An explicit MUSCL scheme on staggered grids with kinetic–like fluxes for the barotropic and full Euler system

Thierry Goudon, Julie Llobell, Sebastian Minjeaud

Université Côte d'Azur, Inria, CNRS, LJAD

Abstract. We present a second order scheme for the barotropic and full Euler equations. The scheme works on staggered grids, with numerical unknowns stored at dual locations, while the numerical fluxes are derived in the spirit of kinetic schemes. We identify stability conditions ensuring the positivity of the discrete density and energy. We illustrate the ability of the scheme to capture the structure of complex flows with 1D and 2D simulations on MAC grids.

MSC. 65M08, 76M12, 76NXX, 35Q31.

Contents

1 Introduction

This work is concerned with the development of numerical schemes on staggered grids for the Euler equations. Using staggered grids is non standard for the discretization of hyperbolic system (see e.g. [11, 19, 25, 35, 49]), since, when stored on a collocated grid, the unknowns of the system are usually gathered in a single vector-valued unknown allowing to identify the wave structure of the system in order to built upwinding techniques. The motivation of the use of staggered grids comes from the attempt to have an unified approach with an incompressible code, see e.g. [24, 27, 50, 52, 53, 54]. This is particularly relevant when dealing with low-Mach simulations since letting the Mach number go to 0 enforces incompressibility and collocated approaches might lead to numerical difficulties in this regime, and to the development of spurious instabilities due to an "odd-even decoupling", see [20, 21, 56] and the references therein. This is also of interest in multifluid flows simulations that involve additional solenoidal constraints on a velocity field, see e.g. [6, 16, 17, 46]: coupled with a projection approach, the staggered method makes the discretization of the mass conservation equations for all the species interacting in the mixture and the definition of the pressure field (the Lagrange multiplier associated to the solenoidal constraint) compatible.

We first deal with the barotropic Euler system

$$\begin{cases} \partial_t \rho + \boldsymbol{\nabla} \cdot (\rho \mathbf{u}) = 0, \\ \partial_t (\rho \mathbf{u}) + \boldsymbol{\nabla} \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \boldsymbol{\nabla} (p(\rho)) = 0. \end{cases} \tag{1}$$

This model describes the evolution of a compressible fluid (in the absence of external forces). The unknowns ρ and $\mathbf{u}$ stand respectively for the local density and velocity field of the fluid. They depend on the time and space variables, $t \geqslant 0$ and $x \in \mathbb{R}^N$. The model assumes that the pressure p depends on the density ρ only. Here and below, we suppose that the pressure law $\rho \mapsto p(\rho)$ belongs to $\mathcal{C}^2([0,\infty))$ and satisfies

$$p(\rho) > 0, \quad p'(\rho) > 0, \quad p''(\rho) > 0, \qquad \forall \rho > 0.$$

For instance, these properties hold for the classical power-law $p(\rho) = \lambda \rho^\gamma$ with $\lambda > 0$ and $\gamma > 1$. We refer the reader to the classical treatises [11, 19, 25, 35, 49] for a thorough introduction to these equations and for a description of the numerical issues. Our aim is here to extend at the second order and to higher dimension the scheme introduced in [5]. This scheme is characterized by the following two main features:

- first of all, as said previously, it works on staggered grids, meaning that densities and velocities are stored on different grid points,
- second of all, the fluxes are defined with a flavor of kinetic schemes [18, 22, 23, 31, 43, 44].

Consequently, the scheme differs in many aspects from standard approaches, for which we refer the reader e. g. to [11, 49]. In particular, due to the staggered discretization, the system is not treated "as a whole", but each equation are updated successively, which makes the numerical analysis different, see e.g. [4, 27, 29, 48]. Next, the definition of the fluxes involves the characteristic speeds of the system, but, despite the "kinetic" motivation, their evaluation do not require to compute complicated integrals. They are defined by simple formula and they do not require additional computational cost. The scheme can be shown to preserve the positivity

of the density and the entropy dissipation property under a suitable CFL condition [5], it is thus consistent with the Euler system [4].

Next, we address the full Euler model

$$\partial_t \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ \rho E \end{pmatrix} + \boldsymbol{\nabla} \cdot \begin{pmatrix} \rho \mathbf{u} \\ \rho \|\mathbf{u}\|^2 + p \\ \rho E \mathbf{u} + p \mathbf{u} \end{pmatrix} = 0. \tag{2}$$

As previously, the unknowns depend on the time and space variables $(t, x) \in [0, \infty) \times \mathbb{R}^N$; ρ, $\mathbf{u}$, E and p stand for the mass density, the velocity, the total energy and the pressure respectively. The pressure is now related to the independent unknowns $(\rho, \mathbf{u}, E)$ through an equation of state; in what follows we set

$$E = \frac{\|\mathbf{u}\|^2}{2} + e, \qquad p = (\gamma - 1)\rho e, \tag{3}$$

where $\gamma > 1$ is the adiabatic exponent. The staggered approach induces a new difficulty since in the energy equation, the total energy E involves quantities – the velocity and the internal energy – which are defined on different grids. To cope with this issue, it is tempting to work with the internal energy equation, namely

$$\partial_t(\rho e) + \boldsymbol{\nabla} \cdot (\rho e \mathbf{u}) = -p \boldsymbol{\nabla} \cdot \mathbf{u}, \tag{4}$$

instead of the evolution equation for ρE, since discrete densities, pressures, and internal energies are naturally stored at the same locations. Unfortunately, as it is well-known, this non conservative formulation is not equivalent to (2) when the solution presents discontinuities and schemes that use naively this formulation produce wrong solutions [30]. We refer the reader to [1, 32, 33] for a thorough description of numerical difficulties and attempts to design a scheme that use the primitive variables (ρ, u, p) and non conservative formulations. In what follows, we shall adapt the approach discussed in [27, 29] by plugging in the discrete version of (4) correction terms that account for the kinetic energy balance. The scheme introduced in [27, 29] can be shown: a) to be consistent with (a weak form of) the total energy equation as the space step δx goes to zero and b) to conserve the *global* discrete total energy. Even if, these properties ensures that the scheme properly compute the correct weak solution (in particular, with shocks satisfying the Rankine-Hugoniot conditions), the practitioner can be desappointed by the lack of a form of *local* conservation for a total energy. We address this issue in this article by defining averaged total energies that satisfy conservation relations.

Thus, the purpose of the present work can be summarized as follows:

- to adapt the scheme of [5] for dealing with the full Euler system (2),
- to establish *local* conservation relations satisfied by averaged total energies,
- to include reconstructed quantities in the definition of the fluxes, in the spirit of MUSCL schemes [51], in order to improve the accuracy of the scheme,
- to explain how the schemes can be extended to higher dimensions. The staggered framework naturally leads to a MAC-like discretization, in the spirit of the pioneering work [26] for incompressible flows,
- to analyse the properties of the schemes. In particular, we will discuss stability conditions so that the numerical densities and internal energies (for full Euler system) remain positive.

This work is organized as follows. In Section 2, we start by briefly describing the scheme designed in [5] for the barotropic Euler system and by recalling its main features. The remainder of this section is then devoted to the extension of the scheme to deal with the full Euler system in 1D. Finally, we exhibit the stability conditions that ensure the positivity of the density and internal energy and we define averaged total energies that satisfy a *local* conservation equation. Next, we explain in Section 3 the adaptation of the MUSCL procedure to the staggered schemes and justify that the construction reaches formally the second order accuracy. In Section 4, we briefly explain how to extend the 1D scheme to higher dimensions, when working with Cartesian grids. The case of general meshes will be addressed elsewhere, see [38, Chapter 4]. Section 5 is devoted to numerical validations for both barotropic and full Euler systems. We check numerically the gain of accuracy on explicit solutions and on 1D Riemann problems. Then we address 2D cases, like the simulation of falling columns by the Shallow Water system, as proposed in [2], and the forward facing step inspired from [55].

2 First order numerical schemes and their main properties

2.1 Staggered grids and notation

We focus in this section on the one-dimensional case where x lies in the slab $[0, L] \subset \mathbb{R}$. To define the discrete unknowns, we proceed as follows, see Fig. 1:

- we introduce a set of $J+1$ points $x_1 = 0 < x_2 < ... < x_J < x_{J+1} = L$ in the computational domain; we denote by $\mathcal{C}_{j+\frac{1}{2}} = [x_j, x_{j+1}]$, $j \in [\![1, J]\!]$, the cells defined by these points;
- we denote by $x_{j+\frac{1}{2}} = (x_j + x_{j+1})/2$, $j \in [\![1, J]\!]$, the centers of the cells; these points define the dual cells $\mathcal{C}_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $j \in [\![2, J]\!]$;
- we set the following notation for the mesh-sizes

$$\delta x_{j+\frac{1}{2}} = x_{j+1} - x_j, \ j \in [\![1, J]\!], \quad \text{and} \quad \delta x_j = \frac{\delta x_{j-\frac{1}{2}} + \delta x_{j+\frac{1}{2}}}{2}, \ j \in [\![2, J]\!],$$

(with the specific definition for the end-cells: $\delta x_1 = \frac{1}{2}\delta x_{\frac{3}{2}}$ and $\delta x_{J+1} = \frac{1}{2}\delta x_{J+\frac{1}{2}}$).

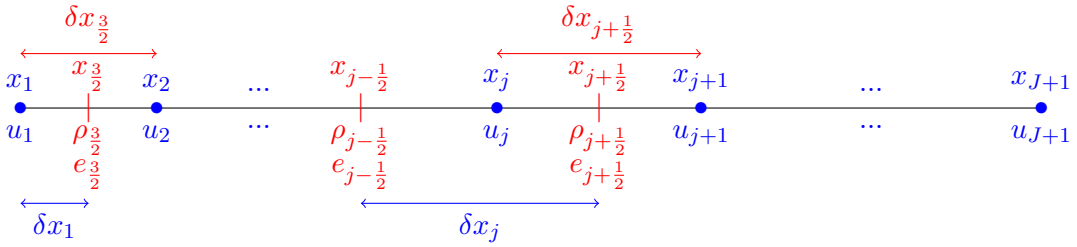


Figure 1: Staggered grid in dimension one.

We have in mind the derivation of Finite Volume schemes where the discrete densities $\rho_{j+\frac{1}{2}}$ (resp. internal energies $e_{j+\frac{1}{2}}$) are thought of as approximation of the density ρ (resp. internal energy e) on the cells $\mathcal{C}_{j+\frac{1}{2}}$ whereas the discrete velocities u_j are thought of as approximation of the velocity u on the cells $\mathcal{C}_j$.

The time discretization is explicit and we use the convention that, with q the evaluation of a certain quantity at time t, $\overline{q}$ stands for its update at time $t + \delta t$.

2.2 Discretization of the barotropic Euler system

The scheme for barotropic Euler equations has the general form

$$\frac{\overline{\rho}_{j+\frac{1}{2}} - \rho_{j+\frac{1}{2}}}{\delta t} + \frac{\mathcal{F}_{j+1} - \mathcal{F}_j}{\delta x_{j+\frac{1}{2}}} = 0, \qquad \forall j \in [\![1, J]\!]. \tag{5}$$

$$\frac{\overline{\rho}_j \overline{u}_j - \rho_j u_j}{\delta t} + \frac{\mathcal{G}_{j+\frac{1}{2}} - \mathcal{G}_{j-\frac{1}{2}}}{\delta x_j} + \frac{\Pi_{j+\frac{1}{2}} - \Pi_{j-\frac{1}{2}}}{\delta x_j} = 0, \qquad \forall j \in [\![2, J]\!]. \tag{6}$$

Equation (5) is the discrete version of the mass balance equation and equation (6) is the discrete version of the momentum balance equation. Of course, the scheme has to be completed by the definition of mass fluxes $\mathcal{F}_j$, momentum fluxes $\mathcal{G}_{j+\frac{1}{2}} + \Pi_{j+\frac{1}{2}}$ (with $\Pi_{j+\frac{1}{2}}$ an approximation of the pressure) and by initial and boundary conditions. The discrete momentum balance involves quantities ρ_j, which stand for the approximations of ρ at the internal edges of the primal mesh and are obtained as averages of quantities $\rho_{j-\frac{1}{2}}$ and $\rho_{j+\frac{1}{2}}$ as follows:

$$\rho_j = \frac{\delta x_{j+\frac{1}{2}} \rho_{j+\frac{1}{2}} + \delta x_{j-\frac{1}{2}} \rho_{j-\frac{1}{2}}}{2\delta x_j}, \qquad \forall j \in [\![2, J]\!]. \tag{7}$$

In contrast to the collocated approach (with the noticeable exception of AUSM schemes [37, 36]), a discretization of each physical variables, ρ and $\mathbf{u}$ separately, is natural on a staggered grid. In particular, the mass flux $\mathcal{F}_j$ at the interface x_j can use directly the material velocity u_j. For instance, it looks tempting to define the flux $\mathcal{F}_j$ based on the Upwinding principles according to the sign of u_j, see [29], but this approach does not use the hyperbolic properties of the system (1) and requires extra-diffusion to reduce spurious oscillations that might appear, see [5, Appendix B]. Instead, the flux designed in [5, 6] makes full use of the characteristic speeds of the system (1), namely

$$\lambda_\pm(c, u) = u \pm c, \text{ where } c \text{ stands for the sound speed.}$$

In the barotropic case, the sound speed depends only on ρ, that is $c = c_b(\rho)$ with

$$c_b(\rho) = \sqrt{p'(\rho)}.$$

The numerical mass fluxes are given by the following formula

$$\mathcal{F}_j = \mathcal{F}_j^+ + \mathcal{F}_j^-, \qquad \forall j \in [\![1, J+1]\!],$$

where

$$\mathcal{F}_j^+ = \mathcal{F}^+(\rho_{j-\frac{1}{2}}, c_j, u_j) \quad \text{and} \quad \mathcal{F}_j^- = \mathcal{F}^-(\rho_{j+\frac{1}{2}}, c_j, u_j), \qquad \forall j \in [\![2, J]\!], \tag{8}$$

with $c_j = c_b(\rho_j)$. The definition of $\mathcal{F}_1^\pm$ and $\mathcal{F}_{J+1}^\pm$ depends on the prescribed boundary conditions. The flux functions $\mathcal{F}^\pm$ are defined as follows

$$\mathcal{F}^+(\rho, c, u) = \begin{cases} 0 & \text{if } u \leqslant -c, \\ \dfrac{\rho}{4c}(u + c)^2 & \text{if } |u| \leqslant c, \\ \rho u & \text{if } u \geqslant c, \end{cases} \tag{9}$$

and

$$\mathcal{F}^-(\rho, c, u) = \begin{cases} \rho u & \text{if } u \leqslant -c, \\ -\dfrac{\rho}{4c}(u - c)^2 & \text{if } |u| \leqslant c, \\ 0 & \text{if } u \geqslant c. \end{cases} \tag{10}$$

We do not explain in this article the derivation of these expressions which is related to kinetic schemes, but we refer the interested reader to [5, 6] where the complete derivation is provided. We note that a symmetry property holds

$$\mathcal{F}^-(\rho, c, u) = -\mathcal{F}^+(\rho, c, -u),$$

and it is clear that the flux–consistency condition is fulfilled

$$\mathcal{F}^+(\rho, c, u) + \mathcal{F}^-(\rho, c, u) = \rho u. \tag{11}$$

It is worth having in mind Fig. 2, which clarifies the correction with respect to the mere UpWind flux based on the sign of the material velocity. As explained in [5], it induces some numerical diffusion which prevents the formation of oscillations in the vicinity of small material velocities.

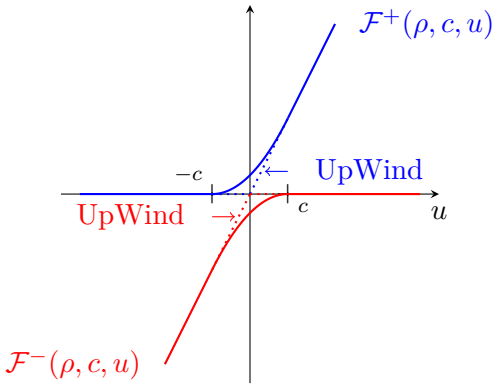


Figure 2: Comparison of the flux (9)–(10) and the UpWind flux for a fixed ρ.

For the momentum flux, the pressure gradient at $x_{j+\frac{1}{2}}$ is naturally centered by using the densities in the neighboring cells with

$$\Pi_{j+\frac{1}{2}} = p(\rho_{j+\frac{1}{2}}), \tag{12}$$

while the convection flux is written by applying the upwinding principle, based on the "sign" of the mass fluxes $\mathcal{F}_j$ and $\mathcal{F}_{j+1}$, to the velocity field. We arrive at the following definition

$$\mathcal{G}_{j+\frac{1}{2}} = u_j \mathcal{F}^+_{j+\frac{1}{2}} + u_{j+1} \mathcal{F}^-_{j+\frac{1}{2}},$$

where the quantities $\mathcal{F}^\pm_{j+\frac{1}{2}}$ are expressed as mean values of $\mathcal{F}^\pm_j$, $\mathcal{F}^\pm_{j+1}$:

$$\mathcal{F}^\pm_{j+\frac{1}{2}} = \frac{\mathcal{F}^\pm_j + \mathcal{F}^\pm_{j+1}}{2}. \tag{13}$$

It is remarkable that a conservation relation holds with the dual quantities ρ_j and $\mathcal{F}^\pm_{j+\frac{1}{2}}$:

$$\overline{\rho}_j - \rho_j + \frac{\delta t}{\delta x_j}(\mathcal{F}_{j+\frac{1}{2}} - \mathcal{F}_{j-\frac{1}{2}}) = 0 \tag{14}$$

where, of course, $\mathcal{F}_{j+\frac{1}{2}} = \mathcal{F}^+_{j+\frac{1}{2}} + \mathcal{F}^-_{j+\frac{1}{2}}$. Due to (11), it is clear that the momentum flux is also consistent.

The scheme has the following properties and abilities, at least in this simple 1D framework:

- stability analysis [5]: up to a (quite standard) stability condition on the numerical parameters, the scheme preserves the positivity of the density, and it makes the total energy of the system decay,
- consistency analysis [4]: the scheme satisfies a Lax-Wendroff type theorem,
- simulations: the scheme has the advantage of algorithmic simplicity (it does not require to solve Riemann problems and the definition of the flux (9)-(10) is fully explicit; despite its "kinetic" flavor, it does not require an additional integration procedure...), it performs well on the standard test cases of Riemann problems and it works for very general pressure laws, like with close-packing pressures, see [5, 6].

We propose in the next section an extension of this scheme to the discretization of the full Euler system.

2.3 Discretization of the full Euler system

The discrete mass and momentum balance equations have been already derived in the previous section. We use here the same definitions except for the sound speed and the pressure. Indeed, as for the barotropic Euler system, the smallest and largest characteristic speeds of the full Euler system are $\lambda_\pm(c,u) = u \pm c$ (the third one being u) where c is the sound speed. The sound speed now depends only on the internal energy, that is $c = c_f(e)$ with

$$c_f(e) = \sqrt{(\gamma-1)\gamma e}.$$

Hence, it leads us to write

$$c_j = c_f(e_j) \quad \text{with} \quad e_j = \frac{e_{j-\frac{1}{2}} + e_{j+\frac{1}{2}}}{2},$$

in the definition of the mass fluxes (8). Moreover, in the full Euler system, the pressure is no longer defined as a function of the density only but instead using the state law (3), so that we set:

$$\Pi_{j+\frac{1}{2}} = (\gamma-1)\rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}},$$

instead of (12).

We now turn to the discrete version of the internal energy equation

$$\frac{\overline{\rho}_{j+\frac{1}{2}}\overline{e}_{j+\frac{1}{2}} - \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}}}{\delta t} + \frac{\mathcal{E}_{j+1} - \mathcal{E}_j}{\delta x_{j+\frac{1}{2}}} + \Pi_{j+\frac{1}{2}} \frac{\overline{u}_{j+1} - \overline{u}_j}{\delta x_{j+\frac{1}{2}}} = \mathcal{S}_{j+\frac{1}{2}}. \tag{15}$$

The left hand side corresponds to the discretization of (4), where the internal energy flux $\mathcal{E}_j$ is given by

$$\mathcal{E}_j = e_{j-\frac{1}{2}} \mathcal{F}_j^+ + e_{j+\frac{1}{2}} \mathcal{F}_j^-. \tag{16}$$

This formula still corresponds to the upwinding principle associated to the transport of ρe with velocity u, according to the definition of the mass fluxes. Note that the discretization of the non conservative term $p\partial_x u$ uses the velocity field $\overline{u}$, just updated in the previous step. Following [27], the right hand side $\mathcal{S}_{j+\frac{1}{2}}$ is designed to account for the remainder term that appears in the discrete kinetic energy balance; it does not vanish when δx goes to zero, precisely because it is intended to capture the correct behavior at discontinuities. To be more specific, the kinetic energy balance is obtained by multiplying (6) by $\overline{u}_j$. We find, see [5, 27]:

$$\frac{1}{2}\frac{\overline{\rho}_j \overline{u}_j^2 - \rho_j u_j^2}{\delta t} + \frac{\mathcal{K}_{j+\frac{1}{2}} - \mathcal{K}_{j-\frac{1}{2}}}{\delta x_j} + \frac{\Pi_{j+\frac{1}{2}} - \Pi_{j-\frac{1}{2}}}{\delta x_j}\overline{u}_j = -\mathcal{R}_j,$$

where the kinetic energy flux is given by

$$\mathcal{K}_{j+\frac{1}{2}} = \frac{u_j^2}{2}\mathcal{F}^+_{j+\frac{1}{2}} + \frac{u_{j+1}^2}{2}\mathcal{F}^-_{j+\frac{1}{2}} \tag{17}$$

and the remainder reads

$$\begin{aligned}\mathcal{R}_j &= \frac{1}{2\delta t}\overline{\rho}_j\,(\overline{u}_j - u_j)^2 + \frac{1}{\delta x_j}\left(\frac{(u_j - u_{j-1})^2}{2}\mathcal{F}^+_{j-\frac{1}{2}} - \frac{(u_{j+1} - u_j)^2}{2}\mathcal{F}^-_{j+\frac{1}{2}}\right)\\ &\quad + \frac{1}{\delta x_j}(\overline{u}_j - u_j)(u_j - u_{j-1})\mathcal{F}^+_{j-\frac{1}{2}} + \frac{1}{\delta x_j}(\overline{u}_j - u_j)(u_{j+1} - u_j)\mathcal{F}^-_{j+\frac{1}{2}}.\end{aligned} \tag{18}$$

It motivates to define the source term for (15) as follows

$$\mathcal{S}_{j+\frac{1}{2}} = \frac{\delta x_{j+1}\mathcal{R}_{j+1} + \delta x_j \mathcal{R}_j}{2\delta x_{j+\frac{1}{2}}}.$$

The scheme shares similarities with the 1D version of the scheme presented in [29, Section 4], see also [48, Chapter 2]. However, it differs by the following two points:

- firstly, the mass fluxes in [27, 29, 48] are upwinded with respect to the material velocity (in other words, it corresponds to the choice $\mathcal{F}^\pm(\rho, c, u) = \pm\rho[u]^\pm$ instead of (9) and (10), see also Fig. 2). The mass flux based (9) and (10) introduces a bit of numerical diffusion [5, Appendix B] which prevents the occurence of spurious oscillations when the material velocity vanishes, see [28, Section 6.1.1] and [48, Section 2.3.5] where an artificial viscosity is added to damp these oscillations.
- secondly, the organization of the time steppings are different: even if both schemes are explicit, the variables are not updated in the same order. We solve the discrete equations in the order $\rho \to u \to e$, as in [1], whereas [29] proceeds according to $\rho \to e \to u$. In particular, here the corrective term $\mathcal{S}_{j+\frac{1}{2}}$ does not need any time shift since the updated velocity $\overline{u}$ is known when solving (15).

2.4 Stability conditions

We now turn to the study of the stability conditions which ensure the positivity of the density and the internal energy. We start by stating a lemma that will be useful in this section and in Section 3.3. This lemma about the flux functions $\mathcal{F}^\pm$ is proved in [5].

Lemma 2.1. *For all $u \in \mathbb{R}$, for all $\rho \geqslant 0$ and for all $c \geqslant 0$, the fluxes $\mathcal{F}^\pm$ satisfy the following inequalities:*

$$0 \leqslant \mathcal{F}^+(\rho, c, u) \leqslant \rho[\lambda_+(c, u)]^+ \quad \text{and} \quad -\rho[\lambda_-(c, u)]^- \leqslant \mathcal{F}^-(\rho, c, u) \leqslant 0. \tag{19}$$

With this lemma at hand, we can prove the following statement.

Proposition 2.2. *Assume that $e_{j+\frac{1}{2}} \geqslant 0$, $\rho_{j+\frac{1}{2}} \geqslant 0$, for any j. If the following CFL-like conditions hold for all j*

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}}\left([u_{j+1}]^+ + \frac{c_f(e_{j+\frac{3}{2}}) + c_f(e_{j+\frac{1}{2}})}{\sqrt{2}} + [u_j]^- + \frac{c_f(e_{j+\frac{1}{2}}) + c_f(e_{j-\frac{1}{2}})}{\sqrt{2}}\right) \leqslant \frac{1}{\gamma}, \tag{20}$$

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}} c_f(e_{j+\frac{1}{2}+k}) \leqslant \frac{(\gamma - 1)}{2\sqrt{2}}, \quad \forall k \in \{-1, 0, 1\}, \tag{21}$$

then $\overline{e}_{j+\frac{1}{2}} \geqslant 0$ and $\overline{\rho}_{j+\frac{1}{2}} \geqslant 0$.

Roughly speaking, the stability condition has the expected form of a constraint governed by the speed $|u| + c$. However, in contrast to what happened for the barotropic case [5, Proposition 3.7], we observe that the expression of the constraint involves additional factors depending on the adiabatic exponent γ. This has to be compared to [29, eq. (39)].

Proof. We assume that $e_{j+\frac{1}{2}} \geqslant 0$, $\rho_{j+\frac{1}{2}} \geqslant 0$ and that (20) and (21) holds for all j.
We start by observing that

$$[\lambda_\pm(c,u)]^\pm \leqslant [u]^\pm + c, \tag{22}$$

$$\sqrt{2}\; c_j \leqslant c_f(e_{j-\frac{1}{2}}) + c_f(e_{j+\frac{1}{2}}). \tag{23}$$

Positivity of the density. As proved in [5], the positivity of $\overline{\rho}_{j+\frac{1}{2}}$ comes from the inequality

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}} \left([\lambda_+(c_{j+1}, u_{j+1})]^+ + [\lambda_-(c_j, u_j)]^- \right) \leqslant 1.$$

It is directly implied by (20) since $\gamma > 1$ and (22), (23) hold.

Positivity of the internal energy. We rewrite the terms $(-1)^i \Pi_{j+\frac{1}{2}} \overline{u}_{j+i},\ i \in \{0,1\}$, which are involved in (15), by making the discrete time derivative $(\overline{u}_{j+i} - u_{j+i})$ appear. Then, we make use of the Young inequality as follows

$$\begin{aligned}(-1)^i \Pi_{j+\frac{1}{2}} \overline{u}_{j+i} &= (-1)^i (\gamma - 1) \left(\rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} \left(\overline{u}_{j+i} - u_{j+i} \right) + \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} u_{j+i} \right) \\ &\geqslant -\rho_{j+\frac{1}{2}} \left(\frac{c_f(e_{j+\frac{1}{2}})}{2\sqrt{2}\gamma} (\overline{u}_{j+i} - u_{j+i})^2 + (\gamma - 1) e_{j+\frac{1}{2}} \left(\frac{c_f(e_{j+\frac{1}{2}})}{\sqrt{2}} - (-1)^i u_{j+i} \right) \right).\end{aligned}$$

Next, we write $\overline{\rho}_{j+\frac{1}{2}} \overline{e}_{j+\frac{1}{2}} \geqslant \mathrm{T}_0 + \mathrm{T}_1^0 + \mathrm{T}_1^1$ where:

$$\begin{aligned}\mathrm{T}_0 &= \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} \left(1 - \frac{\delta t}{\delta x_{j+\frac{1}{2}}} (\gamma - 1) \left(2 \frac{c_f(e_{j+\frac{1}{2}})}{\sqrt{2}} - u_j + u_{j+1} \right) \right) - \delta t \frac{\mathcal{E}_{j+1} - \mathcal{E}_j}{\delta x_{j+\frac{1}{2}}}, \\ \mathrm{T}_1^i &= \frac{\delta t}{2} \frac{\delta x_{j+i}}{\delta x_{j+\frac{1}{2}}} \mathcal{R}_{j+i} - \frac{\delta t}{\delta x_{j+\frac{1}{2}}} \frac{c_f(e_{j+\frac{1}{2}})}{2\sqrt{2}\gamma} \rho_{j+\frac{1}{2}} (\overline{u}_{j+i} - u_{j+i})^2.\end{aligned}$$

In order to guarantee that $\overline{e}_{j+\frac{1}{2}}$ is non negative it is sufficient to ensure that these three terms are non negative. This holds under the assumptions (20) and (21).

Indeed, using the definition of the flux $\mathcal{E}_j$ and owing to (19), we obtain

$$\begin{aligned}\mathrm{T}_0 \geqslant \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} &\left(1 - \frac{\delta t}{\delta x_{j+\frac{1}{2}}} (\gamma - 1) \left([u_j]^- + \frac{c_f(e_{j+\frac{1}{2}})}{\sqrt{2}} + [u_{j+1}]^+ + \frac{c_f(e_{j+\frac{1}{2}})}{\sqrt{2}} \right) \right) \\ &- \frac{\delta t}{\delta x_{j+\frac{1}{2}}} \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} \left([\lambda_+(c_{j+1}, u_{j+1})]^+ + [\lambda_-(c_j, u_j)]^- \right)\end{aligned}$$

where, due to (20), the right hand side is non negative by virtue of (22) and (23) .

Next, we turn to T_1^i. Using twice the Young inequality and bearing in mind the definition of $\overline{\rho}_j$, we observe that

$$\frac{\delta t}{2} \frac{\delta x_{j+i}}{\delta x_{j+\frac{1}{2}}} \mathcal{R}_{j+i} \geqslant \frac{\delta x_{j+i}}{4 \delta x_{j+\frac{1}{2}}} (\overline{u}_{j+i} - u_{j+i})^2 \left(\rho_{j+i} - \frac{\delta t}{\delta x_{j+i}} (\mathcal{F}^+_{j+i+\frac{1}{2}} - \mathcal{F}^-_{j+i-\frac{1}{2}}) \right).$$

Hence, we have

$$\mathrm{T}_1^i \geqslant \frac{\delta x_{j+i}}{4\delta x_{j+\frac{1}{2}}}(\overline{u}_{j+i}-u_{j+i})^2\left(\rho_{j+i}-\frac{\delta t}{\delta x_{j+i}}(\mathcal{F}^+_{j+i+\frac{1}{2}}-\mathcal{F}^-_{j+i-\frac{1}{2}})-\frac{\delta t}{\gamma}\frac{2}{\sqrt{2}}\frac{\rho_{j+\frac{1}{2}}c_f(e_{j+\frac{1}{2}})}{\delta x_{j+i}}\right).$$

Coming back to (7) and (13), we write $T_1^i \geqslant \dfrac{(\overline{u}_{j+i}-u_{j+i})^2}{4\delta x_{j+\frac{1}{2}}}\left(T_2^{i,0}+T_2^{i,1}\right)$ where, for $k=0,1$,

$$\mathrm{T}_2^{i,k}=\frac{\delta x_{j+i+k-\frac{1}{2}}}{2}\rho_{j+i+k-\frac{1}{2}}-\delta t\frac{\mathcal{F}^+_{j+i+k}-\mathcal{F}^-_{j+i+k-1}}{2}-\frac{\delta t}{\gamma}\frac{2}{\sqrt{2}}\rho_{j+i+k-\frac{1}{2}}c_f(e_{j+\frac{1}{2}}).$$

Note that a non negative term has been added to obtain a symmetric formulation in the above inequality. Due to (19) and (20) we get

$$\mathcal{F}^+_{j+i+k}-\mathcal{F}^-_{j+i+k-1}\leqslant\frac{\delta x_{j+i+k-\frac{1}{2}}}{\gamma\delta t}\rho_{j+i+k-\frac{1}{2}},$$

and this allows us to write

$$\mathrm{T}_2^{i,k}\geqslant\frac{\delta x_{j+i+k-\frac{1}{2}}}{2\gamma}\rho_{j+i+k-\frac{1}{2}}\left(\gamma-1-\frac{\delta t}{\delta x_{j+i+k-\frac{1}{2}}}\frac{4}{\sqrt{2}}c_f(e_{j+\frac{1}{2}})\right).$$

We conclude by observing that this term is non negative by virtue of (21). ∎

2.5 Numerical diffusion, contact discontinuities

It is worth discussing the expression of the numerical diffusion produced by the scheme (see [5, Appendix B] for a similar discussion concerning the barotropic case). Let us introduce the following non negative quantity

$$C_j=\begin{cases}-u_j & \text{if } u_j\leqslant -c_j,\\ \dfrac{u_j^2+c_j^2}{4c_j} & \text{if } |u_j|<c_j,\\ u_j & \text{if } u_j>c_j.\end{cases}$$

It is convenient to use the following shorthand notations for averaged quantities

$$\{q\}_j=\frac{q_{j-\frac{1}{2}}+q_{j+\frac{1}{2}}}{2}\quad\text{and}\quad\{q\}_{j+\frac{1}{2}}=\frac{q_j+q_{j+1}}{2}.$$

Finally, we denote

$$\mathcal{F}^{|\cdot|}=\mathcal{F}^+-\mathcal{F}^-,$$

which is a positive quantity. The mass and momentum fluxes can be cast as the sum of a centered term and a diffusion term

$$\mathcal{F}_j=\{\rho\}_j u_j-\frac{C_j}{2}\left(\rho_{j+\frac{1}{2}}-\rho_{j-\frac{1}{2}}\right),$$

$$\mathcal{G}_{j+\frac{1}{2}}=\{\mathcal{F}\}_{j+\frac{1}{2}}\{u\}_{j+\frac{1}{2}}-\frac{\{\mathcal{F}^{|\cdot|}\}_{j+\frac{1}{2}}}{2}\left(u_{j+1}-u_j\right).$$

Concerning the internal energy (16) and kinetic energy fluxes (17), they become:

$$\mathcal{E}_j=\{\rho e\}_j u_j-\frac{C_j}{2}\left(e_{j+\frac{1}{2}}\rho_{j+\frac{1}{2}}-e_{j-\frac{1}{2}}\rho_{j-\frac{1}{2}}\right),$$

$$\mathcal{K}_{j+\frac{1}{2}}=\{\mathcal{F}\}_{j+\frac{1}{2}}\{\frac{u^2}{2}\}_{j+\frac{1}{2}}-\frac{\{\mathcal{F}^{|\cdot|}\}_{j+\frac{1}{2}}}{2}\left(\frac{u_{j+1}^2}{2}-\frac{u_j^2}{2}\right).$$

Remark 2.3. As a by-product, it is remarkable that the scheme properly deals with 1D-*contact discontinuities*. Indeed, let us assume that the discrete velocity and pressure are constant in the neighborhood of $x_{j+\frac{1}{2}}$, which means that $u_{j-1} = u_j = u_{j+1} = u_{j+2} = u$ and $\Pi_{j-\frac{1}{2}} = \Pi_{j+\frac{1}{2}} = \Pi_{j+\frac{3}{2}} = \Pi$ holds. Then the scheme guarantees that they remain constant in the neighborhood of this point at the next time: $\overline{\Pi}_{j+\frac{1}{2}} = \Pi$ and $\overline{u}_{j+1} = u = \overline{u}_j$.

2.6 Conservation of total energy

As said in the introduction, it is far from clear that we can obtain a consistent approximation of the conservation equations (2) when the scheme is defined on the basis of the non conservative formulation (4). In order to analyze this issue, let us now introduce the averaged total energy at $x_{j+\frac{1}{2}}$ and x_j, defined by

$$E_{j+\frac{1}{2}} = e_{j+\frac{1}{2}} + \frac{1}{2}\frac{\delta x_j \rho_j u_j^2 + \delta x_{j+1}\rho_{j+1}u_{j+1}^2}{2\delta x_{j+\frac{1}{2}}\rho_{j+\frac{1}{2}}}$$

and

$$E_j = \frac{u_j^2}{2} + \frac{\delta x_{j+\frac{1}{2}}\rho_{j+\frac{1}{2}}e_{j+\frac{1}{2}} + \delta x_{j-\frac{1}{2}}\rho_{j-\frac{1}{2}}e_{j-\frac{1}{2}}}{2\delta x_j \rho_j}.$$

We wish to obtain conservative equations for those quantities. To this end, we introduce the fluxes

$$\mathcal{T}_j = \mathcal{E}_j + \frac{\mathcal{K}_{j+\frac{1}{2}} + \mathcal{K}_{j-\frac{1}{2}}}{2}$$

and

$$\mathcal{T}^*_{j+\frac{1}{2}} = \frac{\mathcal{E}_{j+1} + \mathcal{E}_j}{2} + \mathcal{K}_{j+\frac{1}{2}} - \frac{\delta x_{j+1}\mathcal{R}_{j+1} - \delta x_j \mathcal{R}_j}{4}$$

which use the quantities defined in (16), (17), (18).

We get the following consistent *local* balance equations for the total energy defined either on the primal mesh $\rho_{j+\frac{1}{2}}E_{j+\frac{1}{2}}$ or the dual mesh $\rho_j E_j$:

$$\frac{\overline{\rho}_{j+\frac{1}{2}}\overline{E}_{j+\frac{1}{2}} - \rho_{j+\frac{1}{2}}E_{j+\frac{1}{2}}}{\delta t} + \frac{\mathcal{T}_{j+1} - \mathcal{T}_j}{\delta x_{j+\frac{1}{2}}} + \frac{\overline{u}_{j+1}\{\Pi\}_{j+1} - \overline{u}_j\{\Pi\}_j}{\delta x_{j+\frac{1}{2}}} = 0,$$

and

$$\frac{\overline{\rho}_j\overline{E}_j - \rho_j E_j}{\delta t} + \frac{\mathcal{T}^*_{j+\frac{1}{2}} - \mathcal{T}^*_{j-\frac{1}{2}}}{\delta x_j} + \frac{\Pi_{j+\frac{1}{2}}\{\overline{u}\}_{j+\frac{1}{2}} - \Pi_{j-\frac{1}{2}}\{\overline{u}\}_{j-\frac{1}{2}}}{\delta x_j} = 0.$$

Note that this construction of a local conservation equation for an averaged total energy also holds for the extension to higher dimensions on MAC grids presented in Section 4 and also for an extension of this scheme on unstructured mesh. It will be explain in a forthcoming work, see [38].

3 A muscl-scheme on staggered grids

In this section we discuss how we adapt the MUSCL principles [51] to the staggered framework. Classicaly, we first reconstruct second order quantities at edges of primal or dual cells depending on the domain of definition of the different variables. Then, concerning the discretization of the mass flux, we keep unchanged the velocity defined at the interface x_j and we shall replace the

Upwind value $\rho_{j\pm\frac12}$ by a MUSCL reconstruction $\rho_j^\pm$ of the density: it defines the upgraded mass flux $\mathcal{F}_j^{ML}$. For the momentum flux, since the discretization of the pressure is centered, we only need to define the convection flux $\mathcal{G}_{j+\frac12}^{ML}$: we shall combine the obtained mass fluxes $\mathcal{F}_j^{ML}$ and $\mathcal{F}_{j+1}^{ML}$ with a MUSCL reconstructed velocity $u_{j+\frac12}^\pm$ at the interfaces $x_{j+\frac12}$. When considering full Euler equation, for the internal energy fluxes, a first attempt would follow the same strategy by combining the upgraded mass fluxes $\mathcal{F}_j^{ML}$ with a MUSCL reconstructed internal energy $e_j^\pm$. However, this approach produces a bad behaviour of the scheme when contact discontinuities occur, see [48] for further comments on this issue. Instead, we combine the upgraded mass fluxes $\mathcal{F}_j^{ML}$ with a MUSCL reconstruction of the internal energy defined from the ratio $\frac{(\rho e)_j^\pm}{\rho_j^\pm}$. We shall see that stability issues might require to strengthen the limitation procedure applied to define the reconstructed velocities $u_{j+\frac12}^\pm$.

3.1 muscl reconstruction at edges of primal or dual mesh

We introduce piecewise linear reconstructions of the mass density ρ and of the density of internal energy ρe, defined, on each cell $\mathcal{C}_{j+\frac12}$, by

$$\hat{q}_{j+\frac12}(x) = q_{j+\frac12} + s^q_{j+\frac12}(x - x_{j+\frac12}), \text{ for any } x \in \mathcal{C}_{j+\frac12} \text{ and with } q = \rho \text{ or } \rho e. \tag{24}$$

The slope $s^q_{j+\frac12} \in \mathbb{R}$ is intended to be an approximation of the gradient of q on the cell $\mathcal{C}_{j+\frac12}$. Classically, it is obtained as a symmetric function of the two discrete derivatives computed using the values of q on the neighboring cells,

$$s^q_{j+\frac12} = \widehat{\Phi}\Big(\frac{q_{j+\frac12} - q_{j-\frac12}}{\delta x_j},\ \frac{q_{j+\frac32} - q_{j+\frac12}}{\delta x_{j+1}}\Big).$$

A suitable adaptation of this formula needs to be introduced at the boundaries of the computational domain; for instance we can simply make the scheme degenerate to first order next to the boundaries ($s^q_{\frac32} = 0$ and $s^q_{J+\frac12} = 0$).

For stability reasons, in order to prevent the formation of over- and undershoots, the value of the reconstructed quantities at an edge should not exceed the values of the quantity in the two neighboring cells and the slope $s^q_{j+\frac12}$ should vanish at extrema. These properties are classically ensured by the definition of the function $\widehat{\Phi}$, the so-called limiter function. It is seen here as a function of two variables (a, b) but it is also customary to use instead a function Φ of the single variable a/b with the following equalities

$$\widehat{\Phi}(a,b) = b\ \Phi\Big(\frac{a}{b}\Big) = a\ \Phi\Big(\frac{b}{a}\Big) = \widehat{\Phi}(b,a),$$

where it is understood that the function Φ satisfies the symmetry property

$$\frac{\Phi(r)}{r} = \Phi\Big(\frac{1}{r}\Big),\ \forall r \neq 0. \tag{25}$$

On uniform grids, the geometric properties stated above are ensured when the limiter function lies in the well-known Sweby TVD region, see [51, 47], which is characterized by the three conditions

$$\underbrace{\Phi(r) = 0}_{(a)},\ \forall r \leqslant 0, \qquad \underbrace{0 \leqslant}_{(b)} \Big(\Phi(r), \frac{\Phi(r)}{r}\Big) \underbrace{\leqslant 2}_{(c)},\ \forall r \geqslant 0.$$

On non-uniform grids, the situation is more intricate as explained in [3]: in condition (c) the upper bound 2 should be replaced by a quantity that depends on the mesh regularity. More precisely the limiter Φ must satisfy

$$\Phi(r) = 0, \ \forall r \leqslant 0, \qquad 0 \leqslant \Big(\Phi(r), \frac{\Phi(r)}{r}\Big) \leqslant \tau, \ \forall r \geqslant 0, \tag{26}$$

where $1 < \tau \leqslant 2$ is the mesh dependent number defined by

$$\tau = \min_{j \in [\![2, J-1]\!]} \left(\frac{2\delta x_j}{\delta x_{j+\frac{1}{2}}}; \frac{2\delta x_{j+1}}{\delta x_{j+\frac{1}{2}}} \right).$$

We will that a strengthened version is needed for the approximation of the full Euler equation, for which we assume that

$$\tau < 2. \tag{27}$$

The role of this last restriction will appear in the stability analysis for proving the positivity of the internal energy (see Section 3.3).

Finally, in order to ensure that the scheme is second order in space (see Section 3.4 below), the limiter function $r \mapsto \Phi(r)$ should be a smooth function – with at least left and right derivatives at the point $r = 1$ – and satisfy

$$\Phi(1) = 1. \tag{28}$$

As discussed in Lemma 3.6 (in Section 3.4 below), if $x \mapsto q(x)$ is a smooth function, the derivatives of which are bounded and remain bounded, then we get

$$s^q_{j+\frac{1}{2}} = q'(x_{j+\frac{1}{2}}) + \mathcal{O}\left(\delta x\right).$$

From classical limiters defined for uniform meshes, we can define τ-limiters that satisfy properties (25), (26) and (28), see [15]:

- the MinMod limiter: $\Phi_{mm}(r) = \max(0, \min(1, r))$, which is actually upper-bounded by 1, and the τ-MinMod limiter:

$$\Phi_{\tau,mm}(r) = \max\left(0, \min(\tau, r)\right)$$

- the SuperBee limiter: $\Phi_{sb}(r) = \max(0, \min\left(2r, 1\right), \min\left(r, 2\right))$, bounded by 2, and the τ-SuperBee limiter:

$$\Phi_{\tau,sb}(r) = \max\left(0, \min\left(\tau r, 1\right), \min\left(r, \tau\right)\right) \tag{29}$$

The affine reconstruction $\hat{q}$ of q in (24) allows us to define the two values $q^-_j = \hat{q}_{j-\frac{1}{2}}(x_j)$ and $q^+_j = \hat{q}_{j+\frac{1}{2}}(x_j)$ at the interface x_j:

$$q^-_j = q_{j-\frac{1}{2}} + \frac{\delta x_{j-\frac{1}{2}}}{2} \ s^q_{j-\frac{1}{2}},$$
$$q^+_j = q_{j+\frac{1}{2}} - \frac{\delta x_{j+\frac{1}{2}}}{2} \ s^q_{j+\frac{1}{2}},$$

which will be used in the numerical fluxes. As discussed in Lemma 3.6, if $x \mapsto q(x)$ is a smooth function, bounded with bounded derivatives, then we get

$$s^q_{j+\frac{1}{2}} = q'(x_{j+\frac{1}{2}}) + \mathcal{O}\left(\delta x\right),$$

which can be used to check that the scheme is formally second-order accurate.

A similar reconstruction is used for the velocity on the dual mesh. We set

$$\widehat{u}_j(x) = u_j + w_j(x - x_j), \quad \forall x \in \mathcal{C}_j, \quad \forall j \in [\![1, J+1]\!].$$

The slopes $w_j \in \mathbb{R}$ are now defined by

$$w_j = \lambda_j \widehat{\Phi}\left(\frac{u_j - u_{j-1}}{\delta x_{j-\frac{1}{2}}}, \frac{u_{j+1} - u_j}{\delta x_{j+\frac{1}{2}}}\right), \quad \forall j \in [\![2, J]\!], \qquad \text{and} \qquad w_1 = 0 = w_{J+1}. \tag{30}$$

Here, the situation is slightly more involved since we have introduced a parameter $\lambda_j \in [0, 1]$. The value $\lambda_j = 1$ corresponds to the usual MUSCL reconstruction. This value is suitable for the discretization of the barotropic Euler equation but, when considering the full Euler equation, the source term $\mathcal{S}_{j+\frac{1}{2}}$ associated to the kinetic energy balance that appears in the internal energy equation induces further constraints in order to preserve the positivity of the internal energy, which might require to strengthen the limitation of the slope by choosing $\lambda_j < 1$.

The affine reconstruction $\widehat{u}$ of the velocity u allows us to define, at the interfaces $x_{j+\frac{1}{2}}$, $u^-_{j+\frac{1}{2}} = \widehat{u}_j(x_{j+\frac{1}{2}})$ and $u^+_{j+\frac{1}{2}} = \widehat{u}_{j+1}(x_{j+\frac{1}{2}})$:

$$u^-_{j+\frac{1}{2}} = u_j + \frac{\delta x_{j+\frac{1}{2}}}{2}\, w_j,$$
$$u^+_{j+\frac{1}{2}} = u_{j+1} - \frac{\delta x_{j+\frac{1}{2}}}{2}\, w_{j+1}.$$

Here, we bear in mind that x_j and x_{j+1} are not necessarily the mid-points of $\mathcal{C}_j$ and $\mathcal{C}_{j+1}$ respectively (see Fig. 1); this is the reason why the formula is not expressed by means of $\delta x_{j+1}/2$ and $\delta x_j/2$.

3.2 Definition of the second order fluxes

With the reconstructed quantities at hand, we can now define the modified fluxes. We update the density by replacing the mass flux $\mathcal{F}_j$ by the MUSCL-flux $\mathcal{F}^{ML}_j$ defined by

$$\mathcal{F}^{ML}_j = \mathcal{F}^+(\rho^-_j, c_j, u_j) + \mathcal{F}^-(\rho^+_j, c_j, u_j), \qquad \forall j \in [\![2, J]\!],$$

(with the corresponding adaptation at the boundary, for instance we set $\mathcal{F}^{ML}_1 = 0 = \mathcal{F}^{ML}_{J+1}$ at the endpoints of the computational domain if the zero-flux condition is imposed). We naturally set

$$\mathcal{F}^{ML,+}_j = \mathcal{F}^+(\rho^-_j, c_j, u_j) \text{ and } \mathcal{F}^{ML,-}_j = \mathcal{F}^-(\rho^+_j, c_j, u_j).$$

We also introduce the notation

$$\mathcal{F}^{ML,\pm}_{j+\frac{1}{2}} = \frac{\mathcal{F}^{ML,\pm}_j + \mathcal{F}^{ML,\pm}_{j+1}}{2}, \quad \text{and} \quad \mathcal{F}^{ML}_{j+\frac{1}{2}} = \mathcal{F}^{ML,+}_{j+\frac{1}{2}} + \mathcal{F}^{ML,-}_{j+\frac{1}{2}}, \tag{31}$$

so that the mass balance on dual mesh (14) remain valid when replacing the mass flux $\mathcal{F}_{j+\frac{1}{2}}$ by the MUSCL-flux $\mathcal{F}^{ML}_{j+\frac{1}{2}}$.

The convection part of the momentum flux is given by

$$\mathcal{G}^{ML}_{j+\frac{1}{2}} = u^-_{j+\frac{1}{2}} \frac{\mathcal{F}^{ML,+}_j + \mathcal{F}^{ML,+}_{j+1}}{2} + u^+_{j+\frac{1}{2}} \frac{\mathcal{F}^{ML,-}_j + \mathcal{F}^{ML,-}_{j+1}}{2}, \qquad \forall j \in [\![2, J-1]\!]. \tag{32}$$

We set $\mathcal{G}^{ML}_{\frac{3}{2}} = \dfrac{u^+_{\frac{3}{2}}}{2}\mathcal{F}^{ML,-}_2$ and $\mathcal{G}^{ML}_{J+\frac{1}{2}} = \dfrac{u^-_{J+\frac{1}{2}}}{2}\mathcal{F}^{ML,+}_J$ for the boundary values.

Finally, when considering the full Euler equations, the internal energy flux is given by

$$\mathcal{E}^{ML}_j = \frac{(\rho e)^-_j}{\rho^-_j}\mathcal{F}^{ML,+}_j + \frac{(\rho e)^+_j}{\rho^+_j}\mathcal{F}^{ML,-}_j.$$

We remind the reader that the MUSCL procedure is not applied directly to the internal energy but to the pressure: we evaluate $(\rho e)^\pm$ and then divide by $\rho^\pm$. This is motivated by the will to obtain a correct treatment of contact discontinuities. Since we wish to satisfy the criterion "if the pressure (and the velocity) is constant in the neigborhood of a point at a certain time, it will be kept constant at the following time", it is quite natural to work on the pressure and not the internal energy. The remainder term reads

$$\begin{aligned}
\delta x_j \mathcal{R}^{ML}_j \;=\; & \frac{\delta x_j}{2\delta t}\overline{\rho}_j\,(\overline{u}_j - u_j)^2 \\
& + \frac{1}{2}\Big((u_j - u^-_{j-\frac{1}{2}})^2\mathcal{F}^{ML,+}_{j-\frac{1}{2}} - (u^+_{j+\frac{1}{2}} - u_j)^2\mathcal{F}^{ML,-}_{j+\frac{1}{2}} \\
& \qquad -(u_j - u^-_{j+\frac{1}{2}})^2\mathcal{F}^{ML,+}_{j+\frac{1}{2}} + (u^+_{j-\frac{1}{2}} - u_j)^2\mathcal{F}^{ML,-}_{j-\frac{1}{2}}\Big) \\
& + (\overline{u}_j - u_j)\Big((u_j - u^-_{j-\frac{1}{2}})\mathcal{F}^{ML,+}_{j-\frac{1}{2}} + (u^+_{j+\frac{1}{2}} - u_j)\mathcal{F}^{ML,-}_{j+\frac{1}{2}} \\
& \qquad -(u_j - u^-_{j+\frac{1}{2}})\mathcal{F}^{ML,+}_{j+\frac{1}{2}} - (u^+_{j-\frac{1}{2}} - u_j)\mathcal{F}^{ML,-}_{j-\frac{1}{2}}\Big).
\end{aligned} \tag{33}$$

3.3 Stability conditions

Firstly, we exhibit a CFL-condition which ensures that the numerical density remains non-negative. This condition should be fulfilled for the approximation of the solutions of both barotropic and full Euler equations. Next, we discuss the condition required to ensure the non-negativity of the internal energy when considering the full Euler system. We will make use of the properties (19) of the flux functions $\mathcal{F}^\pm$ recalled in Lemma 2.1 in Section 2.4.

Proposition 3.1 (Non negativity of the density)**.** *Suppose that the limiter function* Φ *satisfy* (26) *and that* $\rho_{j+\frac{1}{2}} \geqslant 0$ *holds for all* $j \in [\![1, J]\!]$*. We assume the CFL-like condition*

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}}\left([\lambda_-(c_j, u_j)]^- + [\lambda_+(c_{j+1}, u_{j+1})]^+\right) \leqslant \frac{1}{2}, \quad \forall j \in [\![1, J]\!], \tag{34}$$

Then the scheme preserves the non-negativity of the density:

$$\overline{\rho}_{j+\frac{1}{2}} \geqslant 0 \text{ for all } j \in [\![1, J]\!].$$

Proof. We assume that $\rho_{j+\frac{1}{2}} \geqslant 0$ holds for all $j \in [\![1, J]\!]$. Let us introduce the following quantities

$$\alpha_j = \frac{\delta x_{j+\frac{1}{2}}}{2\delta x_j}\;\Phi\left(\frac{\rho_{j+\frac{3}{2}} - \rho_{j+\frac{1}{2}}}{\delta x_{j+1}}\frac{\delta x_j}{\rho_{j+\frac{1}{2}} - \rho_{j-\frac{1}{2}}}\right),$$

and

$$\beta_j = \frac{\delta x_{j+\frac{1}{2}}}{2\delta x_{j+1}}\;\Phi\left(\frac{\rho_{j+\frac{1}{2}} - \rho_{j-\frac{1}{2}}}{\delta x_j}\frac{\delta x_{j+1}}{\rho_{j+\frac{3}{2}} - \rho_{j+\frac{1}{2}}}\right).$$

Owing to property (26), we readily check that $0 \leqslant \alpha_j \leqslant 1$ and $0 \leqslant \beta_j \leqslant 1$. Furthermore, the reconstructed densities can be equivalently recast as

$$\rho_j^+ = (1-\alpha_j)\rho_{j+\frac{1}{2}} + \alpha_j \rho_{j-\frac{1}{2}} \quad \text{and} \quad \rho_{j+1}^- = (1+\alpha_j)\rho_{j+\frac{1}{2}} - \alpha_j \rho_{j-\frac{1}{2}}, \tag{35}$$

or

$$\rho_j^+ = (1+\beta_j)\rho_{j+\frac{1}{2}} - \beta_j \rho_{j+\frac{3}{2}} \quad \text{and} \quad \rho_{j+1}^- = (1-\beta_j)\rho_{j+\frac{1}{2}} + \beta_j \rho_{j+\frac{3}{2}}. \tag{36}$$

In particular, equalities (35) show that

$$\rho_j^+ \geqslant \min\left(\rho_{j-\frac{1}{2}}, \rho_{j+\frac{1}{2}}\right) \geqslant 0, \quad \text{and} \quad \rho_{j+1}^- \leqslant 2\rho_{j+\frac{1}{2}},$$

and equalities (36) show that

$$\rho_j^+ \leqslant 2\rho_{j+\frac{1}{2}}, \quad \text{and} \quad \rho_{j+1}^- \geqslant \min\left(\rho_{j-\frac{1}{2}}, \rho_{j+\frac{1}{2}}\right) \geqslant 0.$$

Reasoning now as in [5, Lemma 3.7], using the sign property of the flux functions $\pm\mathcal{F}^{\pm} \geqslant 0$, we are led to the following estimate

$$\bar{\rho}_{j+\frac{1}{2}} \geqslant \rho_{j+\frac{1}{2}} + \frac{\delta t}{\delta x_{j+\frac{1}{2}}}\left(\mathcal{F}^-(\rho_j^+, c_j, u_j) - \mathcal{F}^+(\rho_{j+1}^-, c_{j+1}, u_{j+1})\right).$$

Owing to equation (19) and since $\rho_j^+ \geqslant 0$ and $\rho_{j+1}^- \geqslant 0$, we obtain

$$\overline{\rho}_{j+\frac{1}{2}} \geqslant \rho_{j+\frac{1}{2}} - \frac{\delta t}{\delta x_{j+\frac{1}{2}}}\left(\rho_j^+ \left[\lambda_-(c_j, u_j)\right]^- + \rho_{j+1}^- \left[\lambda_+(c_{j+1}, u_{j+1})\right]^+\right).$$

Next, bearing in mind that $\rho_j^+ \leqslant 2\rho_{j+\frac{1}{2}}$ and $\rho_{j+1}^- \leqslant 2\rho_{j+\frac{1}{2}}$, we find

$$\overline{\rho}_{j+\frac{1}{2}} \geqslant \rho_{j+\frac{1}{2}}\left(1 - \frac{2\delta t}{\delta x_{j+\frac{1}{2}}}\left(\left[\lambda_-(c_j, u_j)\right]^- + \left[\lambda_+(c_{j+1}, u_{j+1})\right]^+\right)\right).$$

Since it is assumed that $\rho_{j+\frac{1}{2}} \geqslant 0$, the conclusion $\overline{\rho}_{j+\frac{1}{2}} \geqslant 0$ is obtained as a consequence of (34). ∎

Remark 3.2. It is worth pointing out that the CFL condition for the MUSCL scheme is twice more constrained than with the first order scheme in [5, Prop. 3.7]. This is due to the estimate $\rho_j^+ \leqslant 2\rho_{j+\frac{1}{2}}$ and $\rho_{j+1}^- \leqslant 2\rho_{j+\frac{1}{2}}$.

Proposition 3.3 (Non negativity of the internal energy)**.** *Assume that $e_{j+\frac{1}{2}} \geqslant 0$, $\rho_{j+\frac{1}{2}} \geqslant 0$ and that the following CFL-like conditions hold for any j*

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}}\left([u_{j+1}]^+ + \frac{c_f(e_{j+\frac{3}{2}}) + c_f(e_{j+\frac{1}{2}})}{\sqrt{2}} + [u_j]^- + \frac{c_f(e_{j+\frac{1}{2}}) + c_f(e_{j-\frac{1}{2}})}{\sqrt{2}}\right) \leqslant \frac{1}{\gamma+3}, \tag{37}$$

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}} c_f(e_{j+\frac{1}{2}+k}) \leqslant \frac{\gamma-1}{2\sqrt{2}} \cdot \frac{\gamma}{\gamma+3}, \quad \forall k \in \{-1, 0, 1\}. \tag{38}$$

Then, we can find $\lambda_j \in [0,1]$ (see formula (44)) such that $\overline{e}_{j+\frac{1}{2}} \geqslant 0$.

Note that, in comparison to the first order scheme, see Proposition 2.2, the time step is more constrained by a factor $0 < \frac{\gamma}{\gamma+2} < 1$.

Proof. We assume that $e_{j+\frac{1}{2}} \geqslant 0$, $\rho_{j+\frac{1}{2}} \geqslant 0$ and that (37) and (38) holds for all j. Note that, owing to inequalities (22) and (23), the condition (37) implies (34), so that we have $\overline{\rho}_{j+\frac{1}{2}} \geqslant 0$.

We also recall that the reconstructed quantities satisfy the following inequalities (see the proof of Proposition 3.1)

$$0 < \rho^-_{j+1}, \rho^+_j < 2\rho_{j+\frac{1}{2}} \text{ and } 0 < (\rho e)^-_{j+1}, (\rho e)^+_j < 2\rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}}. \tag{39}$$

In order to analyze the positivity of the internal energy, we go back to the evolution of the discrete kinetic energy, which now reads

$$\frac{1}{2}\frac{\overline{\rho}_j \overline{u}_j^2 - \rho_j u_j^2}{\delta t} + \frac{\mathcal{K}^{ML}_{j+\frac{1}{2}} - \mathcal{K}^{ML}_{j-\frac{1}{2}}}{\delta x_j} + \frac{\Pi_{j+\frac{1}{2}} - \Pi_{j-\frac{1}{2}}}{\delta x_j}\overline{u}_j = -\mathcal{R}^{ML}_j,$$

where

$$\mathcal{K}^{ML}_{j+\frac{1}{2}} = \frac{|u^-_{j+\frac{1}{2}}|^2}{2}\mathcal{F}^{ML,+}_{j+\frac{1}{2}} + \frac{|u^+_{j+\frac{1}{2}}|^2}{2}\mathcal{F}^{ML,-}_{j+\frac{1}{2}}.$$

The remainder term $\mathcal{R}^{ML}_j$ is given by (33). The end of the proof of Proposition 3.3 relies on the following claim.

Lemma 3.4. *We can find $\lambda_j \in [0,1]$ such that the following inequality holds:*

$$\delta x_j \mathcal{R}^{ML}_j \geq \frac{\delta x_j}{2\delta t}(\overline{u}_j - u_j)^2 \left(\rho_j - 4\frac{\delta t}{\delta x_j}(\mathcal{F}^{ML,+}_{j+\frac{1}{2}} - \mathcal{F}^{ML,-}_{j-\frac{1}{2}})\right). \tag{40}$$

Let us temporarily assume that Lemma 3.4 holds and defines the coefficients λ_j needed for the construction of the discrete velocities. As in the proof of Proposition 2.2, we write $\overline{\rho}_{j+\frac{1}{2}}\overline{e}_{j+\frac{1}{2}} \geqslant \mathrm{T}_0 + \mathrm{T}^0_1 + \mathrm{T}^1_1$ where:

$$\mathrm{T}_0 = \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} \left(1 - \frac{\delta t}{\delta x_{j+\frac{1}{2}}}(\gamma - 1)\left(2\frac{c(e_{j+\frac{1}{2}})}{\sqrt{2}} - u_j + u_{j+1}\right)\right) - \delta t \frac{\mathcal{E}^{ML}_{j+1} - \mathcal{E}^{ML}_j}{\delta x_{j+\frac{1}{2}}},$$

$$\mathrm{T}^i_1 = \frac{\delta t}{2}\frac{\delta x_{j+i}}{\delta x_{j+\frac{1}{2}}}\mathcal{R}^{ML}_{j+i} - \frac{\delta t}{\delta x_{j+\frac{1}{2}}}\frac{c(e_{j+\frac{1}{2}})}{2\sqrt{2}\gamma}\rho_{j+\frac{1}{2}}(\overline{u}_{j+i} - u_{j+i})^2.$$

Thus, to guarantee that $\overline{e}_{j+\frac{1}{2}}$ is non negative it is sufficient to ensure that these three terms are non negative.

We first consider T_0. Using the definition of the flux $\mathcal{E}^{ML}_j$ and owing to equations (19) and (39), we obtain

$$\begin{aligned}\mathrm{T}_0 \geqslant \rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}} &\left(1 - \frac{\delta t}{\delta x_{j+\frac{1}{2}}}(\gamma - 1)\left([u_j]^- + \frac{c(e_{j+\frac{1}{2}})}{\sqrt{2}} + [u_{j+1}]^+ + \frac{c(e_{j+\frac{1}{2}})}{\sqrt{2}}\right)\right) \\ &- 2\frac{\delta t}{\delta x_{j+\frac{1}{2}}}\rho_{j+\frac{1}{2}} e_{j+\frac{1}{2}}\left([\lambda_+(c_{j+1}, u_{j+1})]^+ + [\lambda_-(c_j, u_j)]^-\right)\end{aligned}$$

where, due to (37), the right hand side is non negative by virtue of (22) and (23).

Next, we turn to T^i_1. Owing to Lemma 3.4, we have

$$\frac{\delta t}{2}\frac{\delta x_j}{\delta x_{j+\frac{1}{2}}}\mathcal{R}^{ML}_j \geq \frac{\delta x_j}{4\delta x_{j+\frac{1}{2}}}(\overline{u}_j - u_j)^2\left(\rho_j - 4\frac{\delta t}{\delta x_j}(\mathcal{F}^{ML,+}_{j+\frac{1}{2}} - \mathcal{F}^{ML,-}_{j-\frac{1}{2}})\right).$$

Hence, we deduce

$$\mathrm{T}^i_1 \geqslant \frac{\delta x_{j+i}}{4\delta x_{j+\frac{1}{2}}}(\overline{u}_{j+i} - u_{j+i})^2\left(\rho_{j+i} - 4\frac{\delta t}{\delta x_{j+i}}(\mathcal{F}^{ML,+}_{j+i+\frac{1}{2}} - \mathcal{F}^{ML,-}_{j+i-\frac{1}{2}}) - \frac{\delta t}{\gamma}\frac{2}{\sqrt{2}}\frac{\rho_{j+\frac{1}{2}} c(e_{j+\frac{1}{2}})}{\delta x_{j+i}}\right).$$

Coming back to (7) and (31), we write $T_1^i \geqslant \dfrac{(\overline{u}_{j+i}-u_{j+i})^2}{4\delta x_{j+\frac12}}\left(T_2^{i,0}+T_2^{i,1}\right)$ where, for $k=0,1$,

$$\mathrm{T}_2^{i,k}=\frac{\delta x_{j+i+k-\frac12}}{2}\rho_{j+i+k-\frac12}-4\frac{\mathcal{F}^{ML,+}_{j+i+k}-\mathcal{F}^{ML,-}_{j+i+k-1}}{2}-\frac{\delta t}{\gamma}\frac{2}{\sqrt{2}}\rho_{j+i+k-\frac12}c(e_{j+\frac12}),$$

Note that a non negative term has been added to obtain a symmetric formulation in the above inequality. Due to equation (19) and (37), we get

$$\mathcal{F}^{ML,+}_{j+i+k}-\mathcal{F}^{ML,-}_{j+i+k-1}\leqslant\frac{\delta x_{j+i+k-\frac12}}{(\gamma+2)\delta t}\rho_{j+i+k-\frac12},$$

and this allows us to write

$$\mathrm{T}_2^{i,k}\geqslant\frac{\delta x_{j+i+k-\frac12}}{2}\rho_{j+i+k-\frac12}\left(1-\frac{4}{\gamma+3}-\frac{\delta t}{\delta x_{j+i+k-\frac12}}\frac{2\sqrt{2}}{\gamma}c(e_{j+\frac12})\right).$$

We conclude by observing that this term is non negative by virtue of (38). ∎

We now go back to the proof of Lemma 3.4.

Proof of Lemma 3.4. We go back to (33). For given coefficients $\alpha_j\geqslant 0$, that will be determined later on, we shall use the following Young inequalities

$$\left|(\overline{u}_j-u_j)(u_j-u^-_{j-\frac12})\mathcal{F}^{ML,+}_{j-\frac12}\right|\leqslant\frac{(\overline{u}_j-u_j)^2}{2}(1+\alpha_j)\mathcal{F}^{ML,+}_{j-\frac12}+\frac{(u_j-u^-_{j-\frac12})^2}{2(1+\alpha_j)}\mathcal{F}^{ML,+}_{j-\frac12}$$

and

$$\left|(\overline{u}_j-u_j)(u^+_{j+\frac12}-u_j)\mathcal{F}^{ML,-}_{j+\frac12}\right|\leqslant-\frac{(\overline{u}_j-u_j)^2}{2}(1+\alpha_j)\mathcal{F}^{ML,-}_{j+\frac12}-\frac{(u_j-u^+_{j+\frac12})^2}{2(1+\alpha_j)}\mathcal{F}^{ML,-}_{j+\frac12}.$$

Using (14), which still holds for the MUSCL version of the scheme, and the standard Young inequality for the last four terms in (33), we are led to

$$\delta x_j\mathcal{R}_j^{ML}\geqslant\frac{\delta x_j}{2\delta t}(\overline{u}_j-u_j)^2R_j^{(1)}+R_j^{(2)},$$

with

$$\begin{aligned}
R_j^{(1)}&=\rho_j-\frac{\delta t}{\delta x_j}\left(\mathcal{F}^{ML,+}_{j+\frac12}-\mathcal{F}^{ML,-}_{j-\frac12}-\mathcal{F}^{ML,+}_{j-\frac12}+\mathcal{F}^{ML,-}_{j+\frac12}\right)\\
&\quad-\frac{\delta t}{\delta x_j}\left((\mathcal{F}^{ML,+}_{j+\frac12}-\mathcal{F}^{ML,-}_{j-\frac12})+(1+\alpha_j)(\mathcal{F}^{ML,+}_{j-\frac12}-\mathcal{F}^{ML,-}_{j+\frac12})\right),\\
R_j^{(2)}&=\frac{\alpha_j}{2(1+\alpha_j)}\left((u_j-u^-_{j-\frac12})^2\mathcal{F}^{ML,+}_{j-\frac12}-(u^+_{j+\frac12}-u_j)^2\mathcal{F}^{ML,-}_{j+\frac12}\right)\\
&\quad+\left((u^+_{j-\frac12}-u_j)^2\mathcal{F}^{ML,-}_{j-\frac12}-(u_j-u^-_{j+\frac12})^2\mathcal{F}^{ML,+}_{j+\frac12}\right).
\end{aligned}$$

We rewrite

$$R_j^{(1)}=\left(\rho_j-4\frac{\delta t}{\delta x_j}\left(\mathcal{F}^{ML,+}_{j+\frac12}-\mathcal{F}^{ML,-}_{j-\frac12}\right)\right)+\frac{\delta t}{\delta x_j}\left(2(\mathcal{F}^{ML,+}_{j+\frac12}-\mathcal{F}^{ML,-}_{j-\frac12})-\alpha_j(\mathcal{F}^{ML,+}_{j-\frac12}-\mathcal{F}^{ML,-}_{j+\frac12})\right),$$

so that we get

$$R_j^{(1)}\geqslant\rho_j-4\frac{\delta t}{\delta x_j}\left(\mathcal{F}^{ML,+}_{j+\frac12}-\mathcal{F}^{ML,-}_{j-\frac12}\right)$$

as soon as

$$\alpha_j \leqslant 2A_j \qquad \text{with } A_j = \frac{\mathcal{F}^{ML,+}_{j+\frac{1}{2}} - \mathcal{F}^{ML,-}_{j-\frac{1}{2}}}{\mathcal{F}^{ML,+}_{j-\frac{1}{2}} - \mathcal{F}^{ML,-}_{j+\frac{1}{2}}}. \tag{41}$$

Note that the quantity A_j is well defined and always positive since

$$\mathcal{F}^{ML,+}_{j\pm\frac{1}{2}} - \mathcal{F}^{ML,-}_{j\mp\frac{1}{2}} \geqslant (\mathcal{F}^{ML,+}_{j} - \mathcal{F}^{ML,-}_{j})/2 > 0.$$

We now turn to the study of $R^{(2)}_j$. Using the following shorthand notation for the discrete derivative

$$a_{j+\frac{1}{2}} = \frac{u_{j+1} - u_j}{\delta x_{j+\frac{1}{2}}},$$

we remark on one hand that

$$\begin{aligned}
(u_j - u^-_{j-\frac{1}{2}})^2 &= \delta x^2_{j-\frac{1}{2}} \left(a_{j-\frac{1}{2}} - \frac{\lambda_{j-1}}{2} \widehat{\Phi}\Big(a_{j-\frac{3}{2}}, a_{j-\frac{1}{2}} \Big) \right)^2 \\
&\geqslant \delta x^2_{j-\frac{1}{2}} a^2_{j-\frac{1}{2}} \left(1 - \frac{\lambda_{j-1}}{2} \tau \right)^2 \\
&\geqslant \delta x^2_{j-\frac{1}{2}} a^2_{j,m} \left(1 - \frac{\tau}{2} \right)^2
\end{aligned}$$

with $a_{j,m} = \min\Big\{ |a_{j-\frac{1}{2}}|, |a_{j+\frac{1}{2}}| \Big\} \geqslant 0$. Similarly, we have

$$(u^+_{j+\frac{1}{2}} - u_j)^2 \geqslant \delta x^2_{j+\frac{1}{2}} a^2_{j,m} \left(1 - \frac{\tau}{2} \right)^2,$$

so that we get

$$\begin{aligned}
&(u_j - u^-_{j-\frac{1}{2}})^2 \mathcal{F}^{ML,+}_{j-\frac{1}{2}} - (u^+_{j+\frac{1}{2}} - u_j)^2 \mathcal{F}^{ML,-}_{j+\frac{1}{2}} \\
&\qquad\qquad \geqslant \Big(\delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,+}_{j-\frac{1}{2}} - \delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,-}_{j+\frac{1}{2}} \Big) a^2_{j,m} \left(1 - \frac{\tau}{2} \right)^2.
\end{aligned} \tag{42}$$

On the other hand, since $|\widehat{\Phi}(a,b)| \leqslant \tau \min(|a|,|b|)$, we have

$$\begin{aligned}
(u^+_{j-\frac{1}{2}} - u_j)^2 &= \frac{\lambda_j^2}{4} \delta x^2_{j-\frac{1}{2}} \widehat{\Phi}\Big(a_{j-\frac{1}{2}}, a_{j+\frac{1}{2}} \Big)^2 \leqslant \frac{\lambda_j^2}{4} \delta x^2_{j-\frac{1}{2}} \tau^2 a^2_{j,m}, \\
(u_j - u^-_{j+\frac{1}{2}})^2 &= \frac{\lambda_j^2}{4} \delta x^2_{j+\frac{1}{2}} \widehat{\Phi}\Big(a_{j-\frac{1}{2}}, a_{j+\frac{1}{2}} \Big)^2 \leqslant \frac{\lambda_j^2}{4} \delta x^2_{j+\frac{1}{2}} \tau^2 a^2_{j,m},
\end{aligned}$$

so that we have

$$(u^+_{j-\frac{1}{2}} - u_j)^2 \mathcal{F}^{ML,-}_{j-\frac{1}{2}} - (u_j - u^-_{j+\frac{1}{2}})^2 \mathcal{F}^{ML,+}_{j+\frac{1}{2}} \geqslant -\frac{\lambda_j^2}{4} \tau^2 a^2_{j,m} \Big(\delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,+}_{j+\frac{1}{2}} - \delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,-}_{j-\frac{1}{2}} \Big). \tag{43}$$

Thus combining (42) and (43) we arrive at

$$\begin{aligned}
R^{(2)}_j \geqslant & \frac{\alpha_j}{2(1+\alpha_j)} \Big(\delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,+}_{j-\frac{1}{2}} - \delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,-}_{j+\frac{1}{2}} \Big) a^2_{j,m} \left(1 - \frac{\tau}{2} \right)^2 \\
& \qquad - \frac{\lambda_j^2}{4} \tau^2 a^2_{j,m} \Big(\delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,+}_{j+\frac{1}{2}} - \delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,-}_{j-\frac{1}{2}} \Big) \\
\geqslant & \Big(\delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,+}_{j-\frac{1}{2}} - \delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,-}_{j+\frac{1}{2}} \Big) a^2_{j,m} \left(\frac{\alpha_j}{2(1+\alpha_j)} \left(1 - \frac{\tau}{2} \right)^2 - \frac{\lambda_j^2}{4} \tau^2 B_j \right),
\end{aligned}$$

with

$$B_j = \frac{\delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,+}_{j+\frac{1}{2}} - \delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,-}_{j-\frac{1}{2}}}{\delta x^2_{j-\frac{1}{2}} \mathcal{F}^{ML,+}_{j-\frac{1}{2}} - \delta x^2_{j+\frac{1}{2}} \mathcal{F}^{ML,-}_{j+\frac{1}{2}}}$$

Then, $R_j^{(2)} \geqslant 0$ provided

$$\lambda_j^2 \Big(\frac{\tau}{2}\Big)^2 B_j \leqslant \frac{\alpha_j}{2(1+\alpha_j)} \Big(\frac{2-\tau}{2}\Big)^2.$$

To relax the condition on λ_j as far as possible, and bearing in mind the condition (41) we set

$$\alpha_j = 2A_j$$

so that we are reduced to check that

$$\lambda_j^2 \leqslant \Big(\frac{2-\tau}{\tau}\Big)^2 \frac{A_j}{(1+2A_j)B_j}.$$

Therefore it is sufficient to set

$$\lambda_j = \frac{2-\tau}{\tau} \frac{1}{\sqrt{1+2A_j}} \sqrt{\frac{A_j}{B_j}}. \tag{44}$$

■

Remark 3.5. The case $\lambda_j = 1$ corresponds to the second order (except for extrema, where the gradient is zero and the scheme degenerates to order 1 due to the limiter) while $\lambda_j = 0$ corresponds to the first order. Note that assumption (27) on the limiter allows us to make use of positive values of λ_j since τ is assumed to be strictly less than 2. On uniform grids, equation (44) becomes

$$\lambda_j = \frac{2-\tau}{\tau} \frac{1}{\sqrt{1+2A_j}},$$

so that the coefficient λ_j satisfies $0 < \lambda_j < 1$. However, setting $\lambda_j < 1$ implies a loss of accuracy compared to the pure second-order MUSCL scheme. Thus, in practice, we adopt the following strategy:

- we first set $\lambda_j = 1$ and compute the discrete internal energies $e_{j+\frac{1}{2}}$.
- if we obtain a negative value for $e_{j+\frac{1}{2}}$, meaning that (40) is not satisfied for $\mathcal{R}_j$ and/or $\mathcal{R}_{j+1}$, we modify the value of λ_j and/or λ_{j+1} following (44) to ensure (40) for both $\mathcal{R}_j$ and $\mathcal{R}_{j+1}$.
- we compute the new values for $\overline{u}_{j-1}$, $\overline{u}_j$, $\overline{u}_{j+1}$, $\mathcal{R}_{j-1}$, $\mathcal{R}_j$, $\mathcal{R}_{j+1}$, $e_{j-\frac{1}{2}}$, $e_{j+\frac{1}{2}}$ and/or $\overline{u}_j$, $\overline{u}_{j+1}$, $\overline{u}_{j+2}$, $\mathcal{R}_j$, $\mathcal{R}_{j+1}$, $\mathcal{R}_{j+2}$, $e_{j+\frac{1}{2}}$, $e_{j+\frac{3}{2}}$.

Nevertheless, we observe that, in practice, it is quite infrequent in the simulations that the criterion that requires to reduce λ_j is activated. Indeed, none the simulations that we are going to show in the last section of this chapter has needed this criterion. Besides, we point out that the scheme designed in [48, Chapter 2] adapts a MUSCL reconstruction on the density and on the internal energy fluxes only, the convection fluxes for the momentum equation remain the first order fluxes (that corresponds here to always set $\lambda_j = 0$). Note also that the stability conditions in [48] are slightly different, precisely because the fluxes in this scheme are based on the material velocity only. However, it requires in certain circumstances (vanishing velocities, low Mach regimes) to introduce artificial diffusion.

3.4 Consistency of the scheme

Let us briefly check the consistency of the scheme, showing it can reach the second order accuracy for smooth solutions, and far away from extrema (since otherwise the limiter reduces the order of the approximation) when λ_j is set to 1 for all j in (30). To this end, we study (at a fixed time) the consistency of the fluxes. The time being fixed, we consider smooth functions $\underline{\rho}$, $\underline{u}$ and $\underline{e}$ (when considering full Euler system) of the space variable x only (say of class C^1 with bounded and not vanishing derivatives). We set $\underline{\rho}_{j+\frac12} = \underline{\rho}(x_{j+\frac12})$, $\underline{e}_{j+\frac12} = \underline{e}(x_{j+\frac12})$ and $\underline{u}_j = \underline{u}(x_j)$ and insert these quantities in the scheme instead of $\rho_{j+\frac12}$, $e_{j+\frac12}$ and u_j. We denote with an underline all the quantities (slopes, reconstructed densities, pressure and velocities, fluxes...) defined in this way from $\underline{\rho}_{j+\frac12}$, $\underline{\rho e}_{j+\frac12}$ and $\underline{u}_j$. The first observation, stated in Lemma 3.6, is that the reconstructed densities $\underline{\rho}_j^{\pm}$ and velocities $\underline{u}_{j+\frac12}^{\pm}$ are second order approximations of $\underline{\rho}(x_j)$ and $\underline{u}(x_{j+\frac12})$, respectively.

Lemma 3.6. *The following equalities hold (when λ_j is set to 1 for all j in (30)):*

$$\underline{\rho}_j^+ = \underline{\rho}(x_j) + \mathcal{O}(\delta x^2),\ \forall j \in [\![1, J-1]\!], \qquad \underline{\rho}_j^- = \underline{\rho}(x_j) + \mathcal{O}(\delta x^2),\ \forall j \in [\![2, J]\!], \tag{45}$$

$$\underline{u}_{j+\frac12}^+ = \underline{u}(x_{j+\frac12}) + \mathcal{O}(\delta x^2),\ \forall j \in [\![1, J-1]\!],\ \underline{u}_{j+\frac12}^- = \underline{u}(x_{j+\frac12}) + \mathcal{O}(\delta x^2),\ \forall j \in [\![2, J]\!], \tag{46}$$

and, when considering full Euler system,

$$\underline{\rho e}_j^+ = \underline{\rho e}(x_j) + \mathcal{O}(\delta x^2) \qquad \text{and} \qquad \underline{\rho e}_j^- = \underline{\rho e}(x_j) + \mathcal{O}(\delta x^2). \tag{47}$$

Proof. We first prove that

$$\underline{s}_{j+\frac12} = \underline{\rho}'(x_{j+\frac12}) + \mathcal{O}(\delta x). \tag{48}$$

Indeed, we clearly have

$$\frac{\underline{\rho}_{j+\frac12} - \underline{\rho}_{j-\frac12}}{\delta x_j} = \underline{\rho}'(x_{j+\frac12}) + \mathcal{O}(\delta x), \quad \text{and} \quad \frac{\underline{\rho}_{j+\frac32} - \underline{\rho}_{j+\frac12}}{\delta x_{j+1}} = \underline{\rho}'(x_{j+\frac12}) + \mathcal{O}(\delta x),$$

so that

$$\frac{\underline{\rho}_{j+\frac32} - \underline{\rho}_{j+\frac12}}{\delta x_{j+1}} \frac{\delta x_j}{\underline{\rho}_{j+\frac12} - \underline{\rho}_{j-\frac12}} = 1 + \mathcal{O}(\delta x).$$

Since $\Phi(1) = 1$ and $r \mapsto \Phi(r)$ admits left and right derivatives at the point $r = 1$ (*cf.* assumption (28)), we get

$$\Phi\Big(\frac{\underline{\rho}_{j+\frac32} - \underline{\rho}_{j+\frac12}}{\delta x_{j+1}} \frac{\delta x_j}{\underline{\rho}_{j+\frac12} - \underline{\rho}_{j-\frac12}}\Big) = 1 + \mathcal{O}(\delta x).$$

This last equality together with the definition of $\underline{s}_{j+\frac12}$

$$\underline{s}_{j+\frac12} = \frac{\underline{\rho}_{j+\frac12} - \underline{\rho}_{j-\frac12}}{\delta x_j} \Phi\Big(\frac{\underline{\rho}_{j+\frac32} - \underline{\rho}_{j+\frac12}}{\delta x_{j+1}} \frac{\delta x_j}{\underline{\rho}_{j+\frac12} - \underline{\rho}_{j-\frac12}}\Big),$$

proves (48). Next, from (48) and the definition of $\underline{\rho}_j^{\pm}$ we readily find, for all $j \in [\![1, J]\!]$,

$$\underline{\rho}_{j+1}^- = \underline{\rho}_{j+\frac12} + \frac{\delta x_{j+\frac12}}{2}\, \underline{\rho}'(x_{j+\frac12}) + \mathcal{O}(\delta x^2),$$
$$\underline{\rho}_j^+ = \underline{\rho}_{j+\frac12} - \frac{\delta x_{j+\frac12}}{2}\, \underline{\rho}'(x_{j+\frac12}) + \mathcal{O}(\delta x^2).$$

The conclusion is then obtained using the following identities, direct consequences of the Taylor-Young expansion,

$$\underline{\rho}(x_{j+1}) = \underline{\rho}_{j+\frac{1}{2}} + \frac{\delta x_{j+\frac{1}{2}}}{2}\ \underline{\rho}'(x_{j+\frac{1}{2}}) + \mathcal{O}(\delta x^2)$$

and

$$\underline{\rho}(x_j) = \underline{\rho}_{j+\frac{1}{2}} - \frac{\delta x_{j+\frac{1}{2}}}{2}\ \underline{\rho}'(x_{j+\frac{1}{2}}) + \mathcal{O}(\delta x^2).$$

The equalities for $\underline{u}^{\pm}_{j+\frac{1}{2}}$ or $\underline{\rho e}^{\pm}_{j}$ can be proved by following the same lines. ∎

With Lemma 3.6 at hand, we can now prove that the approximation of the fluxes can reach second order accuracy in space. Concerning the momentum flux, since the pressure is centered we focus on the convective part $\mathcal{G}^{ML}_{j+\frac{1}{2}}$. We can prove the following result.

Proposition 3.7. *The following equalities hold:*

$$\underline{\mathcal{F}}^{ML}_{j} = \underline{\rho}(x_j)\underline{u}(x_j) + \mathcal{O}(\delta x^2). \tag{49}$$

$$\underline{\mathcal{G}}^{ML}_{j+\frac{1}{2}} = \underline{\rho}(x_{j+\frac{1}{2}})\underline{u}(x_{j+\frac{1}{2}})^2 + \mathcal{O}(\delta x^2) \tag{50}$$

When considering the full Euler system, we have also the following results:

$$\underline{\mathcal{E}}^{ML}_{j} = \underline{\rho}(x_j)\underline{u}(x_j)\underline{e}(x_j) + \mathcal{O}(\delta x^2),$$
$$\underline{\mathcal{K}}^{ML}_{j+\frac{1}{2}} = \underline{\rho}(x_{j+\frac{1}{2}})\frac{\underline{u}(x_{j+\frac{1}{2}})^3}{2} + \mathcal{O}(\delta x^2).$$

Proof. By using (11), namely $\mathcal{F}^+(\rho,u) + \mathcal{F}^-(\rho,u) = \rho u$, we start by rewriting the mass flux as follows

$$\underline{\mathcal{F}}^{ML}_{j} = \frac{\underline{\rho}^+_j + \underline{\rho}^-_j}{2}\underline{u}_j + \frac{\mathcal{F}^{|\cdot|}\left(\underline{\rho}^-_j, \underline{u}_j\right) - \mathcal{F}^{|\cdot|}\left(\underline{\rho}^+_j, \underline{u}_j\right)}{2},$$

where the function $\mathcal{F}^{|\cdot|}$ is defined by $\mathcal{F}^{|\cdot|}(\rho,u) = \mathcal{F}^+(\rho,u) - \mathcal{F}^-(\rho,u) \geqslant 0$. Owing to (45), we readily find that

$$\frac{\underline{\rho}^+_j + \underline{\rho}^-_j}{2}\underline{u}_j = \underline{\rho}(x_j)\underline{u}_j + \mathcal{O}(\delta x^2).$$

Furthermore, since the function $(\rho,u) \mapsto \mathcal{F}^{|\cdot|}(\rho,u)$ is of class $\mathcal{C}^1$ (see [5, Lemma 3.3]), we have

$$\mathcal{F}^{|\cdot|}\left(\underline{\rho}^{\pm}_j, \underline{u}_j\right) = \mathcal{F}^{|\cdot|}\Big(\underline{\rho}(x_j), \underline{u}_j\Big) + \mathcal{O}(\delta x^2).$$

Thus, we find

$$\frac{\mathcal{F}^{|\cdot|}\left(\underline{\rho}^-_j, \underline{u}_j\right) - \mathcal{F}^{|\cdot|}\left(\underline{\rho}^+_j, \underline{u}_j\right)}{2} = \mathcal{O}(\delta x^2)$$

and (49) is proved.

We turn to momentum flux. By using (46), and bearing in mind definition (32) of $\underline{\mathcal{G}}^{ML}_{j+\frac{1}{2}}$, we first observe that

$$\begin{aligned}\underline{\mathcal{G}}^{ML}_{j+\frac{1}{2}} &= \underline{u}(x_{j+\frac{1}{2}})\left(\frac{\underline{\mathcal{F}}^{ML,+}_{j} + \underline{\mathcal{F}}^{ML,+}_{j+1}}{2} + \frac{\underline{\mathcal{F}}^{ML,-}_{j} + \underline{\mathcal{F}}^{ML,-}_{j+1}}{2}\right) + \mathcal{O}(\delta x^2),\\ &= \underline{u}(x_{j+\frac{1}{2}})\left(\frac{\underline{\mathcal{F}}^{ML}_{j} + \underline{\mathcal{F}}^{ML}_{j+1}}{2}\right) + \mathcal{O}(\delta x^2).\end{aligned}$$

We then use (49) to find

$$\underline{\mathcal{G}}^{ML}_{j+\frac{1}{2}} = \underline{u}(x_{j+\frac{1}{2}})\left(\frac{\underline{\rho}(x_j)\underline{u}(x_j) + \underline{\rho}(x_{j+1})\underline{u}(x_{j+1})}{2}\right) + \mathcal{O}(\delta x^2).$$

The conclusion (50) is then obtained since we have

$$\frac{\underline{\rho}(x_j)\underline{u}(x_j) + \underline{\rho}(x_{j+1})\underline{u}(x_{j+1})}{2} = \underline{\rho}(x_{j+\frac{1}{2}})\underline{u}(x_{j+\frac{1}{2}}) + \mathcal{O}(\delta x^2).$$

The results for $\underline{\mathcal{E}}^{ML}_j$ and $\underline{\mathcal{K}}^{ML}_{j+\frac{1}{2}}$ are obtained by following the same lines.

■

The second order accuracy can equally be reached with respect to the time variable, by using the Runge-Kutta discretization (RK2) for approximating the time derivative. Note that this approach may lead to further restriction on the time step in order to preserve the positivity of the density.

4 Higher dimensions on mac grids

As far as we restrict to Cartesian grids, our approach can be easily extended to higher dimensions since the discretization can be interpreted by means of the MAC framework. Dealing with general meshes in higher dimension is much more intricate [7, 8, 15, 13, 45] and beyond the scope of the present paper, see [38].

Remark 4.1. It is worth mentioning here the recent work by C. Berthon, Y. Coudière and V. Desveaux [9, 10] who develop a high order scheme on unstructured meshes for the barotropic Euler system by doubling the set of numerical unknowns: the conserved quantities $U = (\rho, \rho\mathbf{u})$ are stored on both the primal and the dual cells. This approach is very appealing in the multi-dimensional case since it provides naturally a way to define full gradients on the interface of the control volumes of an unstructured mesh. Note also that the definition of limiters on general unstructured meshes gives rise to challenging issues, see [15, 12, 13] and the references therein. Here we are only concerned with the simpler situation of Cartesian grids and the scheme does not need to double all variables. Note also that in the present framework it is more adapted to work with the physical quantities ρ and $\mathbf{u}$.

Let us explain how the scheme works in dimension two for the barotropic Euler system. The application to the full Euler system is left to the reader. The computational domain is the square

$$\Omega = [a_x, b_x] \times [a_y, b_y] \subset \mathbb{R}^2,$$

and we thus aim at writing the scheme for the PDE system

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ \rho v \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 + p(\rho) \\ \rho uv \end{pmatrix} + \partial_y \begin{pmatrix} \rho v \\ \rho vu \\ \rho v^2 + p(\rho) \end{pmatrix} = 0.$$

We define the meshes as follows:

- The primal mesh is defined by the points

$$a_x = x_1 < x_2 < ... < x_{i-1} < x_i < x_{i+1} < ... < x_M < x_{M+1} = b_x,$$

and

$$a_y = y_1 < y_2 < ... < y_{j-1} < y_j < y_{j+1} < ... < y_N < y_{N+1} = b_y.$$

- Then we define the midpoints

$$x_{i+\frac{1}{2}} = \frac{x_i + x_{i+1}}{2},\ \forall i \in [\![1, M]\!], \quad \text{and} \quad y_{j+\frac{1}{2}} = \frac{y_j + y_{j+1}}{2},\ \forall j \in [\![1, N]\!].$$

- We set $\delta x_{i+\frac{1}{2}}$, $\delta y_{j+\frac{1}{2}}$, δx_i and δy_j the length of $[x_i, x_{i+1}]$, $[y_j, y_{j+1}]$, $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ and $[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ respectively.

According to the pioneering approach for incompressible flows in [26], we store the discrete densities, the horizontal and the vertical velocities at different locations, see Fig. 3:

- the density ρ is evaluated at the centers of the primal cells: we are dealing with the numerical unknowns $\rho_{i+\frac{1}{2},j+\frac{1}{2}}$,
- the horizontal velocity u is evaluated at the centers of the cells $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_j, y_{j+1}]$: the numerical unknowns thus reads $u_{i,j+\frac{1}{2}}$
- the vertical velocity v is evaluated at the centers of the cells $[x_i, x_{i+1}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$: the numerical unknowns thus reads $v_{i+\frac{1}{2},j}$.

As in 1D, we need an approximation of ρ at the edges of the primal mesh,

$$\rho_{i,j+\frac{1}{2}} = \frac{\delta x_{i+\frac{1}{2}} \rho_{i+\frac{1}{2},j+\frac{1}{2}} + \delta x_{i-\frac{1}{2}} \rho_{i-\frac{1}{2},j+\frac{1}{2}}}{2\delta x_i} \text{ and } \rho_{i+\frac{1}{2},j} = \frac{\delta y_{j+\frac{1}{2}} \rho_{i+\frac{1}{2},j+\frac{1}{2}} + \delta y_{j-\frac{1}{2}} \rho_{i+\frac{1}{2},j-\frac{1}{2}}}{2\delta y_j}.$$

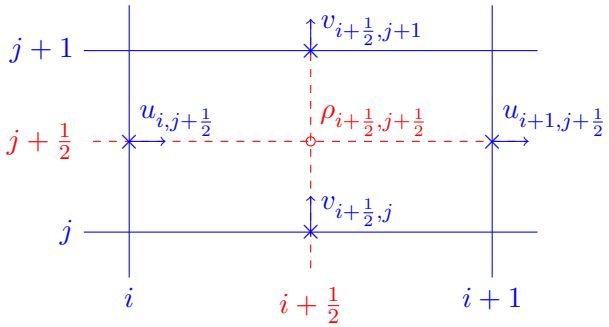


Figure 3: Position of the unknowns on a MAC grid.

The first order scheme is a direct extension of the one proposed in [6] to the 2D framework.

First, the discrete densities $\rho_{i+\frac{1}{2},j+\frac{1}{2}}$, $i \in [\![1, M]\!]$, $j \in [\![1, N]\!]$, are updated using the following explicit scheme

$$\frac{\overline{\rho}_{i+\frac{1}{2},j+\frac{1}{2}} - \rho_{i+\frac{1}{2},j+\frac{1}{2}}}{\delta t} + \frac{\mathcal{F}^x_{i+1,j+\frac{1}{2}} - \mathcal{F}^x_{i,j+\frac{1}{2}}}{\delta x_{i+\frac{1}{2}}} + \frac{\mathcal{F}^y_{i+\frac{1}{2},j+1} - \mathcal{F}^y_{i+\frac{1}{2},j}}{\delta y_{j+\frac{1}{2}}} = 0.$$

Fig. 4 illustrate the following construction: the discrete mass fluxes in the x direction $\mathcal{F}^x_{i,j+\frac{1}{2}}$ are defined, for each value of $j \in [\![1, N]\!]$, as the 1D fluxes, using the values of the horizontal velocity $u_{i,j+\frac{1}{2}}$ to upwind the density in the horizontal direction

$$\begin{aligned} &\mathcal{F}^x_{i,j+\frac{1}{2}} = \mathcal{F}^{x,+}_{i,j+\frac{1}{2}} + \mathcal{F}^{x,-}_{i,j+\frac{1}{2}}, \quad \forall (i,j) \in [\![2, M]\!] \times [\![1, N]\!], \\ &\text{with } \mathcal{F}^{x,+}_{i,j+\frac{1}{2}} = \mathcal{F}^+(\rho_{i-\frac{1}{2},j+\frac{1}{2}}, u_{i,j+\frac{1}{2}}) \quad \text{and} \quad \mathcal{F}^{x,-}_{i,j+\frac{1}{2}} = \mathcal{F}^-(\rho_{i+\frac{1}{2},j+\frac{1}{2}}, u_{i,j+\frac{1}{2}}). \end{aligned} \tag{51}$$

Symmetrically, the mass fluxes $\mathcal{F}^y_{i+\frac{1}{2},j}$ in the y direction are defined using the values of the vertical velocity $v_{i+\frac{1}{2},j}$ to upwind the density in the vertical direction

$$\mathcal{F}^y_{i+\frac{1}{2},j} = \mathcal{F}^{y,+}_{i+\frac{1}{2},j} + \mathcal{F}^{y,-}_{i+\frac{1}{2},j}, \quad \forall (i,j) \in [\![1,M]\!] \times [\![2,N]\!],$$
$$\text{with } \mathcal{F}^{y,+}_{i+\frac{1}{2},j} = \mathcal{F}^{+}(\rho_{i+\frac{1}{2},j-\frac{1}{2}}, v_{i+\frac{1}{2},j}) \quad \text{and} \quad \mathcal{F}^{y,-}_{i+\frac{1}{2},j} = \mathcal{F}^{-}(\rho_{i+\frac{1}{2},j+\frac{1}{2}}, v_{i+\frac{1}{2},j}).$$

For boundary values, we use here zero fluxes: $\mathcal{F}^x_{1,j+\frac{1}{2}} = 0 = \mathcal{F}^x_{M+1,j+\frac{1}{2}}$ and $\mathcal{F}^y_{i+\frac{1}{2},1} = 0 = \mathcal{F}^y_{i+\frac{1}{2},N+1}$.

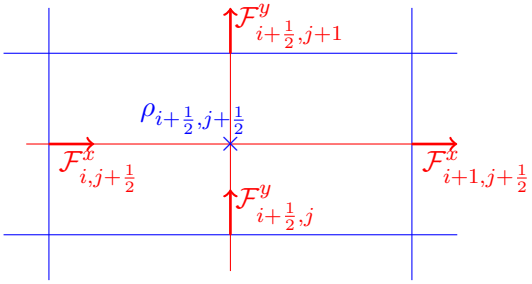


Figure 4: $\mathcal{F}^x$ and $\mathcal{F}^y$ fluxes.

Next, the horizontal velocities $u_{i,j+\frac{1}{2}}$, $i \in [\![2,M]\!]$, $j \in [\![1,N]\!]$ are updated with the following scheme

$$\frac{\overline{\rho}_{i,j+\frac{1}{2}}\overline{u}_{i,j+\frac{1}{2}} - \rho_{i,j+\frac{1}{2}}u_{i,j+\frac{1}{2}}}{\delta t} + \frac{\mathcal{G}^{u,x}_{i+\frac{1}{2},j+\frac{1}{2}} - \mathcal{G}^{u,x}_{i-\frac{1}{2},j+\frac{1}{2}}}{\delta x_i} + \frac{\Pi_{i+\frac{1}{2},j+\frac{1}{2}} - \Pi_{i-\frac{1}{2},j+\frac{1}{2}}}{\delta x_i} + \frac{\mathcal{G}^{u,y}_{i,j+1} - \mathcal{G}^{u,y}_{i,j}}{\delta y_{j+\frac{1}{2}}} = 0. \tag{52}$$

We would define the fluxes $\mathcal{G}^{u,x}_{i+\frac{1}{2},j+\frac{1}{2}}$, resp. $\mathcal{G}^{u,y}_{i,j}$, by upwinding the horizontal momentum $(\rho u)_{i+\frac{1}{2},j+\frac{1}{2}}$, resp. $(\rho u)_{i,j}$, with respect to the value of the horizontal velocity $u_{i+\frac{1}{2},j+\frac{1}{2}}$, resp. the vertical velocity $v_{i,j}$. However, on staggered grids, none of these quantities are obviously defined. As in 1D, we have to bear in mind that, when discretizing the mass conservation equation, we already defined a discrete form of the horizontal, resp. vertical, mass flux based on an upwinding of the density (with respect to the horizontal, resp. vertical, velocity). Thus, the upwinding of horizontal momentum can be next obtained by upwinding the horizontal velocity with respect to the "positive" or "negative" part of the mass fluxes. However, horizontal, resp. vertical, mass fluxes are only defined at points $(x_i, y_{j+\frac{1}{2}})$, resp. $(x_{i+\frac{1}{2}}, y_j)$. The first step is thus to define the "positive" and "negative" parts of the horizontal, resp. vertical, mass flux at points $(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$, resp. (x_i, y_j). This is done by taking the following mean values

$$\mathcal{F}^{x,\pm}_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{2}\left(\mathcal{F}^{x,\pm}_{i,j+\frac{1}{2}} + \mathcal{F}^{x,\pm}_{i+1,j+\frac{1}{2}}\right) \quad \text{and} \quad \mathcal{F}^{y,\pm}_{i,j} = \frac{\delta x_{i+\frac{1}{2}}\mathcal{F}^{y,\pm}_{i+\frac{1}{2},j} + \delta x_{i-\frac{1}{2}}\mathcal{F}^{y,\pm}_{i-\frac{1}{2},j}}{2\delta x_i}.$$

Next, for each $j \in [\![1,N]\!]$, the momentum fluxes $\mathcal{F}^{u,x}_{i+\frac{1}{2},j+\frac{1}{2}}$ are defined, as in 1D, by

$$\mathcal{G}^{u,x}_{i+\frac{1}{2},j+\frac{1}{2}} = u_{i,j+\frac{1}{2}}\mathcal{F}^{x,+}_{i+\frac{1}{2},j+\frac{1}{2}} + u_{i+1,j+\frac{1}{2}}\mathcal{F}^{x,-}_{i+\frac{1}{2},j+\frac{1}{2}}, \quad \forall i \in [\![2,M-1]\!].$$

and

$$\Pi_{i+\frac{1}{2},j+\frac{1}{2}} = p(\rho_{i+\frac{1}{2},j+\frac{1}{2}}), \quad \forall i \in [\![2, M-1]\!].$$

For boundary fluxes, as in 1D, we use slightly different definitions

$$\mathcal{G}^{u,x}_{\frac{3}{2},j+\frac{1}{2}} = \frac{u_{2,j+\frac{1}{2}}}{2}\mathcal{F}^{x,-}_{2,j+\frac{1}{2}}, \text{ and } \mathcal{G}^{u,x}_{M+\frac{1}{2},j+\frac{1}{2}} = \frac{u_{M,j+\frac{1}{2}}}{2}\mathcal{F}^{x,+}_{M,j+\frac{1}{2}}.$$

The fluxes $\mathcal{G}^{u,y}_{i,j}$, for any $(i,j) \in [\![2, M]\!] \times [\![2, N]\!]$ are defined by

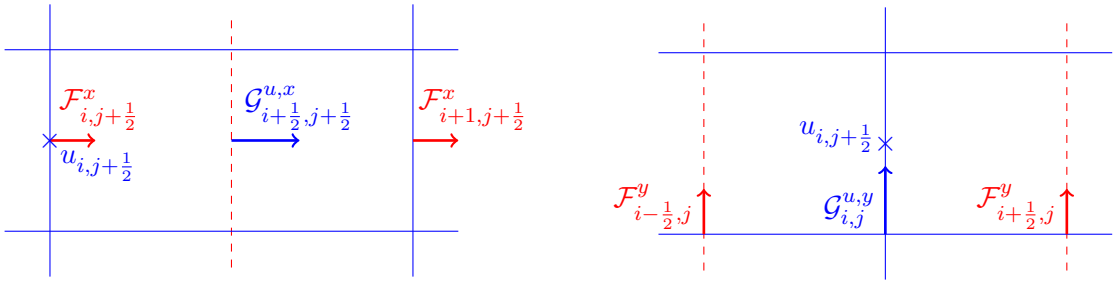
$$\mathcal{G}^{u,y}_{i,j} = u_{i,j-\frac{1}{2}}\mathcal{F}^{y,+}_{i,j} + u_{i,j+\frac{1}{2}}\mathcal{F}^{y,-}_{i,j}.$$

For the boundary values, we set $\mathcal{G}^{u,y}_{i,1} = 0$ and $\mathcal{G}^{u,y}_{i,N+1} = 0$ for all $j \in [\![2, N]\!]$. Fig. 5 illustrate this construction by putting forward the mass fluxes used in the definition of the momentum flux $\mathcal{G}^{u,x}_{i+\frac{1}{2},j+\frac{1}{2}}$ and $\mathcal{G}^{u,y}_{i,j}$.

Finally, symmetrically, the vertical velocity $v_{i+\frac{1}{2},j}$, $i \in [\![1, M]\!]$, $j \in [\![2, N]\!]$ is updated with the following scheme

$$\frac{\overline{\rho}_{i+\frac{1}{2},j}\overline{v}_{i+\frac{1}{2},j} - \rho_{i+\frac{1}{2},j}v_{i+\frac{1}{2},j}}{\delta t} + \frac{\mathcal{G}^{v,x}_{i+1,j} - \mathcal{G}^{v,x}_{i,j}}{\delta x_{i+\frac{1}{2}}} + \frac{\mathcal{G}^{v,y}_{i+\frac{1}{2},j+\frac{1}{2}} - \mathcal{G}^{v,y}_{i+\frac{1}{2},j-\frac{1}{2}}}{\delta y_j} + \frac{\Pi_{i+\frac{1}{2},j+\frac{1}{2}} - \Pi_{i+\frac{1}{2},j-\frac{1}{2}}}{\delta y_j} = 0. \tag{53}$$

The momentum fluxes $\mathcal{G}^{v,x}$ and $\mathcal{G}^{v,y}$ are defined like $\mathcal{G}^{u,x}$ and $\mathcal{G}^{u,y}$ by inverting the roles played by u and v, by x and y, and by i and j.



(a) Flux $\mathcal{G}^{u,x}_{i+\frac{1}{2},j+\frac{1}{2}}$ (b) Flux $\mathcal{G}^{u,y}_{i,j}$

Figure 5: Mass flux used in the definition of momentum fluxes.

It can be shown that under a CFL condition – which can be readily deduced from the 1D statement – the positivity of ρ is preserved. Similarly, strengthened assumptions can be identified to guaranty that the decay of the global entropy under suitable stability constraints is still valid on MAC meshes, see [4].

We now turn to explain how to extend the second order scheme to the 2D framework. We apply the 1D MUSCL method to the rows or the columns of the physical variables.

- To define the upgraded mass flux $\mathcal{F}^{x,ML}$ we use a MUSCL reconstruction only on the columns of the density ρ :

$$\mathcal{F}^{x,ML}_{i,j+\frac{1}{2}} = \mathcal{F}^{+}(\rho^{-}_{i,j+\frac{1}{2}}, u_{i,j+\frac{1}{2}}) + \mathcal{F}^{-}(\rho^{+}_{i,j+\frac{1}{2}}, u_{i,j+\frac{1}{2}}).$$

- To define the upgraded mass flux $\mathcal{F}^{\rho,y,ML}$ we use a MUSCL reconstruction only on the rows of the density ρ.

With this new definition of the mass flux $\mathcal{F}^{x,ML}$ and $\mathcal{F}^{y,ML}$ we define the new convective part of the momentum flux $\mathcal{G}^{u,x,ML}$ and $\mathcal{G}^{u,y,ML}$:

- To define the upgraded momentum flux $\mathcal{G}^{u,x,ML}+p$ we use a MUSCL reconstruction only on the columns of the velocity u in the convection flux:

$$\begin{aligned}\mathcal{G}^{u,x,ML}_{i+\frac{1}{2},j+\frac{1}{2}} &= \frac{u^-_{i+\frac{1}{2},j+\frac{1}{2}}}{2}\left(\mathcal{F}^{x,ML,+}_{i+1,j+\frac{1}{2}}+\mathcal{F}^{x,ML,+}_{i,j+\frac{1}{2}}\right)\\ &\quad+\frac{u^+_{i+\frac{1}{2},j+\frac{1}{2}}}{2}\left(\mathcal{F}^{x,ML,-}_{i+1,j+\frac{1}{2}}+\mathcal{F}^{x,ML,-}_{i,j+\frac{1}{2}}\right)\end{aligned}$$

- To define the upgraded mass flux $\mathcal{G}^{u,y,ML}$ we use a MUSCL reconstruction only on the rows of the velocity u.

The stability and consistency analysis performed in 1D generalize directly to higher dimensions on MAC meshes, a configuration which is appealing when dealing with low Mach regimes, see [38, Chapter 3].

5 Numerical simulations

5.1 Barotropic Euler sytem

5.1.1 Accuracy study using a 1d manufactured solution

In order to numerically validate the abilities of the MUSCL-like approach, we compute the solutions of the 1D problem

$$\partial_t\begin{pmatrix}\rho\\ \rho u\end{pmatrix}+\partial_x\begin{pmatrix}\rho u\\ \rho u^2+p(\rho)\end{pmatrix}=\begin{pmatrix}0\\ f\end{pmatrix},$$

where the force field $(x,t)\mapsto f(x,t)$ is tailored so that the solution reads

$$\begin{cases}\rho(x,t)=\dfrac{\rho_0(x)e^t}{\left(x+e^t(1-x)\right)^2},\\ u(x,t)=x(1-x).\end{cases}$$

In what follows we simply choose $\rho_0(x)=1$. The solution is smooth and we can expect a full benefit of the MUSCL approach. The computational domain is the slab $[0,1]$ and we perform the simulation for $t\in[0,0.5]$. In the definition of the fluxes, we make use of the SuperBee flux limiter (see (29) for the definition).

We first consider the case where the pressure is defined by the perfect gas state law: $p(\rho)=\rho^2$. We give in Table 1 the L^2-norm of the error between the discrete and exact solutions for several numbers J of grid points and $\delta t=10^{-5}$. The small value of the time step ensures that the stability condition is satisfied for all the considered grids. We use here the first order Euler scheme in time and compare the solution produced by the first order scheme of [5] and the solution produced by the proposed MUSCL extension. We clearly observe the gain of accuracy with the MUSCL scheme. Moreover, it reaches the second order for both the density and the velocity, while, as expected, the scheme of [5] approaches the solution at first order only.

	First order scheme				Second order MUSCL scheme			
	Density		Velocity		Density		Velocity	
J	Error	Rate	Error	Rate	Error	Rate	Error	Rate
70	$5.2e^{-3}$	–	$4.27e^{-3}$	–	$7.31e^{-5}$	–	$7.90e^{-5}$	–
80	$4.6e^{-3}$	0.96	$3.77e^{-3}$	0.93	$5.60e^{-5}$	2.00	$6.17e^{-}5$	1.85
90	$4.1e^{-3}$	0.97	$3.38e^{-3}$	0.94	$4.40e^{-5}$	2.05	$4.93e^{-5}$	1.90
100	$3.7e^{-3}$	0.97	$3.06e^{-3}$	0.95	$3.56e^{-5}$	2.01	$4.01e^{-5}$	1.96
110	$3.4e^{-3}$	0.97	$2.79e^{-3}$	0.95	$2.93e^{-5}$	2.04	$3.36e^{-5}$	1.86

Table 1: L^2-norm of the error between approximate and exact solutions for different numbers of grid points. Computations done with $\delta t = 10^{-5}$ on uniform grids using the perfect gas state law.

It is also worth discussing the interest of replacing the Euler scheme by the RK2 method for the time discretization. Hence, for the given mesh size $J = 800$, and using the MUSCL version of the scheme for the space discretization, we make the time step vary $\delta t \in \{10^{-4};\ 5 \times 10^{-5};\ 2.5 \times 10^{-5}; 1.25 \times 10^{-5}\}$. Results are displayed in Table 2: we observe a convergence of order 1 with the Euler scheme, while with RK2 the error due to the time discretization is hidden by the error in space. There is a clear advantage in using the RK2 scheme.

	First order scheme				Second order RK scheme			
	Density		Velocity		Density		Velocity	
δt	Error	Rate	Error	Rate	Error	Rate	Error	Rate
$1e^{-4}$	$8.5e^{-5}$	–	$6.1e^{-5}$	–	$9e^{-7}$	–	$7e^{-7}$	–
$5e^{-5}$	$4.2e^{-5}$	1.01	$3.1e^{-5}$	0.99	$9e^{-7}$	–	$7e^{-7}$	–
$2.5e^{-5}$	$2.1e^{-5}$	1.01	$1.6e^{-5}$	0.98	$9e^{-7}$	–	$7e^{-7}$	–
$1.25e^{-5}$	$1.0e^{-5}$	1.01	$8.0e^{-6}$	0.97	$9e^{-7}$	–	$7e^{-7}$	–

Table 2: L^2-norm of the error between approximate and exact solutions for different time steps. Computation done with $J = 800$ on uniform grids using the perfect gas state law.

We next apply the method on non-uniform meshes for the same tailored solution. We build a sequence of non-uniform nested meshes as follows: we randomly pick $J = 100$ points that define the coarser non-uniform mesh ($\max \delta x_{j+\frac{1}{2}} / \min \delta x_{j+\frac{1}{2}} = 4.18$) ; then, we split each cell in its middle to obtain the next finer mesh – still non-uniform – with twice as many as grid points and so on. Table 3 shows the L^2 error norms for $\delta t = 1e^{-6}$ and four nested meshes. The convergence rate is the same on uniform and non-uniform meshes, in agreement with Lemma 3.6.

	First order scheme				Second order MUSCL scheme			
	Density		Velocity		Density		Velocity	
J	Error	Rate	Error	Rate	Error	Rate	Error	Rate
100	$2.6e^{-3}$	–	$2.1e^{-3}$	–	$2.3e^{-5}$	–	$3.4e^{-5}$	–
200	$1.6e^{-3}$	0.72	$1.3e^{-3}$	0.69	$7.7e^{-6}$	1.56	$9.5e^{-6}$	1.83
400	$8.7e^{-4}$	0.87	$7.3e^{-4}$	0.85	$1.9e^{-6}$	2.01	$2.5e^{-6}$	1.90
800	$4.6e^{-4}$	0.94	$3.8e^{-4}$	0.93	$6.0e^{-7}$	1.68	$1.0e^{-7}$	1.33
	First order scheme				Second order MUSCL scheme			
	Density		Velocity		Density		Velocity	
J	Error	Rate	Error	Rate	Error	Rate	Error	Rate
100	$3.4e^{-3}$	–	$3.3e^{-3}$	–	$3.9e^{-5}$	–	$5.0e^{-5}$	–
200	$2.1e^{-3}$	0.72	$1.7e^{-3}$	0.94	$1.2e^{-5}$	1.63	$1.3e^{-5}$	1.93
400	$1.0e^{-3}$	0.98	$8.7e^{-4}$	0.97	$3.0e^{-6}$	2.05	$2.3e^{-6}$	2.50
800	$5.2e^{-4}$	0.99	$4.4e^{-4}$	0.98	$1.2e^{-6}$	1.32	$1.1e^{-6}$	1.06

Table 3: L^2-norm of the error between approximate and exact solutions for different numbers of grid points. Computation done with $\delta t = 1e^{-6}$ on uniform grids (top) and non uniform grids (bottom) using the perfect gas state law.

We finally check the ability of the scheme in dealing with a more complex pressure law. The tailored solution is still the same, but now we set $p(\rho) = \frac{(\gamma-1)^2}{4\gamma}\left(\frac{\rho}{\rho^*-\rho}\right)^\gamma$. This is a particular case of the Van der Waals state law, it arises in the modeling of dusty gases for instance. We point out that this pressure law does not lead to any difficulty in the design of the scheme and its consistency properties apply equally well to this case. Tests are performed with $\gamma = 0.6$ and $\rho^* = 3$. Note that admissible densities should remain in the domain $0 \leq \rho < \rho^*$; this issue is further discussed in Section 5.1.2 below. Table 4 shows the L^2-error norms: for the smooth solution considered here, the convergence rate is still second order.

	First order scheme				Second order MUSCL scheme			
	Density		Velocity		Density		Velocity	
J	Error	Rate	Error	Rate	Error	Rate	Error	Rate
100	$1.6e^{-3}$	–	$8.2e^{-4}$	–	$2.1e^{-4}$	–	$4.0e^{-5}$	–
200	$8.0e^{-4}$	0.99	$4.2e^{-4}$	0.97	$5.5e^{-5}$	1.94	$1.0e^{-5}$	1.95
400	$4.0e^{-4}$	0.98	$2.1e^{-4}$	0.98	$1.4e^{-5}$	1.98	$2.5e^{-6}$	2.03
800	$2.0e^{-4}$	0.99	$1.1e^{-4}$	0.99	$3.9e^{-6}$	1.83	$8.0e^{-7}$	1.64

Table 4: L^2-norm of the error between approximate and exact solutions for different numbers of grid points. Computation done with $\delta t = 5e^{-6}$ on uniform meshes using the Van der Waals state law.

5.1.2 Simulation of 1d Riemann problems

Perfect gases pressure law We now study the behavior of the scheme with discontinuous solutions. We consider Riemann problems on a computational domain $[a, b]$: the initial data is piecewise constant with a jump located at $x = 0$ and we denote by (ρ_l, u_l) and (ρ_r, u_r) the left and right states for the density/velocity pair, respectively. The pressure law is defined by $p(\rho) = \frac{(\gamma-1)^2}{4\gamma}\rho^\gamma$ with $\gamma = 1.6$. We use the MinMod limiter and the others data are given in Table 5.

	a	b	ρ_l	ρ_r	u_l	u_r
Test 1	-0.7	0.3	0.5	1	-0.5	-0.2
Test 2	-0.2	0.8	1	2	1	0.25
Test 3	-0.7	0.3	1	0.5	-0.5	-0.5

Table 5: Definition of the test case. Computational domain $[a, b]$. Left and right states for the density and velocity (jump located at $x = 0$).

The corresponding Riemann solutions develop two rarefaction waves, two shocks and a rarefaction wave followed by a shock wave respectively. The results obtained at time $T = 0.5$ using the first order scheme and the MUSCL-RK scheme with $J = 200$ and $\delta t = 1e^{-3}$ are plotted in Fig 6, Fig 7 and Fig 8. Of course the solutions of Riemann problems are not smooth and the consistency analysis does not apply. Nevertheless, we clearly see that using the MUSCL-like method provides more accurate results than the first order method. The convergence study of the different test cases is presented in Table 6 and we observe a first order convergence.

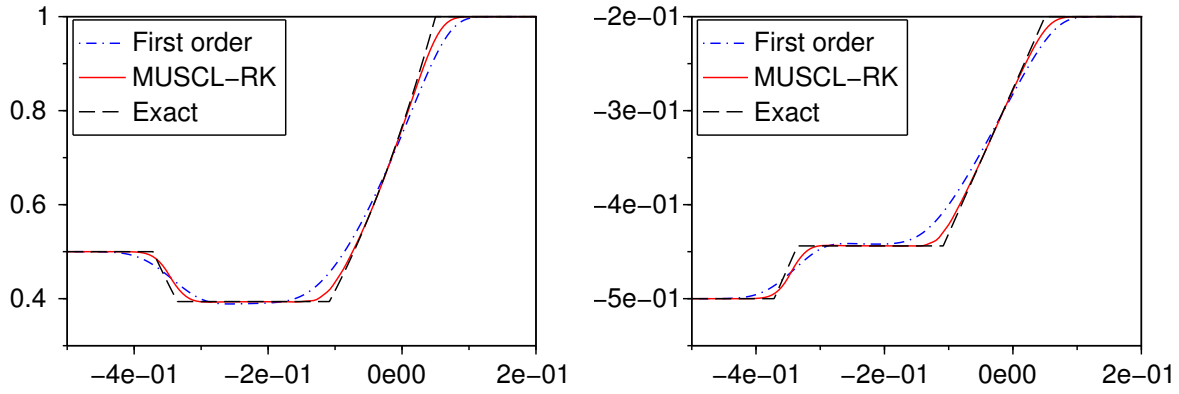


Figure 6: Test 1. Density (at left) and velocity (at right) at time $T = 0.5$. Computation done with $J = 200$ and $\delta t = 1e^{-3}$.

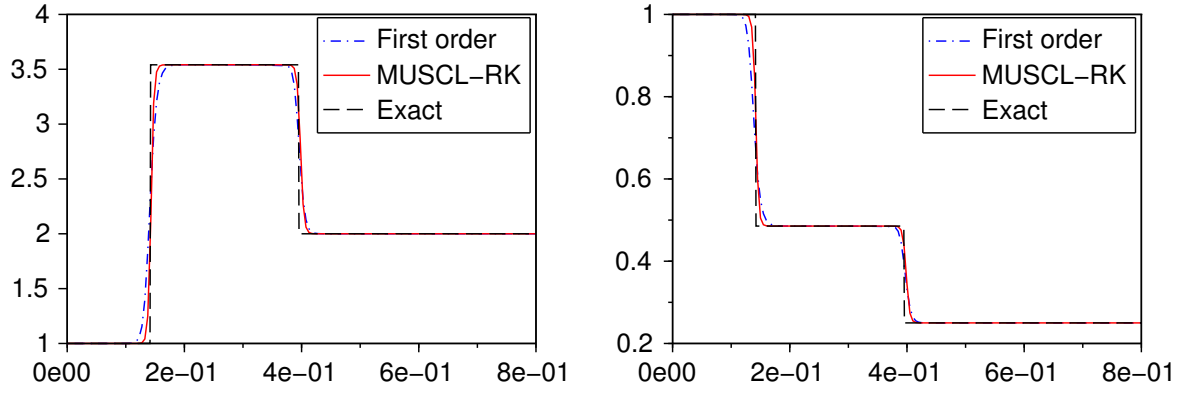


Figure 7: Test 2. Density (at left) and velocity (at right) at time $T = 0.5$. Computation done with $J = 200$ and $\delta t = 1e^{-3}$.

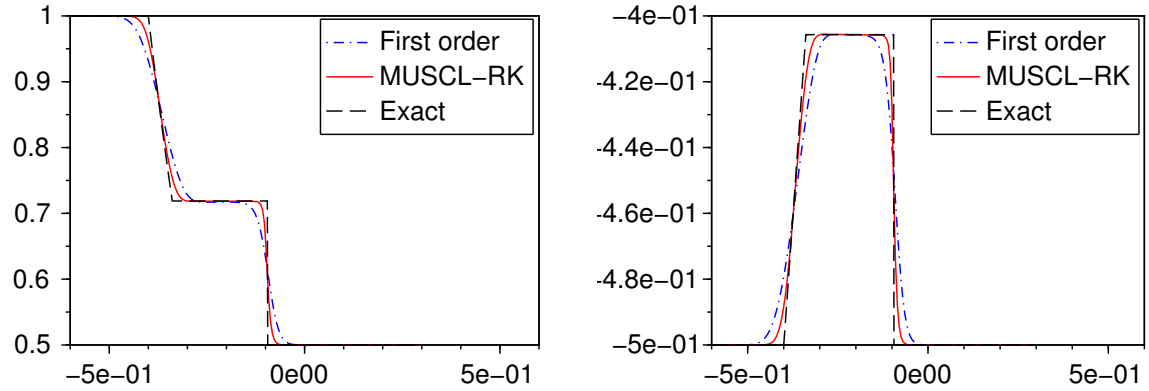


Figure 8: Test 3. Density (at left) and velocity (at right) at time $T = 0.5$. Computation done with $J = 200$ and $\delta t = 1e^{-3}$.

	Test 1			
	Density		Velocity	
J	Error	Rate	Error	Rate
100	$6.7e^{-3}$	–	$3.0e^{-3}$	–
200	$3.4e^{-3}$	0.97	$1.5e^{-3}$	0.98
400	$1.8e^{-3}$	0.97	$7.8e^{-4}$	0.97
800	$8.9e^{-4}$	0.98	$3.9e^{-4}$	0.98
1600	$4.5e^{-4}$	0.99	$2.0e^{-4}$	0.99

	Test 2			
	Density		Velocity	
J	Error	Rate	Error	Rate
100	$3.2e^{-2}$	–	$5.8e^{-3}$	–
200	$1.8e^{-2}$	0.81	$2.6e^{-3}$	1.18
400	$8.0e^{-3}$	1.19	$1.3e^{-3}$	0.94
800	$4.7e^{-3}$	0.77	$6.9e^{-4}$	0.96
1600	$2.2e^{-3}$	1.11	$3.8e^{-4}$	0.87

	Test 3			
	Density		Velocity	
J	Error	Rate	Error	Rate
100	$6.0e^{-3}$	–	$2.3e^{-3}$	–
200	$3.2e^{-3}$	0.89	$1.3e^{-3}$	0.89
400	$1.7e^{-3}$	0.96	$6.4e^{-4}$	0.96
800	$8.3e^{-4}$	1.00	$3.2e^{-4}$	1.00
1600	$4.2e^{-4}$	1.00	$1.6e^{-4}$	1.00

Table 6: Riemann problems. L^1-norm of the error at time $T = 0.5$ between approximate and exact solutions for different numbers of grid points. Computations done with the MUSCL-RK scheme with $\delta t = 0.25\ \delta x$.

Van der Waals pressure law As already said above, our method does not rely on the resolution of Riemann problems, and the numerical fluxes have a simple expression for very general pressure laws, while the scheme is entropy-decaying. For instance the scheme is still efficient for Van der Waals-like laws $p(\rho) = \frac{(\gamma-1)^2}{4\gamma}\frac{\rho^\gamma}{\rho^*-\rho}$. Such a relation is intended to retain some packing effects that prevent the density to exceed the threshold $\rho^\star$, see Fig 9.

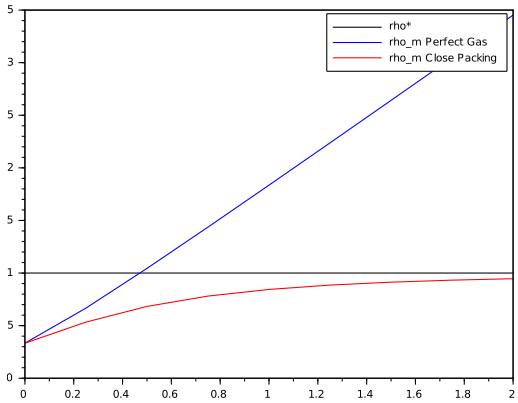


Figure 9: Intermediate volume fraction ρ_m as a fonction of the velocity u.

However the preservation of this constraint by the numerical unknown leads to a strengthened stability condition, see [6, Prop. 5]. For the MUSCL version of the scheme the stability condition takes the following form.

Proposition 5.1 (Close-packing threshold)**.** *Suppose that the initial data satisfies* $\rho^0_{j+\frac{1}{2}} \leq \rho^\star$ *for all* $j \in [\![1, J]\!]$ *and assume that the following CFL-like condition holds*

$$\frac{\delta t}{\delta x_{j+\frac{1}{2}}} \left(\left[\lambda_+(\rho_j^-, u_j) \right]^+ + \left[\lambda_-(\rho_{j+1}^+, u_{j+1}) \right]^- \right) \leqslant \frac{1}{2} \left(1 - \frac{\rho_{j+\frac{1}{2}}}{\rho^\star} \right), \quad \forall j \in [\![1, J]\!].$$

Then $\overline{\rho}_{j+\frac{1}{2}} \leq \rho^\star$ *for all* $j \in [\![1, J]\!]$ *and* $k \in \mathbb{R}^N$.

Proof. We have

$$\begin{aligned}
\overline{\rho}_{j+\frac{1}{2}} &= \rho_{j+\frac{1}{2}} + \frac{\delta t}{\delta x_{j+\frac{1}{2}}} \left(\mathcal{F}^-(\rho_j^+, u_j) - \mathcal{F}^+(\rho_{j+1}^-, u_{j+1}) \right) \\
&\quad + \frac{\delta t}{\delta x_{j+\frac{1}{2}}} \left(\mathcal{F}^+(\rho_j^-, u_j) - \mathcal{F}^-(\rho_{j+1}^+, u_{j+1}) \right) \\
&\leq \rho_{j+\frac{1}{2}} + \frac{\delta t}{\delta x_{j+\frac{1}{2}}} \left(\mathcal{F}^+(\rho_j^-, u_j) - \mathcal{F}^-(\rho_{j+1}^+, u_{j+1}) \right) \\
&\leq \rho_{j+\frac{1}{2}} + \frac{\delta t}{\delta x_{j+\frac{1}{2}}} \left(\rho_j^- \left[\lambda_+(\rho_j^-, u_j) \right]^+ + \rho_{j+1}^+ \left[\lambda_-(\rho_{j+1}^+, u_{j+1}) \right]^- \right).
\end{aligned}$$

Let us assume $\rho_{j+\frac{1}{2}} \leq \rho^\star$ for any j; then $\overline{\rho}_j^\pm \leq 2\rho^\star$ (see the proof of Proposition 3.1). Let us introduce

$$\epsilon_{j+\frac{1}{2}} = 1 - \frac{\rho_{j+\frac{1}{2}}}{\rho^\star}.$$

We get

$$\overline{\rho}_{j+\frac{1}{2}} \leq \rho^\star - \rho^\star \left(\epsilon_{j+\frac{1}{2}} - \frac{2\delta t}{\delta x_{j+\frac{1}{2}}} \left(\left[\lambda_+(\overline{\rho}_j^-, u_j) \right]^+ + \left[\lambda_-(\overline{\rho}_{j+1}^+, u_{j+1}) \right]^- \right) \right).$$

Finally, assuming

$$\epsilon_{j+\frac{1}{2}} \geq \frac{2\delta t}{\delta x_{j+\frac{1}{2}}} \left(\left[\lambda_+(\overline{\rho}_j^-, u_j) \right]^+ + \left[\lambda_-(\overline{\rho}_{j+1}^+, u_{j+1}) \right]^- \right)$$

we obtain $\overline{\rho}_{j+\frac{1}{2}} \leq \rho^\star$ for any j. ∎

As observed for the condition ensuring the positivity of the density, the CFL-like condition for the MUSCL scheme is twice more constrained than with the first order scheme. More important, this condition is much more demanding than the standard CFL condition since the right-hand side vanish when the discrete density become close to $\rho^\star$. Numerical simulations confirm that such a strengthened condition is actually needed to prevent the density to exceed the threshold $\rho^\star$ when using the scheme proposed in this paper. To illustrate the difficulty, we go back to the numerical tests proposed in [6, Section 4.1]. We set $\gamma = 2$ and $\rho^\star = 1$ and we perform 1D simulations of the Riemann problem defined by $\rho_l = \rho_r = \frac{1}{3}$ and $u_l = u_{abs}$ and $u_r = -u_{abs}$ for different values of $u_{abs} \in \{0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5\}$ and different values of the time step δt. The larger the velocity u_{abs} is, the closer to $\rho^\star$ the density ρ_m in the intermediate state is. The simulations are performed up to the time $T = 0.1$ with $J = 200$ on the computational domain $[-0.5, 0.5]$ with a discontinuity initially at $x = 0$. In Fig. 10 we show the solutions obtained with different schemes: for the same numerical conditions δt and δx, the Lax-Friedrichs scheme produces much more numerical diffusion and the solution is poorly captured. In Fig. 11 we zoom on the left part of the density.

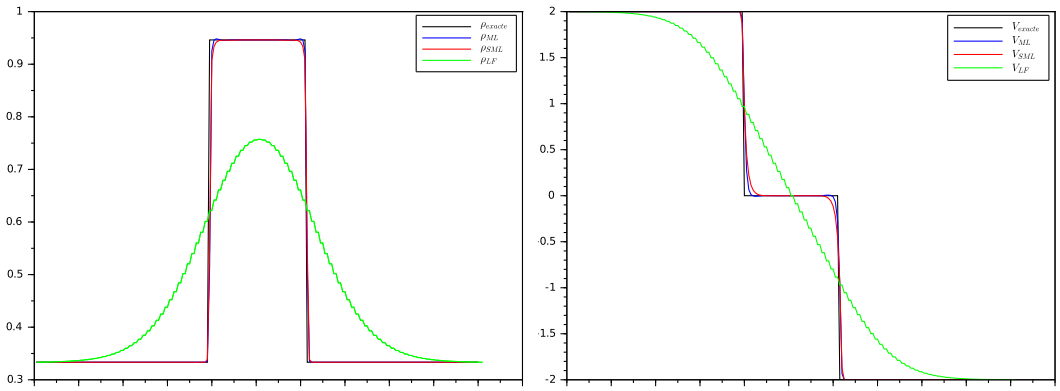


Figure 10: Barotropic gas with Van der Waals law: comparison of the kinetic scheme (1st and 2nd order) with the Lax-Friedrichs scheme. Density and Velocity solutions for $u_{abs} = 2$.

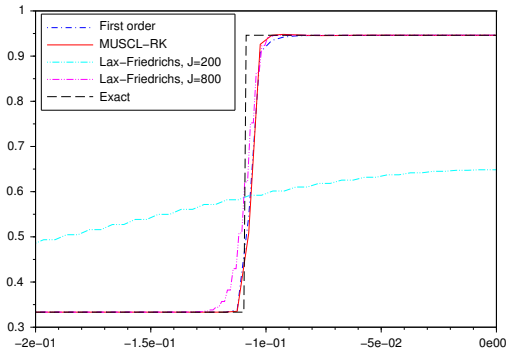


Figure 11: Barotropic gas with Van der Waals law: comparison of the kinetic schemes (1st and 2nd order with $J = 200$) with the Lax-Friedrichs scheme ($J = 200$ and $J = 800$). Zoom on the density solution for $u_{abs} = 2$.

For each values of u_{abs}, we select the largest value of δt which yields an "admissible" result (in the sense that it remains oscillation-free at $T = 0.1$). In Fig. 12, we plot this selected δt as

a function of $1 - \frac{\rho_m}{\rho^\star}$: in a logarithm scale, we obtain a straight line with a slope close to 2, which is consistent with Proposition 5.1 since when ρ_m becomes close to $\rho^\star$, the characteristic speeds behave like $\left(1 - \frac{\rho_m}{\rho^\star}\right)^{-1}$.

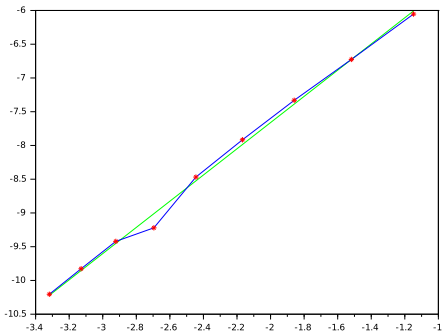


Figure 12: Maximal admissible time step δt as a function of $1 - \frac{\rho_m}{\rho^\star}$ in blue; the green line has a slope 1.94.

Note, nevertheless, that the standard CFL is enough to preserve the bounds on the density when using a Godunov or a Lax-Friedrichs scheme. (This latter result uses crucially the convexity of the invariant domain of the PDE, see [11, Section 2.2.1 & Prop. 2.11], and it does not apply when the invariant regions are non-convex, see e.g. [14].) However, the scheme based on the kinetic fluxes is far less diffusive than Lax-Friedrichs' method, so that it finally competes in terms of numerical effort for a given numerical accuracy.

5.1.3 Numerical simulations in 2d

Falling water columns We turn to 2D simulations, with a test-case inspired from [2]. We simulate three falling columns into a rectangular basin. The computational domain is the two-dimensional square $[-1, 1] \times [-1, 1]$. We are using the dimensionless Shallow Water system which amounts to set $p(\rho) = \rho^2$, without source terms. The PDE system is endowed with zero flux boundary conditions and the following initial data

$$\begin{cases} \rho(0, x, y) = 3 + \mathbb{1}_{(x-0.5)^2+(y-0.5)^2<(0.15)^2} + \mathbb{1}_{(x+0.5)^2+(y+0.5)^2<(0.15)^2} + 2 \cdot \mathbb{1}_{x^2+y^2<(0.2)^2}, \\ u(0, x, y) = 0, \\ v(0, x, y) = 0. \end{cases}$$

The simulation performed in [2] on a 512×512 Cartesian mesh is reproduced in Fig. 13: it is based on the second order Nessayhu-Tadmor scheme [40], coupled to a specific reconstruction procedure which is intended to reduce the numerical diffusion and to capture shocks with an enhanced accuracy. The MUSCL scheme competes with such an approach, as it appears in Fig. 14 on the right (simulations have been performed with the MinMod limiter). Fig. 14 shows the advantages in using the MUSCL method compared to the first order scheme, which, for the same numerical parameters, loses the complex structures of the flow. In these simulations, the mesh is a 512×512 Cartesian mesh, the final time $T = 1.035$ and the time step is $\delta t = 10^{-4}$. As already observed in [2], the simulation is quite sensitive to the time step: some oscillations might appear when δt is not small enough.

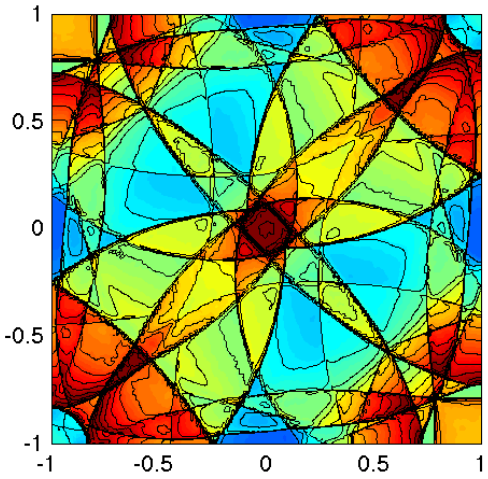


Figure 13: Simulation of the Shallow Water system, by courtesy of N. Aguillon [2].

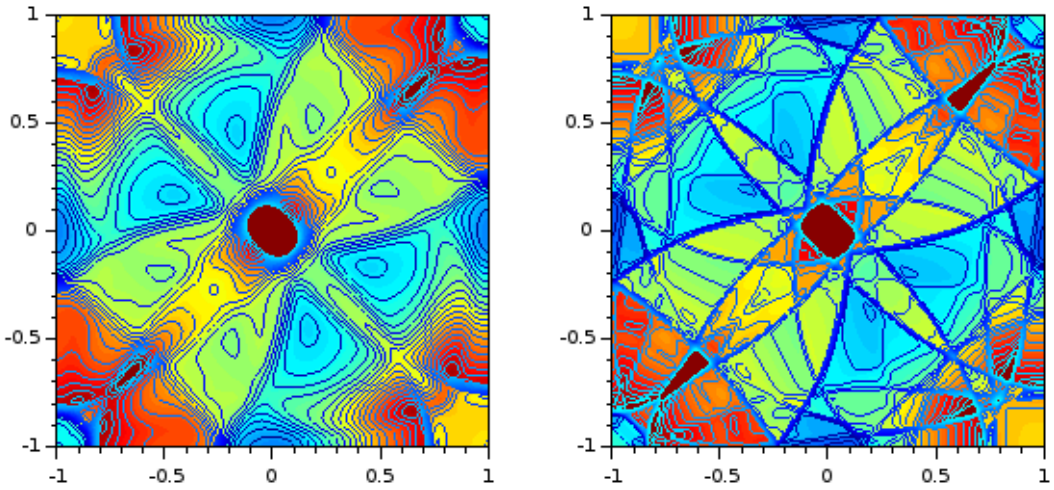


Figure 14: Simulation of the Shallow Water system for $\delta t = 10^{-4}$ on a 512×512 grid: Results with the first order scheme (left) and with the MUSCL scheme (right)

We also compare the results obtained with the first and second order schemes to a simulation performed with a Spectral Element Method (SEM), stabilized with an entropy viscosity method (EVM), see [42, 41]. We also refer the reader to [39] for further details on this method and the test case which is computed to show the ability of this approach to deal with dry-wet transitions and shocks. The SEM-EVM method is driven by two parameters, α, β in the notations of [39]. The color map is identical in Fig. 15 and Fig. 16, but it differs from Fig. 13 and Fig. 14. The left picture shows the result we get when using a first order viscosity everywhere and parameters that imply a $\mathcal{O}(h)$ numerical diffusion equivalent to the numerical diffusion of the upwind scheme (precisely the (α, β) pair is $(0.5,\ \infty)$). The entropy stabilization is strengthened in the right picture (with $(\alpha, \beta) = (1, 3)$). The mesh is of size 100×100 and a fifth order polynomial approximation is used in each quadrangle; this yields 255001 interpolation points in the computational domain whereas our scheme, used for Fig. 16, has 262144 degrees of freedom. In both figures, the results are shown at final time $T = 1.035$, reached with $\delta t = 10^{-4}$.

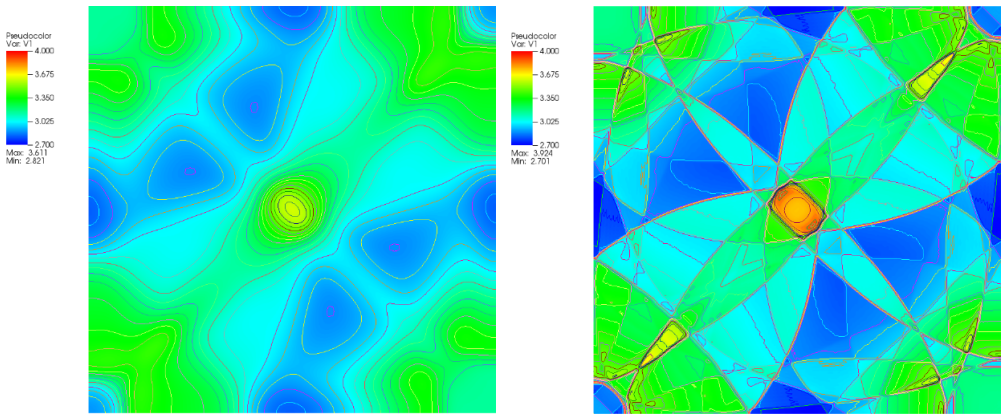


Figure 15: Visualizations of the height at the final time with the stabilized SEM method.

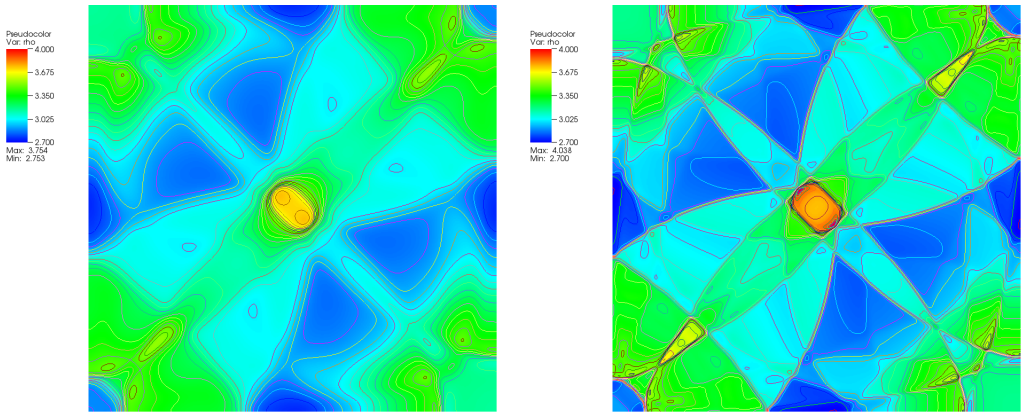


Figure 16: Visualizations of the height at the final time with our first and second order schemes.

Clearly, the result obtained with our first order scheme is smooth and close to the one obtained with the SEM-EVM scheme when adding a first order viscosity. The result obtained with the second order scheme — using the MUSCL procedure — recovers the correct solution (free of spurious oscillations) and is very close to the one obtained with the strengthened EVM stabilization.

Forward facing step This test case is inspired by the standard 2D Mach 3 wind tunnel with a step introduced in [55]. The computational domain Ω is the L-shaped domain

$$\Omega = \Omega_0 \setminus \Omega_{step}, \quad \Omega_0 = [0,3] \times [0,1], \quad \Omega_{step} = [0.6,3] \times [0,0.2].$$

The rectangle Ω_0 is discretized with a $30\sigma \times 10\sigma$ uniform Cartesian grid ($\sigma \in \mathbb{N}^*$). We take the step into account by removing the mesh points corresponding to the step Ω_{step} at the right bottom part of the domain.

The equation of state of the fluid is $p(\rho) = \rho$ and the initial data are given by $\rho = 1$ and $u = (3,0)$. On the top and bottom walls, we use reflection boundary conditions (*ie* zeros flux boundary conditions as described in the previous parts of the article). A Dirichlet boundary condition, $\rho = 1$ and $u = (3,0)$, makes the flow enter through the left boundary whereas a free boundary condition is used for the right section. The free boundary condition is implemented

by assuming *a priori* that the outgoing flow is supersonic. The mass fluxes at the free boundary are thus defined using (51) but with an incoming part set to 0; that is, here, on the right boundary,we set

$$\mathcal{F}^x_{M+1,j+\frac{1}{2}} = \mathcal{F}^+(\rho_{M+\frac{1}{2},j+\frac{1}{2}}, u_{M+1,j+\frac{1}{2}}).$$

The update of the velocity at the free boundary is performed using (52) and (53) which now involve momentum fluxes at the exterior the domain (here at the points $(x_{M+\frac{3}{2}}, y_{j+\frac{1}{2}})$). These ghost fluxes are defined using a constant extrapolation for the pressure and a linear extrapolation for the outgoing part of the mass fluxes; that is here on the right boundary $\left(x_{M+\frac{3}{2}}, y_{j+\frac{1}{2}}\right)$ (using a uniform mesh), we set

$$u_{M+1,j+\frac{1}{2}} \mathcal{F}^{x,+}_{M+\frac{3}{2},j+\frac{1}{2}} + p(\rho_{M+\frac{1}{2},j+\frac{1}{2}}),$$

with $\mathcal{F}^{x,+}_{M+\frac{3}{2},j+\frac{1}{2}} = \frac{3}{2}\mathcal{F}^{x,+}_{M+\frac{1}{2},j+\frac{1}{2}} - \frac{1}{2}\mathcal{F}^{x,+}_{M-\frac{1}{2},j+\frac{1}{2}}$.

Fig. 17 presents the results obtained with the first order scheme and the second order scheme (MUSCL-RK scheme) on a 900×300 grid ($\sigma = 30$) with time steps respectively defined by $\delta t = 1/3000$ and $\delta t = 1/6000$. We observe that the structures are sharper with the MUSCL scheme. Cutlines of the density along the lines $y = 0.3$ obtained using the second order scheme on different grids are plotted in Fig. 18. The results are in agreement with the literature [34].

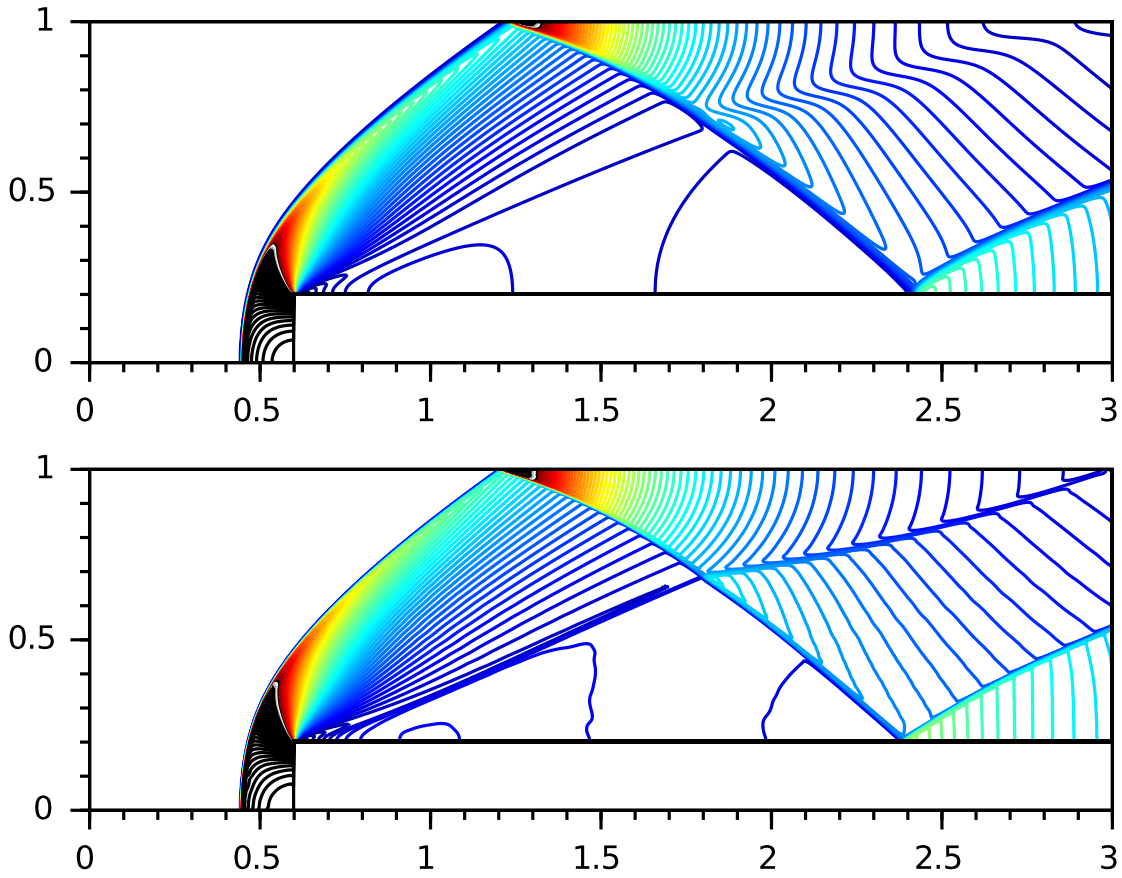


Figure 17: Simulation of the 2D Mach 3 wind tunnel with a step: Density with the first order scheme (up) and with the MUSCL/RK scheme (down)

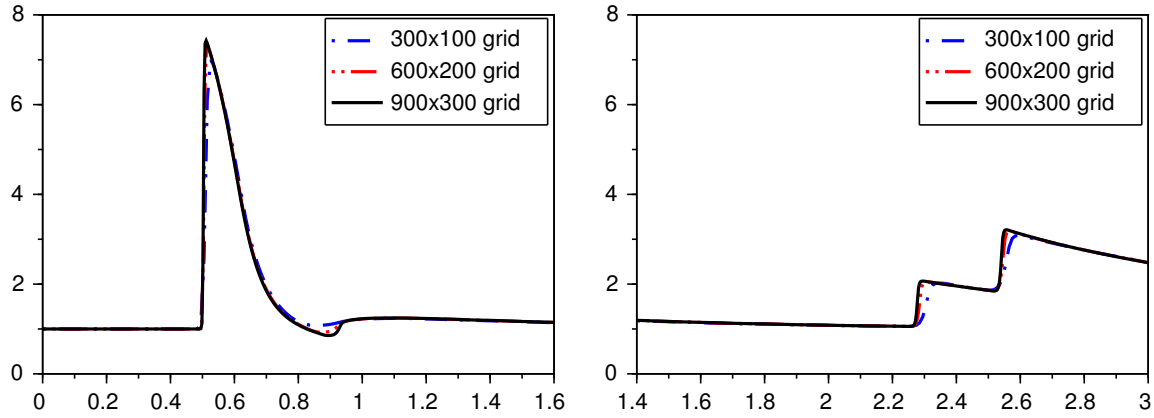


Figure 18: Simulation of the 2D Mach 3 wind tunnel with a step: density cutlines at $y = 0.3$ for different grids with the MUSCL-RK scheme

5.2 Full Euler system

5.2.1 Accuracy study using a 1d manufactured solution

In order to numerically validate the abilities of the MUSCL-like approach, we compute the solutions of the 1D problem

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ \rho e \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 + p(\rho, e) \\ \rho e u \end{pmatrix} + p(\rho, e) \partial_x \begin{pmatrix} 0 \\ 0 \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ f \\ g \end{pmatrix},$$

where the force fields $(t, x) \mapsto f(t, x)$ and $(t, x) \mapsto g(t, x)$ are tailored so that the solution reads

$$\begin{cases} \rho(t, x) = \dfrac{e^t}{(x + e^t(1 - x))^2}, \\ u(t, x) = x(1 - x), \\ p(t, x) = (\gamma - 1) e^t (1 + x)^2. \end{cases}$$

The solution is smooth and we can expect a full benefit of the MUSCL approach. The computational domain is the slab $[0, 1]$ and we perform the simulation for $t \in [0, 0.2]$ on uniform grids with $\gamma = 1, 4$. Table 7 gives the L^1-norm of the error between the discrete and the exact solutions for several numbers J of grid points. For this test, we have $\delta t = 10^{-6}$: the small value of the time step ensures that the stability condition is satisfied for all the considered grids. In fact, we have $\delta t = \mathcal{O}(\delta x^2)$ so that the error can be expected to be dominated by the spatial errors (we use the first order Euler scheme in time). We compare the solution produced by the first order scheme and the solution produced by the proposed MUSCL extension. We clearly observe the gain of accuracy with the MUSCL scheme. It reaches the second order for both the density, the velocity and the internal energy, while, as expected, the first order scheme approaches the solution at first order only.

J	First order scheme			Second order MUSCL scheme		
	Density	Velocity	Internal energy	Density	Velocity	Internal energy
100	–	–	–	–	–	–
200	0.850700	0.949060	0.992880	1.820381	1.874050	1.503473
300	0.880656	0.964206	0.994994	1.821862	1.892868	1.619922
400	0.896963	0.971184	0.995998	1.821327	1.910123	1.566806
500	0.907691	0.975388	0.996509	1.808201	1.927109	1.593161
600	0.914963	0.977970	0.996853	1.803942	1.943983	1.607096

Table 7: Error rate in L^1-norm between approximate and exact solutions for different numbers of grid points.

5.2.2 Simulation of 1d Riemann problems

We perform the numerical resolutions of some Riemann problems inspired from the classical textbook [49, Section 4.3.3, Chapter 4, pages 129-131] on the computational domain $[0, 1]$, see 8. All tests are with $\gamma = 1.4$. The simulations are performed 1000 grid points. The time step is given by $\delta t_1 = \delta x/100$ for the first order scheme and $\delta t_2 = \frac{\delta t_1}{2}$ for the MUSCL scheme. The initial data $\rho,\ u,\ p$ are piecewise constant functions with a discontinuity located at $x_0 = 0.5$, according to the table below.

	ρ_l	ρ_r	u_l	u_r	p_l	p_r	T
Test 1	1	0.125	0	0	1	0.1	0.25
Test 2	1	1	-2	2	0.4	0.4	0.15
Test 3	1	1	0	0	1000	0.01	0.012
Test 4	1	1	0	0	0.01	100	0.35
Test 5	5.99924	5.99242	19.5975	-6.19633	460.894	46.0950	0.035

Table 8: Definition of the test cases: left and right states.

In Figures 19 to 23, we represent at the final time T the density $\rho_{j+\frac{1}{2}}$ and the velocity u_j on the first line and the pressure $p_{j+\frac{1}{2}}$ and the internal energy $e_{j+\frac{1}{2}}$ one the second line. The exact solution is in dotted lines and the numerical solutions are given with the solid blue lines for the SML-scheme and with the solid red lines for the ML-scheme.

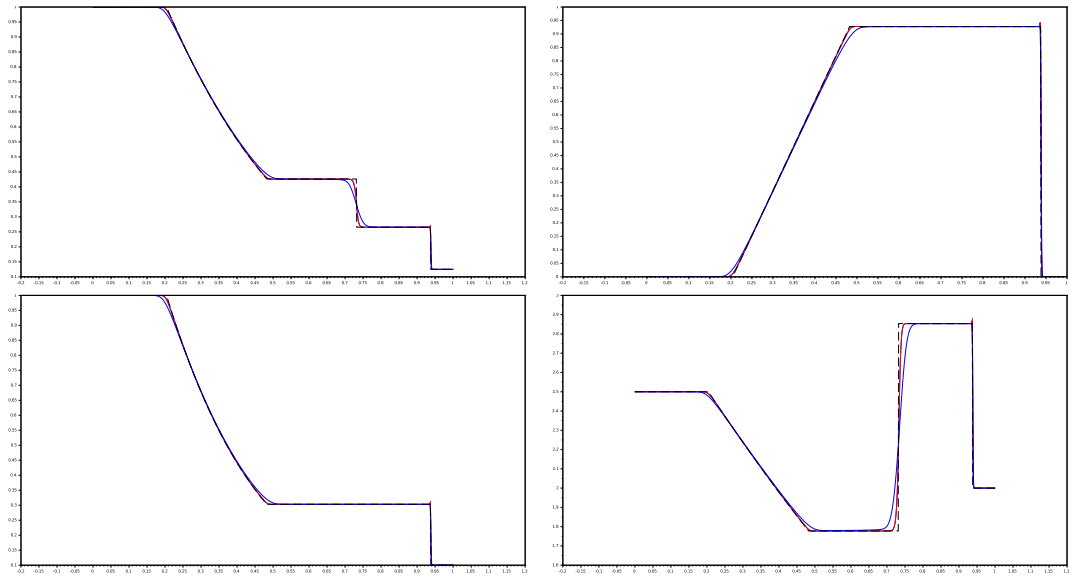


Figure 19: Test 1: the so-called Sod test problem, is a mild test: the solution consists in a left rarefaction, a contact discontinuity and a right shock.

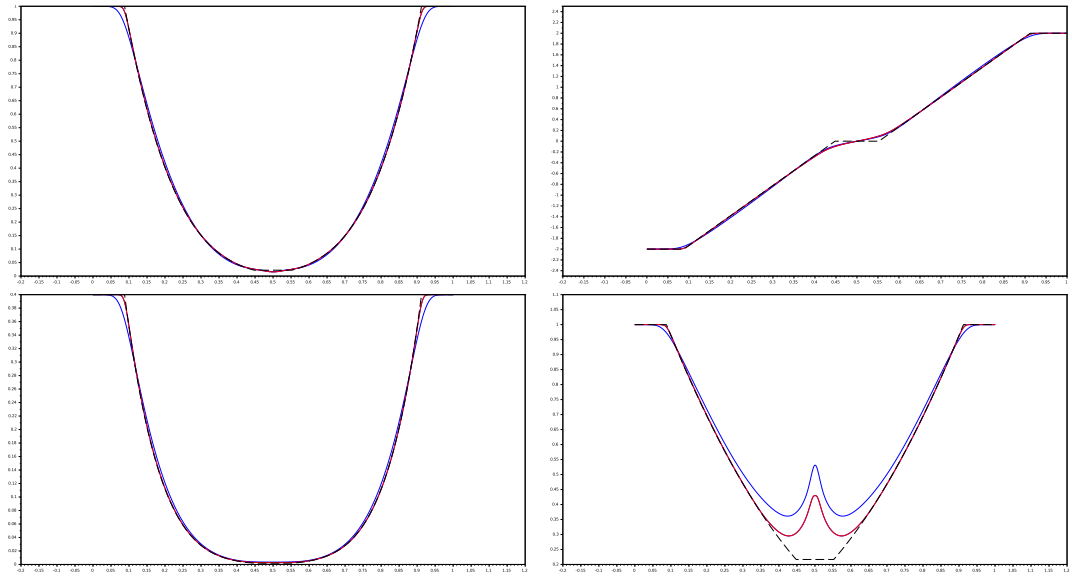


Figure 20: Test 2: the so-called 123 problem, has a solution made of two strong rarefactions and a trivial stationary contact discontinuity.

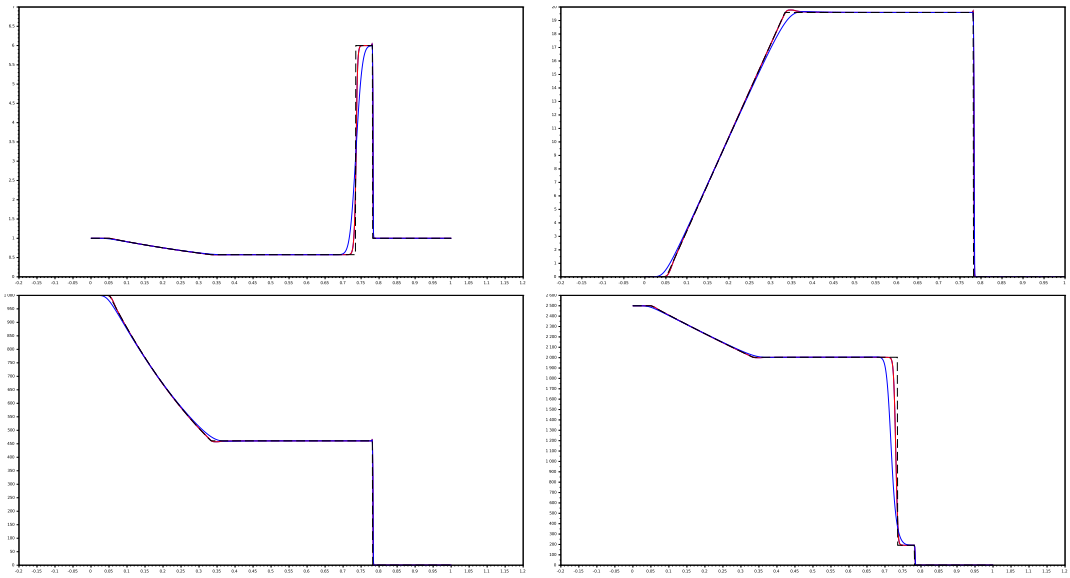


Figure 21: Test 3: inspired from [55], has a solution made of left rarefaction, a contact discontinuity and a right shock.

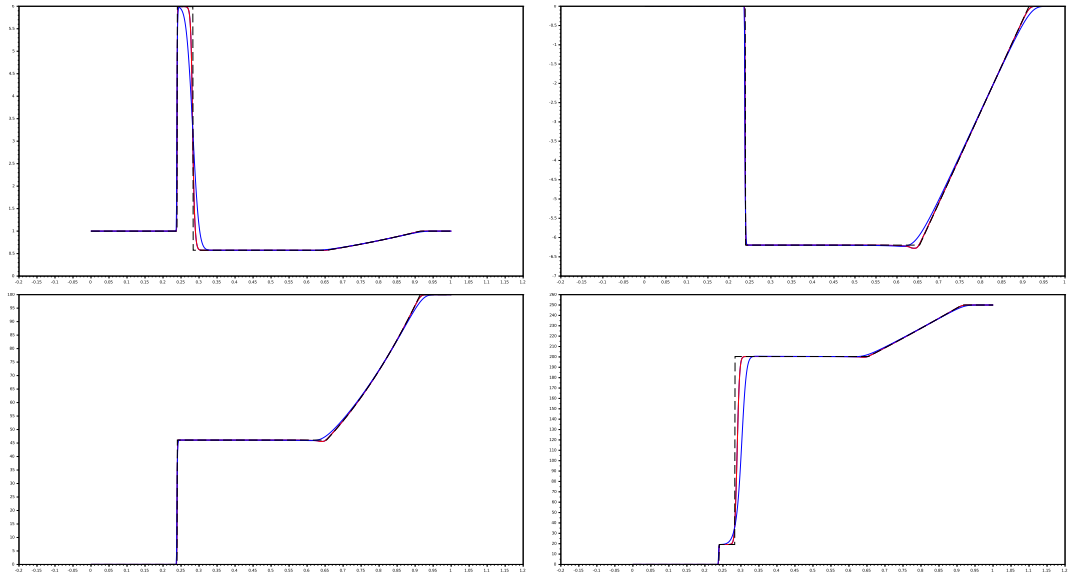


Figure 22: Test 4: has a solution made of a left shock, a contact discontinuity and a right rarefaction.

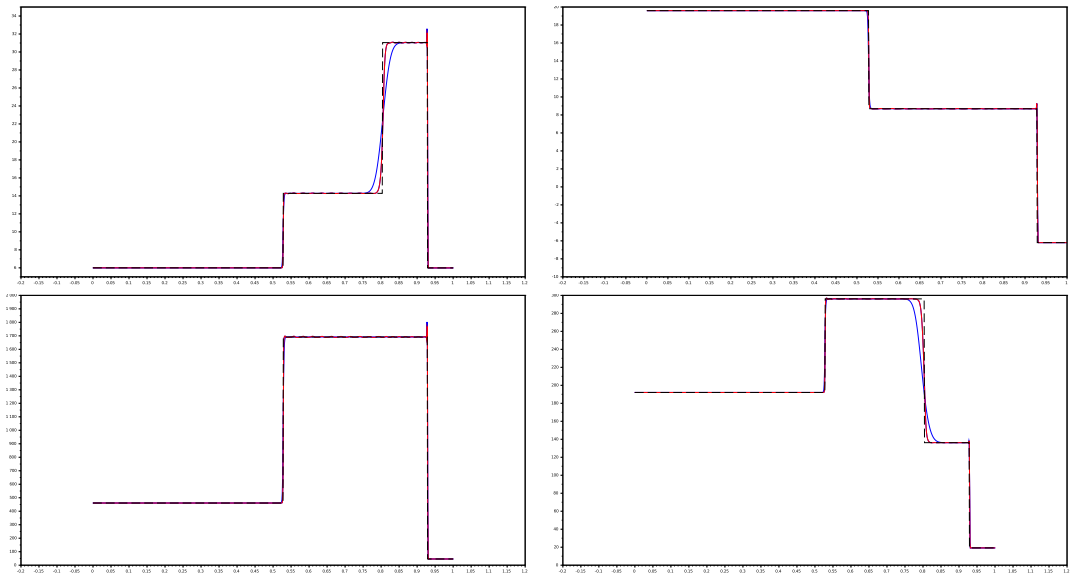


Figure 23: Test 5: corresponds to the collision of two strong shocks, we have a left facing shock, that travels very slowly to the right, a right travelling contact discontinuity and a right travelling shock wave.

Both versions of the scheme produce satisfactory results. As expected, the second order scheme has a reduced numerical diffusion and offers better approximations. Concerning Test #2, we remark that the results obtained for the physical variables ρ, p and u is quite satisfactory while the internal energy presents significative discrepancies. According to [49, Section 6.4, Chapter 6, pages 225-235], this test is indeed known to be particularly challenging for the internal energy e, and even Godunov's method fails on this problem (and the results in Fig. 19 to Fig. 23 are definitely better than what can be obtained with, say, the Lax-Friedrichs scheme). In [49], a detailed description of the test cases is given and the performance of several standard numerical methods is commented. Be aware that from [49, Chapter 5] to the end Test 1 is not the same as in [49, Chapter 4], and that Test 5 from [49, Chapter 4] (which is also our Test #5) becomes Test 4 from [49, Chapter 5] to the end. About the overshoot that may be seen in Test #5 and at a lower level in Test #1 and #3, Toro explains in [49, Chapter 6] that is quite usual, and even unresolved in the case of Test #5. In [49, Section 8.5.5, Chapter 8, pages 282], more details are given about the spurious oscillations and diffusion that appears with most type of resolutions, whereas, similarly to the scheme of Liou and Steffen shown in this part of [49], our scheme sharply solves the fast right shock but at the price of the creation of an overshoot.

5.2.3 Numerical simulations in 2d

This test case is the well-known 2D Mach 3 wind tunnel with a step, introduced in [55]. The computational domain Ω is the L-shaped domain

$$\Omega = \Omega_0 \setminus \Omega_{\text{step}}, \quad \Omega_0 = [0,3] \times [0,1], \quad \Omega_{\text{step}} = [0.6,3] \times [0,0.2].$$

The rectangle Ω_0 is discretized with a $30\sigma \times 10\sigma$ uniform Cartesian grid ($\sigma \in \mathbb{N}^*$). We take the step into account by removing the mesh points corresponding to the domain Ω_{step} at the right bottom part of the domain.

The equation of state of the fluid is $p(\rho, e) = (\gamma - 1)\rho e$ where $\gamma = 1.4$, and the initial data are given by $\rho = 1.4$, $u = (3, 0)$ and $p = 1$. On the top and bottom walls, we use reflection

boundary conditions which means zero flux boundary conditions as described in the previous parts of the paper). A Dirichlet boundary condition, $\rho = 1.4$ and $u = (3, 0)$, makes the flow enter through the left boundary whereas a free boundary condition is used for the right section. The free boundary condition is implemented by assuming *a priori* that the outgoing flow is supersonic. The mass fluxes at the free boundary are thus defined with an incoming part set to 0; that is, here, on the right boundary, we set

$$\mathcal{F}^x_{M+1,j+\frac{1}{2}} = \mathcal{F}^+(\rho_{M+\frac{1}{2},j+\frac{1}{2}}, u_{M+1,j+\frac{1}{2}}).$$

The update of the velocity at the free boundary is performed by involving momentum fluxes at the exterior the domain (here at the points $(x_{M+\frac{3}{2}}, y_{j+\frac{1}{2}})$). These ghost fluxes are defined using a constant extrapolation for the pressure and a linear extrapolation for the outgoing part of the mass fluxes; that is here on the right boundary $\left(x_{M+\frac{3}{2}}, y_{j+\frac{1}{2}}\right)$ (using a uniform mesh)

$$u_{M+1,j+\frac{1}{2}} \mathcal{F}^{x,+}_{M+\frac{3}{2},j+\frac{1}{2}} + p(\rho_{M+\frac{1}{2},j+\frac{1}{2}}),$$

with $\mathcal{F}^{x,+}_{M+\frac{3}{2},j+\frac{1}{2}} = \frac{3}{2}\mathcal{F}^{x,+}_{M+\frac{1}{2},j+\frac{1}{2}} - \frac{1}{2}\mathcal{F}^{x,+}_{M-\frac{1}{2},j+\frac{1}{2}}$.

We present the results obtained with the first order scheme and the second order scheme on a 960×320 grid ($\sigma = 32$) with time steps respectively defined by $\delta t = 1/(100\sigma)$ and $\delta t = 1/(400\sigma)$ at time $T = 4$. We observe that the structures are sharper with the MUSCL scheme in Fig. 24 and cutlines of the density along the lines $y = 0.3$ obtained using the first and second order scheme are plotted in Fig. 25.

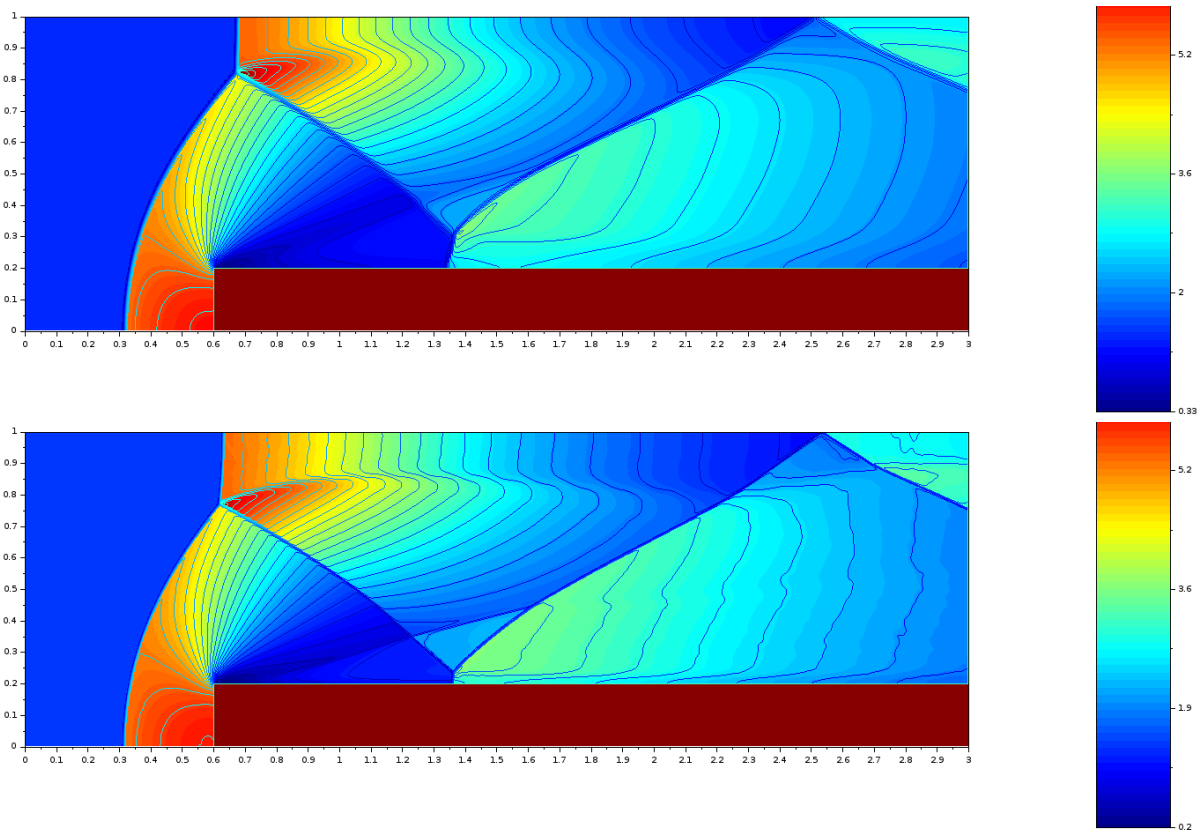


Figure 24: Simulation of the 2D Mach 3 wind tunnel with a step: Density with the first order scheme (up) and with the MUSCL scheme (down).

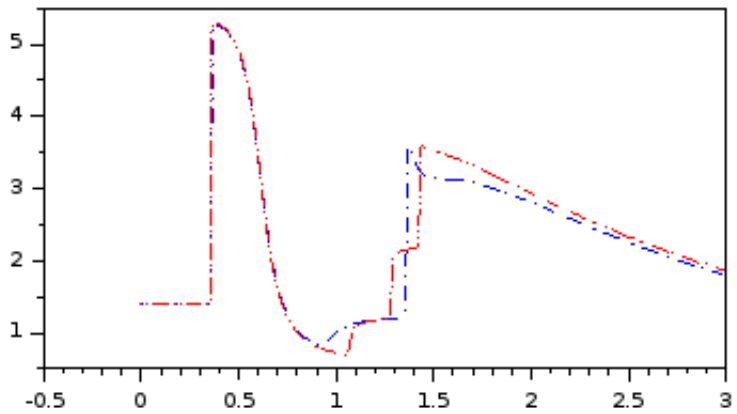


Figure 25: Simulation of the 2D Mach 3 wind tunnel with a step: density cutlines at $y = 0.3$ for the first (blue) and second (red) order scheme.

Acknowledgements
We warmly thank N. Aguillon for invaluable hints about the test cases in her PhD thesis.

References

[1] R. Abgrall, P. Bacigaluppi, and S. Tokareva. A high-order nonconservative approach for hyperbolic equations in fluid dynamics. Technical report, Institute of Mathematics, University of Zurich, Switzerland, 2017. Available as `hal-01476636`.

[2] N. Aguillon. *Problèmes d'interfaces et couplages singuliers dans les systèmes hyperboliques: analyse et analyse numérique*. PhD thesis, Univ. Paris Sud, 2014.

[3] M. Berger, M.-J. Aftosmis, and S.-M. Murman. Analysis of slope limiters on irregular grids. In *AIAA 43rd Aerospace Sciences Meeting, Reno, NV, USA*, 2005.

[4] F. Berthelin, T. Goudon, and S. Minjeaud. Consistency analysis of a 1D finite volume scheme for barotropic Euler models. In J. Fuhrmann, M. Ohlberger, and C. Rohde, editors, *Finite Volumes for Complex Applications VII; Methods, Theoretical Aspects, and Elliptic, Parabolic and Hyperbolic Problems, Berlin*, volume 77 and 78 of *Springer Proceedings in Mathematics & Statistics*, pages 97–106. Springer, 2014.

[5] F. Berthelin, T. Goudon, and S. Minjeaud. Kinetic schemes on staggered grids for barotropic Euler models: entropy-stability analysis. *Math. Comput.*, 84:2221–2262, 2015.

[6] F. Berthelin, T. Goudon, and S. Minjeaud. Multifluid flows: a kinetic approach. *J. Sci. Comput.*, 66(2):792–824, 2016.

[7] C. Berthon. Stability of the MUSCL schemes for the Euler equations. *Comm. Math. Sciences*, 3(2):133–157, 2005.

[8] C. Berthon. Robustness of MUSCL schemes for 2D unstructured meshes. *J. Comput. Physics*, 218(2):495–509, 2006.

[9] C. Berthon, Y. Coudière, and V. Desveaux. Development of DDFV methods for the Euler equations. In J. Fort, J. Fürst, J. Halama, R. Herbin, and F. Hubert, editors, *Finite Volumes for Complex Applications VI; Methods, Theoretical Aspects, and Elliptic, Parabolic*

and Hyperbolic Problems, Prague, volume 4 of *Springer Proceedings in Mathematics & Statistics*, pages 117–124. Springer, 2011.

[10] C. Berthon, Y. Coudière, and V. Desveaux. Second-order MUSCL schemes based on dual mesh gradient reconstruction (DMGR). *ESAIM Math. Model. Numer. Anal.*, 48(2):583–602, 2014.

[11] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources.* Frontiers in Math. Birkhauser, 2004.

[12] C. Calgaro, E. Chane-Kane, E. Creusé, and T. Goudon. L^∞-stability of vertex-based MUSCL finite volume schemes on unstructured grids: simulation of incompressible flows with high density ratios. *J. Comput. Phys.*, 229(17):6027–6046, 2010.

[13] C. Calgaro, E. Creusé, T. Goudon, and Y. Penel. Positivity-preserving schemes for Euler equations: sharp and practical CFL conditions. *J. Comput. Phys.*, 234:417–438, 2013.

[14] C. Chalons and P. Goatin. Transport-equilibrium schemes for computing contact discontinuities in traffic flow modeling. *Commun. Math. Sci.*, 5(3):533–551, 2007.

[15] S. Clain and V. Clauzon. L^∞ stability of the MUSCL methods. *Numerische Mathematik*, 116:31–64, 2010.

[16] F. Clarelli, C. Di Russo, N. Natalini, and M. Ribot. Mathematical models for biofilms on the surface of monuments. In *Applied and Industrial Mathematics in Italy III, 9th Conference SIMAI*, volume 82 of *Advances in Mathematics for Applied Sciences*. World Scientific, 2009.

[17] F. Clarelli, C. Di Russo, N. Natalini, and M. Ribot. A fluid dynamics model of the growth of phototrophic biofilms. *J. Math. Biology*, 66(7):1387–408, 2013.

[18] F. Coron and B. Perthame. Numerical passage from kinetic to fluid equations. *SIAM J. Numer. Anal.*, 28:26–42, 1991.

[19] C. Dafermos. *Hyperbolic conservvation laws in continuum physics*, volume 325 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2010. Third ed.

[20] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number. *J. Comput. Phys.*, 229(4):978–1016, 2010.

[21] S. Dellacherie, P. Omnes, and F. Rieper. The influence of cell geometry on the Godunov scheme applied to the linear wave equation. *J. Comput. Phys.*, 229:5315–5338, 2010.

[22] S. M. Deshpande. Kinetic theory based new upwind methods for inviscid compressible flows. In *AIAA 24th Aerospace Science Meeting, Jan 6-9, 1986, Nevada, USA*, 1986. AIAA paper 86-0275.

[23] S. M. Deshpande. On the Maxwellian distribution, symmetric form and entropy conservation for the Euler equations. Technical report, NASA Langley Research Centre, Hampton, VA, 1986. NASA TP2613.

[24] L. Gastaldo, R. Herbin, W. Kheriji, C. Lapuerta, and J.-C. Latché. Staggered discretizations, pressure correction schemes and all speed barotropic flows. In *Finite Volumes for Complex Applications VI, Problems and Perspectives, Prague, Czech Republic*, volume 4, pages 839–855, 2011.

[25] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer, New-York, 1996.

[26] F. H. Harlow and J. E. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys. Fluids*, 8(12):2182–2189, 1965.

[27] R. Herbin, W. Kheriji, and J.-C. Latché. Staggered schemes for all speed flows. *ESAIM:Proc*, 35:122–150, 2012. Actes du Congrès National de Mathématiques Appliquées et Industrielles.

[28] R. Herbin, J.-C. Latché, and T. T. Nguyen. Consistent explicit staggered schemes for compressible flows; part I: the barotropic Euler equations. Technical report, LATP, Univ. Aix-Marseille & CNRS, 2013.

[29] R. Herbin, J.-C. Latché, and T. T. Nguyen. Explicit staggered schemes for the compressible Euler equations. In *Applied mathematics in Savoie—AMIS 2012: Multiphase flow in industrial and environmental engineering*, volume 40 of *ESAIM Proc.*, pages 83–102. EDP Sci., Les Ulis, 2013.

[30] T. Y. Hou and P. G. Le Floch. Why nonconservative schemes converge to wrong solutions: Error analysis. *Math. Comput.*, 62(206):497–530, 1994.

[31] S. Kaniel and J. Falcovitz. Approximation of the hydrodynamic equations by a transport process. In R. Rautman, editor, *Proceedings of IUTAM Symposium on Approximation Methods for Navier-Stokes Problems*, volume 771 of *Lecture Notes in Math.* Springer-Verlag, 1980.

[32] S. Karni. Multicomponent flow calculations by a consistent primitive algorithm. *J. Comput. Phys.*, 112(1):31–43, 1994.

[33] S. Karni. Hybrid multifluid algorithms. *SIAM J. Sci. Comput.*, 17(5):1019–1039, 1996.

[34] W. Kheriji, R. Herbin, and J.-C. Latché. Pressure correction staggered schemes for barotropic one-phase and two-phase flows. *Computers & Fluids*, 88:524 – 542, 2013.

[35] R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.

[36] M.-S. Liou. A sequel to AUSM: AUSM+. *J. Comput. Phys.*, 129:364–382, 1996.

[37] M.-S. Liou and C. J. Steffen Jr. A new flux splitting scheme. *J. Comput. Phys.*, 107:23–39, 1993.

[38] J. Llobell. *Schémas Volumes Finis à mailles décalées pour la dynamique des gaz*. PhD thesis, Université Côte d'Azur, 2018.

[39] J. Llobell, S. Minjeaud, and R. Pasquetti. High order CG schemes for KdV and Saint-Venant flows. In *XIXth International Conference on Finite Elements in Flow Problems, Roma, Italy*, 2018, submitted.

[40] H. Nessyahu and E. Tadmor. Non-oscillatory central differencing for hyperbolic conservation laws. *J. Comput. Phys.*, 87(2):408–463, 1990.

[41] R. Pasquetti. Viscous stabilizations for high order approximations of saint-venant and boussinesq flows. *Lecture Notes in computational Science and Engineering: Spectral and High Order Methods for Partial Differential Equations - ICOSAHOM 2016*, 119:519–531, 2017.

[42] R. Pasquetti, J. Guermond, and B. Popov. Stabilized spectral element approximation of the saint-venant system using the entropy viscosity technique. *Lecture Notes in computational Science and Engineering: Spectral and High Order Methods for Partial Differential Equations - ICOSAHOM 2014*, 106:397–404, 2015.

[43] B. Perthame. Second order Boltzmann schemes for compressible Euler equations in one and two space dimension. *SIAM J. Numer. Anal.*, 29(1):1–19, 1992.

[44] B. Perthame. *Kinetic formulation of conservation laws.* Oxford Lecture Series in Math. and its Appl. Oxford University Press, 2003.

[45] B. Perthame and C.-W. Shu. On positivity preserving finite volume schemes for Euler equations. *Numer. Math*, 73:119–130, 1996.

[46] B. Polizzi. *Modeling and numerical simulations for fluid mechanics systems with constraints; application to biology and road traffic.* PhD thesis, Univ. Côte d'Azur, 2016.

[47] P. Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, 21(5):995–1011, 1984.

[48] N. Therme. *Schémas numériques pour la simulation de l'explosion.* PhD thesis, Aix-Marseille Univ., 2015.

[49] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics.* Springer-Verlag, Berlin, third edition, 2009.

[50] D. R. van der Heul, C. Vuik, and P. Wesseling. A conservative pressure-correction method for flow at all speeds. *Comput. & Fluids*, 32(8):1113–1132, 2003.

[51] B. van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.*, 135(2):227–248, 1997. With an introduction by Ch. Hirsch, Commemoration of the 30th anniversary of J. Comput. Phys.

[52] B. van't Hof and A. E. P. Veldman. Mass, momentum and energy conserving (MaMEC) discretizations on general grids for the compressible Euler and shallow water equations. *J. Comput. Phys.*, 231(14):4723–4744, 2012.

[53] I. Wenneker, A. Segal, and P. Wesseling. A Mach-uniform unstructured staggered grid method. *Internat. J. Numer. Methods Fluids*, 40(9):1209–1235, 2002.

[54] I. Wenneker, A. Segal, and P. Wesseling. Conservation properties of a new unstructured staggered scheme. *Comput. & Fluids*, 32(1):139–147, 2003.

[55] P. Woodward and P. Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.*, 54(1):115–173, 1984.

[56] C. Zaza. *Contribution à la résolution numérique d'écoulements à tout nombre de Mach et au couplage fluide-poreux en vue de la simulation d'écoulements diphasiques homogénéisés dans les composants nucléaires.* PhD thesis, Aix–Marseille Univ., 2015.