Diese Arbeit wurde vorgelegt am
Lehrstuhl für Mathematik (MathCCES)

A Higher-Order Finite Volume Method for Non-Conservative Equations on 2D Unstructured Quadrilateral Grids

CES Seminar
Computational Engineering Science
RWTH Aachen University

November 2016

Vorgelegt von
Presented by
Yannick Till Düren
Matrikelnummer: 323497

Betreuer
Supervisor
Julian Köllermeier
Lehrstuhl für Mathematik (MathCCES)
RWTH Aachen University

Contents

1 Introduction

This seminar work concerns the development of a second-order accurate numerical scheme for two-dimensional hyperbolic partial-differential equations on quadrilateral grids. In order to obtain numerical solutions with high accuracy one usually has two options: First, very fine grids can be used for the simulation. In case of an optimal scaling classical first-order accurate solver one has to divide the characteristic length of the mesh by two in order to double the accuracy of the obtained solution. Due to that extremly fine grids are required in order to achieve a demanded accuracy which leads to higher risk of numerical problems, high memory consumption and the need of very small time steps. Thus, memory or simulation time constraints may often be exceeded and violated using solution methods of first-order accuracy. Instead, the approach presented in this seminar work uses a numerical scheme with second-order accuracy. It can be easily extended to higher-order. This allows for higher solution accuracy on much coarser grids, which has the potential to save a lot of computation time.

Finally, we want to apply the method not only to problems in conservative form but also non-conservative systems such as shallow water equations and Quadrature-Based Moment Equations. Therefore, the numerical scheme should be able to handle systems of partial-differential eqution in non-conservative form.

On the basis of the first-order arrucate non-conservative PRICE-αC scheme, a second order extension will be presented in chapter 2 using a reconstruction of the solution from cell averages. This reconstruction may introduce spurious oscillations of the solution in non-smooth regions. Therefore, we present a solution limiting scheme (section 2.2.1) and integrate it into the method. The resulting second-order accurate scheme was combined with an existing finite volume solver framework. In chapter 3.1 we validate the solver and prove the claimed higher solution accuracy for smooth problems. Then the scheme is applied to non-smooth application problems in non-conservative form in chapter 3.2. Finally, we summarize the results in chapter 4 and discuss possibilities of future work on the project.

2 Numerics

This chapter concerns the PRICE-αC scheme, a numerical method for non-conservative hyperbolic systems of partial differential equations developed by Toro and Siviglia in [10]. This first-order scheme is then extended to second-order. However, the resulting scheme can easily be adapted to other non-conservative schemes. Since the final aim is to build a numerical method that is able to deal with complex geometries, the scheme is based on an unstructured mesh. We mainly follow the notation of Canestrelli et al., see [2, 3].

We consider a hyperbolic system of partial differential equations in two space dimensions

$$\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{A}_x(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial \mathbf{x}} + \mathbf{A}_y(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial \mathbf{y}} = \mathbf{0}, \tag{2.1}$$

$$\mathbf{Q} = \mathbf{Q}(\mathbf{x}, t), \quad (\mathbf{x}, t) \in \Omega \subset \mathbb{R}^2 \times [0, T], \quad \mathbf{Q} \in \Omega \subseteq \mathbb{R}^N,$$

where Ω is an open convex set, $\mathbf{Q} = (q_1, q_2, \ldots, q_N)^T$ is the vector of unknowns and $\mathbf{A}_x(\mathbf{Q}) \in \mathbb{R}^{N \times N}$ and $\mathbf{A}_y(\mathbf{Q}) \in \mathbb{R}^{N \times N}$ are the coefficient matrices, which are assumed to be smooth and locally bounded. If the system matrices $\mathbf{A}_x(\mathbf{Q})$ and $\mathbf{A}_y(\mathbf{Q})$ can be written as the Jacobi matrices of some flux functions $\mathbf{F}_x(\mathbf{Q})$, $\mathbf{F}_y(\mathbf{Q})$, such that $\mathbf{A}_x = \frac{\partial \mathbf{F}_x}{\partial x}$ and $\mathbf{A}_y = \frac{\partial \mathbf{F}_y}{\partial y}$, then the system of partial differential equations can be expressed in the conservative form

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}_x}{\partial x} + \frac{\partial \mathbf{F}_y}{\partial y} = \mathbf{0}. \tag{2.2}$$

2.1 First-order PRICE Scheme

Toro and Siviglia developed a series of *pri*mitive *ce*ntered (PRICE) numerical schemes for solving hyperbolic systems in the form (2.1) (see [10]). They are based on the FORCE scheme, a first-order centered scheme for conservation laws in the form (2.2). As proposed by Canestrelli et al. in [3] the most promising of these schemes, the so-called PRICE-αC scheme is used as the basis for the second-order centered numerical method used in this seminar.

Consider a conforming tessellation of the computational domain $\mathcal{T}_\Omega \subset \Omega$ by elements T_i such that

$$\mathcal{T}_\Omega = \bigcup_i T_i. \tag{2.3}$$

The $n_f = 4$ faces of each element T_i are given by ∂T_i^j where each face has length S_i^j and an associated outward normal $\mathbf{n}_i^j$. The volume $|T_i|$ of each element T_i is divided

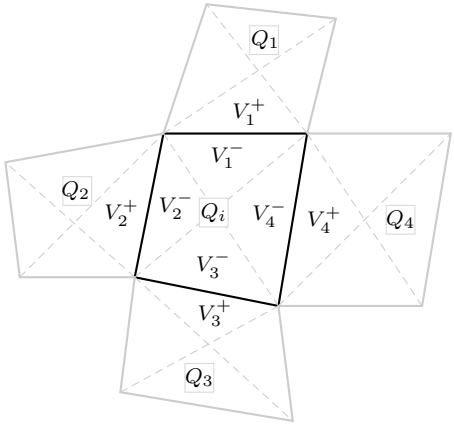


Figure 2.1: Notation of the general configuration on an unstructured two-dimensional quadrilateral mesh

into local sub-volumes V_j^- generated by connecting the barycenter of T_i with the two vertices of face j. The neighboring sub-volume adjacent to face ∂T_i^j is then locally denoted by V_j^+ as shown in figure 2.1.

Now the two-dimensional first-order PRICE-αC scheme can be given as the following one-step scheme:

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{|T_i|} \sum_{j=1}^{n_f} S_i^j \mathbf{A}_{ij}^- \left(\mathbf{Q}_j^n - \mathbf{Q}_i^n \right), \tag{2.4}$$

where

$$\mathbf{A}_{ij}^- = \frac{1}{2} \hat{\mathbf{A}}_{j+\frac{1}{2}} - \frac{V_j^+ V_j^-}{V_j^+ + V_j^-} \frac{1}{\Delta t S_i^j} \mathbf{I} - \frac{1}{4} \frac{\Delta t S_i^j}{V_j^+ + V_j^-} \hat{\mathbf{A}}_{j+\frac{1}{2}}^2. \tag{2.5}$$

Here $\mathbf{I}$ denotes the identity matrix and $V_j^\pm$ the area of the sub-volumes associated with cell T_i. $\hat{\mathbf{A}}_{j+\frac{1}{2}}$ is an approximation of the *Roe matrix* as defined in the following. For the complete derivation of this numerical scheme we refer to [2].

Given a family of paths $\mathbf{\Psi} = \mathbf{\Psi}(\mathbf{Q}_L, \mathbf{Q}_R, s, \mathbf{n})$ connecting $\mathbf{Q}_L$ and $\mathbf{Q}_R$, a matrix $\mathbf{A}_{\mathbf{\Psi}}(\mathbf{Q}_L, \mathbf{Q}_R, \mathbf{n})$ is a Roe matrix if it satisfies the following properties:

- for any $\mathbf{Q}_L, \mathbf{Q}_R \in \Omega$, $\mathbf{n} \in S^1$: $\mathbf{A}_{\mathbf{\Psi}}(\mathbf{Q}_L, \mathbf{Q}_R, \mathbf{n})$ has N real distinct eigenvalues;
- $\mathbf{A}_{\mathbf{\Psi}}(\mathbf{Q}, \mathbf{Q}, \mathbf{n}) = \mathbf{A}(\mathbf{Q}, \mathbf{n})$ for any $\mathbf{Q} \in \Omega$, $\mathbf{n} \in S^1$;
- for any $\mathbf{Q}_L, \mathbf{Q}_R \in \Omega$, $\mathbf{n} \in S^1$:

$$\mathbf{A}_{\mathbf{\Psi}}(\mathbf{Q}_L, \mathbf{Q}_R, \mathbf{n})(\mathbf{Q}_R - \mathbf{Q}_L) = \int_0^1 \mathbf{A}\left(\mathbf{\Psi}(s, \mathbf{Q}_L, \mathbf{Q}_R), \mathbf{n}\right) \frac{\partial \mathbf{\Psi}}{\partial s} ds, \tag{2.6}$$

where $\mathbf{A}(\mathbf{Q}, \mathbf{n}) = (\mathbf{A}_x(\mathbf{Q}), \mathbf{A}_y(\mathbf{Q}))^T \cdot \mathbf{n}$ is the directional system matrix in direction $\mathbf{n}$. In the conservative case, if $\mathbf{A}_x(\mathbf{Q}), \mathbf{A}_y(\mathbf{Q})$ are the Jacobi matrices of some flux

functions $\mathbf{F}_x(\mathbf{Q}), \mathbf{F}_y(\mathbf{Q})$ then 2.6 reduces to

$$\mathbf{A}_{\mathbf{\Psi}}(\mathbf{Q}_L, \mathbf{Q}_R, \mathbf{n})(\mathbf{Q}_R - \mathbf{Q}_L) = \mathbf{F}_{\mathbf{n}}(\mathbf{Q}_R) - \mathbf{F}_{\mathbf{n}}(\mathbf{Q}_L) \tag{2.7}$$

for any $\mathbf{n} \in S^1$ and $\mathbf{F}_{\mathbf{n}}(\mathbf{Q}) = (\mathbf{F}_x(\mathbf{Q}), \mathbf{F}_y(\mathbf{Q}))^T \cdot \mathbf{n}$.

A desirable choice for the matrix $\hat{\mathbf{A}}_{j+\frac{1}{2}}$ in equation (2.5) would be

$$\hat{\mathbf{A}}_{j+\frac{1}{2}} = \mathbf{A}_{\mathbf{\Psi}}\left(\mathbf{Q}_i^n, \mathbf{Q}_j^n, \mathbf{n}\right). \tag{2.8}$$

In this case the PRICE scheme reduces to a modified conservative FORCEα scheme if the system can be written in conservative form (2.2). Since the analytical computation of the integral in (2.6) can be very expensive or impossible, we do not want to compute the Roe matrix explicitly. An alternative approach is to use a simple segment path, e.g. a linear path

$$\mathbf{\Psi}(\mathbf{Q}_L, \mathbf{Q}_R, s) = \mathbf{Q}_L + s(\mathbf{Q}_R - \mathbf{Q}_L), \tag{2.9}$$

which yields the simplification

$$\mathbf{A}_{\mathbf{\Psi}}(\mathbf{Q}_L, \mathbf{Q}_R, \mathbf{n}) = \int_0^1 \mathbf{A}(\Psi(\mathbf{Q}_L, \mathbf{Q}_R, s), \mathbf{n})\, ds. \tag{2.10}$$

The *key idea* of the PRICE scheme as proposed in [2] is now to approximate the integral in (2.10) using a Gaussian quadrature technique. The resulting very accurate numerical approximation of the Roe matrix $\mathbf{A}_{\mathbf{\Psi}}$ is then given by the *centered Roe-type matrix:*

$$\mathbf{A}_{\mathbf{\Psi}}^M(\mathbf{Q}_L, \mathbf{Q}_R, \mathbf{n}) = \sum_{j=1}^{M} \omega_i \mathbf{A}(\mathbf{\Psi}(\mathbf{Q}_L, \mathbf{Q}_R, s_j), \mathbf{n}) \tag{2.11}$$

using an M-point Gaussian quadrature rule with quadrature weights ω_j and quadrature nodes s_j. Finally, we choose

$$\hat{\mathbf{A}}_{j+\frac{1}{2}} = \mathbf{A}_{\mathbf{\Psi}}^M\left(\mathbf{Q}_i^n, \mathbf{Q}_j^n, \mathbf{n}\right), \tag{2.12}$$

to avoid the explicit computation of the Roe averages.

2.2 Second-Order Extension

The first step for the extension of the first-order accurate numerical scheme is the construction of an appropriate local polynomial representation of the solution on each computational cell. Due to local extrema of this reconstruction, a limitation of the slopes becomes necessary in order to avoid over- and undershoots of the numerical solution. A number of solution reconstruction and slope limiting techniques are summarized by Hubbard in [6]. Again, we will closely follow the notation by Canestrelli given in [2].

In order to achieve second-order accuracy in space it is sufficient to consider a linear reconstruction profile $\mathbf{w}_i(\mathbf{x}, t^n)$ for each cell [6], i.e. a plane of the form:

$$\mathbf{w}_i(\mathbf{x}, t^n) = (x - x_i)\,\mathbf{a}_i + (y - y_i)\,\mathbf{b}_i + \mathbf{c}_i \tag{2.13}$$

where (x_i, y_i) are the barycenter coordinates of the ith cell and $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i \in \mathbb{R}^N$ are the parameters of $\mathbf{w}_i$. This reconstruction profile over the cell i should satisfy the following constraints:

1. The cell average value should be preserved:

$$\frac{1}{|T_i|}\int_{T_i} \mathbf{w}_i(\mathbf{x}, t^n)\,d\mathbf{x} = \mathbf{Q}_i^n; \tag{2.14}$$

2. The average of the reconstructed value of $\mathbf{w}_i(\mathbf{x}, t^n)$ in jth adjacent cell to the ith cell should be as close as possible to the adjacent cell average $\mathbf{Q}_j^n$

$$J(\mathbf{w}_i(\mathbf{x}, t^n)) \coloneqq \frac{1}{|T_j|}\int_{T_j} \mathbf{w}_i(\mathbf{x}, t^n)\,d\mathbf{x} \approx \mathbf{Q}_j^n, \qquad \forall j \in \mathcal{N}_i \tag{2.15}$$

   where $\mathcal{N}_i$ denotes the set of cells adjacent to cell i.

With these $|\mathcal{N}_i| + 1$ constraints, i.e. 5 conditions for quadrilateral cells, we are able to determine the coefficients $\mathbf{a}_i$, $\mathbf{b}_i$ and $\mathbf{c}_i$ in (2.13). Note that, substituting (2.13) into (2.14) directly leads to

$$\mathbf{c}_i = \mathbf{Q}_i^n.$$

Since the number of conditions is greater than the number of coefficients, not all these conditions can be satisfied simultaneously in most cases. Therefore, the slopes $\mathbf{a}_i$, $\mathbf{b}_i$ of the linear profile are determined by minimizing the square distance between reconstructed values and cell averages, and so minimizing the total residual function

$$\tilde{\mathbf{R}}_i(\mathbf{a}_i, \mathbf{b}_i) = \sum_{j \in \mathcal{N}_i} \left[J(\mathbf{w}_i(\mathbf{x}_G, t^n)) - Q_j^n\right]^2. \tag{2.16}$$

This least squares procedure can be extended with a weighting of the distances:

$$\mathbf{R}_i(\mathbf{a}_i, \mathbf{b}_i) = \sum_{j \in \mathcal{N}_i} \left[\beta_j \left(J(\mathbf{w}_i(\mathbf{x}_G, t^n)) - Q_j^n\right)\right]^2, \tag{2.17}$$

where β_j are the used weights. These can be for example depending on the distance of the barycenter of the adjacent cell j to the barycenter of the ith cell.

Due to local extrema of this reconstruction, over- and undershoots of the numerical solution may appear resulting in oscillations and instabilities in non-smooth regions. Thus, a slope limiting procedure is needed.

2.2.1 Limiter

We apply a TVD slope limiter to the reconstruction (2.13) according to [2, 9] by limiting the slopes with limitation factors $\gamma_1, \ldots, \gamma_N$ such that for $\Gamma = diag(\gamma_1, \ldots, \gamma_N)$

$$\hat{\mathbf{a}}_i = \Gamma \mathbf{a}_i, \quad \hat{\mathbf{b}}_i = \Gamma \mathbf{b}_i,$$

with the resulting limited reconstruction profile

$$\hat{\mathbf{w}}_i(\mathbf{x}, t^n) = \mathbf{Q}_i^n + (x - x_i)\,\hat{\mathbf{a}}_i + (y - y_i)\,\hat{\mathbf{b}}_i. \tag{2.18}$$

Note that each limitation factor γ_i associated to the cell state q_i^n which is a single entry in $\mathbf{Q}_i^n$ is uniformly applied to the slopes in both directions in order to leave the direction of the slope vector unchanged. Following [6, 2] the limitation factors γ_i for any entry q_i^n of $\mathbf{Q}_i^n$ are determined by

$$\gamma_i = \min_{j \in \mathcal{N}_i} (\gamma_{i,j}),$$

where

$$\gamma_{i,j} = \begin{cases} \Phi\left(\frac{q_{i,j}^{\min} - q_i^n}{w_i(\mathbf{x}_{ij}, t^n) - q_i^n}\right) & \text{,if} \quad w_i(\mathbf{x}_{ij}, t^n) < q_i^n \\ \Phi\left(\frac{q_{i,j}^{\max} - q_i^n}{w_i(\mathbf{x}_{ij}, t^n) - q_i^n}\right) & \text{,if} \quad w_i(\mathbf{x}_{ij}, t^n) > q_i^n \\ 1 & \text{,if} \quad w_i(\mathbf{x}_{ij}, t^n) = q_i^n, \end{cases} \tag{2.19}$$

with $q_{i,j}^{\min} = \min\left(q_i^n, q_j^n\right)$, $q_{i,j}^{\max}\left(q_i^n, q_j^n\right)$ and $w_i(\mathbf{x}_{ij}, t^n)$ the value of the reconstruction polynomial at the midpoint of the interface $\mathbf{x}_{ij} = (x_{ij}, y_{ij})$ of the ith and jth cell. Φ denotes a classical limiter function. A wide variety of limiter functions have been studied. Some standard limiters used here are:

- *Sweby* limiter
$$\Phi(\theta) = \max[0, \min(\beta\theta, 1), \min(\theta, 1)],$$
for $\beta \in [1, 2]$. Famous examples are obtained for $\beta = 1$ (*Minmod*) and $\beta = 2$ (*Superbee*)

- *Monotonized Central (MC)* limiter
$$\Phi(\theta) = \max\left[0, \min\left(2\theta, \frac{1+\theta}{2}, 2\right)\right].$$

- *van Leer* limiter
$$\Phi(\theta) = \frac{\theta + |\theta|}{1 + |\theta|}.$$

Further classical limiters can be found in [6]. Moreover, we want to mention the limiting scheme proposed by Buffard and Clain (see [1]) which uses additional geometry information for the calculation of γ_i. This limiting scheme is used by Stecca for the higher-order extension of a modified PRICE scheme called UPRICE [9], but will not be further discussed here.

2.2.2 Reconstruction in Time

After the reconstruction of the solution by a spatial polynomial $\mathbf{w}_i(\mathbf{x}, t^n)$ for each cell T_i at time t^n, we still need to compute the evolution of this polynomial in time. To be able to increase the order of accuracy in $T_i \times [t^n, t^{n+1}]$ we use the ADER approach of Toro and Titarev [11].

The local solution $\mathbf{Q}_i(\mathbf{x}, t)$ of the system (2.1) in each cell is expanded by a space-time Taylor series with respect to the element barycenter $\mathbf{x}_G$ and the time t^n

$$\mathbf{Q}_i(\mathbf{x}, t) = \mathbf{Q}_i^n + (x - x_G)\frac{\partial \mathbf{Q}_i}{\partial x} + (y - y_G)\frac{\partial \mathbf{Q}_i}{\partial y} + (t - t^n)\frac{\partial \mathbf{Q}_i}{\partial t} + \dots \tag{2.20}$$

where the first order terms of the Taylor expansion are sufficient to achieve a second-order accurate numerical scheme. All derivatives in (2.20) are evaluated at $(\mathbf{x}_G, t^n)$. Now using classical Cauchy-Kovalewski procedure with the PDE system (2.1) we get

$$\frac{\partial \mathbf{Q}_i}{\partial t} = -\left(\mathbf{A}_x(\mathbf{Q}_i^n)\frac{\partial \mathbf{Q}_i}{\partial x} + \mathbf{A}_y(\mathbf{Q}_i^n)\frac{\partial \mathbf{Q}_i}{\partial y}\right),$$

in order to substitute the time derivative. This yields

$$\begin{aligned}\mathbf{Q}_i(\mathbf{x}, t) &= \mathbf{Q}_i^n + (x - x_G)\frac{\partial \mathbf{Q}_i}{\partial x} + (y - y_G)\frac{\partial \mathbf{Q}_i}{\partial y} \\ &\quad - (t - t^n)\left(\mathbf{A}_x(\mathbf{Q}_i^n)\frac{\partial \mathbf{Q}_i}{\partial x} + \mathbf{A}_y(\mathbf{Q}_i^n)\frac{\partial \mathbf{Q}_i}{\partial y}\right) \\ &= \mathbf{Q}_i^n + \left((x - x_G)\mathbf{I} - (t - t^n)\mathbf{A}_x(\mathbf{Q}_i^n)\right)\hat{\mathbf{a}}_i \\ &\quad + \left((y - y_G)\mathbf{I} - (t - t^n)\mathbf{A}_y(\mathbf{Q}_i^n)\right)\hat{\mathbf{b}}_i,\end{aligned} \tag{2.21}$$

where we obtained the spatial derivatives from the limited reconstruction polynomial (2.18).

2.2.3 Second-Order PRICE Scheme

Using the reconstruction techniques in space and time one can obtain the final second-order accurate one-step PRICE scheme by integration of (2.1) over a space-time control volume $T_i \times [t^n, t^{n+1}]$:

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{1}{|T_i|}\mathbf{A} \cdot \nabla \mathbf{Q} - \frac{\Delta t}{|T_i|}\sum_{j=1}^{n_f} \mathbf{D}^-_{j+\frac{1}{2}}, \tag{2.22}$$

where

$$\mathbf{A} \cdot \nabla \mathbf{Q} = \int_{t^n}^{t^{n+1}} \int_{T_i} \mathbf{A}_x(\mathbf{Q}_i(\mathbf{x}, t))\frac{\partial \mathbf{Q}_i}{\partial x} + \mathbf{A}_y(\mathbf{Q}_i(\mathbf{x}, t))\frac{\partial \mathbf{Q}_i}{\partial y} d\mathbf{x} dt, \tag{2.23}$$

and

$$\mathbf{D}^{-}_{j+\frac{1}{2}} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \int_{S_{ij}} \mathbf{A}^{-}_{ij} \left(\mathbf{Q}_i \left(\mathbf{x}, t \right) \right) \left(\mathbf{Q}_{jS} - \mathbf{Q}_{iS} \right) ds dt, \tag{2.24}$$

with

$$\mathbf{Q}_{jS} = \mathbf{Q}_j \left(\mathbf{x}_{S_{ij}}, t \right) \quad , \quad \mathbf{Q}_{iS} = \mathbf{Q}_i \left(\mathbf{x}_{S_{ij}}, t \right). \tag{2.25}$$

Here $\mathbf{A}^{-}_{ij} \left(\mathbf{Q} \left(\mathbf{x}, t \right) \right)$ again denotes the PRICE flux matrix from (2.5). The edge adjacent to both the ith and the jth cell is given by $S_{ij} = \partial T_i \cap \partial T_j$. $\mathbf{x}_{S_{ij}}$ is the vector that defines this edge. The integrals in (2.23) and (2.24) are approximated by Gaussian quadrature. For second-order accuracy at least one integration point in time and three integration points per space-dimension are necessary. Note that for a first-order scheme the $\mathbf{A} \cdot \nabla \mathbf{Q}$ term vanishes, since $\nabla \mathbf{Q} \left(\mathbf{x}, t \right) = 0$.

Moreover, (2.22) gives a second-order accurate one-step update formula for other non-conservative finite volume schemes by substitution of the flux matrix $\mathbf{A}^{-}_{ij} \left(\mathbf{Q} \left(\mathbf{x}, t \right) \right)$ with the corresponding flux matrix of the new scheme. Examples for such schemes are the non-conservative Lax-Friedrichs or the UPRICE scheme [9].

3 Results

3.1 Convergence Study

In this chapter we compute the order of accuracy of the presented second-order PRICE scheme numerically to verify that the expected theoretical order is achieved. The numerical method has been integrated into an existing finite-volume solver developed by Koellermeier, that was also used for [7]. All results and timings presented in this chapter have been generated with this package. The flux matrix (2.12) of the PRICE scheme has been computed using three Gaussian integration points on the interval $[0, 1]$

$$s_1 = \frac{1}{2}, \quad s_{2,3} = \frac{1}{2} \pm \frac{\sqrt{15}}{10}, \qquad \omega_1 = \frac{8}{18}, \quad \omega_{2,3} = \frac{5}{18}, \tag{3.1}$$

as proposed by Canestrelli [2].

3.1.1 Linear Transport Equation

First, we consider the scalar linear transport equation given by

$$\frac{\partial u}{\partial t} + a_x \frac{\partial u}{\partial x} + a_y \frac{\partial u}{\partial y} = 0 \tag{3.2}$$

where a_x and a_y denote the transport speeds in x- and y–direction respectively. For this simple linear and conservative PDE a solution can be given by

$$u(\mathbf{x}, t) = u_0(\mathbf{x} - \mathbf{a}t) \tag{3.3}$$

with $u_0(x)$ the initial condition and $\mathbf{a} = (a_x, a_y)^T$.

We solve the equation for $\mathbf{a} = (1, 1)^T$ in a domain $\Omega = [0, 1]^2$ using smooth initial conditions and periodical boundary conditions prescribed at the four boundaries. We used a sequence of regular equidistant quadrilateral meshes in two dimensions where the number of cells N per space dimension is doubled for each refinement of the mesh. The time step size Δt was chosen such that the CFL number is 0.5 for all meshes.

Figure 3.1 shows the relative L^1-error of the PRICE scheme after $T = 1$. As expected, we can observe first- and second-order accuracy for the first-order and unlimited second-order versions of the PRICE scheme respectively. For $N \geq 40$ all limited schemes (Minmod, Superbee, MC and vanLeer limiter) approximately show a second-order behavior, where the Minmod limiter has the smallest relative L^1-error of the limited schemes.

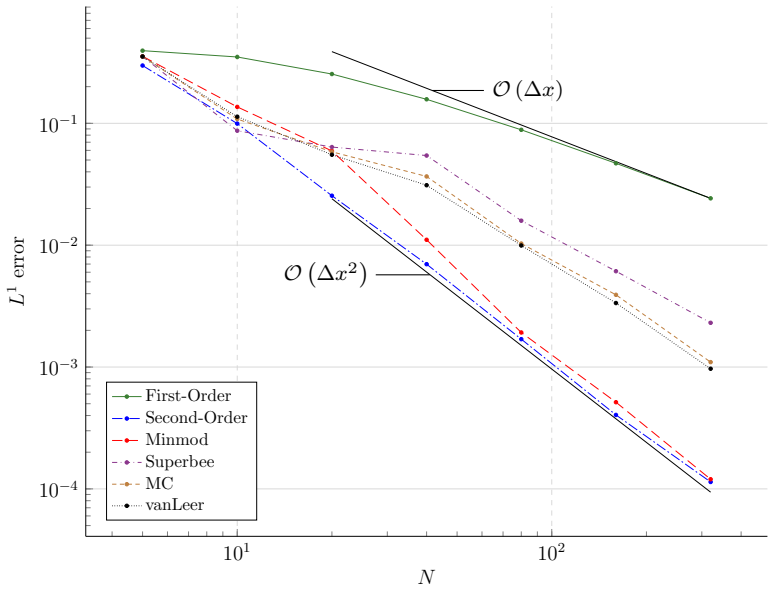


Figure 3.1: Relative L^1-error for linear transport equations test case

3.1.2 Inviscid Shallow Water Equations

We validated the numerical scheme and its second-order accuracy for a linear smooth test case in 2.2.3. The next step is the validation of a non-linear, non-conservative smooth test problem. We consider the 2D inviscid shallow water equations coupled with a bottom evolution equation which can be written in non-conservative form (2.1) with

$$\mathbf{A}_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{q_x^2}{h^2} + gh & \frac{2q_x}{h} & 0 & gh \\ -\frac{q_x q_y}{h^2} & \frac{q_y}{h} & \frac{q_x}{h} & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \mathbf{A}_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -\frac{q_x q_y}{h^2} & \frac{q_y}{h} & \frac{q_x}{h} & 0 \\ -\frac{q_y^2}{h^2} + gh & 0 & \frac{2q_y}{h} & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{3.4}$$

$$\mathbf{Q} = (h, q_x, q_y, b)^T .$$

In order to validate the order of accuracy an exact smooth solution is constructed (see [3]) which reads

$$h(x,y,t) = h_0 + x_0 \sin(kx - \omega t), \tag{3.5}$$
$$q_x(x,y,t) = \frac{\omega}{k} h_0 + c_0 \frac{\omega}{k} \sin(kx - \omega t), \tag{3.6}$$
$$q_y(x,y,t) = 0, \tag{3.7}$$
$$b(x,y,t) = -h(x,y,t), \tag{3.8}$$

and is essentially a 1D test case. Here $k = \frac{2\pi}{L}$ and $\omega = \frac{2\pi}{T_0}$, L and T_0 are the wave length and the period of the sinusoidal oscillation, respectively. The system is solved

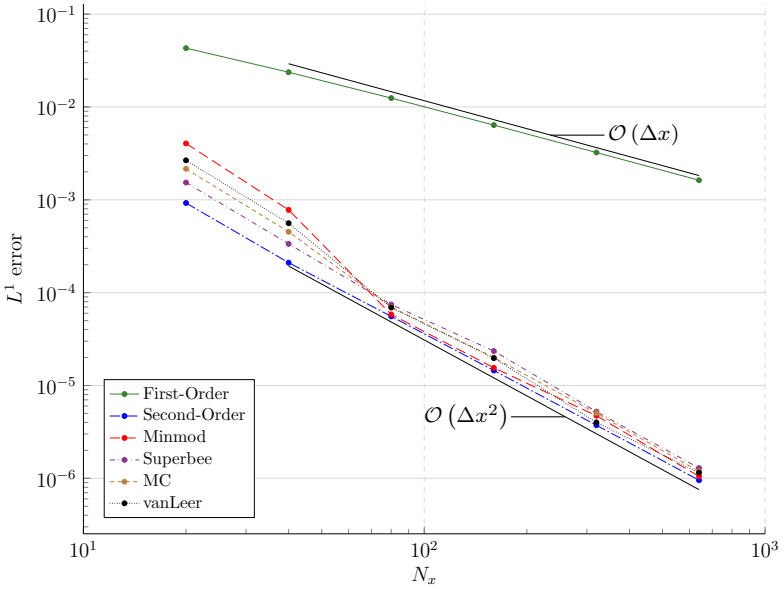


Figure 3.2: L^1-error for inviscid shallow water equations test case

in a domain $\Omega = [0, 800] \times [0, 200]$ using parameters $h_0 = 1$, $c_0 = 0.2$, $T_0 = 100$ and $L = 800$ and the relative error of the numerical solution is computed at time $T = 100s$.

Figure 3.2 depicts the relative L^1-error of the simulation for varying numbers of cells N_x in x-direction for constant CFL number of 0.6. The number of cells in y-direction can be held constant since problem (3.5)-(3.8) is a 1D test case. As for the linear transport test case one can observe near-perfect second-order accuracy for the second-order method 2.2.3 with all tested limiters. Our numerical convergence study also directly allows us to assess the computational efficiency of the second-order scheme in comparison to the first-order PRICE scheme. Table 3.1 lists the total computation time, the average computation time per time step and the relative L^1-error for the first-order and the second-order PRICE method with minmod limiter. All computations were run on a single thread on an Intel i7-2600k (3.4Ghz clock speed) processor. We can observe that the first-order scheme is about 5 to 7 times faster than the second-order scheme. Nevertheless, we can clearly conclude that the second-order scheme is more efficient to use on coarse meshes than low order methods on fine grids (for smooth solutions). For example the computation on the $N_x = 320$ mesh using the first-order scheme took a total of $4.61s$ to reach an error of $0.32 \cdot 10^{-2}$, whereas the second-order scheme on a very coarse grid with $N_x = 20$ reaches an error in the same order of magnitude ($0.40 \cdot 10^{-2}$) in only about $0.12s$.

N_x	1st-order			2nd-order		
	t_{CPU} [s]	$\frac{t_{CPU}}{\text{step}}$ [ms]	L^1-error	t_{CPU} [s]	$\frac{t_{CPU}}{\text{step}}$ [ms]	L^1-error
20	0.019	0.56	$0.43E-01$	0.12	3.75	$0.40E-02$
40	0.062	0.93	$0.24E-01$	0.43	6.44	$0.78E-03$
80	0.268	2.01	$0.12E-01$	1.62	12.15	$0.58E-04$
160	0.939	3.52	$0.65E-02$	6.48	24.28	$0.16E-04$
320	4.608	8.64	$0.32E-02$	25.48	47.78	$0.47E-05$
640	17.225	16.14	$0.16E-02$	107.82	101.08	$0.11E-05$

Table 3.1: Average computation time and relative error for the inviscid shallow water test case.

3.2 Application: Non-Smooth Problems

In this section we apply the scheme (2.22) to two non-smooth non-conservative application problems, namely the circular dam-break problem and the shock tube problem. Again, the centered Roe-type matrix (2.11) is evaluated using three Gaussian quadrature points (3.1) for all computations.

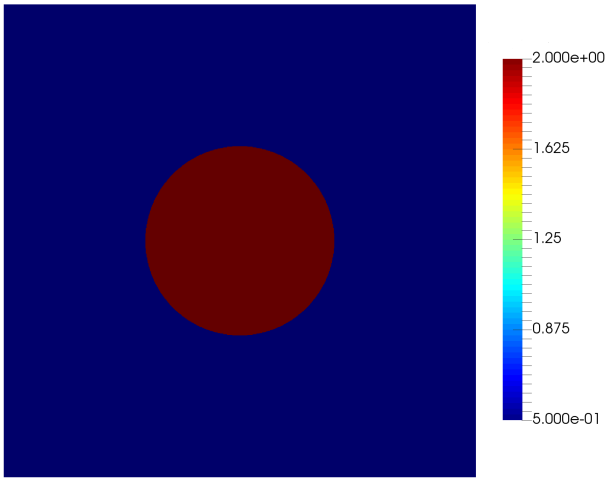
3.2.1 Circular Dam-Break Problem

In this test case we consider a cylindrical tank of diameter $20m$ initially filled with $2m$ water surrounded by water of depth $0.5m$ (see figure 3.3a). At time $t_0 = 0s$ the tank instantaneously breaks generating a shock wave resulting in an increase of water depth in the lower depth region propagated in radial direction [4]. The test can be useful to check the ability of the method to preserve cylindrical symmetry. The physics are modeled using shallow water equations with a fixed bed defined in the form (2.1) for $\mathbf{Q} = (H, q_x, q_y, b)^T$ by

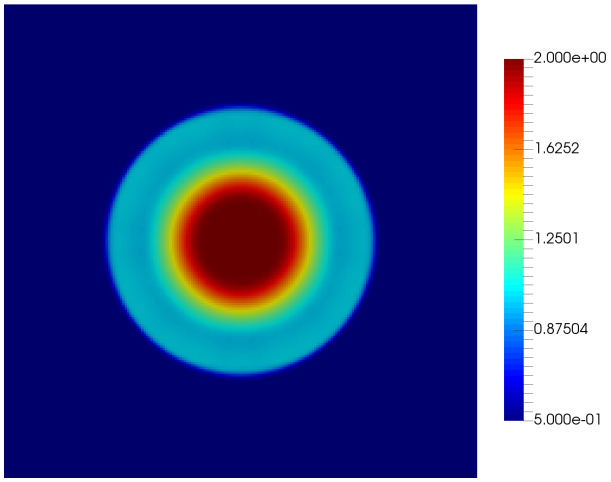
$$\mathbf{A}_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ g\left(H-b\right) - \frac{q_x^2}{(H-b)^2} & \frac{2q_x}{H-b} & 0 & \frac{q_x^2}{(H-b)^2} \\ -\frac{q_x q_y}{(H-b)^2} & \frac{q_y}{H-b} & \frac{q_y}{H-b} & \frac{q_x q_y}{(H-b)^2} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.9}$$

$$\mathbf{A}_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -\frac{q_x q_y}{(H-b)^2} & \frac{q_y}{H-b} & \frac{q_x}{H-b} & \frac{q_x q_y}{(H-b)^2} \\ g\left(H-b\right) - \frac{q_y^2}{(H-b)^2} & 0 & \frac{2q_y}{H-b} & \frac{q_y^2}{(H-b)^2} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.10}$$

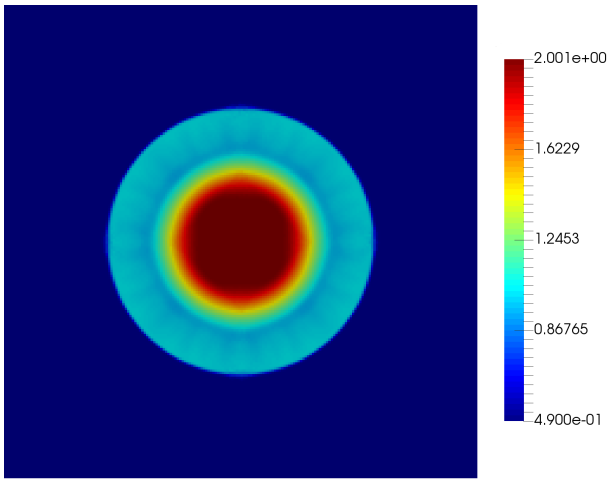
Figures 3.3b and 3.3c show the results of the simulation after $T = 1s$ on a structured quadrilateral mesh with $40,000$ elements. We observe that the method is able to preserve the radial symmetry of the problem. Similar results can be achieved using unstructured grids, compare figure 3.5.



(a) Initial Configuration at $t_0 = 0s$

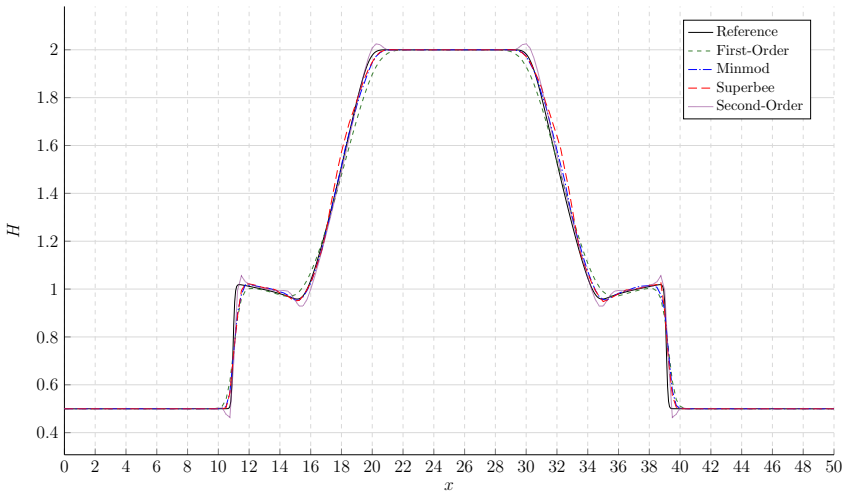


(b) Solution at $T = 1s$ using the first-order method

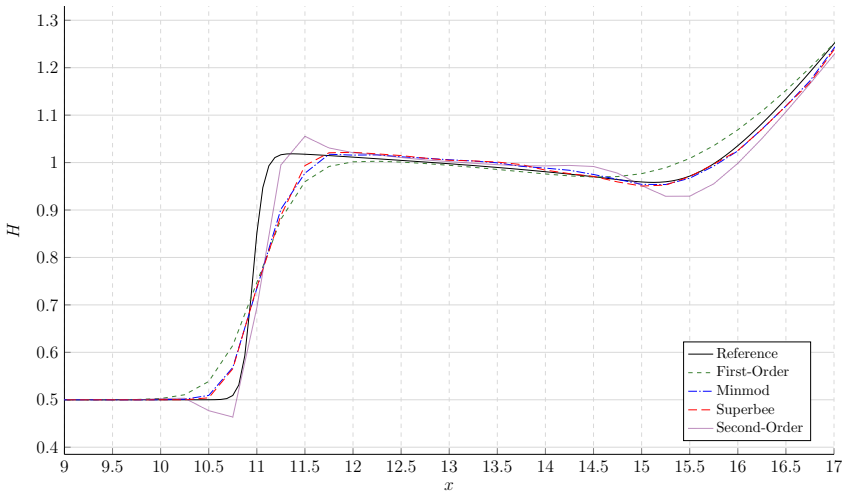


(c) Solution at $T = 1s$ using the second-order method with minmod limiter

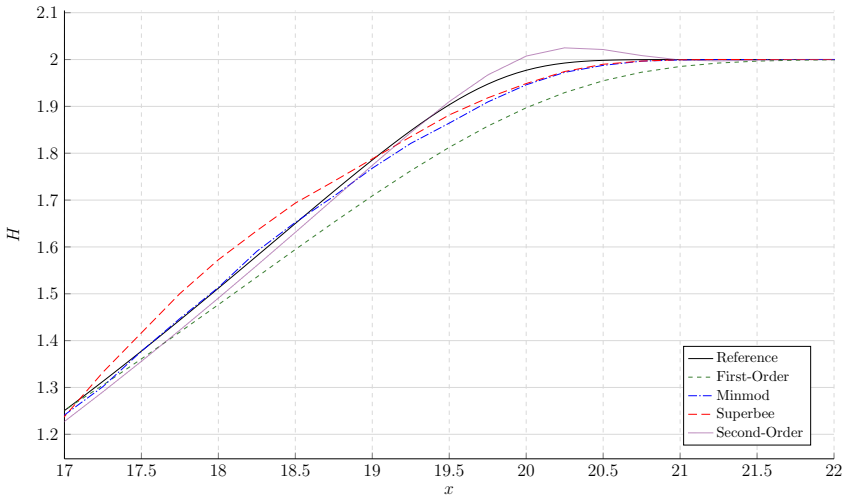
Figure 3.3: Initial condition and results of the dam-break problem after $T = 1s$



(a) Whole domain $[0, 50]$



(b) Cutout for interval $[9, 17]$

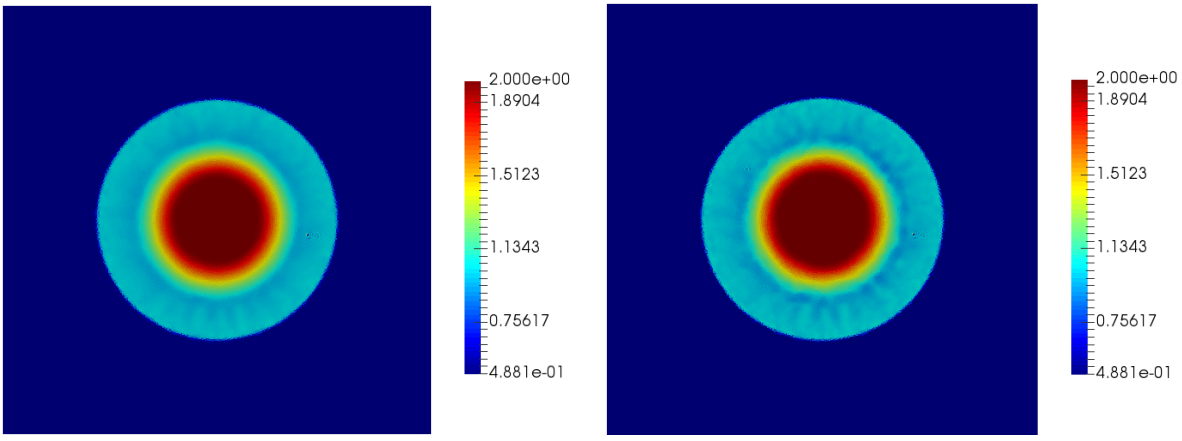


(c) Cutout for interval $[17, 22]$

Figure 3.4: 1D slice of the solution of the dam-break problem at $T = 1s$

N	1st-order		2nd-order	
	t_{CPU} [s]	$\frac{t_{CPU}}{\text{step}}$ [s]	t_{CPU} [s]	$\frac{t_{CPU}}{\text{step}}$ [s]
625	0.04	0.009	0.23	0.046
2,500	0.28	0.028	1.50	0.150
10,000	2.18	0.109	13.07	0.654
40,000	16.59	0.415	105.43	2.636
160,000	123.20	1.540	842.97	10.537

Table 3.2: Average computation time per time step for the dam-break test case.



(a) Minmod (b) Superbee

Figure 3.5: Dam-break problem on unstructured grid with 50,970 cells after $T = 1s$

Now consider a 1D slice in any direction through the center of the domain. These slices should be identical due to the radial symmetry. A slice in x-direction is shown in figure 3.4. The shown reference solution was obtained using the first-order PRICE method on a much finer grid. We observe that the first-order scheme is the most diffusive, while the second-order scheme without any limiter shows the least numerical diffusion. As shortly discussed in section 2.2, we get over- and undershoots of the solution at discontinuities if the slopes are not limited. Now applying limitation with the minmod or superbee limiters yields stable solutions with much higher accuracy without any over- or undershoots.

Similar to the inviscid shallow water test case, we see very good first-order scaling of the simulation runtime with the number of cells as the CFL number is constant, see table 3.2. In addition, one can observe good agreement of the numerical solution and a pseudo-analytical solution of the problem in 1D which is presented in [2] but will not be further discussed here.

3.2.2 Quadrature-Based Moment Equations

Finally, we consider the so-called Quadrature-Based Moment Equations (QBME) which result from a transformation and model reduction of the Boltzmann transport equation [8]. The QBME model in 1D for 5 moments is a non-conservative system of partially-conservative PDEs with a source term and can be given as follows:

$$\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{A}_x \frac{\partial \mathbf{Q}}{\partial x} = -\frac{1}{\tau}\mathbf{B}\mathbf{Q}, \quad \mathbf{Q} = (\rho, u, \theta, f_3, f_4)^T,$$

$$\mathbf{A}_x = \begin{pmatrix} u & \rho & 0 & 0 & 0 \\ \frac{\theta}{\rho} & u & 1 & 0 & 0 \\ 0 & 2\theta & u & \frac{6}{\rho} & 0 \\ 0 & 4f_4 & \frac{\rho\theta}{2} - \frac{10f_4}{\theta} & u & 4 \\ -\frac{f_3\theta}{\rho} & 5f_4 & -f_3 & \theta + \frac{15f_4}{\rho}\theta & u \end{pmatrix}.$$

For a detailed derivation see [5]. Here ρ is the density, u is the velocity, θ is the temperature and f_3 and f_4 are related to higher moments of the system. The relaxation time τ is defined by the Knudsen number Kn of the fluid and given by $\tau = \frac{\mathrm{Kn}}{\rho}$.

We consider a one-dimensional shock tube problem on a domain $[-2, 2]$ with initial data

$$\mathbf{Q}(x, t = 0) = \begin{cases} (7, 0, 1, 0, 0)^T & \text{if } x < 0, \\ (1, 0, 1, 0, 0)^T & \text{if } x > 0, \end{cases}$$

and Knudsen number $\mathrm{Kn} = 0.5$. Figure 3.6 shows a comparison of the results for a simulation up to $T = 0.3s$ generated with the first-order PRICE scheme (2.4) and the second-order PRICE scheme (2.22) with the minmod limiter. Again, we observe that the second-order scheme is less diffusive and the limiter prevents most over- and undershoots of the solution. Note that we can observe some oscillations in the velocity

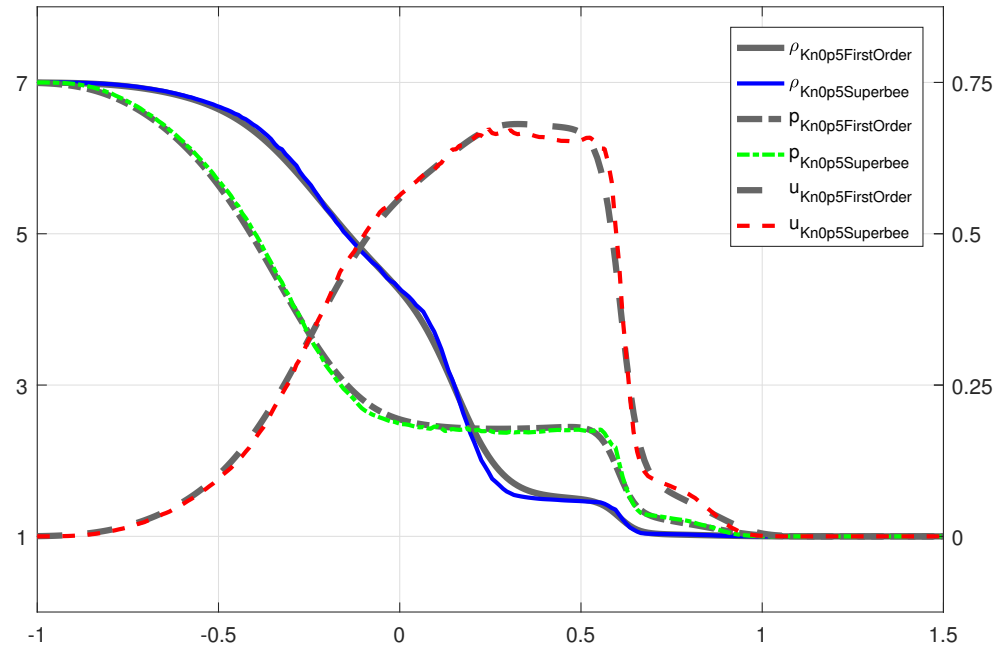


Figure 3.6: Shock tube results after $T = 0.3$

u right before the strong shock for the results of the second-order scheme. These were not sufficiently damped by the limiter but they do not get amplified for $t > 0.3s$.

4 Conclusion

4.1 Summary

After the motivation we have presented the PRICE-αC method which is a first-order monotone centered scheme for non-conservative PDE systems [3]. This scheme was extended to second-order using a polynomial reconstruction and the ADER approach [6, 2]. Slope limitation of the reconstruction became necessary at non-smooth solution regions in order to suppress spurious oscillations [1]. The presented second-order extension can easily be generalized to higher order.

Analyzing the convergence of the second-order scheme we validated its second-order behavior using linear transport and inviscid shallow water equations. We have observed that we can achieve much higher accuracy than with the first-order method. Comparing both schemes we saw that the second-order scheme was also superior in terms of runtime: We achieved a comparable relative error in much less computation time using the extended version of the PRICE scheme which coincides with the results of Canestrelli et al. [3].

Application of the PRICE scheme with second-order extension to a circular dam-break problem using shallow water equations with fixed bed yielded that the scheme preserved the cylindrical symmetry of the initial conditions. We saw the expected over- and undershoots of the numerical solution if no slope limitation was applied. These could be avoided using limiters as introduced in chapter 2.2.1. Finally, we applied the method to a non-conservative PDE system with source term. The Quadrature-Based Moment Equations [5, 8] were used for the physical modeling of a shock tube problem. Again, the second-order scheme showed better resolution of shocks and rarefaction waves and less numerical diffusion.

4.2 Further Work

The work on this topic will be continued by applications of the extended scheme to test and real-world problems. Mainly the QBME and similar systems such as the HME or Grad's (see [5]) system will be considered and analyzed in depth. Additionally, other limiters should be integrated into the solver. This could be for example the limiter presented by Stecca [9]. Moreover, further tests of the schemes for fully unstructured grids will be conducted. As the matter stands, the scheme has only been tested on unstructured grids in case of the dam-break problem. Finally, some runtime optimizations should be performed for our implementation of the scheme.

Bibliography

[1] T. Buffard and S. Clain. Monoslope and multislope muscl methods for unstructured meshes. *Journal of Comptutational Physics*, 229:745–3776, 2010.

[2] A. Canestrelli. *Numerical Modelling of Alluvial Rivers By Shock Capturing Methods*. Phd thesis, Universita' Degli Studi Di Padova, 2008.

[3] A. Canestrelli, A. Siviglia, M. Dumbser, and E. F. Toro. Well-balanced high-order centred schemes for non-conservative hyperbolic systems. applications to shallow water equations with fixed and mobile bed. *Advances in Water Resources*, 32(6): 834 – 844, 2009.

[4] A. Canestrelli, M. Dumbser, A. Siviglia, and E.F. Toro. Well-balanced high-order centered schemes on unstructured meshes for shallow water equations with fixed and mobile bed. *Advances in Water Resources*, 33(3):291 – 303, 2010.

[5] Y. Fan, J. Koellermeier, J. Li, R. Li, and M. Torrilhon. Model reduction of kinetic equations by operator projection. *J. Stat. Phys.*, 162(2):457–486, 2016.

[6] M. E. Hubbard. Multidimensional Slope Limiters for MUSCL-Type Finite Volume Schemes on Unstructured Grids. *Journal of Computational Physics*, 155:54–74, 1999.

[7] J. Koellermeier and M. Torrilhon. Numerical study of partially conservative moment equations in kinetic theory. *accepted in Commun. in Comput. Phys.*

[8] J. Koellermeier, R. P. Schaerer, and M. Torrilhon. A framework for hyperbolic approximation of kinetic equations using quadrature-based projection methods. *Kinet. Relat. Mod.*, 5(3):531–549, 2014.

[9] G. Stecca. *Numerical modelling of gravel-bed river morphodynamics*. Phd thesis, University of Trento, 2012.

[10] E. F. Toro and A. Siviglia. PRICE: primitive centred schemes for hyperbolic systems. *International Journal for Numerical Methods in Fluids*, 42(12):1263–1291, 2003.

[11] E. F. Toro and V. A. Titarev. Solution of the generalized riemann problem for advection–reaction equations. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 458(2018):271–281, 2002.