh width is shown in Fig. 17. The yellow curve is obtained by the active flux with the exact edge update using time steps that satisfy  $CFL \leq 0.9$ , where the CFL number is given by

$$CFL := \frac{2 \max(|a|, |b|)\Delta t}{\Delta x}.$$

A larger error, but still 3rd order accuracy, is observed, if we only include those terms of the Taylor series expansion that are shown in Eq. (40). We call the resulting method ADER reduced. The ADER method which uses all the non-zero derivative terms in the Taylor series expansion will be called ADER full. However, it turns out that ADER reduced is less stable. Therefore, we used smaller time steps corresponding to  $CFL \leq 0.45$ . For comparison, we also show the error curve of the active flux method with exact evolution for the same reduced time step size.

### 6.2 Acoustics

The two-dimensional linear acoustic equations have the form

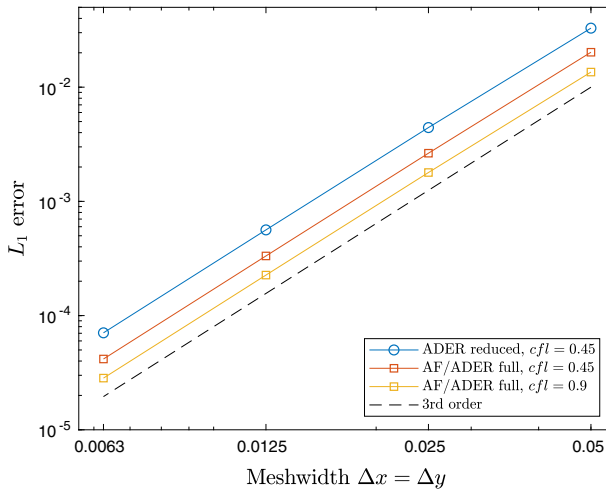
$$p_t + c_0 u_x + c_0 v_y = 0 \tag{41}$$

$$u_t + c_0 p_x = 0 \tag{42}$$

$$v_t + c_0 p_y = 0, \tag{43}$$

where  $c_0$  is the speed of sound,  $p$  is the pressure and  $u, v$  are the velocity components in the  $x$  and  $y$  direction, respectively. Let  $q = (p, u, v)^T$  denote the vector of the conserved quantities and  $q_0 = (p_0, u_0, v_0)^T$  the initial condition. Then, we can write this linear system in the form

$$q_t + \underbrace{\begin{pmatrix} 0 & c_0 & 0 \\ c_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=:A} q_x + \underbrace{\begin{pmatrix} 0 & 0 & c_0 \\ 0 & 0 & 0 \\ c_0 & 0 & 0 \end{pmatrix}}_{=:B} q_y = 0 \tag{44}$$



**Fig. 17** Accuracy study for the two dimensional advection problem. The yellow curve (AF / ADER full) shows the error versus mesh if the exact evolution formula with  $CFL \leq 0.9$  is used to update the interface values. For advection, the ADER method which uses all the nonzero derivative terms for the update of the interface values is equivalent to the method which uses the exact evolution formula. The blue curve shows the error for the ADER update, that uses time steps according to  $CFL \leq 0.45$  and only those derivative values that are necessary in order to obtain third order. The red curve shows the error of the method that uses exact evolution of the interface values and time steps according to  $CFL \leq 0.45$  (Color figure online)

Following our approach, the Taylor series expansion for the acoustic equations is given by

$$q(x, y, t_n + \tau) = q(x, y, t_n) + \tau(-Aq_x - Bq_y) + \frac{\tau^2}{2} (A^2q_{xx} + ABq_{yx} + BAq_{xy} + B^2q_{yy}) + \mathcal{O}(\tau^3). \quad (45)$$

The procedure of solving the two-dimensional Riemann problems is almost the same as it is for the advection equation. The only difference is the treatment of the mixed derivative terms. As  $A$  and  $B$  do not commute, whenever we have to solve a one-dimensional Riemann problem in  $x$ -direction (using matrix  $A$ ) for the term  $ABq_{yx}$ , we first multiply the initial states of the Riemann problem with  $B$ , so they are now  $(Bq_{yx})_L$  and  $(Bq_{yx})_R$ . The solution to this Riemann problem is then plugged back into (45) and only multiplied with  $A$ , and not first with  $B$  and then with  $A$ , as would have been the case if the Riemann problem was solved for  $(q_{yx})_L$  and  $(q_{yx})_R$  directly. The case  $BAq_{xy}$  is treated in analogy. This concept can be extended to higher derivative terms.

For the acoustics problem, the exact two-dimensional solution for data with constant vorticity can be derived by the spherical means formula as was done for example in [4]. This solution formula has the form

$$q(x, y, t) = M_R\{q_0\}(x, y) + R\partial_R M_R\{q_0\}(x, y) - R M_R\{Lq_0\}(x, y) \quad (46)$$

where  $R = c_0t$ ,

$$M_R\{q\}(x, y) = \frac{1}{2\pi R} \int_0^{2\pi} \int_0^R q(x + r \cos \theta, y + r \sin \theta) \frac{r}{\sqrt{R^2 - r^2}} dr d\theta \quad (47)$$

and

$$L = \begin{pmatrix} 0 & \partial_x & \partial_y \\ \partial_x & 0 & 0 \\ \partial_y & 0 & 0 \end{pmatrix}. \quad (48)$$

This is the formula that was proposed to be used to update the interface values in the active flux method in [4]. Instead of  $q_0$  the reconstruction  $u$  at the current time is used and  $t$  is replaced by  $\Delta t$ . The outer integral in (47) is split up into the integrals over the individual cells—two for the edges and four for the corners. This demands that time steps are bounded, such that the circle with radius  $c_0 \Delta t$  around the midpoint of a grid interface is inside the two neighbouring grid cells. Thus we have a time step restriction of the form  $\text{CFL} \leq 1$ , where the CFL number is defined by

$$\text{CFL} := \frac{2c_0 \Delta t}{\min(\Delta x, \Delta y)}.$$

Note that Barsukow et al. [1] have simultaneously to us developed an active flux method for the acoustic equations on Cartesian grids using the spherical mean formula for general initial conditions. In Sect. 6.2, we will compare the active flux method which uses the spherical mean formula with our ADER approach.

Upon inspection, the solution (46) is a polynomial in  $\Delta t$ , as is the Taylor series expansion that we use for our approach. Hence, it is interesting to compare the terms that are obtained by the exact evolution formula with the terms obtained by our approximation. We could not find an approach that was based on the solution of one-dimensional Riemann problems, which matched all the terms that are approximated by the exact evolution formula. Here we do not want to present these long and tedious formulas, which show the differences between the two approaches. Instead we will present some numerical accuracy studies.

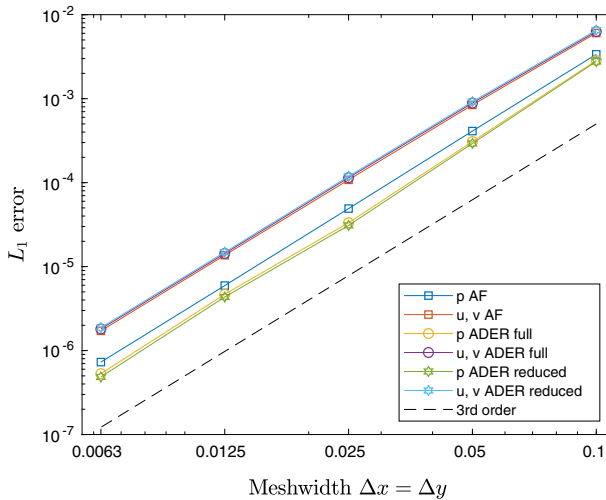
### 6.2.1 Accuracy Study

We consider the acoustic equations (44) with  $c_0 = 1$  and use periodic initial data of the form

$$\begin{aligned} p_0(x, y) &= -\frac{1}{c_0}(\sin(2\pi x) + \sin(2\pi y)) \\ u_0(x, y) &= 0 \\ v_0(x, y) &= 0. \end{aligned}$$

The solution at time  $T = 0.2$  was computed on the rectangular computational domain  $[-1, 1] \times [-1, 1]$ . This test problem was suggested by Lukáčová et al. [9], where the exact solution can also be found. We compare the active flux method with exact evolution with our version of the active flux method that uses the ADER approach to update the interface values. All the other components of the method were handled in the same manner.

In Fig. 18 we show the error as a function of the mesh width. Note that for this example the error in  $u$  is always equal to the error in  $v$ . We again consider two different versions of the ADER approach, one that uses all the non-zero derivative terms in the Taylor series expansion (ADER full) and one version, which uses only those terms of the Taylor series expansion that are needed to get third order accuracy (ADER reduced). In analogy to the advection case, ADER reduced has a reduced stability. Therefore, we used time steps which satisfy  $\text{CFL} \leq 0.45$ . The error in the velocity components is almost identical for the three different methods (i.e., exact evolution, ADER full and ADER reduced). The exact evolution produced



**Fig. 18** Accuracy study for the two-dimensional acoustic equations. We compare results of our ADER approach with results where the exact evolution formula was used to update the interface values

the smallest error, followed by ADER full and ADER reduced. For this test problem, the two ADER versions produced a slightly smaller error in the pressure component than the method which used the exact evolution formula.

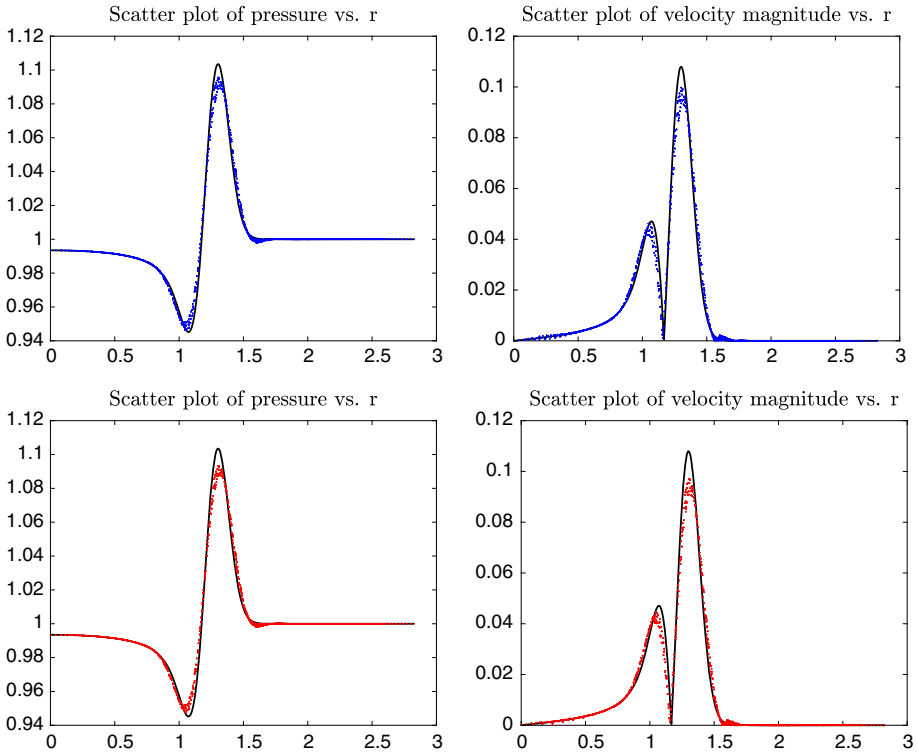
### 6.2.2 Radial Symmetric Test Problem

Now we study a radial symmetric problem, i.e. we solve the acoustic equations with  $c_0 = 1$  and initial values of the form

$$\begin{aligned}
 p(x, y, 0) &= 1 + \exp(-\mu((x - x_0)^2 + (y - y_0)^2)) \\
 u(x, y, 0) &= 0 \\
 v(x, y, 0) &= 0.
 \end{aligned}$$

We approximate the problem on the domain  $[-2, 2] \times [-2, 2]$ , set  $x_0 = y_0 = 0$  and use  $\mu = 50$ . Note that the same problem was studied by Roe and Eymann [4]. There, the active flux method was implemented on a triangular grid and the spherical mean formula was used in order to compute edge values of the conserved quantities during each time step.

Again we compare two versions of the active flux which only vary in the update formula of the interface values. We either use the spherical mean formula or the ADER reduced approach. In Figs. 19 and 20, we show numerical results at time  $t = 1.25$  on grids with  $50 \times 50$  as well as  $150 \times 150$  grid cells. These are scatter plots of the solution, i.e. we plot the interface values over  $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ . A scattering of the data indicates grid effects. Both, on the coarse as well as on the fine grid, the two methods produce results that are comparable in accuracy. Note that the accuracy observed on the Cartesian grid compares well with the accuracy documented in [4], where the active flux method was used on triangular grids with comparable resolution.



**Fig. 19** Scatter plots of pressure and velocity magnitude obtained with two different methods on a grid with  $50 \times 50$  grid cells. The blue dots (first line) show the results obtained by using the exact evolution formula for the update of the edge values, the red dots (second line) indicate the results obtained using the ADER update formula. The black line is the exact solution (Color figure online)

## 7 Conclusion

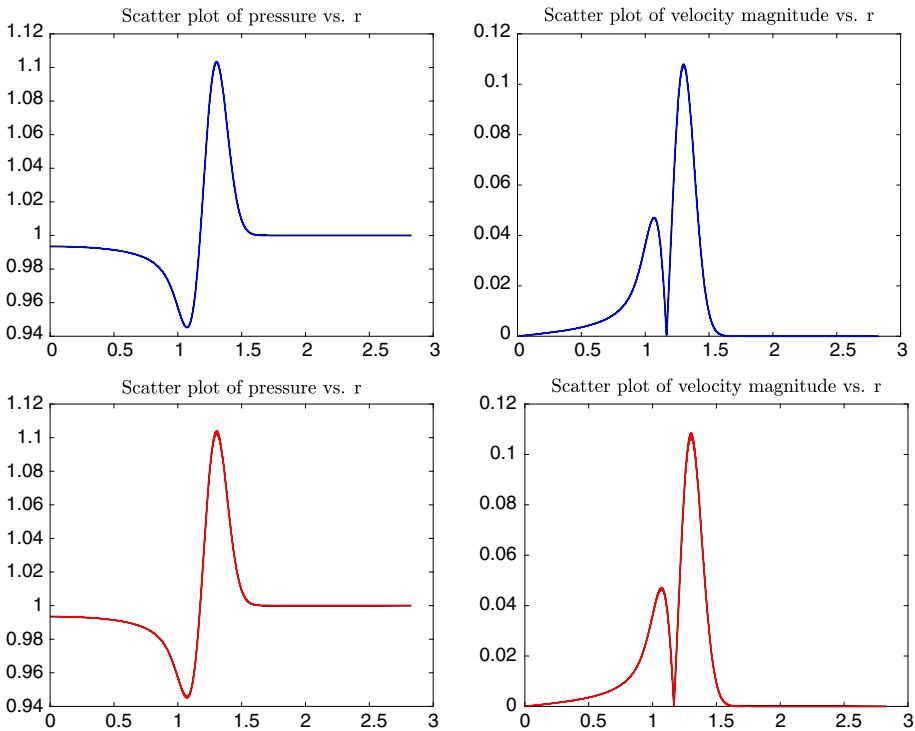
We have studied the active flux method of Eymann and Roe and proposed several variations of this method for one and two-dimensional hyperbolic problems. The active flux method is very attractive, due to its local stencil as well as its high accuracy, which was documented for several test problems.

For one-dimensional linear hyperbolic problems, we showed that the active flux method can be interpreted as a new version of the ADER method. We used this interpretation to propose new active flux methods for nonlinear problems which do not rely on a splitting approach.

While discontinuous reconstructions appear in almost all numerical methods for hyperbolic problems, Roe [12] raised the question as to whether this is really a good idea. Our test computations show that in most situations a continuous reconstruction leads to very accurate results. However, for nonlinear problems we identified a situation where the active flux method in its original version fails to approximate the correct solution structure and becomes unstable. For this situation we propose a discontinuous reconstruction which cures this failure by solving an additional Riemann problem.

Furthermore, we discussed several limiting strategies for the one-dimensional active flux method. For the original active flux method, we proposed two continuous reconstructions,





**Fig. 20** Scatter plots of pressure and velocity magnitude obtained with two different methods on a grid with  $150 \times 150$  grid cells. The blue dots (first line) show the results obtained by using the exact evolution formula for the update of the edge values, the red dots (second line) indicate the results obtained using the ADER update formula. The black line is the exact solution (Color figure online)

which led to accurate approximations of linear problems. Unfortunately, these new reconstructions are not appropriate for the ADER implementation of the active flux method. For the ADER version, limiting was introduced via the use of discontinuous, piecewise quadratic reconstructions. Our limiting does not increase the stencil of the active flux method.

Alternatively, we can use a discontinuous, piecewise quadratic reconstruction that is bound preserving as described by Zhang and Shu [22]. Such a limiting approach can be extended to multi-dimensional problems and can be incorporated into the original version of the active flux method as well as into our ADER approach. We are currently investigating this approach for one and two-dimensional nonlinear problems.

In the two-dimensional case, there are many possibilities to define ADER methods which use the same degrees of freedom as the original active flux method. We constructed ADER methods, which use the underlying continuous, piecewise quadratic reconstruction as much as possible in order to reduce the number of Riemann solves. For the advection equation, the resulting ADER method is equivalent to the active flux method that uses the exact evolution formula for the interface update. For the two-dimensional acoustic equations, the exact evolution of interface values could not be obtained by an ADER type approach. However, our accuracy studies show, that we can still obtain very accurate results.

Roe's motivation to introduce the active flux method was the construction of truly multidimensional methods which do not rely on the solution of one-dimensional Riemann problems.

If an exact evolution formula exists, very efficient and accurate active flux methods can be derived. However, for most hyperbolic problems exact evolution formulas are not available. Therefore, we believe that a good understanding of the connection between the active flux method and the ADER method can lead to new third order accurate methods for conservation laws.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### Appendix A: Implementation of the Active Flux Method for the Euler Equations

In this appendix we present some more details, which are needed in order to implement the ADER version of the active flux method for the Euler equations.

The flux Jacobian matrix  $f'(q)$  has the components

$$f'(q) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma - 3)v^2 & (3 - \gamma)v & (\gamma - 1) \\ \frac{1}{2}(\gamma - 1)v^3 - vH & H - (\gamma - 1)v^2 & \gamma v \end{pmatrix}, \tag{49}$$

with  $H = (E + p)/\rho$ .

In order to compute the interface values  $Q_x(x_{i+\frac{1}{2}}, t_n)$  and  $Q_{xx}(x_{i+\frac{1}{2}}, t_n)$  with the ADER version of the active flux method with continuous, piecewise quadratic reconstruction, we evaluate the flux Jacobian matrix at  $Q^n_{i+\frac{1}{2}}$ . At each time step and each grid cell interface, we then solve Riemann problems of the form (28) and (29). These Riemann problem can be solved exactly in analogy to the solution of Riemann problems for linear hyperbolic systems, see for example [8].

At interfaces with discontinuous reconstruction, we solve an additional Riemann problem of the form

$$Q_t + f'(\hat{Q}^n_{i+\frac{1}{2}})Q_x = 0$$

$$Q(x, t_n) = \begin{cases} u_i(1) : & x < x_{i+\frac{1}{2}} \\ u_{i+1}(0) : & x \leq x_{i+\frac{1}{2}}, \end{cases}$$

where  $\hat{Q}^n_{i+\frac{1}{2}}$  is the Roe averaged state based on the values  $u_i(1)$  and  $u_{i+1}(0)$ .

The tensor  $f''(q)$  has the components

$$\left(\frac{\partial^2 f_1}{\partial q_i \partial q_j}(q)\right)_{i,j=1,\dots,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \left(\frac{\partial^2 f_2}{\partial q_i \partial q_j}(q)\right)_{i,j=1,\dots,3} = \begin{pmatrix} (3 - \gamma)\frac{v^2}{\rho} & (\gamma - 3)\frac{v}{\rho} & 0 \\ (\gamma - 3)\frac{v}{\rho} & (3 - \gamma)\frac{1}{\rho} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\left(\frac{\partial^2 f_3}{\partial q_i \partial q_j}(q)\right)_{i,j=1,\dots,3} = \begin{pmatrix} 2\gamma E \frac{v}{\rho^2} - 3(\gamma - 1)\frac{v^3}{\rho} & -\gamma \frac{E}{\rho^2} + 3(\gamma - 1)\frac{v^2}{\rho} & -\gamma \frac{v}{\rho} \\ -\gamma \frac{E}{\rho^2} + 3(\gamma - 1)\frac{v^2}{\rho} & -3(\gamma - 1)\frac{v}{\rho} & \gamma \frac{1}{\rho} \\ -\gamma \frac{v}{\rho} & \gamma \frac{1}{\rho} & 0 \end{pmatrix}.$$

In regions with continuous reconstruction, we compute  $f''(Q_{i+\frac{1}{2}}^n)$ , i.e. we evaluate the tensor at the interface value of the conserved quantities, which were computed in the previous time step. At interfaces with a discontinuous reconstruction, we first recompute  $Q_{i+\frac{1}{2}}^n$  by solving an additional Riemann problem as described in Sect. 5.2. Now the tensor  $f''$  is evaluated at this recomputed interface value.

## References

1. Barsukow, W., Hohm, J., Klingenberg, C., Roe, P.L.: The active flux scheme on cartesian grids and its low mach number limit. (2018)
2. Eymann, T.A., Roe, P.L.: Active flux schemes. AIAA 2011-382
3. Eymann, T.A., Roe, P.L.: Active flux schemes for systems. AIAA 2011-3840
4. Eymann, T.A., Roe, P.L.: Multidimensional active flux schemes. In: AIAA Conference Paper (2013)
5. Hesthaven, J.S.: Numerical Methods for Conservation Laws: From Analysis to Algorithm (Computational Science and Engineering). SIAM, Philadelphia (2018)
6. Kaiser, K., Schütz, J., Schöbel, R., Noelle, S.: A new stable splitting for the isentropic Euler equations. *J. Sci. Comput.* **70**, 1390–1407 (2017)
7. Karabasov, S.A., Goloviznin, V.M.: Compact accurately boundary-adjusting high-resolution technique for fluid dynamics. *J. Comput. Phys.* **228**, 7426–7451 (2009)
8. LeVeque, R.J.: Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, Cambridge (2002)
9. Lukáčová-Medvid'ová, M., Morton, K.W., Warnecke, G.: Evolution Galerkin methods for hyperbolic systems in two space dimensions. *Math. Comput.* **69**, 1355–1384 (2000)
10. Marquina, A.: Local piecewise hyperbolic reconstruction of numerical fluxes for nonlinear scalar conservation laws. *SIAM J. Sci. Comput.* **15**, 892–915 (1994)
11. Roe, P.: A simple explanation of superconvergence for discontinuous Galerkin solutions to  $u_t + u_x = 0$ . *Commun. Comput. Phys.* **21**, 905–912 (2017)
12. Roe, P.: Is discontinuous reconstruction really a good idea? *J. Sci. Comput.* **73**, 1094–1114 (2017)
13. Roe, P.L., Lung, T., Maeng, J.: New approaches to limiting. In: 22nd AIAA Computational Fluid Dynamics Conference (2015)
14. Roe, P.L., Maeng, J., Fan, D.: Comparing active flux and discontinuous Galerkin methods for compressible flow. In: 2018 AIAA Aerospace Science Meeting (2018)
15. Shu, C.-W.: High order weighted essentially nonoscillatory schemes for convection dominated problems. *SIAM Rev.* **51**, 82–126 (2009)
16. Shu, C.-W., Osher, S.: Efficient implementation of essentially non-oscillatory shock-capturing schemes, II. *J. Comput. Phys.* **83**, 32–78 (1989)
17. Titarev, V.A., Toro, E.F.: ADER: arbitrary high order Godunov approach. *J. Sci. Comput.* **17**, 609–618 (2002)
18. Toro, E.F., Vázquez-Cendón, M.E.: Flux splitting schemes for the Euler equations. *Comput. Fluids* **70**, 1–12 (2012)
19. Trangenstein, J.A.: Numerical Solution of Hyperbolic Partial Differential Equations. Cambridge University Press, Cambridge (2009)
20. van Leer, B.: Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection. *J. Comput. Phys.* **23**, 276–299 (1977)
21. Zeng, X.: A high-order hybrid finite difference–finite volume approach with application to inviscid compressible flow problems: a preliminary study. *Comput. Fluids* **98**, 91–110 (2014)
22. Zhang, X., Shu, C.-W.: Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proc. R. Soc. A* **467**, 2752–2776 (2011)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.