Consistent segregated staggered schemes with explicit steps for the isentropic and full Euler equations

Raphaele Herbin, J.-C Latché, Trung Tan Nguyen

CONSISTENT SEGREGATED STAGGERED SCHEMES WITH EXPLICIT STEPS FOR THE ISENTROPIC AND FULL EULER EQUATIONS.

R. Herbin[1], J.-C. Latché[2] and T.T. Nguyen[3]

Abstract. In this paper, we build and analyze the stability and consistency of decoupled schemes, involving only explicit steps, for the isentropic Euler equations and for the full Euler equations. These schemes are based on staggered space discretizations, with an upwinding performed with respect to the material velocity only. The pressure gradient is defined as the transpose of the natural velocity divergence, and is thus centered. The velocity convection term is built in such a way that the solutions satisfy a discrete kinetic energy balance, with a remainder term at the left-hand side which is shown to be non-negative under a CFL condition. In the case of the full Euler equations, we solve the internal energy balance, to avoid the space discretization of the total energy, whose expression involves cell-centered and face-centered variables. However, since the residual terms in the kinetic energy balance (probably) do not tend to zero with the time and space steps when computing shock solutions, we compensate them by corrective terms in the internal energy equation, to make the scheme consistent with the conservative form of the continuous problem. We then show, in one space dimension, that, if the scheme converges, the limit is indeed an entropy weak solution of the system. In any case, the discretization preserves by construction the convex of admissible states (positivity of the density and, for Euler equations, of the internal energy), under a CFL condition. Finally, we present numerical results which confort this theory.

2010 AMS Subject Classification. 35Q31,65N12,65M12,76M10,76M12.

December 13, 2016.

Contents

1. Introduction

The objective pursued in this work is to develop and study, both theoretically and numerically, a decoupled scheme for the simulation of non viscous compressible flows, modeled either by the isentropic Euler equations or by the full Euler equations for an ideal gas. More precisely, we intend to build a variant involving only explicit time-steps (*i.e.* without any linear system solution) of implicit and semi-implicit schemes that were developed and studied recently in the framework of the simulation of compressible flows at all speeds [13,17]. In this latter works, the implicit scheme is studied as a first step in the mathematical analysis of some pressure correction schemes; these are obtained by extending some algorithms which are classical in the incompressible framework;

Keywords and phrases: finite volumes, staggered grid, Euler equations, isentropic barotropic, compressible flows, shallo water, analysis.

[1] 'Aix-Marseille Université, I2M, UMR CNRS 7373, France (raphaele.herbin@univ-amu.fr)

[2] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), France (jean-claude.latche]@irsn.fr)

[3] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), France (tan-trung.nguyen@centralesupelec.fr)

they are based on (inf-sup stable) staggered discretizations. In our approach, the upwinding techniques which are implemented for stability reasons are performed for each equation separately and with respect to the material velocity only. This is in contradiction with the most common strategy adopted for hyperbolic systems, where upwinding is built from the wave structure of the system (see *e.g.* [2, 7, 23] for surveys). However, it yields algorithms which are used in practice (see e.g. the so-called AUSM family of schemes [19, 20]), because of their generality (a closed-form solution of Riemann problems is not needed), their implementation simplicity and their efficiency, thanks to an easy construction of the fluxes at the cell faces. Up to now, these schemes have scarcely been studied from a theoretical point of view; one of our main concerns here will thus be to bring, as far as possible, theoretical arguments supporting our numerical developments.

We first deal with the isentropic Euler equations:

$$\partial_t \rho + \mathrm{div}(\rho \boldsymbol{u}) = 0, \tag{1a}$$

$$\partial_t (\rho \boldsymbol{u}) + \mathrm{div}(\rho \boldsymbol{u} \otimes \boldsymbol{u}) + \boldsymbol{\nabla} p = 0, \tag{1b}$$

$$\rho \geq 0,\ p = \wp(\rho) = \rho^\gamma, \tag{1c}$$

where t stands for the time, ρ, $\boldsymbol{u}$, and p are the density, velocity and pressure respectively, and $\gamma \geq 1$ is a coefficient specific to the considered fluid. Note that, for $\gamma = 2$, this system is identical to the usual shallow water (or Saint-Venant) equations in the case of no source term (no topography, no Coriolis force), up to a multiplicative coefficient $1/2$ at the right-hand side of the equation of state (and replacing the density ρ by the fluid height h). Of course, this minor change in the equation of state does not bring any additional difficulty, and, more generally, present results may probably be extended to the barotropic case, *i.e.* to general equations of state of the form $p = \phi(\rho)$ with ϕ a strictly increasing function.

We then address the full Euler equations, which read:

$$\partial_t \rho + \mathrm{div}(\rho\, \boldsymbol{u}) = 0, \tag{2a}$$

$$\partial_t (\rho\, \boldsymbol{u}) + \mathrm{div}(\rho\, \boldsymbol{u} \otimes \boldsymbol{u}) + \boldsymbol{\nabla} p = 0, \tag{2b}$$

$$\partial_t (\rho\, E) + \mathrm{div}(\rho\, E\, \boldsymbol{u}) + \mathrm{div}(p\, \boldsymbol{u}) = 0, \tag{2c}$$

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

where E and e are the total energy and internal energy respectively. For this system, the coefficient γ is now supposed to be strictly greater than 1. Problems (1) and (2) are posed over $\Omega \times (0, T)$, where Ω is an open bounded connected subset of $\mathbb{R}^d$, $1 \leq d \leq 3$, and $(0, T)$ is a finite time interval. They are complemented by initial conditions for ρ, e and $\boldsymbol{u}$, denoted by ρ_0 e_0, and $\boldsymbol{u}_0$ respectively, with $\rho_0 > 0$ and $e_0 > 0$, and by a boundary condition which we suppose to be $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ at any time and *a.e.* on $\partial\Omega$, where $\boldsymbol{n}$ stands for the normal vector to the boundary.

The organization and main results of this paper are as follows.

- The space discretization is given in Section 2.
- Section 3 is devoted to the isentropic Euler equations (*i.e.* System (1)). The proposed scheme is decoupled in time (the mass and momentum equations are solved one after the other) and only involves explicit steps. The scheme is based on a staggered mesh, and obtained by writing a finite volume discretization on the primal cells for the mass balance equation and a finite volume scheme on the dual cells for the momentum balance equation. Upwinding is performed with respect to the material velocity (by opposition to the speed of the waves of the system), and the pressure gradient is defined as the transpose of the natural velocity divergence, so is thus centered. We prove that the solutions of this scheme satisfy a discrete kinetic energy balance (on dual cells) and an elastic potential balance (on primal cells). Then, in one space dimension, we show that the algorithm is consistent in the Lax-Wendroff sense: passing to the limit in the scheme, we prove that, if a sequence of discrete solutions obtained with vanishing time and space steps converges and is uniformly bounded in suitable norms, then its limit satisfies a weak formulation of the continuous problem. Then, passing now to the limit in the discrete kinetic energy and elastic potential equations, we show that the limit of such a converging sequence also satisfies the weak form of the entropy balance.
- Section 4 is devoted to the full Euler equations. The scheme is obtained by complementing the algorithm developed for the isentropic case by an explicit finite volume discretization of the internal energy balance equation on the primal mesh. This offers two main advantages: first, we avoid the space discretization of the total energy, the expression of which involves cell-centered and face-centered variables; second,

the discretization ensures by construction the positivity of the internal energy, under a CFL condition. However, since this scheme does not use the original (total) energy conservative equation, in order to obtain correct weak solutions (in particular, with shocks satisfying the Rankine-Hugoniot conditions), we need to introduce corrective terms in the internal energy balance. These corrective terms are found from the discrete kinetic energy balance (already derived in the isentropic case), observing that this relation contains residual terms which do not tend to zero (at least, under reasonable stability assumptions) and, finally, compensating them in the discrete internal energy balance. With this correction, we are once again able to prove, in 1D, the consistency of the scheme in the Lax-Wendroff sense; more precisely speaking, passing to the limit separately in the discrete kinetic and internal energy balances (which are not posed on the same mesh), we obtain that the limit of a convergent sequence of discrete solutions satisfies a weak form of the total energy equation.

- Finally, we present some numerical tests for both the isentropic case and the full Euler case in Section 5.

In several theoretical developments, we are lead to use a derived form of a discrete finite volume convection operator (for instance, typically, a convection operator for the kinetic energy, possibly with residual terms, obtained from the finite volume discretization of the convection of the velocity components); the proofs of various related discrete identities are given in the Appendix A. Note that some of the results of the work which we present here were announced in the proceedings [15], but without any proof.

2. Meshes and unknowns

In this section, we recall some staggered discretizations which were already used for implicit schemes for compressible flows, see e.g. [13]; we focus here on the discretization of a multi-dimensional domain (*i.e.* $d = 2$ or $d = 3$); the extension to the one-dimensional case is straightforward (see sections 3.3 and 4.2).

Let $\mathcal{M}$ be a mesh of the domain Ω, supposed to be regular in the usual sense of the finite element literature (*e.g.* [4]). The cells of the mesh are assumed to be:

- for a general domain Ω, either non-degenerate quadrilaterals ($d = 2$) or hexahedra ($d = 3$) or simplices; in two space dimensions, both types of cells may possibly be combined in a same mesh;
- for a domain whose boundaries are hyperplanes normal to a coordinate axis, rectangles ($d = 2$) or rectangular parallelepipeds ($d = 3$) (the faces of which, of course, are then also necessarily normal to a coordinate axis).

By $\mathcal{E}$ and $\mathcal{E}(K)$ we denote the set of all $(d-1)$-faces σ of the mesh and of the element $K \in \mathcal{M}$ respectively. The set of faces included in the boundary of Ω is denoted by $\mathcal{E}_{\rm ext}$ and the set of internal faces (*i.e.* $\mathcal{E} \setminus \mathcal{E}_{\rm ext}$) is denoted by $\mathcal{E}_{\rm int}$; a face $\sigma \in \mathcal{E}_{\rm int}$ separating the cells K and L is denoted by $\sigma = K|L$. The outward normal vector to a face σ of K is denoted by $\boldsymbol{n}_{K,\sigma}$. For $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}$, we denote by $|K|$ the measure of K and by $|\sigma|$ the $(d-1)$-measure of the face σ. For $1 \leq i \leq d$, we denote by $\mathcal{E}^{(i)} \subset \mathcal{E}$ and $\mathcal{E}^{(i)}_{\rm ext} \subset \mathcal{E}_{\rm ext}$ the subset of the faces of $\mathcal{E}$ and $\mathcal{E}_{\rm ext}$ respectively which are perpendicular to the i^{th} unit vector of the canonical basis of $\mathbb{R}^d$.

The space discretization is staggered, using either the Marker-And Cell (MAC) scheme [11,12], or the degrees of freedom (*i.e.* the discrete unknowns) of nonconforming low-order finite element approximations, namely the Rannacher and Turek element (RT) [21] for quadrilateral or hexahedric meshes, or the lowest degree Crouzeix-Raviart element (CR) [5] for simplicial meshes.

For all these space discretizations, the degrees of freedom for the pressure, the density and the internal energy are associated to the cells of the mesh $\mathcal{M}$, and are denoted by:

$$\big\{p_K,\ \rho_K,\ e_K,\ K \in \mathcal{M}\big\}.$$

Let us then turn to the degrees of freedom for the velocity (*i.e.* the discrete velocity unknowns).

- **Rannacher-Turek** or **Crouzeix-Raviart** discretizations – In this case, all the components of the velocity are approximated on each face of the mesh, so the degrees of freedom for the velocity components are located at their center. The set of degrees of freedom reads:

$$\{u_{\sigma,i},\ \sigma \in \mathcal{E},\ 1 \leq i \leq d\}.$$

- **MAC** discretization – Only the normal components of the velocities are approximated, and the degrees of freedom for the i^{th} component of the velocity are defined at the centre of the faces $\sigma \in \mathcal{E}^{(i)}$, and the set of discrete velocity unknowns is:

$$\big\{u_{\sigma,i},\ \sigma \in \mathcal{E}^{(i)},\ 1 \leq i \leq d\big\}.$$

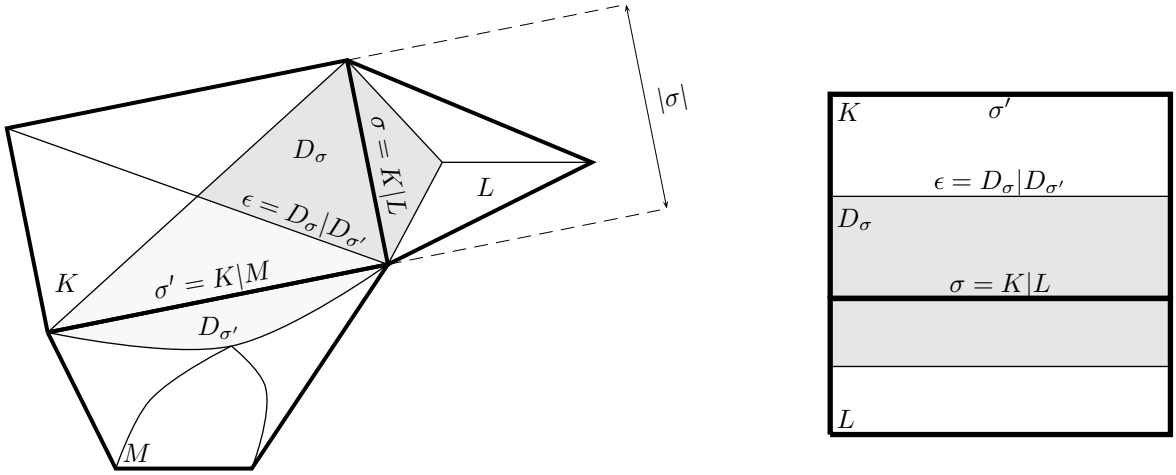

FIGURE 1. Notations for control volumes and dual cells – Left: Finite Elements (the present sketch illustrates the possibility, implemented in our software CALIF^{3}S [3], of mixing simplicial (Crouzeix-Raviart) and quadrangular (Rannacher-Turek) cells) – Right: MAC discretization, dual cell for the y-component of the velocity.

We now introduce a dual mesh, which will be used for the finite volume approximation of the time derivative and convection terms in the momentum balance equation.

- **Rannacher-Turek** or **Crouzeix-Raviart** discretizations – For the RT or CR discretizations, the dual mesh is the same for all the velocity components. When $K \in \mathcal{M}$ is a simplex, a rectangle or a cuboid, for $\sigma \in \mathcal{E}(K)$, we define $D_{K,\sigma}$ as the cone with basis σ and with vertex the mass center of K (see Figure 1). We thus obtain a partition of K in m sub-volumes, where m is the number of faces of the mesh, each sub-volume having the same measure $|D_{K,\sigma}| = |K|/m$. We extend this definition to general quadrangles and hexahedra, with a (virtual) partition with sub-cells which are still of equal-volume, and with the same connectivities. The volume $D_{K,\sigma}$ is referred to as the half-diamond cell associated to K and σ.
  For $\sigma \in \mathcal{E}_{\rm int}$, $\sigma = K|L$, we now define the diamond cell D_σ associated to σ by $D_\sigma = D_{K,\sigma} \cup D_{L,\sigma}$; for an external face $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}(K)$, D_σ is just the same volume as $D_{K,\sigma}$.
- **MAC** discretization – For the MAC scheme, the definition of the dual mesh depends on the component of the velocity. For each component, the MAC dual mesh only differs from the RT or CR dual mesh by the choice of the half-diamond cell, which, for $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}(K)$, is now the rectangle or rectangular parallelepiped of basis σ and of measure $|D_{K,\sigma}| = |K|/2$.

We denote by $|D_\sigma|$ the measure of the dual cell D_σ, and by $\epsilon = D_\sigma|D_{\sigma'}$ the face separating two diamond cells D_σ and $D_{\sigma'}$. The set of the faces of a dual cell D_σ is denoted by $\tilde{\mathcal{E}}(D_\sigma)$.

Finally, we need to deal with the impermeability (*i.e.* $\boldsymbol{u} \cdot \boldsymbol{n} = 0$) boundary condition. As in [13] we suppose throughout this paper that the boundary is *a.e.* normal to a coordinate axis which allows to simply set to zero the corresponding velocity unknowns:

$$\text{for } i = 1, \ldots, d, \ \forall \sigma \in \mathcal{E}_{\rm ext}^{(i)}, \qquad u_{\sigma,i} = 0. \tag{3}$$

Therefore, there are no degrees of freedom for the velocity on the boundary for the MAC scheme, and there are only $d-1$ degrees of freedom on each boundary face for the CR and RT discretizations, which depend on the orientation of the face. We again use the notations of [13] to be able to write a unique expression of the discrete equations for both MAC and CR/RT schemes: we introduce the sets of faces $\mathcal{E}_{\mathcal{S}}^{(i)}$ associated to the degrees of freedom of each component of the velocity ($\mathcal{S}$ stands for "scheme"):

$$\mathcal{E}_{\mathcal{S}}^{(i)} = \left| \begin{array}{l} \mathcal{E}^{(i)} \setminus \mathcal{E}_{\rm ext}^{(i)} \text{ for the MAC scheme,} \\ \mathcal{E} \setminus \mathcal{E}_{\rm ext}^{(i)} \text{ for the CR or RT schemes.} \end{array} \right.$$

For both schemes, we define $\tilde{\mathcal{E}}^{(i)}$, for $1 \leq i \leq d$, as the set of faces of the dual mesh associated to the i^{th} component of the velocity. For the RT or CR discretizations, the sets $\tilde{\mathcal{E}}^{(i)}$ does not depend on the component (*i.e.* of i), up to the elimination of some unknowns (and so some dual cells and, finally, some external faces) to

take the boundary conditions into account. For the MAC scheme, $\tilde{\mathcal{E}}^{(i)}$ depends on i; note that each face of $\tilde{\mathcal{E}}^{(i)}$ is perpendicular to a unit vector of the canonical basis of $\mathbb{R}^d$, but not necessarily to the i^{th} one.

General domains can be addressed with the CR or RT discretizations by redefining, through linear combinations, the degrees of freedom at the external faces, so as to introduce the normal velocity as a new degree of freedom.

3. The isentropic Euler equations

We address in this section the numerical solution of the isentropic Euler equations (*i.e.* System (1)). The presentation is organized as follows. We first describe the proposed scheme (Section 3.1). Then, we study its stability properties in Section 3.2 (precisely speaking, we show that the solutions satisfy a discrete kinetic energy and an elastic potential balance). Finally, consistency properties of the scheme, in 1D, are studied in Section 3.3.

3.1. **The scheme**

Let us consider a discretization $0 = t_0 < t_1 < \ldots < t_N = T$ of the time interval $(0, T)$, which we suppose uniform for the sake of simplicity, and let $\delta t = t_{n+1} - t_n$ for $n = 0, 1, \ldots, N-1$ be the (constant) time step. We consider a decoupled-in-time scheme, which reads in its fully discrete form, for $0 \leq n \leq N-1$:

$$\forall K \in \mathcal{M}, \qquad \frac{|K|}{\delta t}(\rho_K^{n+1} - \rho_K^n) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^n = 0, \tag{4a}$$

$$\forall K \in \mathcal{M}, \qquad p_K^{n+1} = \wp(\rho_K^{n+1}) = (\rho_K^{n+1})^\gamma, \tag{4b}$$

For $1 \leq i \leq d$, $\forall \sigma \in \mathcal{E}_{\rm s}^{(i)}$,

$$\frac{|D_\sigma|}{\delta t}(\rho_{D_\sigma}^{n+1} u_{\sigma,i}^{n+1} - \rho_{D_\sigma}^n u_{\sigma,i}^n) + \sum_{\epsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\epsilon}^n u_{\epsilon,i}^n + |D_\sigma| \, (\boldsymbol{\nabla} p)_{\sigma,i}^{n+1} = 0, \tag{4c}$$

where the terms introduced for each discrete equation are defined herafter.

Equation (4a) is obtained by the discretization of the mass balance equation (1a) over the primal mesh, and $F_{K,\sigma}^n$ stands for the mass flux across σ outward K, which, because of the impermeability condition, vanishes on external faces and is given on the internal faces by:

$$\forall \sigma = K|L \in \mathcal{E}_{\rm int}, \qquad F_{K,\sigma}^n = |\sigma| \; \rho_\sigma^n \; u_{K,\sigma}^n. \tag{5}$$

In this relation, $u_{K,\sigma}^n$ is an approximation of the normal velocity to the face σ outward K, defined by:

$$u_{K,\sigma}^n = \begin{cases} u_{\sigma,i}^n \, \boldsymbol{e}^{(i)} \cdot \boldsymbol{n}_{K,\sigma} \text{ for } \sigma \in \mathcal{E}^{(i)} \text{ in the MAC case,} \\ \boldsymbol{u}_\sigma^n \cdot \boldsymbol{n}_{K,\sigma} \text{ in the CR and RT cases,} \end{cases} \tag{6}$$

where $\boldsymbol{e}^{(i)}$ denotes the i-th vector of the orthonormal basis of $\mathbb{R}^d$. The density at the face $\sigma = K|L$ is approximated by the upwind technique:

$$\rho_\sigma^n = \begin{cases} \rho_K^n & \text{if } u_{K,\sigma}^n \geq 0, \\ \rho_L^n & \text{otherwise.} \end{cases} \tag{7}$$

We now turn to the discrete momentum balance (4c), which is obtained by discretizing the momentum balance equation (1b) on the dual cells associated to the faces of the mesh. The first task is to define the values $\rho_{D_\sigma}^{n+1}$ and $\rho_{D_\sigma}^n$, which approximate the density over the dual cell D_σ at time t^{n+1} and t^n respectively, and the discrete mass flux through the dual face ϵ outward D_σ, denoted by $F_{\sigma,\epsilon}^n$; the guideline for their construction is that we want a finite volume discretization of the mass balance equation over the diamond cells, of the form

$$\forall \sigma \in \mathcal{E}, \qquad \frac{|D_\sigma|}{\delta t} \, (\rho_{D_\sigma}^{n+1} - \rho_{D_\sigma}^n) + \sum_{\epsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\epsilon}^n = 0, \tag{8}$$

to hold in order to be able to derive a discrete kinetic energy balance (see Section 3.2 below). The density on the dual cells is given by the following weighted average:

$$\text{for } \sigma = K|L \in \mathcal{E}_{\rm int},\ \text{for } k = n \text{ and } k = n+1, \quad |D_\sigma|\ \rho^k_{D_\sigma} = |D_{K,\sigma}|\ \rho^k_K + |D_{L,\sigma}|\ \rho^k_L. \tag{9}$$

For the MAC scheme, the flux through a dual face which is located on two primal faces is the mean value of the sum of fluxes on the two primal faces, and the flux through a dual face located between two primal faces is again the mean value of the sum of fluxes on the two primal faces [14]. In the case of the CR and RT schemes, for a dual face ϵ included in the primal cell K, this flux is computed as a linear combination (with constant coefficients, *i.e.* independent of the cell) of the mass fluxes through the faces of K, *i.e.* the quantities $(F^n_{K,\sigma})_{\sigma\in\mathcal{E}(K)}$ appearing in the discrete mass balance (4a). We refer to [1, 6] for a detailed construction of this approximation. Let us remark that a dual face lying on the boundary is then either also a primal face or the union of the half-part of two primal faces, and, in both cases, the flux across this face is zero. Therefore, the values $u^n_{\epsilon,i}$ are only needed at the internal dual faces, and are upwinded:

$$\text{for } \epsilon = D_\sigma|D_{\sigma'}, \quad u^n_{\epsilon,i} = \begin{vmatrix} u^n_{\sigma,i} \text{ if } F^n_{\sigma,\epsilon} \geq 0, \\ u^n_{\sigma',i} \text{ otherwise.} \end{vmatrix} \tag{10}$$

The last term $(\boldsymbol{\nabla} p)^{n+1}_{\sigma,i}$ stands for the i-th component of the discrete pressure gradient at the face σ. The gradient operator is built as the transpose of the discrete operator for the divergence of the velocity, the discretization of which is based on the primal mesh. Let us denote the divergence of $\boldsymbol{u}^{n+1}$ over $K \in \mathcal{M}$ by $(\mathrm{div}\boldsymbol{u})^{n+1}_K$; its natural approximation reads:

$$\text{for } K \in \mathcal{M}, \qquad (\mathrm{div}\boldsymbol{u})^{n+1}_K = \frac{1}{|K|} \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\ u^{n+1}_{K,\sigma}. \tag{11}$$

Consequently, we choose the components of the pressure gradient as:

$$\text{for } \sigma = K|L \in \mathcal{E}_{\rm int}, \qquad (\boldsymbol{\nabla} p)^{n+1}_{\sigma,i} = \frac{|\sigma|}{|D_\sigma|}\ (p^{n+1}_L - p^{n+1}_K)\ \boldsymbol{n}_{K,\sigma} \cdot \boldsymbol{e}^{(i)}, \tag{12}$$

in order that the following duality relation (with respect to the L^2 inner product) be satisfied:

$$\sum_{K\in\mathcal{M}} |K|\ p^{n+1}_K\ (\mathrm{div}\boldsymbol{u})^{n+1}_K + \sum_{i=1}^{d} \sum_{\sigma\in\mathcal{E}^{(i)}_{\rm s}} |D_\sigma|\ u^{n+1}_{\sigma,i}\ (\boldsymbol{\nabla} p)^{n+1}_{\sigma,i} = 0. \tag{13}$$

Note that, because of the impermeability boundary conditions, the discrete gradient is not defined at the external faces.

Finally, the initial approximations for ρ and $\boldsymbol{u}$ are given by the average of the initial conditions ρ_0 and $\boldsymbol{u}_0$ on the primal and dual cells respectively:

$$\begin{aligned} &\forall K \in \mathcal{M}, \quad \rho^0_K = \frac{1}{|K|} \int_K \rho_0(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \\ &\text{for } 1 \leq i \leq d,\ \forall \sigma \in \mathcal{E}^{(i)}_{\rm s}, \quad u^0_{\sigma,i} = \frac{1}{|D_\sigma|} \int_{D_\sigma} (\boldsymbol{u}_0(\boldsymbol{x}))_i\, \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{14}$$

The following positivity result is a classical consequence of the upwind choice in the mass balance equation.

Lemma 3.1 (Positivity of the density)**.** *Let ρ^0 be given by* (14)*. For $a \in \mathbb{R}$, let us define $a^+ = \max(a, 0)$. Then, since ρ_0 is assumed to be a positive function, $\rho^0 > 0$ and, under the* CFL *condition:*

$$\delta t \leq \frac{|K|}{\displaystyle\sum_{\sigma\in\mathcal{E}(K)} |\sigma|\ (u^n_{K,\sigma})^+}, \quad \forall K \in \mathcal{M} \text{ and for } 0 \leq n \leq N-1, \tag{15}$$

the solution to the scheme satisfies $\rho^n > 0$, for $1 \leq n \leq N$.

3.2. **Discrete kinetic energy and elastic potential balances**

We begin by deriving a discrete kinetic energy balance equation, as was already done in [13] in the implicit and fractional time step cases. Let us denote by E_k the kinetic energy $E_k = \frac{1}{2}|\boldsymbol{u}|^2$. Let us recall that, taking the inner product of (1b) by $\boldsymbol{u}$ yields, after formal compositions of partial derivatives and using the mass balance (1a):

$$\partial_t(\rho E_k) + \mathrm{div}\big(\rho\, E_k\, \boldsymbol{u}\big) + \boldsymbol{\nabla} p \cdot \boldsymbol{u} = 0. \tag{16}$$

This relation is referred to as the kinetic energy balance, and we recover its discrete analogue from the scheme by some equivalent discrete computations.

Lemma 3.2 (Discrete kinetic energy balance)**.** *A solution to the system* (4) *satisfies the following equality, for* $1 \le i \le d$, $\sigma \in \mathcal{E}_{\rm S}^{(i)}$ *and* $0 \le n \le N-1$*:*

$$\frac{1}{2}\frac{|D_\sigma|}{\delta t}\Big[\rho_{D_\sigma}^{n+1}(u_{\sigma,i}^{n+1})^2 - \rho_{D_\sigma}^{n}(u_{\sigma,i}^{n})^2\Big] + \frac{1}{2}\sum_{\epsilon\in\tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\epsilon}^n\ (u_{\epsilon,i}^n)^2 + |D_\sigma|\,(\boldsymbol{\nabla}p)_{\sigma,i}^{n+1}\ u_{\sigma,i}^{n+1} = -R_{\sigma,i}^{n+1}, \tag{17}$$

with:

$$R_{\sigma,i}^{n+1} = \frac{1}{2}\frac{|D_\sigma|}{\delta t}\rho_{D_\sigma}^{n+1}(u_{\sigma,i}^{n+1} - u_{\sigma,i}^{n})^2 + \frac{1}{2}\sum_{\epsilon=D_\sigma|D_{\sigma'}\in\tilde{\mathcal{E}}(D_\sigma)} (F_{\sigma,\epsilon}^n)^-(u_{\sigma',i}^n - u_{\sigma,i}^n)^2 \\ - \sum_{\epsilon=D_\sigma|D_{\sigma'}\in\tilde{\mathcal{E}}(D_\sigma)} (F_{\sigma,\epsilon}^n)^-(u_{\sigma',i}^n - u_{\sigma,i}^n)\,(u_{\sigma,i}^{n+1} - u_{\sigma,i}^n), \tag{18}$$

where, for $a \in \mathbb{R}$, $a^- \ge 0$ *is defined by* $a^- = -\min(a,0)$*. This remainder term is non-negative under the following CFL condition:*

$$\forall \sigma \in \mathcal{E}_{\rm S}^{(i)}, \qquad \delta t \le \frac{|D_\sigma|\ \rho_{D_\sigma}^{n+1}}{\displaystyle\sum_{\epsilon\in\tilde{\mathcal{E}}(D_\sigma)} (F_{\sigma,\epsilon}^n)^-}. \tag{19}$$

Proof. The proof of this lemma is simply obtained by multiplying the (i^{th} component of the) momentum balance equation (4c) associated to the face σ by the unknown $u_{\sigma,i}^{n+1}$, and invoking Lemma A.2 of the appendix A. □

We now derive a balance equation (with remainder terms) for the so-called elastic potential. This quantity is the function $\mathcal{P}$, from $(0,+\infty)$ to $\mathbb{R}$, defined as a primitive of $s \mapsto \wp(s)/s^2$; as in [13], we also introduce $\mathcal{H}$, defined by $\mathcal{H}(s) = s\mathcal{P}(s)$, $\forall s \in (0,+\infty)$. For the specific equation of state $\wp$ used here, we obtain:

$$\mathcal{H}(s) = s\,\mathcal{P}(s) = \begin{cases} \dfrac{s^\gamma}{\gamma-1} \text{ if } \gamma > 1, \\ s\ln(s) \text{ if } \gamma = 1. \end{cases} \tag{20}$$

As soon as $\wp$ is an increasing function, which is true here, $\mathcal{H}$ is convex. In addition, it may easily be checked that $\rho\mathcal{H}'(\rho) - \mathcal{H}(\rho) = \wp(\rho)$. Therefore, by a formal computation detailed in [13, Appendix A], multiplying (1a) by $\mathcal{H}'(\rho)$ yields:

$$\partial_t\big(\mathcal{H}(\rho)\big) + \mathrm{div}\big(\mathcal{H}(\rho)\,\boldsymbol{u}\big) + p\,\mathrm{div}(\boldsymbol{u}) = 0. \tag{21}$$

The solution to the scheme (4) satisfies a discrete version of this relation, which we now state.

Lemma 3.3 (Discrete potential balance)**.** *Let* $\mathcal{H}$ *be defined by* (20)*. A solution to the system* (4) *satisfies the following equality, for* $K \in \mathcal{M}$ *and* $0 \le n \le N-1$*:*

$$\frac{|K|}{\delta t}\ \Big[\mathcal{H}(\rho_K^{n+1}) - \mathcal{H}(\rho_K^n)\Big] + \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\ \mathcal{H}(\rho_\sigma^n)\ u_{K,\sigma}^n + |K|\ p_K^n(\mathrm{div}\boldsymbol{u}^n)_K = -R_K^{n+1}. \tag{22}$$

In this relation, the remainder term is defined by:

$$R_K^{n+1} = \frac{1}{2}\frac{|K|}{\delta t}\ \mathcal{H}''(\overline{\rho}_{K,1}^n)\ (\rho_K^{n+1} - \rho_K^n)^2 + \frac{1}{2}\sum_{\sigma=K|L\in\mathcal{E}(K)} |\sigma|\,(u_{K,\sigma}^n)^-\ \mathcal{H}''(\overline{\rho}_\sigma^n)\ (\rho_K^n - \rho_L^n)^2 \\ + \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,u_{K,\sigma}^n\ \mathcal{H}''(\overline{\rho}_{K,2}^n)\ \rho_\sigma^n\,(\rho_K^{n+1} - \rho_K^n), \tag{23}$$

with $\overline{\rho}^n_{K,1},\ \overline{\rho}^n_{K,2} \in [\![\rho^{n+1}_K, \rho^n_K]\!]$, *and* $\overline{\rho}^n_\sigma \in [\![\rho^n_K, \rho^n_\sigma]\!]$ *for all* $\sigma \in \mathcal{E}(K)$, *where, for* $a,\ b \in \mathbb{R}$, *we denote by* $[\![a,b]\!]$ *the interval* $\{\theta a + (1-\theta)b,\ \theta \in [0,1]\}$.

Proof. The proof of this lemma is obtained by multiplying the discrete mass balance equation (4a) by $\mathcal{H}'(\rho^{n+1}_K)$ and invoking Lemma A.1 of the appendix A. □

Summing (16) and (21), we get: $\partial_t \eta + \operatorname{div}\big((\eta + p)\,\boldsymbol{u}\big) = 0$, where $\eta = \rho E_k + \mathcal{H}(\rho)$. In fact this computation can only be done for regular functions; for irregular functions, one gets the following entropy inequality (see *e.g.* [7, Introduction, Section 3.2]):

$$\partial_t \eta + \operatorname{div}\big((\eta + p)\,\boldsymbol{u}\big) \leq 0. \tag{24}$$

The quantity η is an entropy of the system, and an entropy solution to (1) is thus required to satisfy:

$$\int_0^T \int_\Omega \big[-\eta \partial_t \varphi - (\eta + p)\,\boldsymbol{u} \cdot \boldsymbol{\nabla}\varphi\big]\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t - \int_\Omega \eta_0\ \varphi(\boldsymbol{x},0)\,\mathrm{d}\boldsymbol{x} \leq 0, \quad \forall \varphi \in \mathbb{C}^\infty_c\big(\Omega \times [0,T)\big),\ \varphi \geq 0, \tag{25}$$

with $\eta_0 = \frac{1}{2}\rho_0 |\boldsymbol{u}_0|^2 + \mathcal{H}(\rho_0)$. Then, since the normal velocity is prescribed to zero at the boundary, integrating (24) over Ω yields:

$$\frac{d}{dt}\int_\Omega \big[\frac{1}{2}\,\rho\,|\boldsymbol{u}|^2 + \mathcal{H}(\rho)\big]\,\mathrm{d}\boldsymbol{x} \leq 0. \tag{26}$$

Since $\rho \geq 0$ by Lemma 3.1 and the function $s \mapsto \mathcal{H}(s)$ is bounded by below and increasing at least for s large enough, Inequality (26) provides an estimate on the solution. In [13, Proposition 3.3 and 3.13], we gave a discrete equivalent of this latter estimate for implicit and semi-implicit schemes. Unfortunately, we are not able to do so for the explicit scheme since the remainder term R^{n+1}_K defined by (23) is not always positive; therefore we are not able to prove a discrete counterpart of the total entropy estimate (26), which would yield a stability estimate for the present explicit scheme. However, under a condition for a time step which is only slightly more restrictive than a CFL-condition, and under some stability assumptions for the solutions to the scheme, we are able to show, in one space dimension, that the possible non-positive part of this remainder term tends to zero in $\mathrm{L}^1(\Omega \times (0,T))$ with the space and time steps; this allows to conclude, still in the 1D case, that a convergent sequence of solutions satisfies the entropy inequality (25): this is the result stated in Lemma 3.6 below.

3.3. **Passing to the limit in the scheme**

The objective of this section is to show, in the one dimensional case, that if a sequence of solutions is controlled in suitable norms and converges to a limit, this latter necessarily satisfies a (part of the) weak formulation of the continuous problem.

As in [13, Sections 3.1.3 and 3.3.2], the 1D version of the scheme which is studied in this section may be obtained from Scheme (4) by taking the MAC variant of the scheme, using only one horizontal stripe of grid cells, supposing that the vertical component of the velocity (the degrees of freedom of which are located on the top and bottom boundaries) vanishes, and that the measure of the vertical faces is equal to 1. For the sake of readability, however, we completely rewrite this 1D scheme, and, to this purpose, we first introduce some adaptations of the notations to the one dimensional case. For any face $\sigma \in \mathcal{E}$, let x_σ be its abscissa. For $K \in \mathcal{M}$, we denote by h_K its length (so $h_K = |K|$); when we write $K = [\sigma\sigma']$, this means that either $K = (x_\sigma, x_{\sigma'})$ or $K = (x_{\sigma'}, x_\sigma)$; if we need to specify the order, *i.e.* $K = (x_\sigma, x_{\sigma'})$ with $x_\sigma < x_{\sigma'}$, then we write $K = [\overrightarrow{\sigma\sigma'}]$. For an interface $\sigma = K|L$ between two cells K and L, we define $h_\sigma = (h_K + h_L)/2$, so, by definition of the dual mesh, $h_\sigma = |D_\sigma|$. If we need to specify the order of the cells K and L, say K is left of L, then we write $\sigma = \overrightarrow{K|L}$. With these notations, the explicit scheme (4) may be written as follows in the one dimensional setting:

$$\begin{aligned} &\forall K \in \mathcal{M}, \quad \rho^0_K = \frac{1}{|K|}\int_K \rho_0(x)\,\mathrm{d}x, \\ &\forall \sigma \in \mathcal{E}_{\mathrm{int}}, \quad u^0_\sigma = \frac{1}{|D_\sigma|}\int_{D_\sigma} u_0(x)\,\mathrm{d}x, \end{aligned} \tag{27a}$$

$$\forall K = [\overrightarrow{\sigma\sigma'}] \in \mathcal{M}, \quad \frac{|K|}{\delta t}(\rho^{n+1}_K - \rho^n_K) + F^n_{\sigma'} - F^n_\sigma = 0, \tag{27b}$$

$$\forall K \in \mathcal{M}, \quad p^{n+1}_K = \wp(\rho^{n+1}_K) = (\rho^{n+1}_K)^\gamma, \tag{27c}$$

$$\forall \sigma = \overrightarrow{K|L} \in \mathcal{E}_{\rm int},$$
$$\frac{|D_\sigma|}{\delta t}(\rho_{D_\sigma}^{n+1}u_\sigma^{n+1} - \rho_{D_\sigma}^n u_\sigma^n) + F_L^n u_L^n - F_K^n u_K^n + p_L^{n+1} - p_K^{n+1} = 0. \tag{27d}$$

The mass flux in the discrete mass balance equation is given, for $\sigma \in \mathcal{E}_{\rm int}$, by $F_\sigma^n = \rho_\sigma^n u_\sigma^n$, where the upwind approximation for the density at the face, ρ_σ^n, is defined by (7). In the momentum balance equation, the density associated to the dual cell D_σ, with $\sigma = K|L$, reads

$$\text{for } k = n \text{ and } k = n+1, \quad \rho_{D_\sigma}^k = \frac{1}{2\,|D_\sigma|}\,(|K|\,\rho_K^k + |L|\,\rho_L^k), \tag{28}$$

and the application of the procedure described in Section 3.1 yields, for the mass fluxes at the dual face located at the center of the mesh $K = [\overrightarrow{\sigma\sigma'}]$:

$$F_K^n = \frac{1}{2}\,(F_\sigma^n + F_{\sigma'}^n). \tag{29}$$

The approximation of the velocity at this face is upwind: $u_K^n = u_\sigma^n$ if $F_K^n \geq 0$ and $u_K^n = u_{\sigma'}^n$ otherwise.

Let a sequence of discretizations $(\mathcal{M}^{(m)},\ \delta t^{(m)})_{m\in\mathbb{N}}$ be given. We define the size $h^{(m)}$ of the mesh $\mathcal{M}^{(m)}$ by $h^{(m)} = \sup_{K\in\mathcal{M}^{(m)}} h_K$. Let $\rho^{(m)}$, $p^{(m)}$ and $u^{(m)}$ be the solution given by the scheme (27) with the mesh $\mathcal{M}^{(m)}$ and the time step $\delta t^{(m)}$. To the discrete unknowns, we associate piecewise constant functions on time intervals and on primal or dual meshes, so the density $\rho^{(m)}$, the pressure $p^{(m)}$ and the velocity $u^{(m)}$ are defined almost everywhere on $\Omega \times (0,T)$ by:

$$\begin{aligned}
\rho^{(m)}(x,t) &= \sum_{n=0}^{N-1}\sum_{K\in\mathcal{M}} (\rho^{(m)})_K^n\,\mathfrak{X}_K(x)\,\mathfrak{X}_{[n,n+1)}(t),\\
p^{(m)}(x,t) &= \sum_{n=0}^{N-1}\sum_{K\in\mathcal{M}} (p^{(m)})_K^n\,\mathfrak{X}_K(x)\,\mathfrak{X}_{[n,n+1)}(t),\\
u^{(m)}(x,t) &= \sum_{n=0}^{N-1}\sum_{\sigma\in\mathcal{E}} (u^{(m)})_\sigma^n\,\mathfrak{X}_{D_\sigma}(x)\,\mathfrak{X}_{[n,n+1)}(t),
\end{aligned} \tag{30}$$

where $\mathfrak{X}_K$, $\mathfrak{X}_{D_\sigma}$ and $\mathfrak{X}_{[n,n+1)}$ stand for the characteristic function of the intervals K, D_σ and $[t^n, t^{n+1})$ respectively.

For discrete functions q and v defined on the primal and dual mesh, respectively, we define a discrete $\mathrm{L}^1((0,T);\mathrm{BV}(\Omega))$ norm by:

$$\|q\|_{\mathfrak{T},x,\mathrm{BV}} = \sum_{n=0}^{N}\delta t \sum_{\sigma=K|L\in\mathcal{E}_{\rm int}} |q_L^n - q_K^n|, \quad \|v\|_{\mathfrak{T},x,\mathrm{BV}} = \sum_{n=0}^{N}\delta t \sum_{\epsilon=D_\sigma|D_{\sigma'}\in\tilde{\mathcal{E}}_{\rm int}} |v_{\sigma'}^n - v_\sigma^n|,$$

and a discrete $\mathrm{L}^1(\Omega;\mathrm{BV}((0,T)))$ norm by:

$$\|q\|_{\mathfrak{T},t,\mathrm{BV}} = \sum_{K\in\mathcal{M}} |K| \sum_{n=0}^{N-1} |q_K^{n+1} - q_K^n|, \qquad \|v\|_{\mathfrak{T},t,\mathrm{BV}} = \sum_{\sigma\in\mathcal{E}} |D_\sigma| \sum_{n=0}^{N-1} |v_\sigma^{n+1} - v_\sigma^n|.$$

For the consistency result that we are seeking (Theorem 3.5 below), it is assumed that a sequence of discrete solutions $\left(\rho^{(m)}, p^{(m)}, u^{(m)}\right)_{m\in\mathbb{N}}$ satisfies $\rho^{(m)} > 0$ and $p^{(m)} > 0$, $\forall m \in \mathbb{N}$ (which may be a consequence of the fact that the CFL stability condition (15) is satisfied), and is uniformly bounded in $\mathrm{L}^\infty((0,T)\times\Omega)^3$, *i.e.*:

$$0 < (\rho^{(m)})_K^n \leq C,\ 0 < (p^{(m)})_K^n \leq C,\ \text{for } K \in \mathcal{M}^{(m)}, 0 \leq n \leq N^{(m)},\ m \in \mathbb{N}, \tag{31}$$

and

$$|(u^{(m)})_\sigma^n| \leq C, \quad \forall \sigma \in \mathcal{E}^{(m)},\ \text{for } 0 \leq n \leq N^{(m)},\ \forall m \in \mathbb{N}, \tag{32}$$

where C is a positive real number. By definition of the initial conditions of the scheme, these inequalities imply that the functions ρ_0 and u_0 belong to $\mathrm{L}^\infty(\Omega)$. We also assume that a sequence of discrete solutions satisfies the following uniform bounds in the discrete BV-norms:

$$\|\rho^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} + \|u^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} \leq C, \quad \forall m \in \mathbb{N}. \tag{33}$$

We are not able to prove the estimates (31)–(33) for the solutions of the scheme; however, such inequalities are satisfied by the "interpolates" (for instance, by taking the cell average) of the solution to a Riemann problem, and are observed in computations (of course, as far as possible, *i.e.* in a limited number of cases and with a limited sequence of meshes and time steps).

A weak solution to the continuous problem satisfies, for any $\varphi \in \mathbb{C}_c^\infty\big(\Omega \times [0,T)\big)$:

$$-\int_0^T \int_\Omega \Big[\rho\, \partial_t \varphi + \rho\, u\, \partial_x \varphi\Big]\, \mathrm{d}x\, \mathrm{d}t - \int_\Omega \rho_0(x)\, \varphi(x,0)\, \mathrm{d}x = 0, \tag{34a}$$

$$-\int_0^T \int_\Omega \Big[\rho\, u\, \partial_t \varphi + (\rho\, u^2 + p)\, \partial_x \varphi\Big]\, \mathrm{d}x\, \mathrm{d}t - \int_\Omega \rho_0(x)\, u_0(x)\, \varphi(x,0)\, \mathrm{d}x = 0, \tag{34b}$$

$$p = \rho^\gamma. \tag{34c}$$

Even though these relations do not take into account the boundary conditions, they allow to derive the Rankine-Hugoniot conditions; hence, if they are shown to be satisfied by the limit of a sequence of solutions to the scheme, this implies, loosely speaking, that the scheme computes correct shocks. This is the result stated in Theorem 3.5. In order to prove this theorem, the following definitions of interpolates of regular test functions on the primal and dual mesh are useful.

Definition 3.4 (Interpolates on one-dimensional meshes)**.** Let Ω be an open bounded interval of $\mathbb{R}$, let $\varphi \in C_c^\infty(\Omega \times [0,T))$, and let $\mathcal{M}$ be a mesh over Ω. The interpolate $\varphi_\mathcal{M}$ of φ on the primal mesh $\mathcal{M}$ is defined by:

$$\varphi_\mathcal{M} = \sum_{n=0}^{N-1} \sum_{K\in\mathcal{M}} \varphi_K^n\ \mathfrak{X}_K\ \mathfrak{X}_{[t^n,t^{n+1})},$$

where, for $0 \le n \le N$ and $K \in \mathcal{M}$, $\varphi_K^n = \varphi(x_K, t^n)$, with x_K the mass center of K. The time and space discrete derivatives of the discrete function $\varphi_\mathcal{M}$ are defined by:

$$\eth_t \varphi_\mathcal{M} = \sum_{n=0}^{N-1} \sum_{K\in\mathcal{M}} \frac{\varphi_K^{n+1} - \varphi_K^n}{\delta t}\ \mathfrak{X}_K\ \mathfrak{X}_{[t^n,t^{n+1})}, \text{ and } \eth_x \varphi_\mathcal{M} = \sum_{n=0}^{N-1} \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \frac{\varphi_L^n - \varphi_K^n}{h_\sigma}\ \mathfrak{X}_{D_\sigma}\ \mathfrak{X}_{[t^n,t^{n+1})}.$$

Let $\varphi_\mathcal{E}$ be an interpolate of φ on the dual mesh, defined by:

$$\varphi_\mathcal{E} = \sum_{n=0}^{N-1} \sum_{\sigma\in\mathcal{E}} \varphi_\sigma^n\ \mathfrak{X}_{D_\sigma}\ \mathfrak{X}_{[t^n,t^{n+1})},$$

where, for $0 \le n \le N$ and $\sigma \in \mathcal{E}$, $\varphi_\sigma^n = \varphi(x_\sigma, t^n)$, with x_σ the abscissa of the interface σ. We also define the time and space discrete derivatives of this discrete function by:

$$\eth_t \varphi_\mathcal{E} = \sum_{n=0}^{N-1} \sum_{\sigma\in\mathcal{E}} \frac{\varphi_\sigma^{n+1} - \varphi_\sigma^n}{\delta t}\ \mathfrak{X}_{D_\sigma}\ \mathfrak{X}_{[t^n,t^{n+1})}, \text{ and } \eth_x \varphi_\mathcal{E} = \sum_{n=0}^{N-1} \sum_{K=\overrightarrow{[\sigma\sigma']}\in\mathcal{M}} \frac{\varphi_{\sigma'}^n - \varphi_\sigma^n}{h_K}\ \mathfrak{X}_K\ \mathfrak{X}_{[t^n,t^{n+1})}.$$

Finally, let $\eth_x \varphi_{\mathcal{M},\mathcal{E}}$ be defined by:

$$\eth_x \varphi_{\mathcal{M},\mathcal{E}} = \sum_{n=0}^{N-1} \sum_{K=\overrightarrow{[\sigma\sigma']}\in\mathcal{M}} \frac{\varphi_K^{n+1} - \varphi_\sigma^{n+1}}{h_K/2}\ \mathfrak{X}_{D_{K,\sigma}}\ \mathfrak{X}_{[t^n,t^{n+1})} + \frac{\varphi_{\sigma'}^{n+1} - \varphi_K^{n+1}}{h_K/2}\ \mathfrak{X}_{D_{K,\sigma'}}\ \mathfrak{X}_{[t^n,t^{n+1})}.$$

Theorem 3.5 (Consistency of the one-dimensional scheme)**.**
Let Ω be an open bounded interval of $\mathbb{R}$. We suppose that the initial data satisfies $\rho_0 \in \mathrm{L}^\infty(\Omega)$ and $u_0 \in \mathrm{L}^\infty(\Omega)$. Let $(\mathcal{M}^{(m)},\, \delta t^{(m)})_{m\in\mathbb{N}}$ be a sequence of discretizations such that both the time step $\delta t^{(m)}$ and the size $h^{(m)}$ of the mesh $\mathcal{M}^{(m)}$ tend to zero as $m \to +\infty$, and let $(\rho^{(m)}, p^{(m)}, u^{(m)})_{m\in\mathbb{N}}$ be the corresponding sequence of solutions. We suppose that this sequence satisfies the estimates (31)–(33) and converges in $\mathrm{L}^r(\Omega\times(0,T))^3$, for $1 \le r < \infty$, to $(\bar\rho,\, \bar p,\, \bar u) \in \mathrm{L}^\infty(\Omega \times (0,T))^3$.

Then the limit $(\bar\rho,\, \bar p,\, \bar u)$ satisfies the system (34).

Proof. It is clear that, with the assumed convergence for the sequence of solutions, the limit satisfies the equation of state. The proof of this theorem is thus obtained by passing to the limit in the scheme for the mass balance equation first, and then for the momentum balance equation.

Mass balance equation – Let $\varphi \in \mathbb{C}_c^\infty(\Omega \times [0,T))$. Let $m \in \mathbb{N}$, $\mathcal{M}^{(m)}$ and $\delta t^{(m)}$ be given. Dropping for short the superscript $^{(m)}$, let $\varphi_\mathcal{M}$ be the interpolate of φ on the primal mesh and let $\eth_t\varphi_\mathcal{M}$ and $\eth_x\varphi_\mathcal{M}$ be its time and space discrete derivatives in the sense of Definition 3.4. Thanks to the regularity of φ, these functions respectively converge in $\mathrm{L}^r(\Omega \times (0,T))$, for $r \geq 1$ (including $r = +\infty$), to φ, $\partial_t\varphi$ and $\partial_x\varphi$ respectively. In addition, $\varphi_\mathcal{M}(\cdot,0)$ (which, for $K \in \mathcal{M}$ and $x \in K$, is equal to $\varphi_K^0 = \varphi(x,0)$) converges to $\varphi(\cdot,0)$ in $\mathrm{L}^r(\Omega)$ for $r \geq 1$. Since the support of φ is compact in $\Omega \times [0,T)$, for m large enough, the interpolate of φ vanishes at the boundary cells and at the last time step(s); this is always assumed in the sequel.

Let us multiply the first equation (27b) of the scheme by $\delta t\, \varphi_K^{n+1}$, and sum the result for $0 \leq n \leq N-1$ and $K \in \mathcal{M}$, to obtain $T_1^{(m)} + T_2^{(m)} = 0$ with

$$T_1^{(m)} = \sum_{n=0}^{N-1} \sum_{K\in\mathcal{M}} |K|(\rho_K^{n+1} - \rho_K^n)\, \varphi_K^{n+1}, \qquad T_2^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} (F_{\sigma'}^n - F_\sigma^n)\, \varphi_K^{n+1}.$$

Reordering the sums in $T_1^{(m)}$ yields:

$$T_1^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} |K|\, \rho_K^n\, \frac{\varphi_K^{n+1} - \varphi_K^n}{\delta t} - \sum_{K\in\mathcal{M}} |K|\, \rho_K^0\, \varphi_K^0,$$

so:

$$T_1^{(m)} = -\int_0^T \int_\Omega \rho^{(m)} \eth_t\varphi_\mathcal{M}\, \mathrm{d}x\, \mathrm{d}t - \int_\Omega (\rho^{(m)})^0(x)\ \varphi_\mathcal{M}(x,0)\, \mathrm{d}x.$$

The boundedness of ρ_0 and the definition (27a) of the initial conditions for the scheme ensures that the sequence $((\rho^{(m)})^0)_{m\in\mathbb{N}}$ converges to ρ_0 in $\mathrm{L}^r(\Omega)$ for $r \geq 1$. Since, by assumption, the sequence of discrete solutions and of the interpolate time derivatives converge in $\mathrm{L}^r\big(\Omega \times (0,T)\big)$ for $r \geq 1$, we obtain:

$$\lim_{m\to+\infty} T_1^{(m)} = -\int_0^T \int_\Omega \bar\rho\, \partial_t\varphi\, \mathrm{d}x\, \mathrm{d}t - \int_\Omega \rho_0(x)\, \varphi(x,0)\, \mathrm{d}x.$$

Using the expression of the mass flux F_σ^n and reordering the sums in $T_2^{(m)}$, we get, remarking that $|D_\sigma| = h_\sigma$:

$$T_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} |D_\sigma|\ \rho_\sigma^n u_\sigma^n\ \frac{\varphi_L^{n+1} - \varphi_K^{n+1}}{h_\sigma}.$$

Since $|D_\sigma| = (|K| + |L|)/2$ and ρ_σ^n is the upwind approximation of ρ^n at the face σ, we can rewrite $T_2^{(m)} = \mathfrak{T}_2^{(m)} + \mathfrak{R}_2^{(m)}$ with

$$\mathfrak{T}_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} \Big(\frac{|K|}{2}\rho_K^n + \frac{|L|}{2}\rho_L^n\Big) u_\sigma^n\ \frac{\varphi_L^{n+1} - \varphi_K^{n+1}}{h_\sigma},$$

$$\mathfrak{R}_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} (\rho_K^n - \rho_L^n)\ \Big[\frac{|K|}{2}(u_\sigma^n)^- + \frac{|L|}{2}(u_\sigma^n)^+\Big]\ \frac{\varphi_L^{n+1} - \varphi_K^{n+1}}{h_\sigma},$$

where, for $a \in \mathbb{R}$, $a^+ = \max(a,0)$ and $a^- = -\min(a,0)$ (so $a = a^+ - a^-$). We have, for the term $\mathfrak{T}_2^{(m)}$:

$$\mathfrak{T}_2^{(m)} = -\int_0^T \int_\Omega \rho^{(m)} u^{(m)} \eth_x\varphi_\mathcal{M}\, \mathrm{d}x\, \mathrm{d}t$$

and therefore

$$\lim_{m\to+\infty} \mathfrak{T}_2^{(m)} = -\int_0^T \int_\Omega \bar\rho\, \bar u\, \partial_x\varphi\, \mathrm{d}x\, \mathrm{d}t.$$

The remainder term $\mathfrak{R}_2^{(m)}$ is bounded as follows, with $C_\varphi = \|\partial_x \varphi\|_{\mathrm{L}^\infty(\Omega\times(0,T))}$:

$$|\mathfrak{R}_2^{(m)}| \le C_\varphi \sum_{n=0}^{N-1} \delta t \sum_{\sigma=K|L\in\mathcal{E}} |\rho_K^n - \rho_L^n|\ |D_\sigma|\ |u_\sigma^n| \le C_\varphi\ \|u^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|\rho^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}}\ h^{(m)},$$

and therefore tends to zero when m tends to $+\infty$, by the assumed boundedness of the sequence of solutions.

Momentum balance equation – Let $\varphi_\mathcal{E}$, $\eth_t\varphi_\mathcal{E}$ and $\eth_x\varphi_\mathcal{E}$ be the interpolate of φ on the dual mesh and its discrete time and space derivatives, in the sense of Definition 3.4, which converge in $\mathrm{L}^r(\Omega\times(0,T))$, for $r\ge 1$, to φ, $\partial_t\varphi$ and $\partial_x\varphi$ respectively. Let us multiply Equation (27d) by $\delta t\,\varphi_\sigma^{n+1}$, and sum the result for $0\le n\le N-1$ and $\sigma\in\mathcal{E}_{\rm int}$. We obtain $T_1^{(m)}+T_2^{(m)}+T_3^{(m)}=0$ with

$$\begin{aligned}
T_1^{(m)} &= \sum_{n=0}^{N-1}\sum_{\sigma\in\mathcal{E}_{\rm int}} |D_\sigma|\ (\rho_{D_\sigma}^{n+1}u_\sigma^{n+1} - \rho_{D_\sigma}^n u_\sigma^n)\,\varphi_\sigma^{n+1},\\
T_2^{(m)} &= \sum_{n=0}^{N-1}\delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \Big[F_L^n\,u_L^n - F_K^n\,u_K^n\Big]\ \varphi_\sigma^{n+1},\\
T_3^{(m)} &= \sum_{n=0}^{N-1}\delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} (p_L^{n+1}-p_K^{n+1})\,\varphi_\sigma^{n+1}.
\end{aligned}$$

Reordering the sums, we get for $T_1^{(m)}$:

$$T_1^{(m)} = -\sum_{n=0}^{N-1}\delta t \sum_{\sigma\in\mathcal{E}_{\rm int}} |D_\sigma|\,\rho_{D_\sigma}^n u_\sigma^n\ \frac{\varphi_\sigma^{n+1}-\varphi_\sigma^n}{\delta t} - \sum_{\sigma\in\mathcal{E}_{\rm int}} |D_\sigma|\,\rho_{D_\sigma}^0\,u_\sigma^0\,\varphi_\sigma^0.$$

Thanks to the definition of the quantity ρ_{D_σ} (namely the fact that $|D_\sigma|\,\rho_{D_\sigma}^n = (|K|\,\rho_K^n + |L|\,\rho_L^n)/2$), we have:

$$T_1^{(m)} = -\int_0^T\int_\Omega \rho^{(m)}\,u^{(m)}\,\eth_t\varphi_\mathcal{E}\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega (\rho^{(m)})^0(x)\ (u^{(m)})^0(x)\ \varphi_\mathcal{E}(x,0)\,\mathrm{d}x.$$

By the same arguments as for the mass balance equation, we therefore obtain:

$$\lim_{m\to+\infty} T_1^{(m)} = -\int_0^T\int_\Omega \bar\rho\,\bar u\,\partial_t\varphi\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \rho_0(x)\,u_0(x)\,\varphi(x,0)\,\mathrm{d}x.$$

Let us now turn to $T_2^{(m)}$. Reordering the sums and using the definition of the mass fluxes at the dual faces, we get:

$$T_2^{(m)} = -\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} F_K^n\,u_K^n\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}) = -\frac12\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} (\rho_\sigma^n u_\sigma^n + \rho_{\sigma'}^n u_{\sigma'}^n)\,u_K^n\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}).$$

Using the relation

$$\int_0^T\int_\Omega \rho^{(m)}\,(u^{(m)})^2\,\eth_x\varphi_\mathcal{E}\,\mathrm{d}x\,\mathrm{d}t = \frac12\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} \rho_K^n\left[(u_\sigma^n)^2 + (u_{\sigma'}^n)^2\right]\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}),$$

we can rewrite the term $T_2^{(m)}$ as

$$T_2^{(m)} = -\int_0^T\int_\Omega \rho^{(m)}\,{u^{(m)}}^2\,\eth_x\varphi_\mathcal{E}\,\mathrm{d}x\,\mathrm{d}t + \mathfrak{R}_2^{(m)},$$

where:

$$\mathfrak{R}_2^{(m)} = -\frac12\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} \Big[(\rho_\sigma^n u_\sigma^n + \rho_{\sigma'}^n u_{\sigma'}^n)\,u_K^n - \rho_K^n\left((u_\sigma^n)^2 + (u_{\sigma'}^n)^2\right)\Big]\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}).$$

Let us split this latter expression as $\mathcal{R}_2^{(m)} = \mathcal{R}_{21}^{(m)} + \mathcal{R}_{22}^{(m)}$, with:

$$\mathcal{R}_{21}^{(m)} = -\frac{1}{2}\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} u_\sigma^n\ (\rho_\sigma^n u_K^n - \rho_K^n u_\sigma^n)\ (\varphi_{\sigma'}^{n+1} - \varphi_\sigma^{n+1}),$$

$$\mathcal{R}_{22}^{(m)} = -\frac{1}{2}\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} u_{\sigma'}^n\ (\rho_{\sigma'}^n u_K^n - \rho_K^n u_{\sigma'}^n)\ (\varphi_{\sigma'}^{n+1} - \varphi_\sigma^{n+1}).$$

Applying the identity $2(ab-cd) = (a-c)(b+d) + (a+c)(b-d),\ \forall (a,b,c,d)\in\mathbb{R}^4$, to the term $\rho_\sigma^n u_K^n - \rho_K^n u_\sigma^n$ and using the fact that the quantities $\rho_\sigma^n - \rho_K^n$ and $u_\sigma^n - u_K^n$ are either zero or differences of the density at two neighbouring cells and of the velocity at two neighbouring faces respectively, we obtain for $\mathcal{R}_{21}^{(m)}$:

$$|\mathcal{R}_{21}^{(m)}| \le C_\varphi\ \left[\|u^{(m)}\|^2_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|\rho^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} + \|u^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|u^{(m)}\|_{\mathcal{T},x,\mathrm{BV}}\ \|\rho^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\right]\ h^{(m)},$$

where the real number C_φ only depends on φ. Since the same estimate holds for $\mathcal{R}_{22}^{(m)}$, the remainder term $\mathcal{R}_2^{(m)}$ tends to zero when m tends to $+\infty$ and:

$$\lim_{m\to+\infty} T_2^{(m)} = -\int_0^T\int_\Omega \bar\rho\,\bar u^2\,\partial_x\varphi\,\mathrm{d}x\,\mathrm{d}t.$$

Let us finally study $T_3^{(m)}$. Reordering the sums, we obtain $T_3^{(m)} = \mathcal{T}_3^{(m)} + \mathcal{R}_3^{(m)}$ with:

$$\mathcal{T}_3^{(m)} = -\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} p_K^n\ (\varphi_{\sigma'}^{n+1} - \varphi_\sigma^{n+1}),$$

$$\mathcal{R}_3^{(m)} = -\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} (p_K^{n+1} - p_K^n)\ (\varphi_{\sigma'}^{n+1} - \varphi_\sigma^{n+1}).$$

The remainder term reads:

$$\mathcal{R}_3^{(m)} = \sum_{n=1}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} p_K^n\ \left[(\varphi_{\sigma'}^{n+1} - \varphi_\sigma^{n+1}) - (\varphi_{\sigma'}^{n} - \varphi_\sigma^{n})\right] + \delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} p_K^0\ (\varphi_{\sigma'}^1 - \varphi_\sigma^1),$$

and thus:

$$|\mathcal{R}_3^{(m)}| \le C_\varphi\ (\delta t^{(m)} + h^{(m)})\ \|p\|_{\mathrm{L}^\infty(\Omega\times(0,T))},$$

where the real number C_φ only depends on (the first and second derivatives of) φ. Thus $\mathcal{R}_3^{(m)}$ tends to zero when m tends to $+\infty$ and, since

$$\mathcal{T}_3^{(m)} = -\int_0^T\int_\Omega p^{(m)}\,\eth_x\varphi_\mathcal{M}\,\mathrm{d}x\,\mathrm{d}t,$$

we obtain that:

$$\lim_{m\to+\infty} T_3^{(m)} = \int_0^T\int_\Omega \bar p\,\partial_x\varphi\,\mathrm{d}x\,\mathrm{d}t.$$

Conclusion – Gathering the limits of all the terms of the mass and momentum balance equations concludes the proof. □

We now turn to the entropy balance (25). To this purpose, we need to introduce the following additional condition for a sequence of discretizations:

$$\lim_{m\to+\infty} \frac{\delta t^{(m)}}{\min_{K\in\mathcal{M}^{(m)}} h_K} = 0. \tag{35}$$

Note that this condition is more restrictive than a standard CFL condition. It allows to bound the remainder term in the discrete elastic potential balance as stated in the following lemma.

Lemma 3.6. *Let Ω be an open bounded interval of $\mathbb{R}$. Let $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m\in\mathbb{N}}$ be a sequence of discretizations such that the time step $\delta t^{(m)}$ tends to zero as $m \to +\infty$, and let $(\rho^{(m)}, p^{(m)}, u^{(m)})_{m\in\mathbb{N}}$ be the corresponding sequence of solutions. We suppose that this sequence satisfies the estimates* (31) *and* (32)*. In addition, we assume that $(\rho^{(m)})_{m\in\mathbb{N}}$ satisfies the following uniform BV estimate:*

$$\|\rho^{(m)}\|_{\mathfrak{T},t,\mathrm{BV}} \leq C, \quad \forall m \in \mathbb{N}, \tag{36}$$

and, for $\gamma < 2$ only, is uniformly bounded by below, i.e. *that there exists $c > 0$ such that:*

$$c \leq (\rho^{(m)})_K^n, \quad \forall K \in \mathcal{M}^{(m)}, \ \textit{for } 0 \leq n \leq N^{(m)}, \ \forall m \in \mathbb{N}. \tag{37}$$

Let us suppose that the condition (35) *holds. Let $\mathcal{R}^{(m)}$ be defined by:*

$$\mathcal{R}^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} (R_K^{n+1})^-,$$

with R_K^{n+1} given by (23)*. Then:*

$$\lim_{m\to+\infty} \mathcal{R}^{(m)} = 0.$$

Proof. For $K = [\overrightarrow{\sigma\sigma'}] \in \mathcal{M}$, with $\sigma = \overrightarrow{M|K}$ and $\sigma' = \overrightarrow{K|L}$, we write $R_K^{n+1} = (T_1)_K^{n+1} + (T_2)_K^{n+1} + (T_3)_K^{n+1}$, with:

$$\begin{aligned}
(T_1)_K^{n+1} &= \frac{1}{2}\frac{|K|}{\delta t}\ \mathcal{H}''(\overline{\rho}_{K,1}^n)\ (\rho_K^{n+1} - \rho_K^n)^2,\\
(T_2)_K^{n+1} &= \frac{1}{2}\Big[(u_{\sigma'}^n)^-\ \mathcal{H}''(\overline{\rho}_{\sigma'}^n)\ (\rho_K^n - \rho_L^n)^2 + (-u_\sigma^n)^-\ \mathcal{H}''(\overline{\rho}_\sigma^n)\ (\rho_K^n - \rho_M^n)^2\Big],\\
(T_3)_K^{n+1} &= \Big[\rho_{\sigma'}^n\, u_{\sigma'}^n - \rho_\sigma^n\, u_\sigma^n\Big]\ \mathcal{H}''(\overline{\rho}_{K,2}^n)\ (\rho_K^{n+1} - \rho_K^n),
\end{aligned}$$

where $\overline{\rho}_{K,1}^n,\ \overline{\rho}_{K,2}^n \in [\![\rho_K^{n+1}, \rho_K^n]\!]$, $\overline{\rho}_{\sigma'}^n \in [\![\rho_K^n, \rho_L^n]\!]$ and $\overline{\rho}_\sigma^n \in [\![\rho_K^n, \rho_M^n]\!]$. The first two terms are non-negative, and thus $(R_K^{n+1})^- \leq |(T_3)_K^{n+1}|$. Since both ρ, u and, for $\gamma < 2$, $1/\rho$ are supposed to be bounded, we have:

$$\sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} |(T_3)_K^{n+1}| \leq C \frac{\delta t^{(m)}}{\min_{K\in\mathcal{M}} h_K}\ \|\rho^{(m)}\|_{\mathfrak{T},t,\mathrm{BV}},$$

which yields the conclusion by the assumption (35). □

Theorem 3.7 (Entropy consistency of the one dimensional scheme)**.**
Let the assumptions of Theorem 3.5 hold. Let us suppose in addition that the considered sequence of discretizations satisfies (35)*, and that $(\rho^{(m)})_{m\in\mathbb{N}}$ satisfies the BV estimate* (36) *and, for $\gamma < 2$, the uniform control* (37) *of $1/\rho^{(m)}$. Then the limit $(\bar\rho, \bar p, \bar u)$ satisfies the entropy condition* (25)*.*

Proof. Let $\varphi \in \mathbb{C}_c^\infty\big(\Omega \times [0,T)\big)$, $\varphi \geq 0$. As in the previous proofs, we suppose that the space and times steps are small enough for φ to vanish at the boundary cells and at the last time step. With the notations for the interpolate of φ given in Definition 3.4, we multiply the kinetic balance equation (17)-(18) by φ_σ^{n+1}, and the elastic potential balance (22)-(23) by φ_K^{n+1}, sum over the edges and cells respectively and over the time steps, to obtain the discrete version of (25):

$$T_1^{(m)} + T_2^{(m)} + T_3^{(m)} + \tilde T_1^{(m)} + \tilde T_2^{(m)} + \tilde T_3^{(m)} = -R^{(m)} - \tilde R^{(m)} \tag{38}$$

where:

$$\begin{aligned}
T_1^{(m)} &= \sum_{n=0}^{N-1} \sum_{K\in\mathcal{M}} |K|\ \big[\mathcal{H}(\rho_K^{n+1}) - \mathcal{H}(\rho_K^n)\big]\ \varphi_K^{n+1},\\
T_2^{(m)} &= \sum_{n=0}^{N-1} \delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} \big[\mathcal{H}(\rho_{\sigma'}^n)\, u_{\sigma'}^n - \mathcal{H}(\rho_\sigma^n)\, u_\sigma^n\big]\varphi_K^{n+1},
\end{aligned}$$

$$T_3^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} \left[p_K^n(u_{\sigma'}^n - u_\sigma^n)\right]\ \varphi_K^{n+1},$$

$$\tilde{T}_1^{(m)} = \frac{1}{2}\sum_{n=0}^{N-1} \sum_{\sigma\in\mathcal{E}_{\rm int}} |D_\sigma|\ \left[\rho_{D_\sigma}^{n+1}(u_\sigma^{n+1})^2 - \rho_{D_\sigma}^n(u_\sigma^n)^2\right]\ \varphi_\sigma^{n+1},$$

$$\tilde{T}_2^{(m)} = \frac{1}{2}\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \left[F_L^n\,(u_L^n)^2 - F_K^n\,(u_K^n)^2\right]\ \varphi_\sigma^{n+1},$$

$$\tilde{T}_3^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} (p_L^{n+1} - p_K^{n+1})\,u_\sigma^{n+1}\,\varphi_\sigma^{n+1},$$

$$R^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} R_K^{n+1}\ \varphi_K^{n+1}, \qquad \tilde{R}^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{\sigma\in\mathcal{E}_{\rm int}} R_\sigma^{n+1}\ \varphi_\sigma^{n+1},$$

and the quantities R_K^{n+1} and R_σ^{n+1} are given by (the one-dimensional version of) Equation (23) and (18) respectively.

The fact that

$$\lim_{m\to+\infty} T_1^{(m)} = -\int_0^T \int_\Omega \mathcal{H}(\bar\rho)\ \partial_t\varphi\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \mathcal{H}(\rho_0)(x)\ \varphi(x,0)\,\mathrm{d}x,$$

is proven by the same technique as in the passage to the limit in the term $T_1^{(m)}$ of the discrete mass balance equation in the proof Theorem 3.5, thanks to the fact that, with the assumed convergence of the sequence $(\rho^{(m)})_{m\in\mathbb{N}}$, the sequence $(\mathcal{H}(\rho^{(m)}))_{m\in\mathbb{N}}$ converge to $\mathcal{H}(\bar\rho)$ in $\mathrm{L}^r(\Omega\times(0,T))$, for $r\geq 1$. For $T_2^{(m)}$, we have, reordering the sums:

$$T_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \mathcal{H}(\rho_\sigma^n)\,u_\sigma^n\ (\varphi_L^{n+1} - \varphi_K^{n+1}).$$

Let us write $T_2^{(m)} = \mathfrak{T}_2^{(m)} + \mathfrak{R}_2^{(m)}$, with

$$\mathfrak{T}_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \left(|D_{K,\sigma}|\ \mathcal{H}(\rho_K^n) + |D_{L,\sigma}|\ \mathcal{H}(\rho_L^n)\right) u_\sigma^n\ \frac{\varphi_L^{n+1} - \varphi_K^{n+1}}{h_\sigma}.$$

We have:

$$\mathfrak{T}_2^{(m)} = -\int_0^T \int_\Omega \mathcal{H}(\rho^{(m)})\ u^{(m)}\ \eth_x\varphi_\mathcal{M}\,\mathrm{d}x\,\mathrm{d}t,$$

so

$$\lim_{m\to+\infty} T_2^{(m)} = -\int_0^T \int_\Omega \mathcal{H}(\bar\rho)\ \bar u\ \partial_x\varphi\,\mathrm{d}x\,\mathrm{d}t.$$

The remainder term $\mathfrak{R}_2^{(m)}$ reads:

$$\mathfrak{R}_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \left[|D_\sigma|\ \mathcal{H}(\rho_\sigma^n) - |D_{K,\sigma}|\ \mathcal{H}(\rho_K^n) - |D_{L,\sigma}|\ \mathcal{H}(\rho_L^n)\right] u_\sigma^n\ \frac{\varphi_L^{n+1} - \varphi_K^{n+1}}{h_\sigma}.$$

This term satisfies:

$$|\mathfrak{R}_2^{(m)}| \leq \sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} |\mathcal{H}(\rho_K^n) - \mathcal{H}(\rho_L^n)|\ u_\sigma^n\ |\varphi_L^{n+1} - \varphi_K^{n+1}|,$$

and so

$$|\mathfrak{R}_2^{(m)}| \leq C_\varphi\ h^{(m)}\ \|u^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|\rho^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}},$$

provided that a uniform (with respect to the faces, the time steps and the meshes) Lipschitz condition holds for $|\mathcal{H}(\rho_K^n) - \mathcal{H}(\rho_L^n)|$ which, in view of the expression of $\mathcal{H}$, requires that the sequence $(\rho^{(m)})_{m\in\mathbb{N}}$ be bounded by below away from zero when $\gamma = 1$.

The other terms at the left-hand side of (38) are similar to the samely-named terms in the proof of consistency of the scheme for the full Euler equations, and their treatment is detailed in the proof of Theorem 4.2 below.

Finally, the remainder term $R^{(m)}$ is non-negative under the CFL condition (19), while the positive part of $\tilde{R}^{(m)}$ tends to zero in $\mathrm{L}^1(\Omega \times (0,T))$ under the assumption (35) by Lemma 3.6. The proof is thus complete. □

Remark 3.8 (On BV-stability assumptions)**.**
The proof of Theorem 3.5 shows that the scheme is consistent under a BV-stability assumption much weaker than (33), namely:

$$\lim_{m\to+\infty} h^{(m)} \left[\|\rho^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} + \|u^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} \right] = 0.$$

The situation is completely different when proving that the limit of convergent sequences is an entropy solution (*i.e.* when proving Theorem 3.7); indeed, in the preliminary lemma 3.6, we need:

$$\lim_{m\to+\infty} \frac{\delta t^{(m)}}{\min_{K\in\mathcal{M}^{(m)}} h_K} \|\rho^{(m)}\|_{\mathfrak{T},t,\mathrm{BV}} = 0.$$

4. The full Euler equations

We build in this section a scheme for the solution of the full Euler equations (2). Let us recall that the (conservative) energy equation in this system is the total energy balance, which reads:

$$\partial_t(\rho\, E) + \mathrm{div}(\rho\, E\, \boldsymbol{u}) + \mathrm{div}(p\, \boldsymbol{u}) = 0.$$

If we subtract to this relation the kinetic energy balance (see Section 3.2)

$$\partial_t(\rho E_k) + \mathrm{div}\big(\rho\, E_k\, \boldsymbol{u}\big) + \boldsymbol{\nabla} p \cdot \boldsymbol{u} = 0,$$

we obtain the so-called internal energy balance equation:

$$\partial_t(\rho e) + \mathrm{div}(\rho e \boldsymbol{u}) + p\, \mathrm{div}\boldsymbol{u} = 0. \tag{39}$$

Since,

- thanks to the mass balance equation, the first two terms in the left-hand side of (39) may be recast as a transport operator: $\partial_t(\rho e) + \mathrm{div}(\rho e \boldsymbol{u}) = \rho\, [\partial_t e + \boldsymbol{u} \cdot \boldsymbol{\nabla} e]$,
- and, from the equation of state, the pressure vanishes when $e = 0$,

this equation implies, if $e \geq 0$ at $t = 0$ and with suitable boundary conditions, that e remains non-negative at all times. As mentioned in the introduction, solving this latter equation instead of the total energy balance is appealing, to preserve by construction of the scheme this positivity property. In addition, it avoids to introduce a space discretization for the total energy which, for a staggered discretization, combines cell-centered (the internal energy and the density) and face-centered (the velocity) variables. However, a raw discretization of a non-conservative equation derived (formally, *i.e.* supposing unrealistic regularity properties of the solution) from a conservative system may be non-consistent (and numerical experiments show that, for the problem at hand, the so-derived scheme is unable to capture shock solutions). We circumvent here this problem by correcting the internal energy balance discretization, following a strategy already implemented in [13] for pressure correction schemes: the remainder terms obtained in the kinetic energy balance (term defined by Equation (18)) are compensated in the internal energy one, in order to make the scheme consistent with the total energy balance, in a sense which will be clarified in Section 4.2 below.

The paper is organized as follows. We first introduce the scheme in Section 4.1, and the above-mentionned corrective terms in the discrete internal energy balance are given; their expression is justified in Section 4.2 by proving a Lax-Wendroff consistency property for the algorithm in one space dimension (if a sequence of discrete solution converges in suitable norms, the limit necessarily satisfies a weak formulation of the continuous problem). The fact that the scheme keeps the internal energy positive under a CFL condition is demonstrated in Section 4.1; since ρ is positive thanks to the upwind discretization of the mass balance (Lemma 3.1), the proposed algorithm thus preserves the convex of admissible states.

4.1. The scheme

Let us consider a partition $0 = t_0 < t_1 < \ldots < t_N = T$ of the time interval $(0,T)$, which we suppose uniform for the sake of simplicity, and let $\delta t = t_{n+1} - t_n$ for $n = 0, 1, \ldots, N-1$ be the (constant) time step. We consider a decoupled-in-time scheme, which reads in its fully discrete form, for $0 \leq n \leq N-1$:

$$\forall K \in \mathcal{M},\ \frac{|K|}{\delta t}(\rho_K^{n+1} - \rho_K^n) + \sum_{\sigma\in\mathcal{E}(K)} F_{K,\sigma}^n = 0, \tag{40a}$$

$$\forall K \in \mathcal{M},\ \frac{|K|}{\delta t}(\rho_K^{n+1} e_K^{n+1} - \rho_K^n e_K^n) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^n e_\sigma^n + |K|\ p_K^n\,(\mathrm{div}\boldsymbol{u})_K^n = S_K^n, \tag{40b}$$

$$\forall K \in \mathcal{M},\ p_K^{n+1} = (\gamma - 1)\ \rho_K^{n+1}\ e_K^{n+1}, \tag{40c}$$

For $1 \le i \le d,\ \forall \sigma \in \mathcal{E}_{\mathrm{S}}^{(i)}$,

$$\frac{|D_\sigma|}{\delta t}(\rho_{D_\sigma}^{n+1} u_{\sigma,i}^{n+1} - \rho_{D_\sigma}^n u_{\sigma,i}^n) + \sum_{\epsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\epsilon}^n u_{\epsilon,i}^n + |D_\sigma|\,(\boldsymbol{\nabla} p)_{\sigma,i}^{n+1} = 0. \tag{40d}$$

The discrete mass balance and momentum balance equations (40a) and (40d) have already been derived in the previous section. Equation (40b) is an approximation of the internal energy balance (39) over the primal cell K. The positivity of the convection operator is ensured if we use an upwinding technique for this term [18]:

$$\text{for } \sigma = K|L \in \mathcal{E}_{\mathrm{int}}, \qquad e_\sigma^n = \begin{vmatrix} e_K^n & \text{if } F_{K,\sigma}^n \ge 0, \\ e_L^n & \text{otherwise.} \end{vmatrix}$$

The discrete divergence of the velocity, $(\mathrm{div}\boldsymbol{u})_K^n$, is defined by (11). The right-hand side, S_K^n, is derived below, using consistency arguments; at the first time step, it is simply set to zero:

$$\forall K \in \mathcal{M}, \qquad S_K^0 = 0.$$

The initial approximations for ρ, e and $\boldsymbol{u}$ are given by the average of the initial conditions ρ_0 and e_0 on the primal cells and of $\boldsymbol{u}_0$ on the dual cells:

$$\begin{aligned} &\forall K \in \mathcal{M}, \quad \rho_K^0 = \frac{1}{|K|}\int_K \rho_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad \text{and } e_K^0 = \frac{1}{|K|}\int_K e_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \\ &\text{for } 1 \le i \le d,\ \forall \sigma \in \mathcal{E}_{\mathrm{S}}^{(i)}, \quad u_{\sigma,i}^0 = \frac{1}{|D_\sigma|}\int_{D_\sigma} (\boldsymbol{u}_0(\boldsymbol{x}))_i\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{41}$$

Let us now detail how we choose the corrective term S_K in the internal energy balance, with the aim to recover a consistent discretization of the total energy balance. We wish to build these corrective terms so as to compensate the remainder terms in the kinetic energy balance (16), which we suspect not to tend to zero (for instance, the piecewise constant function associated to these terms for a shock solution - precisely speaking, to the terms obtained by applying (18) to the interpolate of a discontinuous function, on a sequence of discretizations with vanishing time and space steps - does not tend to zero in L^1). The first idea to do this could be just to sum the (discrete) kinetic energy balance with the internal energy balance: it would indeed be possible for a collocated discretization. But here, we face the fact that the kinetic energy balance is associated to the dual mesh, while the internal energy balance is discretized on the primal mesh. The way to circumvent this difficulty is to remark that we do not really need a discrete total energy balance; in fact, we only need to recover (a weak form of) this equation when the mesh and time steps tend to zero. To this purpose, we choose the quantities (S_K^{n+1}) in such a way as to somewhat compensate the terms $(R_{\sigma,i}^{n+1})$ given by (18). For $K \in \mathcal{M}$, we obtain $S_K^{n+1} = \sum_{i=1}^d S_{K,i}^{n+1}$ with:

$$S_{K,i}^{n+1} = \frac{1}{2}\,\rho_K^{n+1} \sum_{\sigma \in \mathcal{E}(K) \cap \mathcal{E}_{\mathrm{S}}^{(i)}} \frac{|D_{K,\sigma}|}{\delta t}\,\left(u_{\sigma,i}^{n+1} - u_{\sigma,i}^n\right)^2 + \sum_{\epsilon \in \tilde{\mathcal{E}}_{\mathrm{S}}^{(i)},\ \epsilon \cap \bar{K} \neq \emptyset} \alpha_{K,\epsilon}\ R_{\epsilon,i}^{n+1} \tag{42}$$

where $R_{\epsilon,i}^{n+1}$ is defined as follows. Let σ_ϵ^U and σ_ϵ^D be the two primal faces such that $\epsilon = D_{\sigma_\epsilon^D}|D_{\sigma_\epsilon^U}$ and that $F_{\sigma_\epsilon^D,\epsilon}^n \le 0$, which means that $D_{\sigma_\epsilon^D}$ is the dual cell located downstream ϵ. Then:

$$R_{\epsilon,i}^{n+1} = \frac{|F_{\sigma_\epsilon^D,\epsilon}^n|}{2}\,(u_{\sigma_\epsilon^D,i}^n - u_{\sigma_\epsilon^U,i}^n)^2 + F_{\sigma_\epsilon^D,\epsilon}^n\left(u_{\sigma_\epsilon^D,i}^{n+1} - u_{\sigma_\epsilon^D,i}^n\right)\,(u_{\sigma_\epsilon^U,i}^n - u_{\sigma_\epsilon^D,i}^n).$$

The coefficient $\alpha_{K,\epsilon}$ allows to distribute the remainder term $R_{\epsilon,i}^{n+1}$ over the neighbouring primal cells. If the face ϵ is included in K, $\alpha_{K,\epsilon} = 1$, which means that $R_{\epsilon,i}^{n+1}$ is totally affected to K; this is the only situation to consider for the RT and CR discretizations. For the MAC scheme, some dual faces also are included in the

primal cells (and for these faces, $\alpha_{K,\epsilon}$ is still set to 1), but some lie on the boundary of the primal cells. In such a case, let us denote by $\mathcal{N}_\epsilon^D$ the set of the two primal control volumes separated by σ_ϵ^D. The coefficient $\alpha_{K,\epsilon}$ is then given by:

$$\alpha_{K,\epsilon} = \begin{cases} \dfrac{|K|}{\sum_{L\in\mathcal{N}_\epsilon^D} |L|} \text{ if } K \in \mathcal{N}_\epsilon^D, \\ 0 \text{ otherwise.} \end{cases} \tag{43}$$

For a uniform grid, this formula yields $\alpha_{K,\epsilon} = 1/2$ if $K \in \mathcal{N}_\epsilon^D$.

The expression of the correction terms $S_K^{n+1}, K \in \mathcal{M}$ is justified by the passage to the limit in the scheme (for a one-dimensional problem) performed in the next section. We note however here that:

$$\sum_{K\in\mathcal{M}} S_K^{n+1} - \sum_{i=1}^{d} \sum_{\sigma\in\mathcal{E}_g^{(i)}} R_{\sigma,i}^{n+1} = 0. \tag{44}$$

Indeed, the first part of $S_{K,i}^{n+1}$, thanks to the expression (9) of the density at the face $\rho_{D_\sigma}^{n+1}$, results from dispatching the first part of the residual over the two adjacent cells:

$$\frac{1}{2}\,\frac{|D_\sigma|}{\delta t}\,\rho_{D_\sigma}^{n+1}\,\left(u_{\sigma,i}^{n+1}-u_{\sigma,i}^{n}\right)^2 = \underbrace{\frac{1}{2}\,\frac{|D_{K,\sigma}|}{\delta t}\,\rho_K^{n+1}\,\left(u_{\sigma,i}^{n+1}-u_{\sigma,i}^{n}\right)^2}_{\text{affected to K}} + \underbrace{\frac{1}{2}\,\frac{|D_{L,\sigma}|}{\delta t}\,\rho_L^{n+1}\,\left(u_{\sigma,i}^{n+1}-u_{\sigma,i}^{n}\right)^2}_{\text{affected to L}}.$$

The same argument holds for the terms associated to the dual faces, which explains, in particular, the definition of the coefficients $\alpha_{K,\epsilon}$. The scheme thus conserves the integral of the total energy over the computational domain. In the scheme itself, we shall use the term S_K^n rather than S_K^{n+1}, because we want an explicit scheme, but this does not hinder the consistency of the scheme, as shown in the proof of Theorem 4.2.

The definition (42) of $(S_K^{n+1})_{K\in\mathcal{M}}$ allows to prove that, under a CFL condition, the scheme preserves the positivity of e.

Lemma 4.1. *Let us suppose that, for $0 \le n \le N-1$, for all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}(K)$, we have:*

$$\delta t \le \min\Big(\frac{|K|}{\gamma \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,(u_{K,\sigma}^n)^+}, \frac{|D_{K,\sigma}|\,\rho_K^{n+1}}{\sum_{\epsilon\in\tilde{\mathcal{E}}(D_\sigma),\ \epsilon\cap\bar{K}\neq\emptyset} \alpha_{K,\epsilon}\,(F_{\sigma,\epsilon}^n)^-}\Big). \tag{45}$$

Then the internal energy $(e^n)_{1\le n\le N}$ given by the scheme (40) is positive.

Proof. Let n such that $0 \le n \le N$ be given, and let us assume that $e_K^n \ge 0$ and $S_K^n \ge 0$ for all $K \in \mathcal{M}$. Since, by assumption, $\gamma > 1$, the CFL condition (45) implies that the CFL condition (15) is satisfied, and by Lemma 3.1 we thus have $\rho_K^n > 0$ and $\rho_K^{n+1} > 0$, for all $K \in \mathcal{M}$. In the internal energy equation (40b), let us express the pressure thanks to the equation of state (40c) to obtain:

$$\begin{aligned}\frac{|K|}{\delta t}\rho_K^{n+1}e_K^{n+1} = \Big[\frac{|K|}{\delta t}\rho_K^n - \sum_{\sigma\in\mathcal{E}(K)} (F_{K,\sigma}^n)^+ - (\gamma-1)\,\rho_K^n \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,(u_{K,\sigma}^n)^+\Big]\,e_K^n \\ + \sum_{\sigma\in\mathcal{E}(K)} (F_{K,\sigma}^n)^- e_L^n + (\gamma-1)\,\rho_K^n\,e_K^n \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,(u_{K,\sigma}^n)^- + S_K^n.\end{aligned} \tag{46}$$

Using the fact that, when $u_{K,\sigma}^n \ge 0$, the upwind density at the face is ρ_K^n, we have:

$$(F_{K,\sigma}^n)^+ + (\gamma-1)\,|\sigma|\,\rho_K^n\,(u_{K,\sigma}^n)^+ = \gamma\,|\sigma|\,\rho_K^n\,(u_{K,\sigma}^n)^+,$$

and hence Relation (46) reads:

$$\begin{aligned}\frac{|K|}{\delta t}\rho_K^{n+1}e_K^{n+1} = \Big[\frac{|K|}{\delta t} - \gamma \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,(u_{K,\sigma}^n)^+\Big]\,\rho_K^n\,e_K^n \\ + \sum_{\sigma\in\mathcal{E}(K)} (F_{K,\sigma}^n)^- e_L^n + (\gamma-1)\,\rho_K^n\,e_K^n \sum_{\sigma\in\mathcal{E}(K)} |\sigma|\,(u_{K,\sigma}^n)^- + S_K^n.\end{aligned}$$

Then we get $e_K^{n+1} > 0$ under the following CFL condition:

$$\delta t \leq \frac{|K|}{\gamma \ \sum_{\sigma\in\mathcal{E}(K)} |\sigma|(u_{K,\sigma}^n)^+}.$$

Let us now derive a condition for the non-negativity of the source term. To this purpose, for $K \in \mathcal{M}$, let us recall the definition (42) of $S_{K,i}^{n+1}$:

$$S_{K,i}^{n+1} = \frac{1}{2}\,\rho_K^{n+1} \sum_{\sigma\in\mathcal{E}(K)\cap\mathcal{E}_{\rm S}^{(i)}} \frac{|D_{K,\sigma}|}{\delta t}\,\left(u_{\sigma,i}^{n+1} - u_{\sigma,i}^n\right)^2 + \sum_{\substack{\epsilon\in\tilde{\mathcal{E}}_{\rm S}^{(i)},\ \epsilon\cap\bar{K}\neq\emptyset,\\ \epsilon=D_\sigma|D_{\sigma'},\, F_{\sigma,\epsilon}^n\leq 0}} \alpha_{K,\epsilon}\,\Big[\frac{|F_{\sigma,\epsilon}^n|}{2}\,(u_{\sigma,i}^n - u_{\sigma',i}^n)^2 + F_{\sigma,\epsilon}^n\left(u_{\sigma,i}^{n+1} - u_{\sigma,i}^n\right)\,(u_{\sigma',i}^n - u_{\sigma,i}^n)\Big].$$

In the indexes of the last sum, the purpose of the second line is to define the notations used in the sum: the diamond cells separated by ϵ are denoted by D_σ and $D_{\sigma'}$, and D_σ is the cell downstream ϵ. We also recall that the coefficient $\alpha_{K,\epsilon}$ is different from zero only if σ is a face of K (and, of course, $\sigma \in \mathcal{E}_{\rm S}^{(i)}$, so, if $\alpha_{K,\epsilon} \neq 0$, σ appears in the first sum of the expression). Applying Young's inequality to the last term of $S_{K,i}^{n+1}$, denoted by $(S_{K,i}^{n+1})_3$, we obtain

$$(S_{K,i}^{n+1})_3 \geq -\sum_{\substack{\epsilon\in\tilde{\mathcal{E}}_{\rm S}^{(i)},\ \epsilon\cap\bar{K}\neq\emptyset,\\ \epsilon=D_\sigma|D_{\sigma'},\, F_{\sigma,\epsilon}^n\leq 0}} \alpha_{K,\epsilon}\,\frac{|F_{\sigma,\epsilon}^n|}{2}\,\left(u_{\sigma,i}^{n+1} - u_{\sigma,i}^n\right)^2 - \sum_{\substack{\epsilon\in\tilde{\mathcal{E}}_{\rm S}^{(i)},\ \epsilon\cap\bar{K}\neq\emptyset,\\ \epsilon=D_\sigma|D_{\sigma'},\, F_{\sigma,\epsilon}^n\leq 0}} \alpha_{K,\epsilon}\,\frac{|F_{\sigma,\epsilon}^n|}{2}\,(u_{\sigma',i}^n - u_{\sigma,i}^n)^2.$$

Gathering all terms of $S_{K,i}^{n+1}$ yields:

$$S_{K,i}^{n+1} \geq \sum_{\sigma\in\mathcal{E}(K)} \frac{1}{2}\big(u_{\sigma,i}^{n+1} - u_{\sigma,i}^n\big)^2\Big[\frac{|D_{K,\sigma}|}{\delta t}\rho_K^{n+1} - \sum_{\epsilon\in\tilde{\mathcal{E}}(D_\sigma),\ \epsilon\cap\bar{K}\neq\emptyset} \alpha_{K,\epsilon}\,(F_{\sigma,\epsilon}^n)^-\Big],$$

thus $S_{K,i}^{n+1}$ is non-negative under the CFL condition:

$$\delta t \leq \frac{|D_{K,\sigma}|\ \rho_K^{n+1}}{\displaystyle\sum_{\epsilon\in\tilde{\mathcal{E}}(D_\sigma),\ \epsilon\cap\bar{K}\neq\emptyset} \alpha_{K,\epsilon}\,(F_{\sigma,\epsilon}^n)^-}, \qquad \forall\sigma\in\mathcal{E}(K),$$

which concludes the proof. □

4.2. **Passing to the limit in the scheme**

As in the isentropic case, we are now going to show in the one dimensional case that if a sequence of solutions is controlled in suitable norms and converges to a limit, this latter necessarily satisfies a (part of the) weak formulation of the continuous problem. We again write a 1D version of the scheme and use the same notations as in Section 3.3. The explicit scheme (40) may be written as follows in the one dimensional setting:

$$\begin{aligned} &\forall K\in\mathcal{M}, \qquad \rho_K^0 = \frac{1}{|K|}\int_K \rho_0(x)\,\mathrm{d}x, \qquad e_K^0 = \frac{1}{|K|}\int_K e_0(x)\,\mathrm{d}x,\\ &\forall\sigma\in\mathcal{E}_{\rm int}, \qquad u_\sigma^0 = \frac{1}{|D_\sigma|}\int_{D_\sigma} u_0(x)\,\mathrm{d}x, \end{aligned} \tag{47a}$$

$$\forall K = [\overrightarrow{\sigma\sigma'}]\in\mathcal{M}, \quad \frac{|K|}{\delta t}(\rho_K^{n+1} - \rho_K^n) + F_{\sigma'}^n - F_\sigma^n = 0, \tag{47b}$$

$$\forall K = [\overrightarrow{\sigma\sigma'}]\in\mathcal{M}, \quad \frac{|K|}{\delta t}(\rho_K^{n+1}e_K^{n+1} - \rho_K^n e_K^n) + F_{\sigma'}^n e_{\sigma'}^n - F_\sigma^n e_\sigma^n + p_K^n(u_{\sigma'}^n - u_\sigma^n) = S_K^n, \tag{47c}$$

$$\forall K \in \mathcal{M}, \quad p_K^{n+1} = (\gamma - 1)\, \rho_K^{n+1}\, e_K^{n+1}, \tag{47d}$$

$$\forall \sigma = \overrightarrow{K|L} \in \mathcal{E}_{\rm int},$$
$$\frac{|D_\sigma|}{\delta t}(\rho_{D_\sigma}^{n+1} u_\sigma^{n+1} - \rho_{D_\sigma}^n u_\sigma^n) + F_L^n u_L^n - F_K^n u_K^n + p_L^{n+1} - p_K^{n+1} = 0, \tag{47e}$$

where the corrective term S_K^n reads, for $1 \le n \le N$ and $\forall K = [\sigma' \to \sigma]$:

$$S_K^n = \frac{|K|}{4\,\delta t}\,\rho_K^n \left[(u_\sigma^n - u_\sigma^{n-1})^2 + (u_{\sigma'}^n - u_{\sigma'}^{n-1})^2\right] + \frac{|F_K^{n-1}|}{2}\,(u_\sigma^{n-1} - u_{\sigma'}^{n-1})^2$$
$$- |F_K^{n-1}|(u_\sigma^n - u_\sigma^{n-1})\,(u_{\sigma'}^{n-1} - u_\sigma^{n-1}), \tag{48}$$

where the notation $K = [\sigma' \to \sigma]$ means that the flow goes from σ' to σ (*i.e.*, if $F_K^n \ge 0$, $K = [\overrightarrow{\sigma'\sigma}]$ and, if $F_K^n \le 0$, $K = [\overrightarrow{\sigma\sigma'}]$). At the first time step, we set $S_K^0 = 0,\ \forall K \in \mathcal{M}$.

We again consider a sequence of discretizations $(\mathcal{M}^{(m)},\ \delta t^{(m)})_{m\in\mathbb{N}}$ with $h^{(m)} = \sup_{K\in\mathcal{M}^{(m)}} h_K$. Let $\rho^{(m)}$, $p^{(m)}$, $e^{(m)}$ and $u^{(m)}$ be the solution given by the scheme (47) with the mesh $\mathcal{M}^{(m)}$ and the time step $\delta t^{(m)}$. As in the isentropic case, to the discrete unknowns, we associate piecewise constant functions on time intervals and on primal or dual meshes, so the density $\rho^{(m)}$, the pressure $p^{(m)}$, the internal energy $e^{(m)}$ and the velocity $u^{(m)}$ are defined almost everywhere on $\Omega \times (0,T)$ by (30) and

$$e^{(m)}(x,t) = \sum_{n=0}^{N-1} \sum_{K\in\mathcal{M}} (e^{(m)})_K^n\, \mathfrak{X}_K(x)\, \mathfrak{X}_{[n,n+1)}(t). \tag{49}$$

For the consistency result that we are seeking (Theorem 4.2 below), we have to assume that a sequence of discrete solutions $(\rho^{(m)}, p^{(m)}, e^{(m)}, u^{(m)})_{m\in\mathbb{N}}$ satisfies $\rho^{(m)} > 0$, $p^{(m)} > 0$ and $e^{(m)} > 0$, $\forall m \in \mathbb{N}$ (which may be a consequence of the fact that the CFL stability condition (15) is satisfied), and is uniformly bounded in $\mathrm{L}^\infty(\Omega \times (0,T))^4$, *i.e.*, for $m \in \mathbb{N}$ and $0 \le n \le N^{(m)}$:

$$0 < (\rho^{(m)})_K^n \le C,\ 0 < (p^{(m)})_K^n \le C,\ 0 < (e^{(m)})_K^n \le C, \quad \forall K \in \mathcal{M}^{(m)}, \tag{50}$$

and

$$|(u^{(m)})_\sigma^n| \le C, \quad \forall \sigma \in \mathcal{E}^{(m)}, \tag{51}$$

where C is a positive real number. Note that, by definition of the initial conditions of the scheme, these inequalities imply that the functions ρ_0, e_0 and u_0 belong to $\mathrm{L}^\infty(\Omega)$. We also have to assume that a sequence of discrete solutions satisfies the following uniform bounds with respect to the discrete BV-norms:

$$\|\rho^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} + \|p^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} + \|e^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} + \|u^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}} \le C, \quad \forall m \in \mathbb{N}, \tag{52}$$

and

$$\|u^{(m)}\|_{\mathfrak{T},t,\mathrm{BV}} \le C, \quad \forall m \in \mathbb{N}. \tag{53}$$

Again, we are not able to prove such estimates for the solutions of the scheme; however, such inequalities are satisfied by the "interpolates" (for instance, by taking the cell average) of the solution to a Riemann problem, and are observed in computations (of course, as far as possible, *i.e.* with a limited sequence of meshes and time steps).

A weak solution to the continuous problem satisfies, for any $\varphi \in C_c^\infty\big(\Omega \times [0,T)\big)$:

$$-\int_0^T \int_\Omega \Big[\rho\,\partial_t\varphi + \rho\,u\,\partial_x\varphi\Big]\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \rho_0(x)\,\varphi(x,0)\,\mathrm{d}x = 0, \tag{54a}$$

$$-\int_0^T \int_\Omega \Big[\rho\,u\,\partial_t\varphi + (\rho\,u^2 + p)\,\partial_x\varphi\Big]\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \rho_0(x)\,u_0(x)\,\varphi(x,0)\,\mathrm{d}x = 0, \tag{54b}$$

$$-\int_0^T \int_\Omega \Big[\rho\,E\,\partial_t\varphi + (\rho\,E + p)\,u\,\partial_x\varphi\Big]\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \rho_0(x)\,E_0(x)\,\varphi(x,0)\,\mathrm{d}x = 0, \tag{54c}$$

$$p = (\gamma - 1)\rho\,e, \quad E = \frac{1}{2}u^2 + e, \quad E_0 = \frac{1}{2}u_0^2 + e_0. \tag{54d}$$

As in the isentropic case, these relations are not sufficient to define a weak solution to the problem, since they do not imply anything about the boundary conditions, but allow to derive the Rankine-Hugoniot conditions. We show hereafter that they are satisfied by the limit of a sequence of solutions to the discrete problem (Theorem 4.2). This result thus proves that the introduction of corrective terms in the internal energy balance indeed yields a consistent scheme; conversely, without these terms, we may anticipate that the algorithm will compute uncorrect shocks (*i.e.* shocks where the jumps of the unknowns and of the fluxes are *not* linked to the shock speed by the Rankine-Hugoniot conditions); this is confirmed by numerical experiments.

Theorem 4.2 (Consistency of the one-dimensional explicit scheme)**.**
Let Ω be an open bounded interval of $\mathbb{R}$. We suppose that the initial data satisfies $\rho_0 \in \mathrm{L}^\infty(\Omega)$, $p_0 \in \mathrm{BV}(\Omega)$, $e_0 \in \mathrm{L}^\infty(\Omega)$ and $u_0 \in \mathrm{L}^\infty(\Omega)$. Let $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m\in\mathbb{N}}$ be a sequence of discretizations such that both the time step $\delta t^{(m)}$ and the size $h^{(m)}$ of the mesh $\mathcal{M}^{(m)}$ tend to zero as $m \to \infty$, and let $(\rho^{(m)}, p^{(m)}, e^{(m)}, u^{(m)})_{m\in\mathbb{N}}$ be the corresponding sequence of solutions. We suppose that this sequence satisfies the estimates (50)–(53) and converges in $\mathrm{L}^r(\Omega\times(0,T))^4$, for $1 \le r < \infty$, to $(\bar\rho, \bar p, \bar e, \bar u) \in \mathrm{L}^\infty(\Omega\times(0,T))^4$.

Then the limit $(\bar\rho, \bar p, \bar e, \bar u)$ satisfies the system (54).

Proof. As in the isentropic case, it is clear that with the assumed convergence for the sequence of solutions, the limit satisfies the equation of state. The fact that the limit satisfies the weak mass balance equation (54a) and the weak momentum balance equation (54b) was shown in the previous section. There only remains to prove that (54c) holds, by passing to the limit in the scheme, in the internal and the kinetic energy balance equations.

Let $\varphi \in \mathbb{C}_c^\infty(\Omega\times[0,T))$. Let $m \in \mathbb{N}$, $\mathcal{M}^{(m)}$ and $\delta t^{(m)}$ be given. Dropping for short the superscript $^{(m)}$, let $\varphi_\mathcal{M}$ be the interpolate of φ on the primal mesh and let $\eth_t\varphi_\mathcal{M}$ and $\eth_x\varphi_\mathcal{M}$ be its time and space discrete derivatives in the sense of Definition 3.4. Thanks to the regularity of φ, these functions respectively converge in $\mathrm{L}^r(\Omega\times(0,T))$, for $r \ge 1$ (including $r = +\infty$), to φ, $\partial_t\varphi$ and $\partial_x\varphi$ respectively. In addition, $\varphi_\mathcal{M}(\cdot,0)$ (which, for $K \in \mathcal{M}$ and $x \in K$, is equal to $\varphi_K^0 = \varphi(x_K,0)$) converges to $\varphi(\cdot,0)$ in $\mathrm{L}^r(\Omega)$ for $r \ge 1$. We also define $\varphi_\mathcal{E}$, $\eth_t\varphi_\mathcal{E}$ and $\eth_x\varphi_\mathcal{E}$, as, respectively, the interpolate of φ on the dual mesh and its discrete time and space derivatives, still in the sense of Definition 3.4; once again thanks to the regularity of φ, these functions converge in $\mathrm{L}^r(\Omega\times(0,T))$, for $r \ge 1$, to φ, $\partial_t\varphi$ and $\partial_x\varphi$ respectively. As for the primal mesh interpolate, the dual mesh interpolate $\varphi_\mathcal{E}(\cdot,0)$ (which, for $\sigma \in \mathcal{E}$ and $x \in D_\sigma$, is equal to $\varphi_\sigma^0 = \varphi(x_\sigma,0)$) converges to $\varphi(\cdot,0)$ in $\mathrm{L}^r(\Omega)$ for $r \ge 1$.

Since the support of φ is compact in $\Omega\times[0,T)$, for m large enough, the interpolates of φ vanish on the boundary cells and at the last time step(s); hereafter, we assume that we are in this case.

On one hand, let us multiply the one dimensional discrete internal energy balance equation (47c) by $\delta t\,\varphi_K^{n+1}$, and sum the result for $0 \le n \le N-1$ and $K \in \mathcal{M}$. On the other hand, let us multiply the one-dimensional version of the discrete kinetic energy balance (17) by $\delta t\,\varphi_\sigma^{n+1}$, and sum the result for $0 \le n \le N-1$ and $\sigma \in \mathcal{E}_{\rm int}$. Finally, adding the two obtained relations, we get:

$$T_1^{(m)} + T_2^{(m)} + T_3^{(m)} + \tilde T_1^{(m)} + \tilde T_2^{(m)} + \tilde T_3^{(m)} = S^{(m)} - \tilde R^{(m)}, \tag{55}$$

where:

$$T_1^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} \frac{|K|}{\delta t} \left[\rho_K^{n+1}\, e_K^{n+1} - \rho_K^n\, e_K^n\right] \varphi_K^{n+1},$$

$$T_2^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K=\overrightarrow{[\sigma\sigma']}\in\mathcal{M}} \left[\rho_{\sigma'}^n\, e_{\sigma'}^n\, u_{\sigma'}^n - \rho_\sigma^n\, e_\sigma^n\, u_\sigma^n\right] \varphi_K^{n+1},$$

$$T_3^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K=\overrightarrow{[\sigma\sigma']}\in\mathcal{M}} p_K^n\,(u_{\sigma'}^n - u_\sigma^n)\,\varphi_K^{n+1},$$

$$\tilde T_1^{(m)} = \frac12 \sum_{n=0}^{N-1} \delta t \sum_{\sigma\in\mathcal{E}_{\rm int}} \frac{|D_\sigma|}{\delta t} \left[\rho_{D_\sigma}^{n+1}(u_\sigma^{n+1})^2 - \rho_{D_\sigma}^n (u_\sigma^n)^2\right] \varphi_\sigma^{n+1},$$

$$\tilde T_2^{(m)} = \frac12 \sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} \left[F_L^n\,(u_L^n)^2 - F_K^n\,(u_K^n)^2\right] \varphi_\sigma^{n+1},$$

$$\tilde T_3^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\rm int}} (p_L^{n+1} - p_K^{n+1})\, u_\sigma^{n+1}\, \varphi_\sigma^{n+1},$$

$$S^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} S_K^n \ \varphi_K^{n+1}, \qquad \tilde{R}^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{\sigma\in\mathcal{E}_{\rm int}} R_\sigma^{n+1} \ \varphi_\sigma^{n+1},$$

and the quantities S_K^n and R_σ^{n+1} are given by Equation (48) and (the 1D version of) Equation (18) respectively.

Reordering the sums in $T_1^{(m)}$ yields:

$$T_1^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{K\in\mathcal{M}} |K|\, \rho_K^n\, e_K^n\ \frac{\varphi_K^{n+1}-\varphi_K^n}{\delta t} - \sum_{K\in\mathcal{M}} |K|\, \rho_K^0\, e_K^0\ \varphi_K^0,$$

so that:

$$T_1^{(m)} = -\int_0^T \int_\Omega \rho^{(m)}\, e^{(m)}\, \eth_t \varphi_{\mathcal{M}}\, \mathrm{d}x\, \mathrm{d}t - \int_\Omega (\rho^{(m)})^0(x)\ (e^{(m)})^0(x)\ \varphi_{\mathcal{M}}(x,0)\, \mathrm{d}x.$$

The boundedness of ρ_0, e_0 and the definition (47a) of the initial conditions for the scheme ensures that the sequences $((\rho^{(m)})^0)_{m\in\mathbb{N}}$ and $((e^{(m)})^0)_{m\in\mathbb{N}}$ converge to ρ_0 and e_0 respectively in $\mathrm{L}^r(\Omega)$ for $r\geq 1$. Since, by assumption, the sequence of discrete solutions and of the interpolate time derivatives converge in $\mathrm{L}^r\big(\Omega\times(0,T)\big)$ for $r\geq 1$, we thus obtain:

$$\lim_{m\to+\infty} T_1^{(m)} = -\int_0^T \int_\Omega \bar\rho\, \bar e\, \partial_t\varphi\, \mathrm{d}x\, \mathrm{d}t - \int_\Omega \rho_0(x)\ e_0(x)\ \varphi(x,0)\, \mathrm{d}x.$$

Reordering the sums in $T_2^{(m)}$, we get:

$$T_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} \rho_\sigma^n\, e_\sigma^n\, u_\sigma^n\ (\varphi_L^{n+1}-\varphi_K^{n+1}).$$

Using the fact that $h_\sigma = |D_\sigma|$, this relation reads:

$$T_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} |D_\sigma|\ \rho_\sigma^n\, e_\sigma^n\, u_\sigma^n\ \frac{\varphi_L^{n+1}-\varphi_K^{n+1}}{h_\sigma},$$

thus $T_2^{(m)} = \mathcal{T}_2^{(m)} + \mathcal{R}_2^{(m)}$ with:

$$\mathcal{T}_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} \Big[|D_{K,\sigma}|\ \rho_K^n\, e_K^n + |D_{L,\sigma}|\ \rho_L^n\, e_L^n\Big]\, u_\sigma^n\ \frac{\varphi_L^{n+1}-\varphi_K^{n+1}}{h_\sigma},$$

$$\mathcal{R}_2^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}} \Big[|D_\sigma|\ \rho_\sigma^n\, e_\sigma^n - |D_{K,\sigma}|\ \rho_K^n\, e_K^n - |D_{L,\sigma}|\ \rho_L^n\, e_L^n\Big] u_\sigma^n\ \frac{\varphi_L^{n+1}-\varphi_K^{n+1}}{h_\sigma}.$$

The first expression reads:

$$\mathcal{T}_2^{(m)} = -\int_0^T \int_\Omega \rho^{(m)}\, e^{(m)}\, u^{(m)}\ \eth_x \varphi_{\mathcal{M}}\, \mathrm{d}x\, \mathrm{d}t,$$

and thus, thanks to the convergence assumptions:

$$\lim_{m\to+\infty} \mathcal{T}_2^{(m)} = -\int_0^T \int_\Omega \bar\rho\, \bar e\, \bar u\ \partial_x\varphi\, \mathrm{d}x\, \mathrm{d}t.$$

Let us choose σ in such a way that $\rho_\sigma^n = \rho_K^n$ and $e_\sigma^n = e_K^n$ (in other words, we choose to call K the upwind cell to σ instead of the left cell, which we denote by $\sigma = K\to L$). We thus get, with $C_\varphi = \|\partial_x\varphi\|_{\mathrm{L}^\infty(\Omega\times(0,T))}$:

$$|\mathcal{R}_2^{(m)}| \leq C_\varphi \sum_{n=0}^{N-1} \delta t \sum_{\sigma=K\to L\in\mathcal{E}} |D_{L,\sigma}|\ \Big|\rho_K^n\, e_K^n - \rho_L^n\, e_L^n\Big|\ |u_\sigma^n|.$$

Applying the identity $2\,(ab-cd) = (a-c)(b+d) + (a+c)(b-d)$, which holds for any $\{a,b,c,d\} \subset \mathbb{R}$, to the quantity $\rho_K^n\, e_K^n - \rho_L^n\, e_L^n$, we obtain:

$$|\mathcal{R}_2^{(m)}| \leq C_\varphi\ h^{(m)}\ \|u^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \Big[\|\rho^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}}\ \|e^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))} + \|\rho^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|e^{(m)}\|_{\mathfrak{T},x,\mathrm{BV}}\Big],$$

and thus $|\mathcal{R}_2^{(m)}|$ tends to zero when m tends to $+\infty$.

For the term $\tilde{T}_1^{(m)}$, the definition (9) of ρ_{D_σ} and a reordering in the summation yield:

$$\tilde{T}_1^{(m)} = -\frac{1}{2}\sum_{n=0}^{N-1}\delta t \sum_{\sigma=K|L\in\mathcal{E}}\Big[|D_{K,\sigma}|\,\rho_K^n + |D_{L,\sigma}|\,\rho_L^n\Big]\,(u_\sigma^n)^2\ \frac{\varphi_K^{n+1}-\varphi_K^n}{\delta t} - \frac{1}{2}\sum_{\sigma=K|L\in\mathcal{E}}\Big[|D_{K,\sigma}|\,\rho_K^0 + |D_{L,\sigma}|\,\rho_L^0\Big]\,(u_\sigma^0)^2\ \varphi_K^0,$$

so that, by similar arguments as for the term $T_1^{(m)}$, we get:

$$\lim_{m\to+\infty}\tilde{T}_1^{(m)} = -\int_0^T\int_\Omega \frac{1}{2}\ \bar{\rho}\ \bar{u}^2\ \partial_t\varphi\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \frac{1}{2}\ \rho_0(x)\ u_0(x)^2\ \varphi(x,0)\,\mathrm{d}x.$$

Let us now turn to the term $\tilde{T}_2^{(m)}$. Reordering the sums, we get:

$$\tilde{T}_2^{(m)} = -\frac{1}{2}\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} F_K^n\,(u_K^n)^2\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}),$$

and, by definition of the mass flux at the dual edges:

$$\tilde{T}_2^{(m)} = -\frac{1}{4}\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} (\rho_\sigma^n u_\sigma^n + \rho_{\sigma'}^n u_{\sigma'}^n)\,(u_K^n)^2\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}),$$

where we recall that u_K^n is equal to either u_σ^n or $u_{\sigma'}^n$, depending on the sign of F_K^n. Let us write $\tilde{T}_2^{(m)} = \tilde{\mathfrak{T}}_2^{(m)} + \tilde{\mathfrak{R}}_2^{(m)}$, with:

$$\tilde{\mathfrak{T}}_2^{(m)} = -\frac{1}{4}\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} \rho_K^n\,\big[(u_\sigma^n)^3 + (u_{\sigma'}^n)^3\big]\ (\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}).$$

We have:

$$\tilde{\mathfrak{T}}_2^{(m)} = -\int_0^T\int_\Omega \frac{1}{2}\ \rho^{(m)}\ (u^{(m)})^3\ \eth_x\varphi_{\mathcal{E}}\,\mathrm{d}x\,\mathrm{d}t,$$

and hence:

$$\lim_{m\to+\infty}\tilde{\mathfrak{T}}_2^{(m)} = -\int_0^T\int_\Omega \frac{1}{2}\ \bar{\rho}\ \bar{u}^3\ \partial_x\varphi\,\mathrm{d}x\,\mathrm{d}t.$$

The remainder term reads:

$$\tilde{\mathfrak{R}}_2^{(m)} = -\frac{1}{4}\sum_{n=0}^{N-1}\delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} \Big[(\rho_\sigma^n u_\sigma^n + \rho_{\sigma'}^n u_{\sigma'}^n)\,(u_K^n)^2 - \rho_K^n\,\Big((u_\sigma^n)^3 + (u_{\sigma'}^n)^3\Big)\Big](\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}).$$

Using the notation $K = \sigma\to\sigma'$ in the above summation in order to have $u_K^n = u_\sigma^n$, we obtain, with $\varepsilon = \pm 1$:

$$\tilde{\mathfrak{R}}_2^{(m)} = -\frac{\varepsilon}{4}\sum_{n=0}^{N-1}\delta t \sum_{K=\sigma\to\sigma'\in\mathcal{M}} \Big[(\rho_\sigma^n u_\sigma^n + \rho_{\sigma'}^n u_{\sigma'}^n)\,(u_\sigma^n)^2 - \rho_K^n\,\Big((u_\sigma^n)^3 + (u_{\sigma'}^n)^3\Big)\Big](\varphi_{\sigma'}^{n+1}-\varphi_\sigma^{n+1}).$$

For $0\leq n\leq N-1$ and $K\in\mathcal{M}$, we have:

$$(\rho_\sigma^n u_\sigma^n + \rho_{\sigma'}^n u_{\sigma'}^n)\,(u_\sigma^n)^2 - \rho_K^n\,\Big((u_\sigma^n)^3 + (u_{\sigma'}^n)^3\Big) = (\rho_\sigma^n - \rho_K^n)\,(u_\sigma^n)^3 + (\rho_{\sigma'}^n - \rho_K^n)\,u_{\sigma'}^n\,(u_\sigma^n)^2 + \rho_K^n\,u_{\sigma'}^n\,(u_\sigma^n + u_{\sigma'}^n)\,(u_\sigma^n - u_{\sigma'}^n),$$

Since, in this expression, ρ_σ^n and $\rho_{\sigma'}^n$ are the density either in K or in a neighbouring cell of K, we get:

$$|\tilde{\mathcal{R}}_2^{(m)}| \leq C_\varphi \; h^{(m)} \; \Big[\|u^{(m)}\|^3_{\mathrm{L}^\infty(\Omega\times(0,T))} \; \|\rho\|_{\mathcal{T},x,\mathrm{BV}} + \|\rho^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))} \; \|u^{(m)}\|^2_{\mathrm{L}^\infty(\Omega\times(0,T))} \; \|u^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} \Big],$$

where the real number C_φ only depends on φ. Hence $|\tilde{\mathcal{R}}_2^{(m)}|$ tends to zero when m tends to $+\infty$.

We now turn to $T_3^{(m)}$ and $\tilde{T}_3^{(m)}$. By a change in the notation of the time exponents, using the fact that φ_σ vanishes at the last time step(s), we get:

$$\tilde{T}_3^{(m)} = \sum_{n=1}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\mathrm{int}}} (p_L^n - p_K^n) \, u_\sigma^n \, \varphi_\sigma^n = \tilde{\mathcal{T}}_3^{(m)} + \tilde{\mathcal{R}}_3^{(m)},$$

with:

$$\tilde{\mathcal{T}}_3^{(m)} = \sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\mathrm{int}}} (p_L^n - p_K^n) \, u_\sigma^n \, \varphi_\sigma^{n+1},$$

$$\tilde{\mathcal{R}}_3^{(m)} = -\delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\mathrm{int}}} (p_L^0 - p_K^0) \, u_\sigma^0 \, \varphi_\sigma^0 + \sum_{n=0}^{N-1} \delta t \sum_{\sigma=\overrightarrow{K|L}\in\mathcal{E}_{\mathrm{int}}} (p_L^n - p_K^n) \, u_\sigma^n \, (\varphi_\sigma^n - \varphi_\sigma^{n+1}).$$

We have, thanks to the regularity of φ:

$$|\tilde{\mathcal{R}}_3^{(m)}| \leq C_\varphi \; \delta t^{(m)} \Big[\|(u^{(m)})^0\|_{\mathrm{L}^\infty(\Omega)} \; \|(p^{(m)})^0\|_{\mathrm{BV}(\Omega)} + \|u^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))} \; \|p^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} \Big].$$

Therefore, invoking the regularity of the initial conditions, this term tends to zero when m tends to $+\infty$. We now have for the other terms, reordering the summations:

$$T_3^{(m)} + \tilde{\mathcal{T}}_3^{(m)} = -\sum_{n=0}^{N-1} \delta t \sum_{K=[\overrightarrow{\sigma\sigma'}]\in\mathcal{M}} p_K^n \, u_\sigma^n \, (\varphi_K^{n+1} - \varphi_\sigma^{n+1}) + p_K^n \, u_{\sigma'}^n \, (\varphi_{\sigma'}^{n+1} - \varphi_K^{n+1})$$

$$= -\int_0^T \int_\Omega p^{(m)} \; u^{(m)} \; \eth_x \varphi_{\mathcal{M},\mathcal{E}} \, \mathrm{d}x \, \mathrm{d}t.$$

Since $\eth_x \varphi_{\mathcal{M},\mathcal{E}}$ converges to $\partial_x \varphi$ in $\mathrm{L}^r(\Omega \times (0,T))$ for any $r \geq 1$, we get:

$$\lim_{m\to+\infty} (T_3^{(m)} + \tilde{\mathcal{T}}_3^{(m)}) = -\int_0^T \int_\Omega \bar{p} \; \bar{u} \; \partial_x \varphi \, \mathrm{d}x \, \mathrm{d}t.$$

It now remains to check that $\lim_{m\to+\infty}(S^{(m)} - \tilde{R}^{(m)}) = 0$. Let us write this quantity as $S^{(m)} - \tilde{R}^{(m)} = \mathcal{R}_1^{(m)} + \mathcal{R}_2^{(m)}$ where, using that,$\forall K \in \mathcal{M}$, $S_K^0 = 0$:

$$\mathcal{R}_1^{(m)} = \sum_{n=0}^{N-1} \delta t \big[\sum_{K\in\mathcal{M}} S_K^{n+1} \varphi_K^{n+1} - \sum_{\sigma\in\mathcal{E}} R_\sigma^{n+1} \varphi_\sigma^{n+1} \big],$$

$$\mathcal{R}_2^{(m)} = \sum_{n=1}^{N-1} \delta t \sum_{K\in\mathcal{M}} S_K^n \; (\varphi_K^{n+1} - \varphi_K^n).$$

First, we prove that $\lim_{m\to+\infty} \mathcal{R}_1^{(m)} = 0$. Gathering and reordering the sums, we obtain $\mathcal{R}_1^{(m)} = \mathcal{R}_{1,1}^{(m)} + \mathcal{R}_{1,2}^{(m)} + \mathcal{R}_{1,3}^{(m)}$ with

$$\begin{aligned}
\mathcal{R}_{1,1}^{(m)} &= \frac{1}{2}\sum_{n=0}^{N-1}\delta t \sum_{\sigma=K|L\in\mathcal{E}} \Big[\frac{|D_{K,\sigma}|}{\delta t}\rho_K^{n+1}(u_\sigma^{n+1}-u_\sigma^n)^2(\varphi_K^{n+1}-\varphi_\sigma^{n+1}) \\
&\qquad\qquad + \frac{|D_{L,\sigma}|}{\delta t}\rho_L^{n+1}(u_\sigma^{n+1}-u_\sigma^n)^2(\varphi_L^{n+1}-\varphi_\sigma^{n+1})\Big], \\
\mathcal{R}_{1,2}^{(m)} &= \frac{1}{2}\sum_{n=0}^{N-1}\delta t \sum_{K=[\sigma'\to\sigma]\in\mathcal{M}} |F_K^n|\ (u_\sigma^n - u_{\sigma'}^n)^2\ (\varphi_K^{n+1}-\varphi_\sigma^{n+1}), \\
\mathcal{R}_{1,3}^{(m)} &= \sum_{n=0}^{N-1}\delta t \sum_{K=[\sigma'\to\sigma]\in\mathcal{M}} |F_K^n|\ (u_{\sigma'}^n - u_\sigma^n)\ (u_\sigma^{n+1}-u_\sigma^n)\ (\varphi_K^{n+1}-\varphi_\sigma^{n+1}).
\end{aligned}$$

We thus obtain:

$$|\mathcal{R}_{1,1}^{(m)}| \le h^{(m)}\ C_\varphi\ \|\rho^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|u^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|u^{(m)}\|_{\mathcal{T},t,\mathrm{BV}},$$

and

$$|\mathcal{R}_{1,2}^{(m)}| + |\mathcal{R}_{1,3}^{(m)}| \le h^{(m)}\ C_\varphi\ \|\rho^{(m)}\|_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|u^{(m)}\|^2_{\mathrm{L}^\infty(\Omega\times(0,T))}\ \|u^{(m)}\|_{\mathcal{T},x,\mathrm{BV}},$$

so all these terms tend to zero. The fact that $|\mathcal{R}_2^{(m)}|$ behaves as $\delta t^{(m)}$ may be proven by similar arguments.

Gathering the limits of all terms concludes the proof.

□

Remark 4.3 (On BV-stability assumptions)**.**
The proof of theorems 3.5 and 4.2 shows that the scheme is consistent under a BV-stability assumption that is much weaker than (52)-(53), namely:

$$\lim_{m\to+\infty}\ (h^{(m)}+\delta t^{(m)})\left[\|\rho^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} + \|p^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} + \|e^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} + \|u^{(m)}\|_{\mathcal{T},x,\mathrm{BV}} + \|u^{(m)}\|_{\mathcal{T},t,\mathrm{BV}}\right] = 0.$$

Remark 4.4 (Convergence to the entropy weak solution)**.** An entropy function for the incompressible Euler equations with the perfect gas EOS may be defined as:

$$\left|\begin{array}{ll} \eta: & \mathbb{R}_+^*\times\mathbb{R}_+^*\to\mathbb{R} \\ & (\rho,e)\mapsto \eta(\rho,e) = \dfrac{1}{\rho}\left(\phi(\rho)+\rho\psi(e)\right) \end{array}\right.$$

where for $s>0$, $\phi(s) = s\ln(s)$ and $\psi(s) = \dfrac{1}{1-\gamma}\ln(s)$. Under the assumptions of Theorem 4.2, assuming the additional time BV estimates on the approximate densities and internal energies:

$$\|\rho^{(m)}\|_{\mathcal{T},t,\mathrm{BV}} \le C,\ \|e^{(m)}\|_{\mathcal{T},t,\mathrm{BV}} \le C, \quad \forall m\in\mathbb{N},$$

and provided that the following stronger CFL condition holds:

$$\lim_{m\to+\infty} \frac{\delta t^{(m)}}{\min\limits_{K\in\mathcal{M}^{(m)}} h_K} = 0,$$

it can be shown that the limit of approximate solutions (up to a subsequence) is an entropy weak solution, in the sense that it also satisfies a weak entropy inequality, which reads

$$\forall\varphi\in C_c^\infty\big(\Omega\times[0,T),\mathbb{R}_+\big),\ -\int_0^T\int_\Omega\Big[\rho\,\eta(\rho,e)\,\partial_t\varphi + \rho\,u\,\eta(\rho,e)\,\partial_x\varphi\Big]\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \eta(\rho_0(x),e_0(x))\,\varphi(x,0)\,\mathrm{d}x \le 0.$$

The proof of this result may be found in [22, Chapter 4] (see also [16]), where the general multi-dimensional case and higher order schemes are also studied.

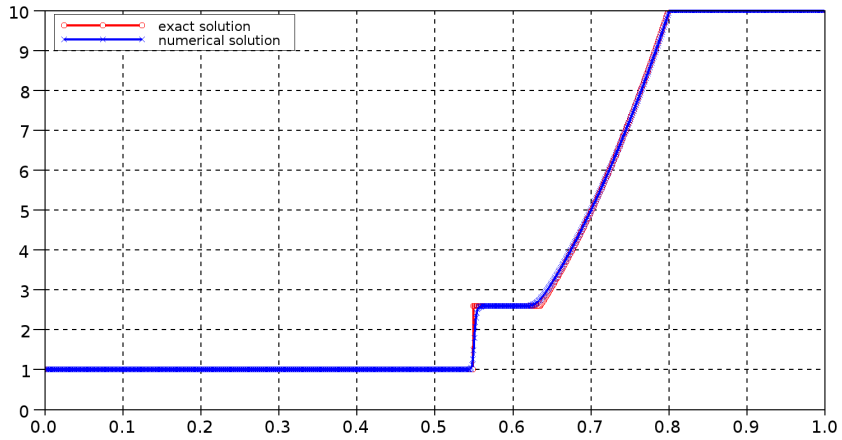

FIGURE 2. Shallow water eq., first Riemann problem – $h = 0.001$, $\delta t = h/12$ – Density at $t = 0.025$.

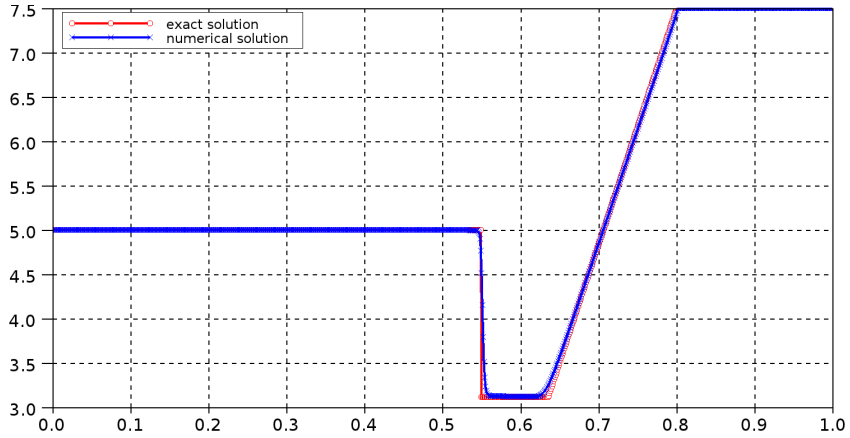

FIGURE 3. Shallow water eq., first Riemann problem – $h = 0.001$, $\delta t = h/12$ –Velocity at $t = 0.025$.

5. Numerical results

5.1. **The isentropic Euler equations**

We assess in this section the behaviour of the scheme on various test cases. For all these tests, we chose $p = \rho^2$ for the equation of state, so the solved system turns out to be the so-called shallow water equations, and we solve Riemann problems, *i.e.* 1D problems the initial conditions of which consist in two constant states separated by a discontinuity.

5.1.1. *A first Riemann problem*

In this first test, the chosen left and right states are given by:

$$\text{left state: } \begin{bmatrix} \rho_L = 1 \\ u_L = 5 \end{bmatrix}; \qquad \text{right state: } \begin{bmatrix} \rho_R = 10 \\ u_R = 7.5 \end{bmatrix}.$$

The computational domain is $\Omega = (0, 1)$ and the final time is $T = 0.025$. The (known) analytical solution of this problem consists, from the left to the right, in a shock wave and a rarefaction wave, both traveling to the right, separated by constant states.

Results - The density and velocity obtained at $t = 0.025 = T$ are shown on Figures 2 and 3 respectively; these results have been obtained with $h = 0.001$ and $\delta t = h/12$ (the maximum velocity and sound speed computed from the analytical solution being $u_{\max} = 7.5$ and $c_{\max} \simeq 4.5$, respectively). In addition, we performed a convergence study, successively dividing by two the space and time steps (so keeping the CFL number constant). The difference between the computed and analytical solution at $t = 0.025$, measured in discrete $\mathrm{L}^1(\Omega)$ norm, are reported in the following table:

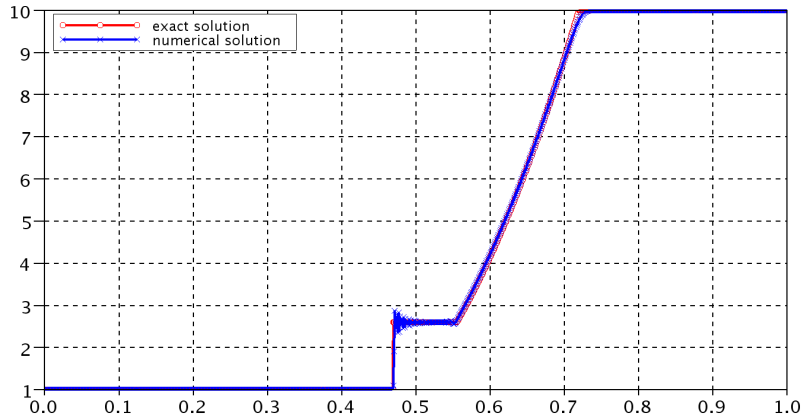

FIGURE 4. Shallow water eq., first Riemann problem modified to obtain a nearly vanishing velocity at the intermediate state – Viscosity= 0 – $h = 0.001$, $\delta t = h/12$ – Density at $t = 0.025$.

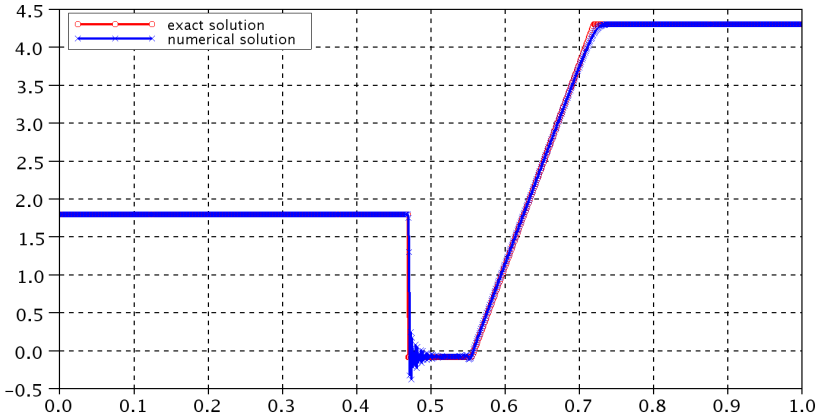

FIGURE 5. Shallow water eq., first Riemann problem modified to obtain a nearly vanishing velocity at the intermediate state – Viscosity= 0 – $h = 0.001$, $\delta t = h/12$ – Velocity at $t = 0.025$.

space step	$h_0 = 1/250$	$h_0/2$	$h_0/4$	$h_0/8$	$h_0/16$
$\|\rho - \bar{\rho}\|_{\mathrm{L}^1(\Omega)}$	0.0449	0.0256	0.0135	0.00775	0.00429
$\|u - \bar{u}\|_{\mathrm{L}^1(\Omega)}$	0.0411	0.0233	0.0119	0.00696	0.00384

We observe an approximatively first-order convergence rate.

A problem with a vanishing velocity in the intermediate state - To complete the study, we perform a computation of a Riemann problem obtained from the former one by subtracting a constant real number to the left and right velocity, in such a way that the velocity on the intermediate state approximatively vanishes. In this case, we observe spurious oscillations on the solution (see Figures 4 and 5), probably due to the fact that the numerical diffusion in the scheme vanishes. However, adding an artificial viscosity term in the discrete momentum balance equation, with a viscosity equal to $0.5\,\rho\,h$ (so equal to the upwind viscosity which would be associated to a velocity equal to 1) completely cures the problem (see Figures 6 and 7). This observation strongly supports the idea to build a higher order scheme using an *a posteriori* fitted viscosity technique, as in the so-called entropy viscosity method [9, 10]; this work is underway.

When we subtract once again a constant to the velocity at both left and right state, and so the velocity at the intermediate becomes negative, we recover a wiggle-free solution without adding any viscosity (Figures 8 and 9).

On a naive scheme - We also test the "naive" explicit scheme obtained by evaluating all the terms, except of course the time-derivative one, at time t^n. In the one dimensional setting and with the same notations as in

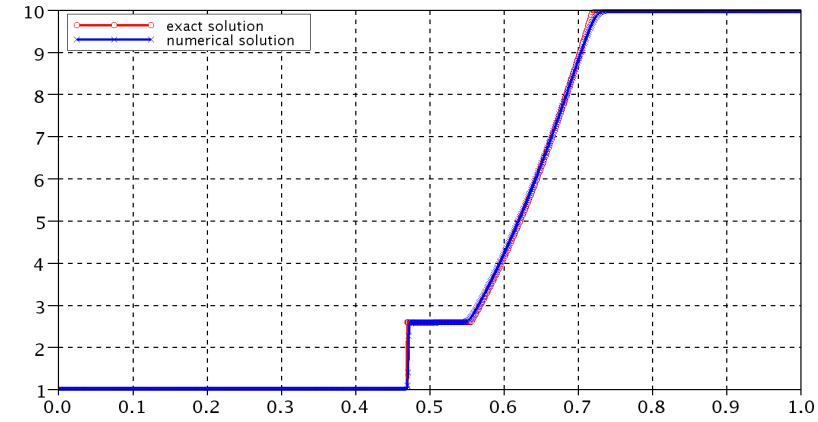

FIGURE 6. Shallow water eq., first Riemann problem modified to obtain a nearly vanishing velocity at the intermediate state – Viscosity= $0.5\,\rho\,h$ – $h = 0.001$, $\delta t = h/12$ – Density at $t = 0.025$.

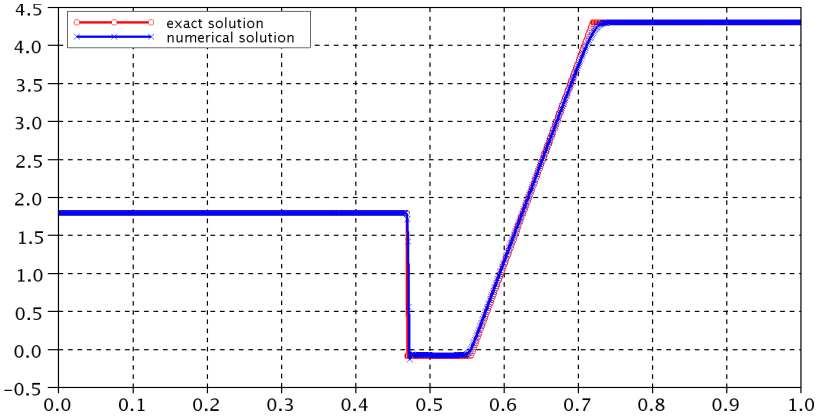

FIGURE 7. Shallow water eq., first Riemann problem modified to obtain a nearly vanishing velocity at the intermediate state – Viscosity= $0.5\,\rho\,h$ – $h = 0.001$, $\delta t = h/12$ – Velocity at $t = 0.025$.

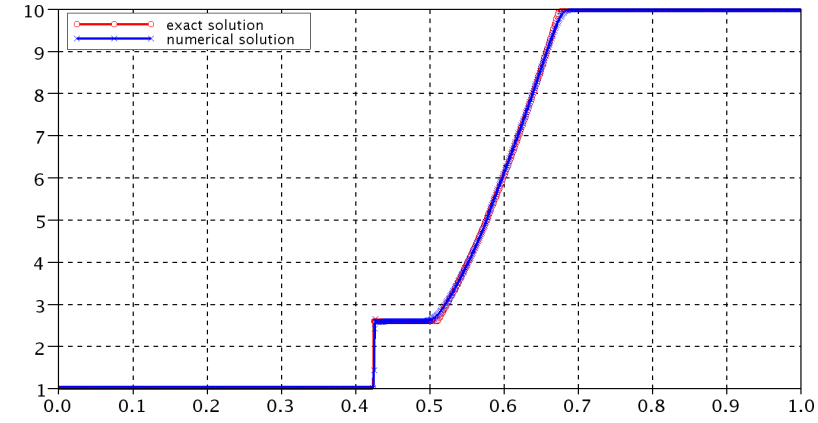

FIGURE 8. Shallow water eq., first Riemann problem modified to obtain a negative velocity at the intermediate state – $h = 0.001$, $\delta t = h/12$ – Density at $t = 0.025$.

Section 3.3, this scheme thus reads:

$$\forall K = [\overrightarrow{\sigma\sigma'}] \in \mathcal{M}, \qquad \frac{|K|}{\delta t}(\rho_K^{n+1} - \rho_K^n) + F_{\sigma'}^n - F_\sigma^n = 0, \tag{56a}$$

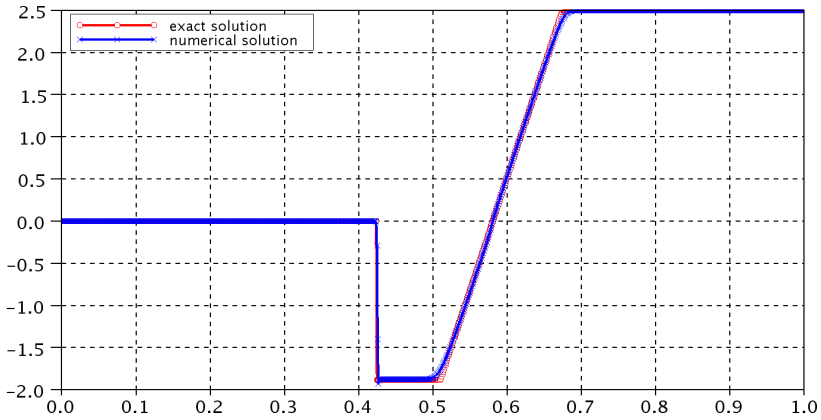

FIGURE 9. Shallow water eq., first Riemann problem modified to obtain a negative velocity at the intermediate state – $h = 0.001$, $\delta t = h/12$ – Velocity at $t = 0.025$.

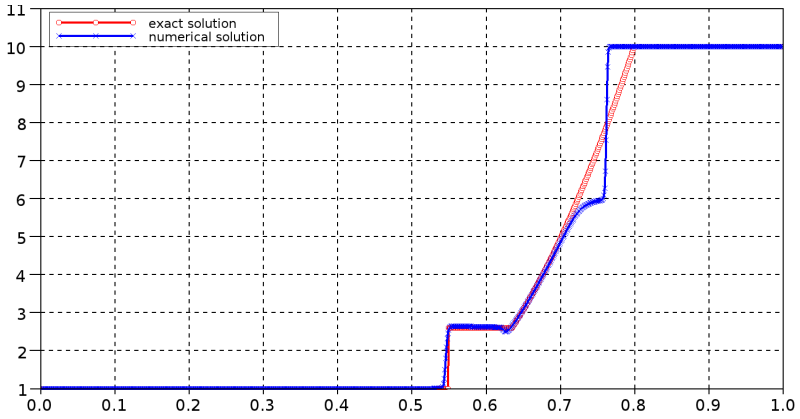

FIGURE 10. Shallow water eq., first Riemann problem – $\rho \rightsquigarrow u \rightsquigarrow p$ scheme – $h = 0.001$, $\delta t = h/12$ – Density at $t = 0.025$.

$$\forall \sigma = \overrightarrow{K|L} \in \mathcal{E}_{\rm int}, \qquad \frac{|D_\sigma|}{\delta t}(\rho_{D_\sigma}^{n+1} u_\sigma^{n+1} - \rho_{D_\sigma}^n u_\sigma^n) + F_L^n u_L^n - F_K^n u_K^n + p_L^n - p_K^n = 0, \tag{56b}$$

$$\forall K \in \mathcal{M}, \qquad p_K^{n+1} = \wp(\rho_K^{n+1}) = (\rho_K^{n+1})^\gamma. \tag{56c}$$

Hereafter and on the figure captions, this scheme is referred to as the "$\rho \rightsquigarrow u \rightsquigarrow p$ scheme" (since the pressure is updated after the computation of the velocity rather than after the computation of the density).

The computed density and velocity at time $T = 0.025$ are plotted on figures 10 and 11 respectively. From these results, it appears clearly that the $\rho \rightsquigarrow u \rightsquigarrow p$ scheme generates discontinuities in the rarefaction wave, and further experiments show that this phenomenon is not cured by a decrease of the time and space steps; this seems to be connected to the fact that, for this variant, we cannot prove that the limits of converging sequences satisfy the entropy condition (in fact, they probably do not). When trying to do so, in our proof and from a purely technical point of view, the trouble comes from the fact that the pressure gradient term which appears in the kinetic energy balance reads $\boldsymbol{u}^{n+1}\boldsymbol{\nabla} p^n$ and it seems difficult to make the counterpart (*i.e.* $p^n \mathrm{div}(\boldsymbol{u}^{n+1})$) appear, with the corresponding time levels, in the elastic potential balance, starting from a mass balance with a convection term written with $\boldsymbol{u}^n$; hence a dicretization of the momentum balance equation with an updated pressure gradient term $\boldsymbol{\nabla} p^{n+1}$, and thus the inversion of steps in the algorithm, to get the actual scheme proposed in this paper.

5.1.2. *Problems involving vacuum zones in the flow*

The objective of the two tests presented in this section is to check that the time step does not have to be drastically reduced in the presence of vacuum. Both are Riemann problems, posed on $\Omega = (0, 1)$.

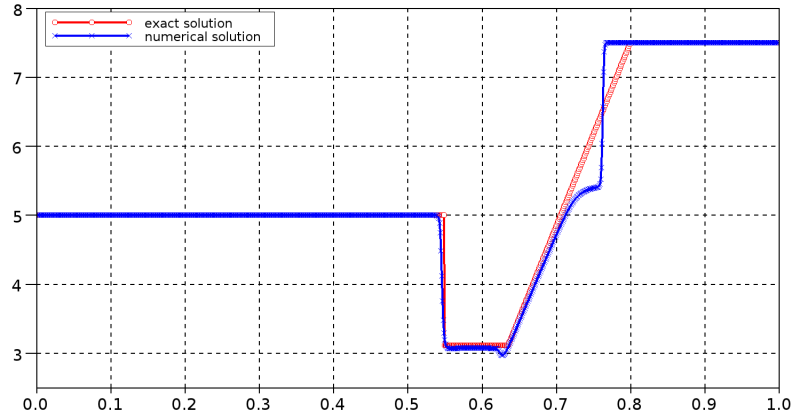

FIGURE 11. Shallow water eq., first Riemann problem – $\rho \rightsquigarrow u \rightsquigarrow p$ scheme – $h = 0.001$, $\delta t = h/12$ – Velocity at $t = 0.025$.

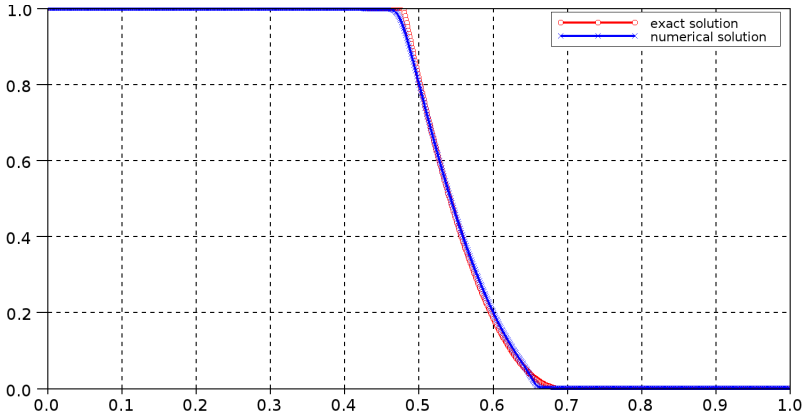

FIGURE 12. Shallow water eq., Riemann problem with vacuum at the right state – $h = 0.001$, $\delta t = h/8$ – Density at $t = 0.05$.

We first begin with a case where the vacuum is initially present, at the right initial state:

$$\text{left state: } \begin{bmatrix} \rho_L = 1 \\ u_L = 1 \end{bmatrix}; \qquad \text{right state: } \begin{bmatrix} \rho_R = 0 \\ u_R = 0 \end{bmatrix}.$$

In the computer code, ρ_R is fixed as $\rho_R = 10^{-20}$, to prevent divisions by zero due to imprudent programming. The results obtained at $t = 0.05$ are plotted on Figure 12 (density) and Figure 13 (velocity); they have been obtained with $h = 0.001$ and a constant time step equal to $\delta t = h/8$, which seems to be near to the stability limit (the maximum velocity and sound speed computed from the analytical solution being given by $u_{\max} \simeq 3.8$ and $c_{\max} \simeq 1.4$, respectively). We observe that the accuracy of the velocity computation is rather poor near to the vacuum front; we however check on Figure 14 that the scheme converges to the right solution. Moreover, Figure 15 shows that the quantity $\rho\, u$ (which is, in this case, the quantity of physical interest) is in fact obtained with a reasonable accuracy with the coarsest meshes of this study.

We now turn to a case where the chosen left and right states are given by:

$$\text{left state: } \begin{bmatrix} \rho_L = 1 \\ u_L = -8 \end{bmatrix}; \qquad \text{right state: } \begin{bmatrix} \rho_R = 1 \\ u_R = 8 \end{bmatrix}.$$

In this case, the solution consists in an intermediate state corresponding to vacuum connected to the left and right initial states by rarefaction waves. The computed density and velocity at $t = 0.03$, with $h = 0.001$ and $\delta t = h/12$ (while, in the analytical solution, $u_{\max} = 8$ and $c_{\max} \simeq 1.4$), are plotted on Figures 16 and 17 respectively. Once again, the behaviour of the scheme is satisfactory.

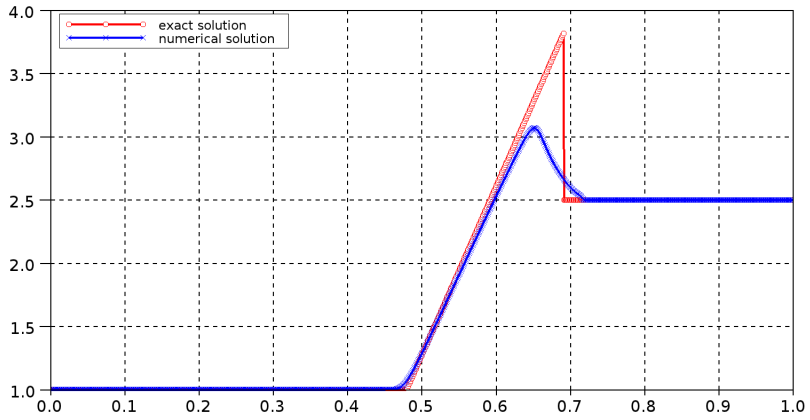

FIGURE 13. Shallow water eq., Riemann problem with vacuum at the right state – $h = 0.001$, $\delta t = h/8$ – Velocity at $t = 0.05$.

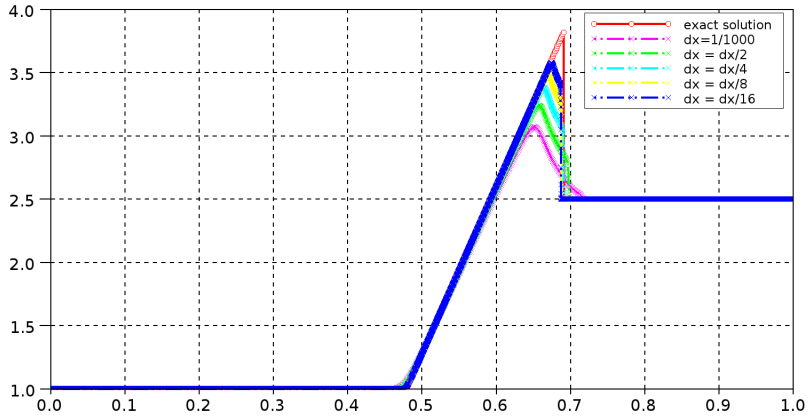

FIGURE 14. Shallow water eq., Riemann problem with vacuum at the right state – $h = h_0 = 0.001$ to $h = h_0/16$, $\delta t = h/8$ – Velocity at $t = 0.05$.

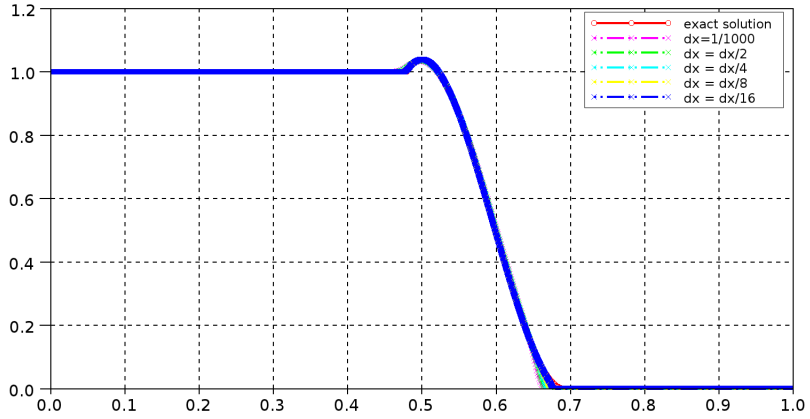

FIGURE 15. Shallow water eq., Riemann problem with vacuum at the right state – $h = h_0 = 0.001$ to $h = h_0/16$, $\delta t = h/8$ – Mass flow rate at $t = 0.05$.

5.2. **The full Euler equations**

5.2.1. *Riemann Problems*

We first assess in this section the behaviour of the scheme on a Riemann problem referred to as Test 3 in [23, Chapter 4], which is stiff enough to evidence consistency and stability properties of the scheme. The left

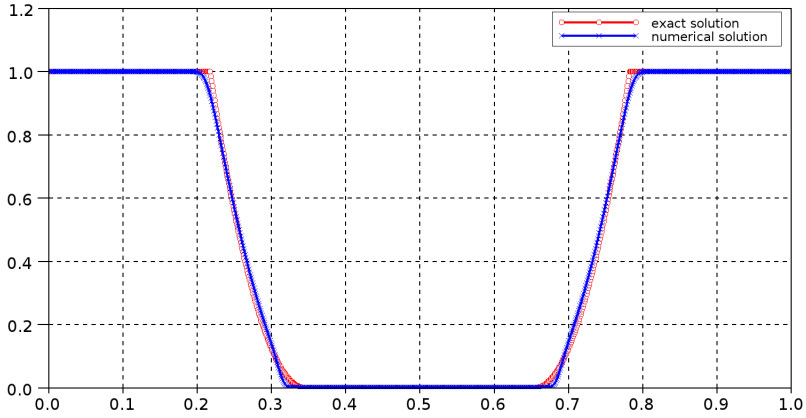

FIGURE 16. Shallow water eq., Riemann problem with vacuum appearance – $h = 0.001$, $\delta t = h/12$ – Density at $t = 0.03$.

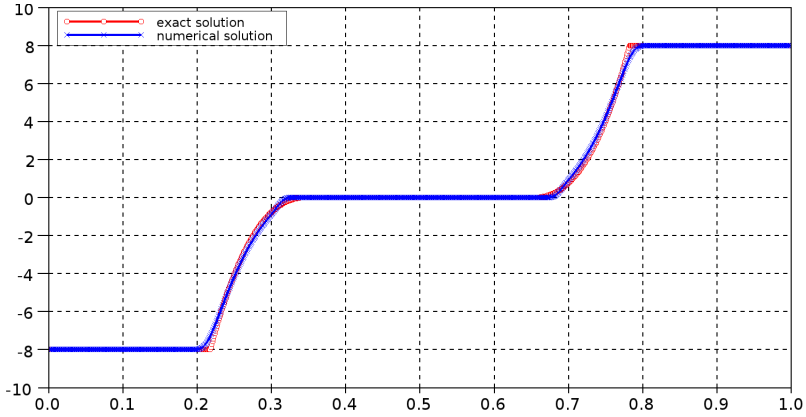

FIGURE 17. Shallow water eq., Riemann problem with vacuum appearance – $h = 0.001$, $\delta t = h/12$ – Mass flow rate at $t = 0.03$.

and right states are given by:

$$\text{left state: } \begin{bmatrix} \rho_L = 1 \\ u_L = 0 \\ p_L = 1000 \end{bmatrix}; \qquad \text{right state: } \begin{bmatrix} \rho_R = 1 \\ u_R = 0 \\ p_R = 0.001 \end{bmatrix}.$$

The computational domain is $\Omega = (0, 1)$ and the final time is $T = 0.012$. The (known) analytical solution of this type of problem consists in two genuinely nonlinear waves (*i.e.* rarefaction or shock waves) separated by a contact discontinuity. For the initial data chosen in this section, the left wave is a rarefaction wave, traveling to the left, and the right wave is a shock wave, traveling to the right.

The density, pressure, internal energy and velocity obtained at $t = 0.012 = T$ with $h = 0.001$ and $\delta t = h/100$ (as the maximal celerity of waves is close to 60) are shown on Figures 18, 19, 20 and 21 respectively. We observe that the scheme is rather diffusive especially for contact discontinuities for which the beneficial compressive effect of the shocks does not apply. More accurate variants may certainly be derived, using for instance MUSCL-like techniques; this work is underway.

We also observe that the scheme keeps the velocity and pressure constant through the contact discontinuity; this may be checked directly from the expression of the discrete balance equations (precisely speaking, one may prove that, if p^n and u^n are constant, so are p^{n+1} and u^{n+1}).

In addition, we perform a convergence study, successively dividing by two the space and time steps (so keeping the CFL number constant). The differences between the computed and analytical solution at $t = 0.025$, measured in discrete $\mathrm{L}^1(\Omega)$ norm, are reported in the following table.

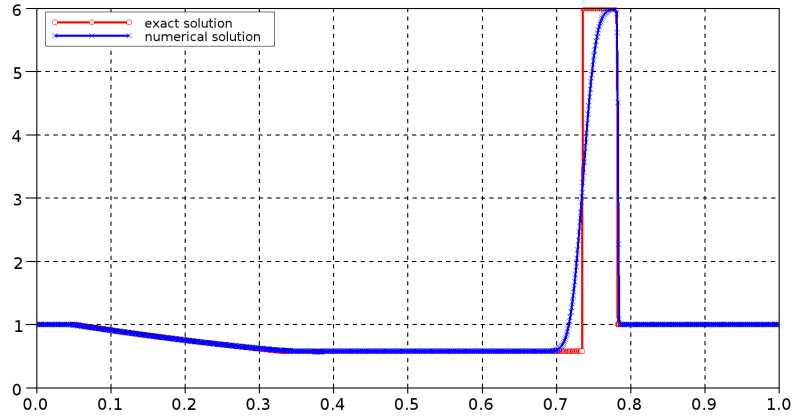

FIGURE 18. Euler equations, Riemann problem 3 of [23, Chapter 4] – $h = 0.001$ and $\delta t = h/100$ – Density at $t = 0.012$.

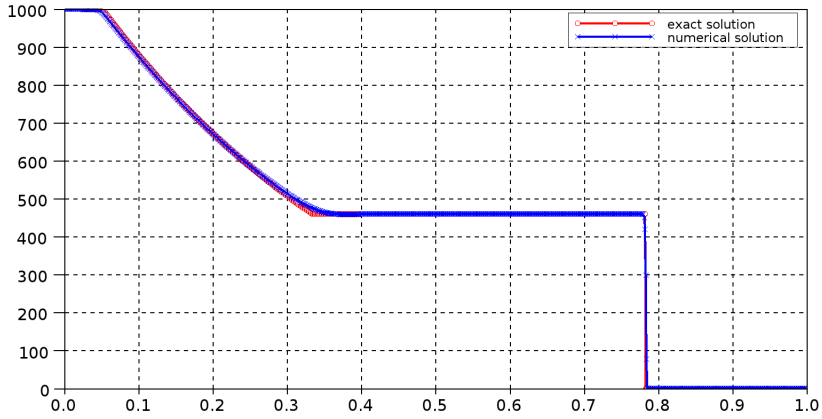

FIGURE 19. Euler equations, Riemann problem 3 of [23, Chapter 4] – $h = 0.001$ and $\delta t = h/100$ – Pressure at $t = 0.012$.

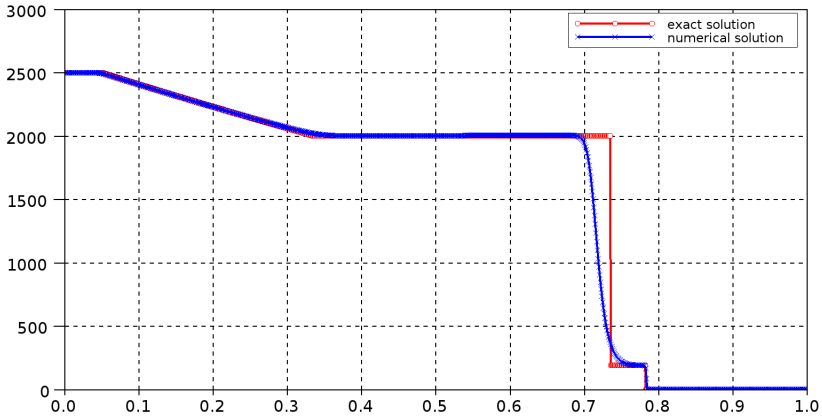

FIGURE 20. Euler equations, Riemann problem 3 of [23, Chapter 4] – $h = 0.001$ and $\delta t = h/100$ – Internal energy at $t = 0.012$.

space step	$h_0 = 0.001$	$h_0/2$	$h_0/4$	$h_0/8$	$h_0/16$
$\|\rho - \bar{\rho}\|_{\mathrm{L}^1(\Omega)}$	0.0651	0.0455	0.0310	0.0217	0.0153
$\|p - \bar{p}\|_{\mathrm{L}^1(\Omega)}$	1.87	1.05	0.530	0.284	0.164
$\|u - \bar{u}\|_{\mathrm{L}^1(\Omega)}$	0.0967	0.0536	0.0258	0.0134	0.00795

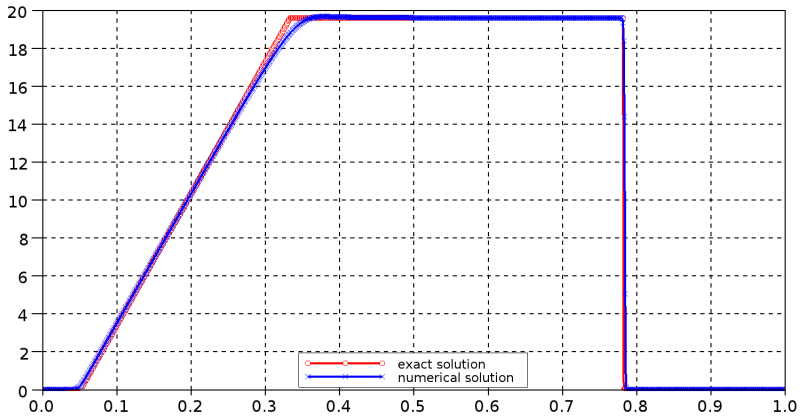

FIGURE 21. Euler equations, Riemann problem 3 of [23, Chapter 4] – $h = 0.001$ and $\delta t = h/100$ – Velocity at $t = 0.012$.

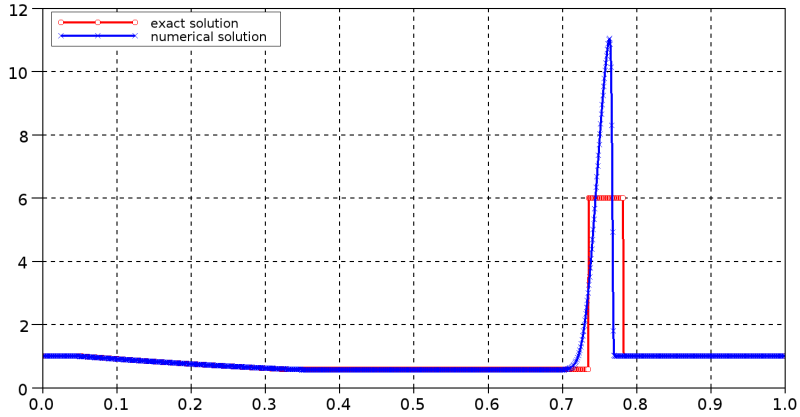

FIGURE 22. Euler equations, Riemann problem 3 of [23, Chapter 4] – Scheme without corrective terms – $h = 0.001$ and $\delta t = h/100$ – Density at $t = 0.012$.

We measure a convergence rate which is slightly lower to 1 for the variables which are constant through the contact discontinuity (*i.e.* p and u), and equal to 1/2 for ρ.

Finally, we test the behaviour of the scheme obtained when setting to zero the corrective terms in the internal energy balance. The density obtained with $h = 0.001$ and $\delta t = h/100$ is reported on Figure 22. From this result and from further numerical experiments with more and more refined meshes, it seems that the scheme converge, but to a limit which is not a weak solution to the Euler system: indeed, the Rankine-Hugoniot condition applied to the total energy balance, with the states obtained numerically, yields a right shock velocity slightly greater than the analytical solution one, while the same shock velocity obtained numerically is clearly lower.

Importance of the order of the equations in the decoupling – We also test the "naive" explicit scheme obtained by evaluating all the terms, except in time-derivative one, at time t^n. In the one dimensional setting and with the same notations as in Section 4.2, this scheme thus reads:

$$\forall K = [\overrightarrow{\sigma\sigma'}] \in \mathcal{M}, \quad \frac{|K|}{\delta t}(\rho_K^{n+1} - \rho_K^n) + F_{\sigma'}^n - F_\sigma^n = 0, \tag{57a}$$

$$\forall \sigma = \overrightarrow{K|L} \in \mathcal{E}_{\text{int}},$$
$$\frac{|D_\sigma|}{\delta t}(\rho_{D_\sigma}^{n+1} u_\sigma^{n+1} - \rho_{D_\sigma}^n u_\sigma^n) + F_L^n u_L^n - F_K^n u_K^n + p_L^n - p_K^n = 0, \tag{57b}$$

$$\forall K = [\overrightarrow{\sigma\sigma'}] \in \mathcal{M},$$
$$\frac{|K|}{\delta t}(\rho_K^{n+1} e_K^{n+1} - \rho_K^n e_K^n) + F_{\sigma'}^n e_{\sigma'}^n - F_\sigma^n e_\sigma^n + p_K^n (u_{\sigma'}^{n+1} - u_\sigma^{n+1}) = S_K^{n+1}, \tag{57c}$$

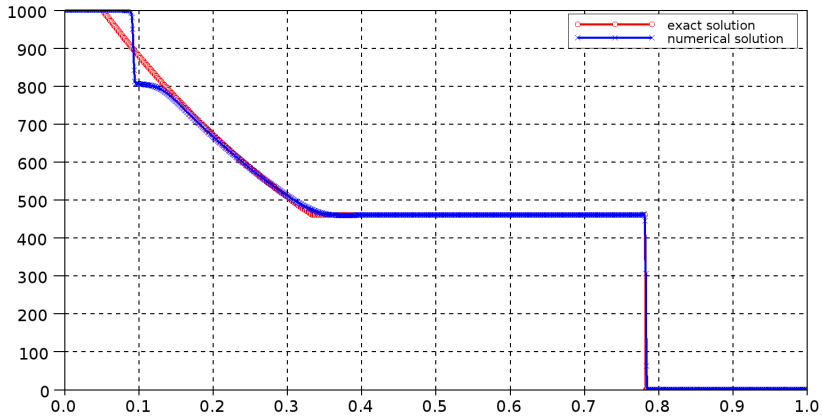

FIGURE 23. Euler equations, Riemann problem 3 of [23, Chapter 4] – $\rho \rightsquigarrow u \rightsquigarrow e \rightsquigarrow p$ scheme – $h = 0.001$ and $\delta t = h/100$ – Pressure at $t = 0.012$.

$$\forall K \in \mathcal{M}, \quad p_K^{n+1} = (\gamma - 1)\ \rho_K^{n+1}\ e_K^{n+1}. \tag{57d}$$

Hereafter and on the figure captions, this scheme is referred to by the $\rho \rightsquigarrow u \rightsquigarrow e \rightsquigarrow p$ scheme (according to the order of update of the unknowns). Note that we are able, for this scheme also, to prove a consistency result similar to Theorem 4.2.

The computed pressure at time $T = 0.012$ is plotted on figures 23. From this result, it appears clearly that, as in the isentropic case, the $\rho \rightsquigarrow u \rightsquigarrow e \rightsquigarrow p$ scheme generates discontinuities in the rarefaction wave, and further experiments show that this phenomenon is not cured by a reduction of the time and space step.

5.2.2. *A two-dimensional problem*

We now turn to a two-dimensional problem, consisting in the interaction of a shock wave with an obstacle. The initial data is $(\rho, \boldsymbol{u}, p) = (\rho_L, \boldsymbol{u}_L, p_L)$ (resp. $(\rho, \boldsymbol{u}, p) = (\rho_R, \boldsymbol{u}_R, p_R)$) for $x_1 \leq 0.7$ (resp. $x_1 > 0.7$), with:

$$\begin{bmatrix} \rho_L \\ \boldsymbol{u}_L \\ p_L \end{bmatrix} = \begin{bmatrix} 8 \\ 8.25\ (1,0)^t \\ 116.5 \end{bmatrix}, \qquad \begin{bmatrix} \rho_R \\ \boldsymbol{u}_R \\ p_R \end{bmatrix} = \begin{bmatrix} 1.4 \\ (0,0)^t \\ 1 \end{bmatrix},$$

Without any obstacle, this initial condition would yield a pure shock wave travelling to the left at the speed $v = 10$; since the speed of sound in the right state is $c = 1$, this wave is often referred to in the literature as a "Mach= 10 shock wave". The obstacle is the square $(1, 3) \times (-1, 1)$. Thanks to the symmetry with respect to the axis $x_2 = 0$, the chosen computational domain is $\Omega = (-1, 6) \times (0, 4)$. The final time is $t = 0.5$ (so that in the absence of an obstacle, the shock line would be defined by $x_1 = 5.7$ at the final time).

We present two computations. The first one is a uniform 1400×800 grid from which the cells corresponding to the interior of the obstacle have been removed, leading to a total number of cells close to 10^6; the time step is equal to 10^{-4}. For the second one, the mesh is built from a 10000×5700 grid, for a total number of cells close to $53\,10^6$, and a time step equal to 10^{-5}. In both cases, the MAC scheme is used for the space discretization, and the numerical viscosity is set to ρh. Both computations are performed in parallel (the CALIF^{3}S software uses PETSc primitives), with a multi-domain technique: the domain is split into subdomains (using the open-access software METIS), and each subdomain is treated by a processor. The second computation involves 120 Intel Xeon X5660 2.8GHz processors on an InfiniBand Linux cluster, for about 190 hours of restitution time (for 50000 time steps, so that each time step takes about 13.7 seconds; note that the software is designed for general meshes, and thus is by construction not optimized for structured grid). We observe here the beneficial effects of the simplicity of the convective flux construction; indeed, solving a Riemann problem at each interface would lead to a much more CPU time-consuming algorithm.

The obtained density at $t = 0.5$ is shown on Figure 24. One observes a strong reflection upstream the obstacle; behind this reflection, a shock-to-shock interaction occurs, which does not seem to generate an irregular reflection. As expected, the second computation shows much more details, especially in the wake of the obstacle. A closer view of this zone is provided on Figure 25

Finally, we also test the algorithm on unstructured meshings, using the RT discretization. The initial condition is the same, the obstacle is now a disk centered on $(2, 0)^t$ and of radius equal to 1, $\Omega = (0, 5) \times (0, 3.3)$

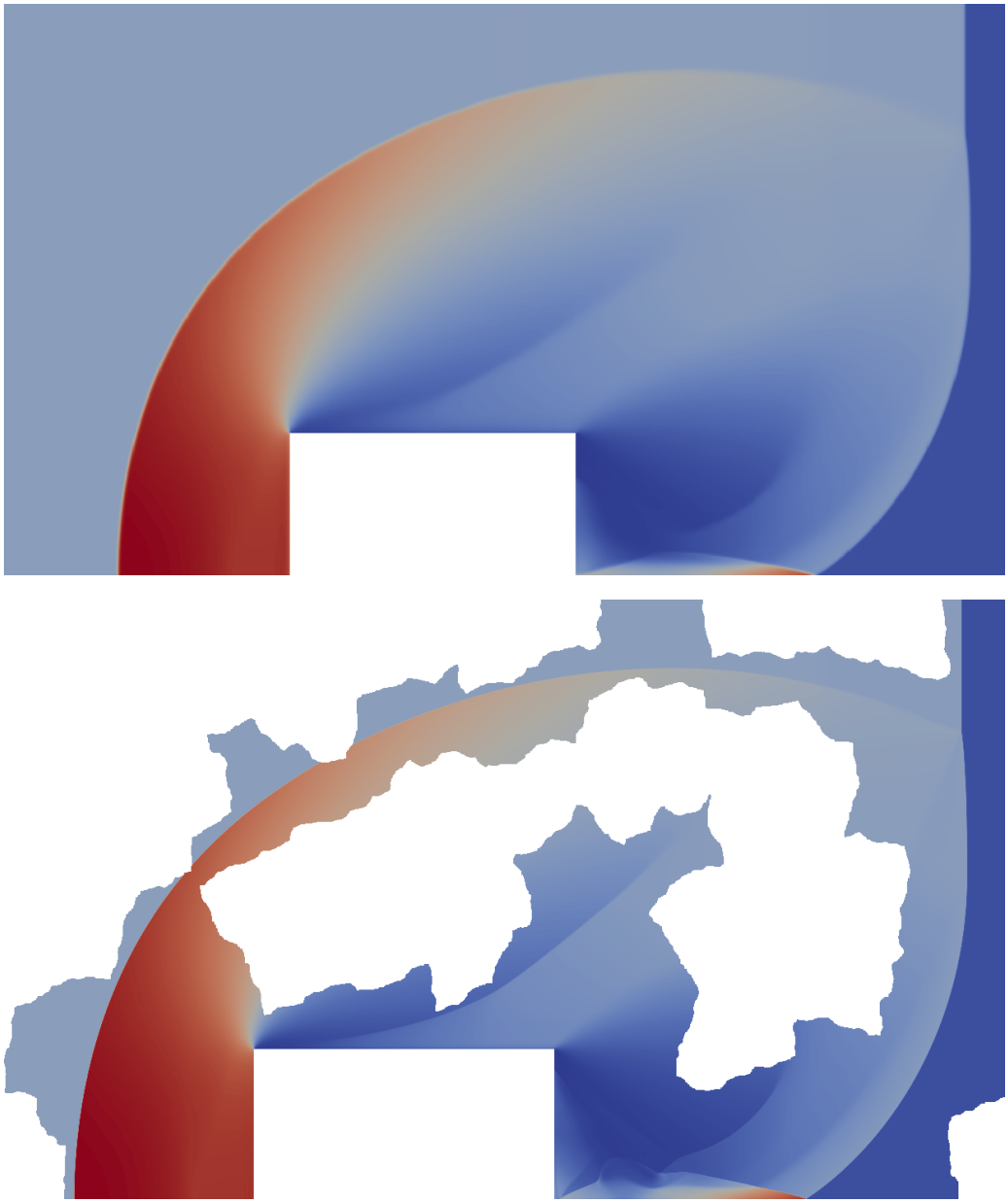

FIGURE 24. Euler equations, shock-square interaction – Density at $t = 0.5$ – Results obtained with $1\,10^6$ (top) and $53\,10^6$ cells (bottom). For this last computation, only a part of the subdomains (about one half of the 120 subdomains) are drawn.

and the final time is $t = 0.4$. The number of primal cells is close to $11\,10^6$ (so each velocity component is approximated on about $22\,10^6$ dual cells), and the time step is $5\,10^{-6}$. The numerical viscosity is set to $2\,\rho\,h$. The density fields obtained at time $t = 0.2$ and $t = 0.4$ are shown on Figure 26.

Finally, note that other 2D and 3D experiments have been conducted and may be found in the more numerical paper [8].

6. Conclusion

In this paper we presented a decoupled scheme based on staggered meshes for the isentropic and full Euler equations. This algorithm uses a very simple first-order upwinding strategy which consists, equation by equation, to implement an upwind discretization with respect to the material velocity of the convection term. The pressure gradient is defined as the transpose of the natural velocity divergence, and is thus centered. In the case of the full Euler equations, the scheme solves the internal energy balance instead of the total energy balance, to ensure the positivity of the internal energy by the above-mentionned upwinding technique; because of the staggered nature of the scheme, the total energy balance is only recovered at the limit of vanishing time and space steps, thanks to the addition of corrective source terms in the discrete internal energy balance. Under CFL-like conditions

FIGURE 25. Euler equations, shock-square interaction – Density at $t = 0.5$ – Results obtained with $53\,10^6$ cells, in the area where shock-to-shock interaction occurs, behind the obstacle.

which based on the material velocity only (by opposition to the celerity of waves which constrains classical hyperbolic schemes), this scheme preserves the positivity of the density, the pressure and, for Euler equations, of the internal energy (in other words, the scheme preserves the convex set of admissible states). Finally, the scheme has been shown to be consistent for 1D problems, in the sense that, if a sequence of numerical solutions obtained with more and more refined meshes (and, accordingly, smaller and smaller time steps) converges, then the limit is a weak solution to the continuous problem.

These theoretical results may be extended in two directions: first, in the full Euler case, the limits of convergent sequences may be shown to be entropy solutions; second, the scheme may be shown to be consistent in the multi-dimensional case. This is the object of a paper that will soon be submitted [16]. Another point of further investigation concerns the design of a discretization scheme that would be able to cope with non-conforming locally refined meshes. This work is now being undertaken.

Numerical studies show that the proposed algorithm is stable, even if the largest time step before blow-up is smaller than suggested by the above-mentioned CFL conditions. This behaviour was to be expected, since these CFL conditions only involve the velocity (and not the celerity of the acoustic waves): indeed, were they the only limitation, we would obtain an explicit scheme stable up to the incompressible limit. However, the mechanisms leading to the blow-up of the scheme (or, conversely, the way to fix the time step to ensure stability) remain to be clarified, even if one may anticipate from qualitative arguments (the scheme should allow a "transport of the information" at the same speed as the continuous problem) that the time step should be small enough to avoid that the waves cross more than one cell per time step. In addition, still as expected, the scheme is rather diffusive, especially at contact discontinuities; MUSCL-like extensions have recently been developed [22] to cure this problem, combined with a strategy similar to the so-called entropy-viscosity technique [9, 10] to damp spurious oscillations which are sometimes observed when the velocity is small.

Since the proposed scheme uses very simple numerical fluxes, it is well suited to large multi-dimensional parallel computing applications, and such studies are now starting at IRSN. Still for the same reasons (and, in particular, because the construction of the discretization does not require the solution of the Riemann problem), it seems that the presented approach offers natural extensions to more complex problems, such as reacting flows; this is under development at IRSN, for applications to explosion hazards.

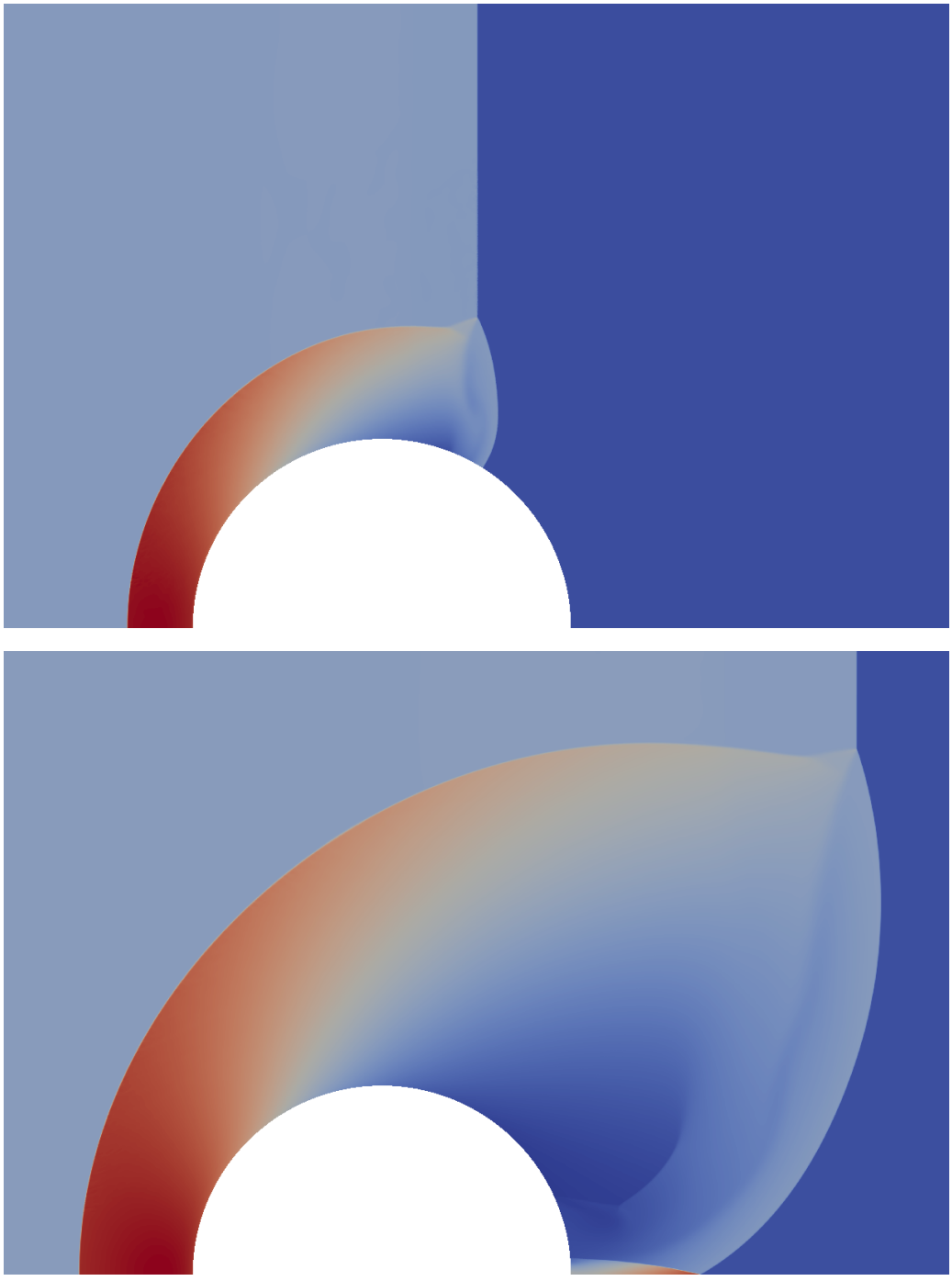

FIGURE 26. Euler equations, shock-disk interaction – Density at $t = 0.2$ (top) and $t = 0.4$ (bottom).

Appendix A. Some results concerning explicit finite volume convection operators

The convection operator appearing in the mass balance equation reads, in the continuous problem, $\rho \to \mathcal{C}(\rho) = \partial_t \rho + \mathrm{div}(\rho \boldsymbol{u})$, where $\boldsymbol{u}$ stands for a given velocity field, which is not assumed to satisfy any divergence constraint. We recall [13, Appendix A] that if ψ is a regular function from $(0,+\infty)$ to $\mathbb{R}$; then:

$$\psi'(\rho)\ \mathcal{C}(\rho) = \partial_t\big(\psi(\rho)\big) + \mathrm{div}\big(\psi(\rho)\boldsymbol{u}\big) + \big(\rho\psi'(\rho) - \psi(\rho)\big)\ \mathrm{div}\boldsymbol{u}. \tag{58}$$

This computation is of course completely formal and only valid for regular functions ρ and $\boldsymbol{u}$. The following lemma states a discrete analogue to (58) for the decoupled scheme studied in this paper (see. [13, Appendix A] for an implicit scheme).

Lemma A.1. *Let P be a polygonal (resp. polyhedral) bounded set of $\mathbb{R}^2$ (resp. $\mathbb{R}^3$), and let $\mathcal{E}(P)$ be the set of its edges (resp. faces). Let ψ be a twice continuously differentiable function defined over $(0,+\infty)$. Let $\rho_P^* > 0$, $\rho_P > 0$, $\delta t > 0$; consider three families $(\rho_\eta^*)_{\eta\in\mathcal{E}(P)} \subset \mathbb{R}_+ \setminus \{0\}, (V_\eta^*)_{\eta\in\mathcal{E}(P)} \subset \mathbb{R}$ and $(F_\eta^*)_{\eta\in\mathcal{E}(P)} \subset \mathbb{R}$ such that*

$$\forall \eta \in \mathcal{E}(P), \qquad F_\eta^* = \rho_\eta^*\ V_\eta^*.$$

Let $R_{P,\delta t}$ be defined by:

$$R_{P,\delta t} = \Big[\frac{|P|}{\delta t}\ (\rho_P - \rho_P^*) + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\Big]\ \psi'(\rho_P) \\ - \frac{|P|}{\delta t}\ [\psi(\rho_P) - \psi(\rho_P^*)] + \sum_{\eta\in\mathcal{E}(P)} \psi(\rho_\eta^*)V_\eta^* + [\rho_P^*\psi'(\rho_P^*) - \psi(\rho_P^*)] \sum_{\eta\in\mathcal{E}(P)} V_\eta^*.$$

Then this quantity may be expressed as follows:

$$R_{P,\delta t} = \frac{1}{2}\frac{|P|}{\delta t}\,(\rho_P - \rho_P^*)^2\,\psi''(\overline{\rho}_P^{(1)}) - \frac{1}{2}\sum_{\eta\in\mathcal{E}(P)} V_\eta^*\,(\rho_P^* - \rho_\eta^*)^2\,\psi''(\overline{\rho}_\eta^*) + \sum_{\eta\in\mathcal{E}(P)} V_\eta^*\rho_\eta^*\,(\rho_P - \rho_P^*)\,\psi''(\overline{\rho}_P^{(2)}),$$

where $\overline{\rho}_P^{(1)},\ \overline{\rho}_P^{(2)} \in [\![\rho_P, \rho_P^*]\!]$ and $\forall \eta \in \mathcal{E}(P)$, $\overline{\rho}_\eta^* \in [\![\rho_P^*, \rho_\eta^*]\!]$. We recall that, for $a,\ b \in \mathbb{R}$, we denote by $[\![a,b]\!]$ the interval $[\![a,b]\!] = \{\theta a + (1-\theta)b,\ \theta \in [0,1]\}$.

Proof. By the definition of F_η^*, we have:

$$\Big[\frac{|P|}{\delta t}(\rho_P - \rho_P^*) + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\Big]\ \psi'(\rho_P) = \frac{|P|}{\delta t}(\rho_P - \rho_P^*)\,\psi'(\rho_P) \\ + \sum_{\eta\in\mathcal{E}(P)} \rho_\eta^* V_\eta^* \psi'(\rho_P^*) + \sum_{\eta\in\mathcal{E}(P)} \rho_\eta^* V_\eta^* \big[\psi'(\rho_P) - \psi'(\rho_P^*)\big]. \tag{59}$$

By Taylor expansions of ψ, there exist two real numbers $\overline{\rho}_P^{(1)}$ and $\overline{\rho}_P^{(2)}\ \in [\![\rho_P^*, \rho_P]\!]$ and a family of real numbers $(\overline{\rho}_\eta^*)_{\eta\in\mathcal{E}(P)}$ satisfying, $\forall \eta \in \mathcal{E}(P)$, $\overline{\rho}_\eta^* \in [\![\rho_P^*, \rho_\eta^*]\!]$, and such that:

$$(\rho_P - \rho_P^*)\,\psi'(\rho_P) = \psi(\rho_P) - \psi(\rho_P^*) + \frac{1}{2}\,(\rho_P - \rho_P^*)^2\,\psi''(\overline{\rho}_P^{(1)}),$$
$$\rho_\eta^*\psi'(\rho_P^*) = \psi(\rho_\eta^*) + [\rho_P^*\psi'(\rho_P^*) - \psi(\rho_P^*)] - \frac{1}{2}\,(\rho_\eta^* - \rho_P^*)^2\,\psi''(\overline{\rho}_\eta^*),$$
$$\psi'(\rho_P) - \psi'(\rho_P^*) = (\rho_P - \rho_P^*)\,\psi''(\overline{\rho}_P^{(2)}).$$

Substituting in (59) yields the result we are seeking. □

We now turn to the convection operator appearing in the momentum balance equation, which reads, in the continuous setting, $z \to \mathcal{C}_\rho(z) = \partial_t(\rho z) + \mathrm{div}(\rho z \boldsymbol{u})$, where ρ (resp. $\boldsymbol{u}$) stands for a given scalar (resp. vector) field; we wish to obtain some property of C_ρ under the assumption that ρ and $\boldsymbol{u}$ satisfy the mass balance

equation, *i.e.* $\partial_t\rho + \mathrm{div}(\rho\boldsymbol{u}) = 0$. Formally, using twice the mass balance yields:

$$\psi'(z)\ \mathcal{C}_\rho(z) = \psi'(z)\big[\partial_t(\rho\, z) + \mathrm{div}(\rho\, z\, \boldsymbol{u})\big] = \psi'(z)\rho\big[\partial_t z + \boldsymbol{u}\cdot\boldsymbol{\nabla} z\big] \\ = \rho\big[\partial_t\psi(z) + \boldsymbol{u}\cdot\boldsymbol{\nabla}\psi(z)\big] = \partial_t\big(\rho\,\psi(z)\big) + \mathrm{div}\big(\rho\,\psi(z)\,\boldsymbol{u}\big).$$

Taking for z a component of the velocity field, this relation is the central argument used to derive the kinetic energy balance. The following lemma states a discrete counterpart of this identity, for a finite volume first-order explicit convection operator.

Lemma A.2. *Let P be a polygonal (resp. polyhedral) bounded set of $\mathbb{R}^2$ (resp. $\mathbb{R}^3$) and let $\mathcal{E}(P)$ be the set of its edges (resp. faces). Let $\rho_P^* > 0$, $\rho_P > 0$, $\delta t > 0$, and $(F_\eta^*)_{\eta\in\mathcal{E}(P)} \subset \mathbb{R}$ be such that*

$$\frac{|P|}{\delta t}\ (\rho_P - \rho_P^*) + \sum_{\eta\in\mathcal{E}(P)} F_\eta^* = 0. \tag{60}$$

Let ψ be a twice continuously differentiable function defined over $(0,+\infty)$. For $u_P^* \in \mathbb{R}$, $u_P \in \mathbb{R}$ and $(u_\eta^*)_{\eta\in\mathcal{E}(P)} \subset \mathbb{R}$ let us define:

$$R_{P,\delta t} = \Big[\frac{|P|}{\delta t}\ \big(\rho_P\, u_P - \rho_P^*\, u_P^*\big) + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ u_\eta^*\Big]\ \psi'(u_P) - \Big[\frac{|P|}{\delta t}\ \big[\rho_P\,\psi(u_P) - \rho_P^*\,\psi(u_P^*)\big] + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ \psi(u_\eta^*)\Big].$$

Then:

(*i*) *the remainder term $R_{P,\delta t}$ reads:*

$$R_{P,\delta t} = \frac{1}{2}\,\frac{|P|}{\delta t}\,\rho_P\,(u_P - u_P^*)^2\psi''(\overline{u}_P^{(1)}) - \frac{1}{2}\sum_{\eta\in\mathcal{E}(P)} F_\eta^*\,(u_\eta^* - u_P^*)^2\psi''(\overline{u}_\eta^*) \\ + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\,(u_\eta^* - u_P^*)\ (u_P - u_P^*)\ \psi''(\overline{u}_P^{(2)}) \tag{61}$$

with $\overline{u}_P^{(1)}, \overline{u}_P^{(2)} \in [\![u_P, u_P^*]\!]$, and $\forall\eta\in\mathcal{E}(P)$, $\overline{u}_\eta^* \in [\![u_P^*, u_\eta^*]\!]$.

(*ii*) *If we suppose that the function ψ is convex and that $u_\eta^* = u_P^*$ as soon as $F_\eta^* \geq 0$, then $R_{P,\delta t}$ is non-negative under the CFL condition:*

$$\delta t \leq \frac{|P|\,\rho_P\,\underline{\psi}_P''}{\sum_{\eta\in\mathcal{E}(P)}(F_\eta^*)^-\ (\overline{\psi}_P'')^2/\underline{\psi}_\eta''}, \tag{62}$$

where $\underline{\psi}_P'' = \min_{s\in[\![u_P,u_P^*]\!]}\psi''(s)$, $\overline{\psi}_P'' = \max_{s\in[\![u_P,u_P^*]\!]}\psi''(s)$ and $\underline{\psi}_\eta'' = \min_{s\in[\![u_P^*,u_\eta^*]\!]}\psi''(s)$.
For $\psi(s) = s^2/2$ (and therefore $\psi''(s) = 1,\ \forall s\in(0,+\infty)$), this CFL condition simply reads:

$$\delta t \leq \frac{|P|\,\rho_P}{\sum_{\eta\in\mathcal{E}(P)}(F_\eta^*)^-}. \tag{63}$$

Proof. Let T_P be defined by:

$$T_P = \Big[\frac{|P|}{\delta t}\ \big(\rho_P\, u_P - \rho_P^*\, u_P^*\big) + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ u_\eta^*\Big]\ \psi'(u_P).$$

Using equation (60) multiplied by u_P^*, we obtain:

$$T_P = \Big[\frac{|P|}{\delta t}\ \rho_P\ \big(u_P - u_P^*\big) + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ (u_\eta^* - u_P^*)\Big]\ \psi'(u_P).$$

We now define the remainder terms r_P and $(r_\eta^*)_{\eta\in\mathcal{E}(P)}$ by:

$$r_P = (u_P - u_P^*)\ \psi'(u_P) - \big[\psi(u_P) - \psi(u_P^*)\big], \quad r_\eta^* = (u_P^* - u_\eta^*)\ \psi'(u_P^*) - \big[\psi(u_P^*) - \psi(u_\eta^*)\big].$$

With these notations, we get:

$$T_P = \frac{|P|}{\delta t}\ \rho_P\ \big[\psi(u_P) - \psi(u_P^*)\big] + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\big[\psi(u_\eta^*) - \psi(u_P^*)\big] + \frac{|P|}{\delta t}\ \rho_P\ r_P - \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ r_\eta^* + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ (u_\eta^* - u_P^*)\ \big(\psi'(u_P) - \psi'(u_P^*)\big).$$

Using once again equation (60), this time multiplied by $\psi(u_P^*)$, we obtain:

$$T_P = \frac{|P|}{\delta t}\ \big[\rho_P\psi(u_P) - \rho_P^*\psi(u_P^*)\big] + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\psi(u_\eta^*) + \frac{|P|}{\delta t}\ \rho_P\ r_P - \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ r_\eta^* + \sum_{\eta\in\mathcal{E}(P)} F_\eta^*\ (u_\eta^* - u_P^*)\ \big(\psi'(u_P) - \psi'(u_P^*)\big).$$

The expression (61) of the remainder term $R_{P,\delta t}$ follow by remarking that, by a Taylor expansion, there exist $\overline{u}_P^{(1)}, \overline{u}_P^{(2)} \in [\![u_P, u_P^*]\!]$, and $\forall\eta\in\mathcal{E}(P)$, $\overline{u}_\eta^* \in [\![u_P^*, u_\eta^*]\!]$ such that:

$$r_P = \frac{1}{2}\ \psi''(\overline{u}_P^{(1)})\,(u_P - u_P^*)^2, \quad r_\eta^* = \frac{1}{2}\ \psi''(\overline{u}_\eta^*)\,(u_\eta^* - u_p^*)^2$$

and

$$\psi'(u_P) - \psi'(u_P^*) = \psi''(\overline{u}_P^{(2)})\,(u_P - u_P^*).$$

If ψ is convex, r_P is non-negative. If, in addition, $u_P^* - u_\eta^*$ vanishes $\forall\eta\in\mathcal{E}(P)$ when F_η^* is non-negative, $-r_\eta^*$ is non-negative. By Young's inequality, the last term in $R_{P,\delta t}$ may be bounded as follows:

$$\Big|\sum_{\eta\in\mathcal{E}(P)} (F_\eta^*)^-\ (u_\eta^* - u_P^*)\ (u_P - u_P^*)\ \psi''(\overline{u}_P^{(2)})\Big| \le \frac{\psi''(\overline{u}_P^{(2)})^2}{2}\Big[\sum_{\eta\in\mathcal{E}(P)} (F_\eta^*)^-\ \frac{1}{\psi''(\overline{u}_\eta^*)}\Big]\ (u_P - u_P^*)^2 + \frac{1}{2}\sum_{\eta\in\mathcal{E}(P)} (F_\eta^*)^-\ (u_\eta^* - u_P^*)^2\ \psi''(\overline{u}_\eta^*),$$

so this term may be absorbed in the first two ones under the CFL condition (62). □

References

[1] G. Ansanay-Alex, F. Babik, J.-C. Latché, and D. Vola. An L^2-stable approximation of the Navier-Stokes convection operator for low-order non-conforming finite elements. *International Journal for Numerical Methods in Fluids*, 66:555–580, 2011.

[2] F. Bouchut. *Nonlinear Stability of finite volume methods for hyperbolic conservation laws.* Birkhauser, 2004.

[3] CALIF^{3}S. A software components library for the computation of reactive turbulent flows. `https://gforge.irsn.fr/gf/project/isis`.

[4] P. G. Ciarlet. Basic error estimates for elliptic problems. In P. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Volume II*, pages 17–351. North Holland, 1991.

[5] M. Crouzeix and P.A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. *RAIRO Série Rouge*, 7:33–75, 1973.

[6] 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 2, pages 39–56, 2011.

[7] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws.* Springer, 1996.

[8] D. Grapsas, R. Herbin, W. Kheriji, and J.-C. Latché. An unconditionally stable staggered pressure correction scheme for the compressible Navier-Stokes equations. *SMAI-Journal of computational mathematics*, 2:51–97, 2016.

[9] J.L Guermond and R. Pasquetti. Entropy-based nonlinear viscosity for Fourier approximations of conservation laws. *Comptes Rendus de l'Académie des Sciences de Paris – Série I – Analyse Numérique*, 346:801–806, 2008.

[10] J.L Guermond, R. Pasquetti, and B. Popov. Entropy viscosity method for nonlinear conservation laws. *Journal of Computational Physics*, 230:4248–4267, 2011.

[11] F.H. Harlow and A.A. Amsden. A numerical fluid dynamics calculation method for all flow speeds. *Journal of Computational Physics*, 8:197–213, 1971.

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

[13] R. Herbin, W. Kheriji, and J.C. Latché. On some implicit and semi-implicit staggered schemes for the shallow water and Euler equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, eFirst, 4 2014.

[14] R. Herbin and J.-C. Latché. Kinetic energy control in the MAC discretization of the compressible Navier-Stokes equations. *International Journal of Finites Volumes*, 7, 2010.
[15] R. Herbin, J.-C. Latché, and T.T. Nguyen. Explicit staggered schemes for the compressible Euler equations. *ESAIM: Proceedings*, 40:83–102, 2013.
[16] R. Herbin, J.-C. Latché, and N. Therme. On a class of consistent staggered schemes for the compressible Euler equations. *submitted*, 2016.
[17] Raphaèle Herbin, Walid Kheriji, and Jean-Claude Latché. Staggered schemes for all speed flows. In *Congrès National de Mathématiques Appliquées et Industrielles*, volume 35 of *ESAIM Proc.*, pages 122–150. EDP Sci., Les Ulis, 2011.
[18] B. Larrouturou. How to preserve the mass fractions positivity when computing compressible multi-component flows. *Journal of Computational Physics*, 95:59–84, 1991.
[19] M.-S. Liou. A sequel to AUSM, part II: AUSM+-up. *Journal of Computational Physics*, 214:137–170, 2006.
[20] M.-S. Liou and C.J. Steffen. A new flux splitting scheme. *Journal of Computational Physics*, 107:23–39, 1993.
[21] R. Rannacher and S. Turek. Simple nonconforming quadrilateral Stokes element. *Numerical Methods for Partial Differential Equations*, 8:97–111, 1992.
[22] N. Therme. *Schémas numériques pour la simulation de l'explosion*. PhD thesis, Aix Marseille Univ, 2015.
[23] E. Toro. *Riemann solvers and numerical methods for fluid dynamics – A practical introduction (third edition)*. Springer, 2009.