



**HAL**  
open science

# Recasting an operator splitting solver into a standard finite volume flux-based algorithm. The case of a Lagrange-Projection-type method for gas dynamic

Remi Bourgeois, Pascal Tremblin, Samuel Kokh, Thomas Padioleau

## ► To cite this version:

Remi Bourgeois, Pascal Tremblin, Samuel Kokh, Thomas Padioleau. Recasting an operator splitting solver into a standard finite volume flux-based algorithm. The case of a Lagrange-Projection-type method for gas dynamic. *Journal of Computational Physics*, 2024, 496, pp.112594. 10.1016/j.jcp.2023.112594 . cea-04331620

**HAL Id: cea-04331620**

**<https://cea.hal.science/cea-04331620v1>**

Submitted on 8 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Recasting an operator splitting solver into a standard finite volume flux-based algorithm. The case of a Lagrange-Projection-type method for gas dynamic.

Rémi Bourgeois<sup>a,\*</sup>, Pascal Tremblin<sup>a</sup>, Samuel Kokh<sup>b</sup>, Thomas Padioleau<sup>a</sup>

<sup>a</sup>Université Paris-Saclay, UVSQ, CNRS, CEA, Maison de la Simulation, 91191, Gif-sur-Yvette, France

<sup>b</sup>Université Paris-Saclay, CEA, Service de Génie Logiciel pour la Simulation, 91191, Gif-sur-Yvette, France.

---

## Abstract

In this paper, we propose a modification of an acoustic-transport operator splitting Lagrange-projection method for simulating compressible flows with gravity. The original method involves two steps that respectively account for acoustic and transport effects. Our work proposes a simple modification of the transport step, and the resulting modified scheme turns out to be a flux-splitting method. This new numerical method is less computationally expensive, more memory efficient, and easier to implement than the original one. We prove stability properties for this new scheme by showing that under classical CFL conditions, the method is positivity preserving for mass, energy and entropy satisfying. The flexible flux-splitting structure of the method enables straightforward extensions of the method to multi-dimensional problems (with respect to space) and high-order discretizations that are presented in this work. We also propose an interpretation of the flux-splitting solver as a relaxation approximation. Both the stability and the accuracy of the new method are tested against one-dimensional and two-dimensional numerical experiments that involve highly compressible flows and low-Mach regimes.

---

## 1. Introduction

In this work, we consider the approximation of the compressible Euler equations in the presence of source terms derived from a smooth potential using a finite volume method. This paper aims to showcase the recasting of an Operator Splitting Lagrange-Projection (OSLP) finite volume algorithm into a corresponding flux-splitting method (FSLP). The flux-splitting method we consider here has several computational and implementation advantages compared to OSLP. It requires a smaller stencil, no intermediate state storage, and can be implemented as a fully explicit flux-based solver. The simplicity of the FSLP method allows us to combine effortlessly with standard means to derive higher-order methods such as MUSCL, ENO, WENO, and MOOD frameworks.

The OSLP algorithm we use as ground material for implementing an FSLP method is presented in [1]. It relies on a separate treatment of acoustic and transport effects, and it enjoys several interesting properties: it is stable under a CFL condition so that it ensures positivity for mass and internal energy and satisfies a discrete entropy inequality. The treatment of the source term in [1] allows us to preserve stationary solution profiles at the discrete level so that the OSLP scheme satisfies a well-balanced property (see *e.g.* [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 1, 21, 22, 23, 24]). Moreover, when the Mach number that characterizes the ratio of the material velocity to the sound velocity is low, cell-centered finite volume methods may suffer an important loss of accuracy [25, 26, 27, 28]. This question is connected to several delicate issues like the influence of the mesh geometry [29, 30], the numerical diffusion (see for example [31, 28, 32, 33, 34, 35]) or the Asymptotic Preserving property with respect to incompressible models [36, 37, 34, 38, 39, 40, 41, 42] and has been extensively investigated in the literature for the past years through several approaches (see also [43, 27, 44, 45, 46, 47, 48, 49]). Although it does not address the full spectrum of problems connected to the simulation of flows in the low Mach regime, a simple modification of the OSLP method ensures a uniform truncation error with respect to the Mach number [32, 1]. The resulting FSLP

---

\*Corresponding author

Email address: remi.bourgeois@cea.fr (Rémi Bourgeois)

algorithm we obtain performs equally concerning these aspects. Moreover, it profits from all the advantages of FSLP methods over OSLP mentioned above. It is also less computationally expensive, requiring fewer sweeps over the numerical solution to reach the same physical time. The derivation of the stability properties of the FSLP method requires novel mathematical developments that we present in this study.

The paper is organized as follows: we first introduce the set of equations with the thermodynamical-related hypotheses that support the stability properties of the model, and we present the stationary profiles and difficult regimes we will be interested in. Then, we will recall the OSLP method we aim to recast into its FSLP version. We will modify the transport step in the original OSLP method so that both steps are revamped into one that can be viewed as a flux-splitting step. We will then provide proof of stability for the FSLP method we obtained. We examine standard ways to extend the FSLP method to higher-order discretizations and multi-dimensional problems. Then we will see that the FSLP method can be connected to a new relaxation approximation of the Euler equations that proposes a single-step but separate treatment of the acoustic and transport effects. Finally, we will present one-dimensional and two-dimensional numerical experiments that demonstrate the good behavior of the scheme.

## 2. Flow model

For the sake of clarity but without loss of generality, we focus on one-dimensional problems. We consider the Euler equations supplemented with a smooth potential source term  $x \mapsto \phi(x)$ ,

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}, \phi), \quad \text{for } x \in \mathbb{R}, t > 0, \quad (1)$$

with  $\mathbf{U} = (\rho, \rho u, \rho E)^T$ ,  $\mathbf{F}(\mathbf{U}) = (\rho u, u\rho u + p, u\rho E + pu)^T$  and  $\mathbf{S}(\mathbf{U}, \phi) = -\rho \partial_x \phi(0, 1, u)^T$  where  $\phi$  is smooth enough so that we can consider that  $\partial_x \phi$  is also regular and bounded.

Although (1) is not strictly limited to flows accounting for gravitational forces, the stationary potential  $x \mapsto \phi(x)$  will be referred to as the gravitational potential. The fields  $\rho$ ,  $u$ ,  $p$ , and  $E$  respectively denote the density, velocity, pressure, and specific total energy of the fluid. If  $e = E - u^2/2$  is the specific internal energy, we define the set of admissible states

$$\Omega = \{(\rho, \rho u, \rho E) \in \mathbb{R}^3 \mid \rho > 0, e > 0\}. \quad (2)$$

Let  $s$  be the specific entropy of the fluid. We consider an Equation Of State (EOS) in the form of a mapping  $(1/\rho, s) \mapsto e^{\text{EOS}}(1/\rho, s)$  that satisfies the classic Weyl assumptions [50, 32]:

$$\frac{\partial e^{\text{EOS}}}{\partial(1/\rho)} < 0, \quad \frac{\partial e^{\text{EOS}}}{\partial s} > 0, \quad \frac{\partial^2 e^{\text{EOS}}}{\partial(1/\rho)^2} > 0, \quad (3a)$$

$$\frac{\partial^2 e^{\text{EOS}}}{\partial s^2} > 0, \quad \left[ \frac{\partial^2 e^{\text{EOS}}}{\partial(1/\rho)^2} \left[ \frac{\partial^2 e^{\text{EOS}}}{\partial s^2} \right] \right] > \left[ \frac{\partial^2 e^{\text{EOS}}}{\partial s \partial(1/\rho)} \right]^2, \quad \frac{\partial^3 e^{\text{EOS}}}{\partial(1/\rho)^3} < 0. \quad (3b)$$

The temperature  $T$  and the pressure  $p$  of the fluids are related to the other parameters, respectively by  $T = T^{\text{EOS}}(1/\rho, s) = \partial e^{\text{EOS}} / \partial s$  and  $p = p^{\text{EOS}}(1/\rho, s) = -\partial e^{\text{EOS}} / \partial(1/\rho)$ . It is possible to define a mapping  $(1/\rho, e) \mapsto s^{\text{EOS}}(1/\rho, e)$  such that  $e = e^{\text{EOS}}(1/\rho, s)$  if  $s = s^{\text{EOS}}(1/\rho, e)$  so that we have the Gibbs relation

$$de + p d(1/\rho) = T ds. \quad (4)$$

Note that (3) imply that  $-s^{\text{EOS}}(1/\rho, e)$  and  $e^{\text{EOS}}(1/\rho, s)$  are strictly convex functions. Relations (3) also ensure that

$$\frac{\partial p^{\text{EOS}}}{\partial(1/\rho)}(1/\rho, s) < 0, \quad (5)$$

so that the sound velocity  $c = \rho^{-1} \sqrt{-\partial p^{\text{EOS}}(1/\rho, s) / \partial(1/\rho)}$  is real valued. Let us recall now that the dimensionless quantity  $\text{Ma} = |u|/c$  is called the Mach number. We also make the classic assumption [51] that

$$\mathcal{M} s(\mathcal{V} \mid \mathcal{M}, \mathcal{E} \mid \mathcal{M}) = S(\mathcal{M}, \mathcal{V}, \mathcal{E}), \quad (6)$$

where the (non-specific) entropy  $(\mathcal{M}, \mathcal{V}, \mathcal{E}) \mapsto S(\mathcal{M}, \mathcal{V}, \mathcal{E})$  is a strictly concave homogeneous first-order function. Let us note that as  $\frac{\partial S}{\partial \mathcal{E}}(\mathcal{M}, \mathcal{V}, \mathcal{E}) = \frac{\partial S}{\partial \mathcal{E}}(\mathcal{V} | \mathcal{M}, \mathcal{E} | \mathcal{M}) = 1/T^{\text{EOS}}(\mathcal{V} | \mathcal{M}, \mathcal{E} | \mathcal{M}) > 0$ , then  $\mathcal{E} \mapsto S(\mathcal{M}, \mathcal{V}, \mathcal{E})$  is a strictly increasing function for a fixed  $\mathcal{M}$  and  $\mathcal{V}$ .

Weak solutions of (1) also satisfy the entropy inequality

$$\partial_t(\rho s) + \partial_x(u\rho s) \geq 0, \quad (7)$$

where the inequality (7) is indeed an equality in the case of smooth solutions (see [52, 53, 54, 55]).

We also are interested in the study of particular steady-state solutions of (1) called the hydrostatic equilibria that are classically defined by

$$\partial_x p = -\rho \partial_x \phi, \quad u = 0. \quad (8)$$

For many years, significant efforts have been dedicated to developing so-called well-balanced numerical methods (see *e.g.* [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 1, 21, 22, 23, 24]) that allow preserving discrete equivalents of equilibrium solutions like (8). In the present work, we intend to investigate well-balanced finite volume approximations of (1) that are compatible with discrete equivalents of (7) and ensure that the fluid states  $(\rho, \rho u, \rho E)$  remain in  $\Omega$ .

Before going any further, let us introduce the notations for our space-time discretization: we consider a strictly increasing sequence  $(x_{j+1/2})_{j \in \mathbb{Z}}$  and divide the real line into cells where the  $j^{\text{th}}$  cell is the interval  $(x_{j-1/2}, x_{j+1/2})$ . The space step of  $j^{\text{th}}$  cell is  $\Delta x_j = x_{j+1/2} - x_{j-1/2} > 0$  that we suppose constant and equal to  $\Delta x$  for the sake of simplicity. We note  $\Delta t > 0$  the time step such that  $t^{n+1} - t^n = \Delta t$  with  $n \in \mathbb{N}$ . For a given initial condition  $x \mapsto \mathbf{U}^0(x)$ , we consider a discrete initial data  $\mathbf{U}_j^0$  defined by  $\mathbf{U}_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}^0(x) dx$ , for  $j \in \mathbb{Z}$ . The algorithm proposed in this paper aims at computing a first-order accurate (in both space and time) approximation of the cell-averaged values  $\mathbf{U}_j^n$  of  $\frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}(x, t^n) dx$  where  $x \mapsto \mathbf{U}(x, t^n)$  is the exact solution of (1) at time  $t^n$  by means of a conservative finite volume discretization of (1) of the form

$$\mathbf{U}_j^{n+1} - \mathbf{U}_j^n + \frac{\Delta t}{\Delta x} (\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}) = \Delta t \mathbf{S}_j. \quad (9)$$

### 3. The original Operator Splitting Lagrange-Projection (OSLP) strategy

Operator splitting strategies allow simpler derivation of numerical methods by solving parts of the system separately and successively. However, this requires storing intermediate state values and may also necessitate specific treatments to implement higher order extension (see, for example [56, 57, 22, 24]).

In this section, we recall the properties of the OSLP method presented in [1]. It combines the all-regime method for gas dynamic proposed by [32] and the well-balanced treatment of source terms introduced in [16] in the context of the shallow water system. We chose to re-introduce all the discretization as the goal of the present paper is to recast this particular OSLP algorithm into a flux-splitting Lagrange-Projection (FSLP) finite volume method, using very similar expressions. We emphasize that the algorithm presented in this section is not new and comes entirely from [32, 16, 1] and that the novelty of our work lies in a modification of this algorithm that will be detailed in section 4. The method is based on the splitting of (1) into an acoustic sub-system:

$$\begin{cases} \partial_t \rho + \rho \partial_x u = 0, & (10a) \\ \partial_t(\rho u) + \rho u \partial_x u + \partial_x p = -\rho \partial_x \phi, & (10b) \\ \partial_t(\rho E) + \rho E \partial_x u + \partial_x(pu) = -\rho u \partial_x \phi, & (10c) \end{cases}$$

and a transport sub-system:

$$\begin{cases} \partial_t \rho + u \partial_x \rho = 0, & (11a) \\ \partial_t(\rho u) + u \partial_x(\rho u) = 0, & (11b) \\ \partial_t(\rho E) + u \partial_x(\rho E) = 0. & (11c) \end{cases}$$

Given a fluid state  $\mathbf{U}^n$ , this operator splitting algorithm can be decomposed as follows.

1. Update the fluid state  $U^n$  to the value  $U^{n+1-}$  by approximating the solution of (10):

$$\begin{cases} L_j \rho_j^{n+1-} = \rho_j^n, & (12a) \\ L_j (\rho u)_j^{n+1-} = (\rho u)_j^n - \frac{\Delta t}{\Delta x} (\Pi_{j+1/2}^{*,\theta} - \Pi_{j-1/2}^{*,\theta}) - \Delta t \{\rho \partial_x \phi\}_j^n, & (12b) \\ L_j (\rho E)_j^{n+1-} = (\rho E)_j^n - \frac{\Delta t}{\Delta x} (\Pi_{j+1/2}^{*,\theta} u_{j+1/2}^* - \Pi_{j-1/2}^{*,\theta} u_{j-1/2}^*) - \Delta t \{\rho u \partial_x \phi\}_j^n, & (12c) \\ L_j = 1 + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* - u_{j-1/2}^*). & (12d) \end{cases}$$

2. Update the fluid state  $U^{n+1-}$  to the value  $U^{n+1}$  by approximating the solution of (11): for  $\varphi \in \{\rho, \rho u, \rho E\}$

$$\varphi_j^{n+1} = \varphi_j^{n+1-} L_j + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* \varphi_{j+1/2}^{n+1-} - u_{j-1/2}^* \varphi_{j-1/2}^{n+1-}) \quad (13)$$

with the upwind choice

$$\varphi_{j+1/2}^{n+1-} = \begin{cases} \varphi_j^{n+1-}, & \text{if } u_{j+1/2}^* \geq 0, \\ \varphi_{j+1}^{n+1-}, & \text{if } u_{j+1/2}^* < 0, \end{cases} \quad (14)$$

and the following formulas for the interface pressures and velocities

$$u_{j+1/2}^* = \frac{(u_{j+1}^n + u_j^n)}{2} - \frac{1}{2a_{j+1/2}} \left( p_{j+1}^n - p_j^n + \frac{\rho_{j+1}^n + \rho_j^n}{2} (\phi_{j+1}^n - \phi_j^n) \right), \quad (15a)$$

$$\Pi_{j+1/2}^{*,\theta} = \frac{(p_{j+1}^n + p_j^n)}{2} - \theta_{j+1/2} \frac{a_{j+1/2}}{2} (u_{j+1}^n - u_j^n), \quad (15b)$$

as well as the source terms discretization:

$$\begin{cases} \{\rho \partial_x \phi\}_j^n = \frac{\{\rho \partial_x \phi\}_{j+1/2} + \{\rho \partial_x \phi\}_{j-1/2}}{2}, & (16a) \end{cases}$$

$$\begin{cases} \{\rho u \partial_x \phi\}_j^n = \frac{u_{j+1/2}^* \{\rho \partial_x \phi\}_{j+1/2} + u_{j-1/2}^* \{\rho \partial_x \phi\}_{j-1/2}}{2}, & (16b) \end{cases}$$

$$\begin{cases} \{\rho \partial_x \phi\}_{j+1/2} = \frac{\rho_{j+1}^n + \rho_j^n}{2} \frac{\phi_{j+1} - \phi_j}{\Delta x}. & (16c) \end{cases}$$

The constant parameter  $a_{j+1/2}$  is a local choice of an approximate acoustic impedance  $a$  associated with each interface  $j + 1/2$ . It should be chosen large enough so that (20) is satisfied, guaranteeing stability for the acoustic step. In practice, we choose

$$a_{j+1/2} = K \max(\rho_j^n c_j^n, \rho_{j+1}^n c_{j+1}^n) \quad \text{with } K > 1. \quad (17)$$

In the tests of section 8 we will use  $K = 1.1$ .

The parameter  $\theta$  enables the implementation of a low-Mach flux correction that ensures a control of the numerical diffusion in the momentum equation. This simple strategy is modeled after [31, 28, 33]. Depending on the choice of  $\theta$ , this correction takes effect whenever  $\text{Ma} < 1$ . In our case, its sole purpose is to help preserve the accuracy in the low-Mach regions of the computational domain by providing a uniform control of the truncation error with respect to  $\text{Ma}$ . We need to emphasize that this approach does not aim at addressing the full complexity of simulating flows in the low-Mach regime that has been widely investigated in the literature and spans for example: from the study of the influence of the grid [29, 30], the potential development of spurious modes [58, 59], the development of asymptotic preserving methods [36, 37, 34, 38, 39, 40, 41, 42], implicit-explicit methods [32, 45, 46, 47, 48] multi-dimensional control of the numerical diffusion [35], use of preconditioning methods [25, 26, 43, 27, 44] to the study of acoustics in low-Mach regime [49].

The discretization of the gravitational source term allows to exactly preserve the following discrete equivalent of the hydrostatic equilibrium (8):

$$\Pi_{j+1}^n - \Pi_j^n = -\frac{\rho_{j+1}^n + \rho_j^n}{2}(\phi_{j+1} - \phi_j), \quad u_j^n = 0, \quad \forall j \in \mathbb{Z}, \forall n \in \mathbb{N}. \quad (18)$$

Note that the resolution of the acoustic system is performed via a Suliciu-type relaxation [60, 61, 62, 63] following [32, 16]. Both steps can be rewritten as a fully conservative update formula:

$$\left\{ \begin{array}{l} \rho_j^{n+1} = \rho_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* \rho_{j+1/2}^{n+1-} - u_{j-1/2}^* \rho_{j-1/2}^{n+1-}), \\ (\rho u)_j^{n+1} = (\rho u)_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* (\rho u)_{j+1/2}^{n+1-} + \Pi_{j+1/2}^{\theta,*} - u_{j-1/2}^* (\rho u)_{j-1/2}^{n+1-} - \Pi_{j-1/2}^{\theta,*}) - \Delta t \{\rho \partial_x \phi\}_j^n, \\ (\rho E)_j^{n+1} = (\rho E)_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* (\rho E)_{j+1/2}^{n+1-} + \Pi_{j+1/2}^{\theta,*} u_{j+1/2}^* - u_{j-1/2}^* (\rho E)_{j-1/2}^{n+1-} - \Pi_{j-1/2}^{\theta,*} u_{j-1/2}^*) - \Delta t \{\rho u \partial_x \phi\}_j^n. \end{array} \right. \quad (19a)$$

$$\left\{ \begin{array}{l} \rho_j^{n+1} = \rho_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* \rho_{j+1/2}^{n+1-} - u_{j-1/2}^* \rho_{j-1/2}^{n+1-}), \\ (\rho u)_j^{n+1} = (\rho u)_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* (\rho u)_{j+1/2}^{n+1-} + \Pi_{j+1/2}^{\theta,*} - u_{j-1/2}^* (\rho u)_{j-1/2}^{n+1-} - \Pi_{j-1/2}^{\theta,*}) - \Delta t \{\rho \partial_x \phi\}_j^n, \end{array} \right. \quad (19b)$$

$$\left\{ \begin{array}{l} \rho_j^{n+1} = \rho_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* \rho_{j+1/2}^{n+1-} - u_{j-1/2}^* \rho_{j-1/2}^{n+1-}), \\ (\rho E)_j^{n+1} = (\rho E)_j^n + \frac{\Delta t}{\Delta x} (u_{j+1/2}^* (\rho E)_{j+1/2}^{n+1-} + \Pi_{j+1/2}^{\theta,*} u_{j+1/2}^* - u_{j-1/2}^* (\rho E)_{j-1/2}^{n+1-} - \Pi_{j-1/2}^{\theta,*} u_{j-1/2}^*) - \Delta t \{\rho u \partial_x \phi\}_j^n. \end{array} \right. \quad (19c)$$

The scheme (19) is proven to be positivity preserving for the density and the internal energy as well as entropy stable when  $\Delta t$  verifies both the acoustic CFL condition:

$$\frac{\Delta t}{\Delta x} \max_{j \in \mathbb{Z}} (\max(1/\rho_j^n, 1/\rho_{j+1}^n) a_{j+1/2}) \leq \frac{1}{2}, \quad (20)$$

and the transport CFL condition:

$$\Delta t \max_{j \in \mathbb{Z}} \left( \left( u_{j-\frac{1}{2}}^* \right)^+ - \left( u_{j+\frac{1}{2}}^* \right)^- \right) < \Delta x, \quad (21)$$

granted that the following inequality:

$$-\frac{1}{2a^2} (p^{\text{EOS}}(\tau_k^{*,\theta}, s_k) - \Pi^*)^2 + \frac{(1-\theta)^2 (u_{j+1} - u_j)^2}{8} \leq 0, \quad k = j, j+1, \quad (22)$$

where  $\tau_j^{*,\theta} = 1/\rho_j^n + \frac{1}{a_{j+1/2}} (u_{j+1/2}^* - u_j^n)$  and  $\tau_{j+1}^{*,\theta} = 1/\rho_{j+1}^n + \frac{1}{a_{j+1/2}} (u_{j+1}^n - u_{j+1/2}^*)$  is satisfied at each interface  $j+1/2$ .

In section 4, we discuss how a simple modification of the transport step allows recasting this two-step OSLP algorithm into a one-step FSLP method while keeping the interesting properties of the original method: the well-balanced property, the accuracy in the low Mach regime, mass, and energy positivity and the discrete entropy inequality.

#### 4. Recasting the OSLP method into a flux-splitting Lagrange-Projection (FSLP) method; a modification of the transport step

In this section, we discuss how a simple modification of the transport step (13) of the OSLP method (19) proposed by [16] leads to a much simpler, FSLP algorithm. Flux-splitting methods have been used in many application contexts thanks to their ease of implementation that relies on building a discrete evaluation of the fluxes (see, for example, [64, 65, 66, 67, 68]). These methods have been extensively developed for several decades (see, for example, [69, 70, 71, 72, 72, 73, 74, 75, 76, 77] and the references therein) yielding efficient simulation tools. Unfortunately, deriving theoretical results that ensure the good behavior of these methods is difficult, which contrasts with their good performance in practice. Before going any further, let us mention that the question of building Eulerian numerical fluxes relying on a Lagrangian approximation of the flow equations has been successfully investigated in the literature with different approaches [78, 79, 80, 76, 81].

A key contribution of the present paper is the derivation of stability properties for the flux-splitting algorithm. These proofs are based on the following observation; let us consider a given hyperbolic problem with a source term for which the set of admissible states is convex (e.g. Euler's equations of gas dynamic or ideal Magneto-hydrodynamics);

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}) \quad (23)$$

We design a separation of the flux and source term into  $N$  parts  $(F_p, S_p)_{1 \leq p \leq N}$  so that:

$$\sum_{p=1}^N F_p(\mathbf{U}) = F(\mathbf{U}), \quad \sum_{p=1}^N S_p(\mathbf{U}) = S(\mathbf{U}), \quad (24)$$

as well as a series of coefficients  $\alpha_j^p \in ]0, 1[$  that sums up to 1;  $\sum_{p=1}^N \alpha_j^p = 1$  for each cell  $j$ . Let us assume that we can build a discretization for each part where the sub-fluxes and sub-source terms are multiplied by the inverses of the coefficients

$$\frac{\mathbf{U}_j^{p,n+1} - \mathbf{U}_j^n}{\Delta t} - \frac{1}{\alpha^p} [\partial_x F_p(\mathbf{U})]_j = \frac{1}{\alpha^p} [S_p(\mathbf{U})]_j \quad \forall p \in [1, N]. \quad (25)$$

Moreover, let us assume that each of these discretizations is stable under their respective local CFL condition:

$$\Delta t < \alpha^p \frac{\Delta x}{v_p^j} \quad (26)$$

where  $v_p^j$  is the local characteristic velocity associated with the discretization of the  $p$ -th flux/source term. By re-assembling the result of each part with the convex combination defined by the coefficients  $\alpha^p$ ,

$$\mathbf{U}_j^{n+1} := \sum_{p=0}^N \alpha^p \mathbf{U}_j^{p,n+1} \quad (27)$$

we obtain a discretization consistent with (23), regardless of the value of the coefficients  $\alpha^p \in ]0, 1[$ . The full update is stable as a convex combination of the stable sub-updates (25). This means we can freely choose the coefficients  $\alpha^p$  to optimize the CFL condition. Indeed, the update (27) is stable as long as each sub-update is stable i.e., the optimum is reached for:

$$\Delta t < \max_{\alpha^1, \dots, \alpha^N} \left[ \min \left( \alpha^1 \frac{\Delta x}{v_1^j}, \dots, \alpha^N \frac{\Delta x}{v_N^j} \right) \right]. \quad (28)$$

that corresponds to the values of  $\alpha^p$  that satisfies  $\alpha^1 \frac{\Delta x}{v_1^j} = \alpha^2 \frac{\Delta x}{v_2^j} = \dots = \alpha^N \frac{\Delta x}{v_N^j}$  and provides the following local CFL condition:

$$\Delta t < \frac{\Delta x}{v_1^j + v_2^j + \dots + v_N^j}. \quad (29)$$

Indeed, if  $\frac{\alpha^p}{v_p} = \frac{1}{v_1^j + v_2^j + \dots + v_N^j}$  for  $p = 1, \dots, N$ , then  $\min_p \left( \frac{\alpha^p}{v_p} \right) = \min_p \left( \frac{1}{v_1^j + v_2^j + \dots + v_N^j} \right) = \frac{1}{v_1^j + v_2^j + \dots + v_N^j}$ . In this work, we separate the system into  $N = 2$  parts corresponding to the pressure and advection terms. This type of splitting is not new and can be found in [70, 82, 77, 83] without entropy stability theorems. Discretization techniques that also feature a separate treatment for the pressure and advection effects have been proposed for fractional step methods [84, 85, 86, 87, 32, 16, 88, 1].

By modifying the transport step of the original operator splitting algorithm (19) by computing the fluxes on the initial states  $n$  instead of the acoustic state  $n + 1 -$ :

$$\varphi_j^{n+1} = \varphi_j^{n+1-} L_j - \frac{\Delta t}{\Delta x} \left( u_{j+1/2}^* \varphi_{j+1/2}^n - u_{j-1/2}^* \varphi_{j-1/2}^n \right) \quad (30)$$

we obtain the following fully conservative update that we refer to as our FSLP method:

$$\begin{cases} \rho_j^{n+1} = \rho_j^n + \frac{\Delta t}{\Delta x} \left( u_{j+1/2}^* \rho_{j+1/2}^n - u_{j-1/2}^* \rho_{j-1/2}^n \right) \\ (\rho u)_j^{n+1} = (\rho u)_j^n + \frac{\Delta t}{\Delta x} \left( u_{j+1/2}^* (\rho u)_{j+1/2}^n + \Pi_{j+1/2}^{\theta,*} - u_{j-1/2}^* (\rho u)_{j-1/2}^n - \Pi_{j-1/2}^{\theta,*} \right) - \Delta t \{ \rho \partial_x \phi \}_j^n, \\ (\rho E)_j^{n+1} = (\rho E)_j^n + \frac{\Delta t}{\Delta x} \left( u_{j+1/2}^* (\rho E)_{j+1/2}^n + \Pi_{j+1/2}^{\theta,*} u_{j+1/2}^* - u_{j-1/2}^* (\rho E)_{j-1/2}^n - \Pi_{j-1/2}^{\theta,*} u_{j-1/2}^* \right) - \Delta t \{ \rho u \partial_x \phi \}_j^n. \end{cases} \quad (31)$$

Note that we keep the upwind choice for the transport scheme:

$$\varphi_{j+1/2}^n = \begin{cases} \varphi_j^n, & \text{if } u_{j+1/2}^* \geq 0, \\ \varphi_{j+1}^n, & \text{if } u_{j+1/2}^* < 0, \end{cases} \quad (32)$$

where  $(u, \Pi)^*$  are given by (15). We provide the CFL condition associated with the new method:

$$\frac{\Delta t}{\Delta x} \max_{j \in \mathbb{Z}} \left( 2 \max(1/\rho_j^n, 1/\rho_{j+1}^n) a_{j+1/2} + \left( u_{j-\frac{1}{2}}^* \right)^+ - \left( u_{j+\frac{1}{2}}^* \right)^- \right) < 1 \quad (33)$$

This CFL condition is indeed of the form (29) with  $N = 2$ . It has the same characteristic speeds as the acoustic condition in (20) and the transport condition in (21), except that they are summed rather than checked separately. The new method has several advantages compared to the original numerical scheme (19):

1. The implementation of the flux-splitting version is much simpler than the operator-splitting version. Indeed, it can be implemented as a standard, simple flux-based finite volume method with the following numerical flux formula:

$$\mathbf{F}^{\text{FSLP}}(U_L, U_R) = \begin{cases} u^* \rho_{LR} \\ u^* (\rho u)_{LR} + \Pi^{*,\theta} \\ u^* (\rho E)_{LR} + \Pi^{*,\theta} u^* \end{cases} \quad (34)$$

with

$$\varphi_{LR} = \begin{cases} \varphi_L & \text{if } u^* > 0, \\ \varphi_R & \text{otherwise.} \end{cases} \quad (35)$$

We can see in (34) that the flux evaluation clearly separates the pressure-related terms from the advection terms so that it can be affiliated with a family of methods proposed in the literature like [70, 82, 77, 83].

2. As the method can be implemented as a simple flux-based solver, it can be seamlessly combined with any existing flux-based high-order algorithm such as MUSCL[89, 90, 91, 92], (W)ENO [93, 94] or MOOD methods [95, 96]. We detail the procedure for the extension to second order in section 6.2 and give some numerical examples in section 8. Note, however, that the well-balanced treatment of gravity is not straightforward to extend to high order and requires a careful examination that is beyond the scope of this paper. Also, using the low-Mach correction  $\theta$  combined with a highly accurate high-order method can amplify numerical instabilities that already exist at first-order (checkerboard modes, for example). We do not address this issue in this paper, as our focus is on demonstrating the recasting of the OSLP method into the FSLP method.
3. The FSLP method is more computationally efficient than the original OSLP method. The OSLP method requires two update loops per time step to compute a time step of size  $\sim \Delta x / \max(v, c)$ , where  $v$  and  $c$  are the velocities associated with transport and acoustic effects, respectively, as they appear in the CFL conditions. In contrast, the FSLP method only requires one loop per time step of size  $\sim \Delta x / (v + c)$ . This means that the FSLP method requires fewer sweeps to reach the same physical time, especially in the low-Mach regime where  $v \ll c$  or in the hypersonic regime where  $v \gg c$ , where it is roughly two times more efficient. If  $v = c$ , both methods are equally efficient.
4. The new update formula eliminates the need to store the intermediate state  $U^{n+1-}$ , as it can be computed in a single sweep. This reduces the algorithm's memory footprint by approximately 2/3, and reduces the stencil radius from two to one cell. The decrease in memory storage requirements can improve performance by reducing the time spent accessing the data arrays.

Despite the update formula being very similar, the mathematical background required to derive the stability properties of (19) is new. It is the object of the next section 5.



## 5. Derivation of the stability properties for our new method

In this section, we focus on deriving the stability properties of our new FSLP scheme (31). To this end, we will perform a Suliciu-type relaxation [60, 61, 62, 63] of the pressure term and introduce a surrogate specific volume. We then isolate two new sub-systems, the advection and pressure sub-systems, for which we derive numerical fluxes. We then re-obtain our new method and derive its stability properties by performing a convex combination of the two fluxes. Note that the proof of stability for the pressure subsystem is similar to the acoustic sub-system in [32]. For this reason, we only recall this proof in the appendix for completeness.

### 5.1. Relaxation and flux-splitting

We first apply a relaxation of the original Euler system. Manipulations of smooth solutions of (1) gives  $\partial_t(\rho p) + \partial_x(u\rho p) + \rho^2 c^2 \partial_x u = 0$ . We choose to perform a Suliciu-type approximation of the system (1) for  $t \in [t^n, t^{n+1})$  by introducing a surrogate pressure  $\Pi$  and considering the relaxed system:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(u\rho u + \Pi) = -\rho \partial_x \phi, \\ \partial_t(\rho E) + \partial_x(u\rho E + \Pi u) = -\rho u \partial_x \phi, \\ \partial_t(\rho \Pi) + \partial_x(u\rho \Pi + a^2 u) = \rho \lambda (p - \Pi). \end{cases} \quad (36)$$

The parameter  $\lambda$  is a frequency that characterizes the strength of the source term that drives  $\Pi$  towards the equilibrium  $\Pi = p$ . In the regime  $\lambda \rightarrow \infty$ , we formally recover (1). In our numerical solver context, we classically mimic the  $\lambda \rightarrow \infty$  regime by enforcing  $\Pi_j^n = p^{\text{EOS}}(1/\rho_j^n, e_j^n)$  at each time step and then solving (36) with  $\lambda = 0$ , which will be the case in all computations below without any ambiguities. We now introduce another auxiliary variable  $\mathcal{T}$  and impose that it verifies

$$\partial_t(\rho \mathcal{T}) = 0. \quad (37)$$

We suppose that  $\mathcal{T}(t=0) = 1/\rho(t=0)$  at the initial instant so that  $\mathcal{T}(x, t)$  is equal to the specific volume  $1/\rho(x, t)$  for all  $x$  and  $t > 0$ . Let us now re-write the system (36)-(37) in order to highlight three different operators that compose the flux and the source term of (36)-(37) following similar lines as [70, 82, 83]

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ \rho E \\ \rho \Pi \\ \rho \mathcal{T} \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 \\ \rho E u \\ \rho \Pi u \\ u \end{pmatrix} + \partial_x \begin{pmatrix} 0 \\ \Pi \\ \Pi u \\ a^2 u \\ -u \end{pmatrix} = - \begin{pmatrix} 0 \\ \rho \\ \rho u \\ 0 \\ 0 \end{pmatrix} \partial_x \phi. \quad (38)$$

Let us underline that both  $\Pi$  and  $\rho \mathcal{T}$  are only mathematical intermediates used to derive the scheme's stability properties. Indeed, these variables do not appear in the update formula (31), so that there is no need to evaluate and store them while implementing the algorithm. Let us introduce the convex combination parameter  $\alpha \in ]0, 1[$  and two sub-systems associated with different parts of the fluxes and source terms featured in (38). The first system gathers the source term and the flux associated with pressure terms ponderated by  $1/\alpha$

$$\begin{cases} \partial_t \rho = 0, \\ \partial_t(\rho u) + \frac{1}{\alpha} \partial_x(\Pi) = -\frac{1}{\alpha} \rho \partial_x \phi, \\ \partial_t(\rho E) + \frac{1}{\alpha} \partial_x(\Pi u) = -\frac{1}{\alpha} \rho u \partial_x \phi, \\ \partial_t(\rho \Pi) + \frac{1}{\alpha} \partial_x(a^2 u) = 0, \\ \partial_t(\rho \mathcal{T}) - \frac{1}{\alpha} \partial_x u = 0. \end{cases} \quad (39)$$

We will refer to (39) as the pressure system. The second sub-system is composed of the remaining terms that pertain to transport effects ponderated by  $1/1 - \alpha$ , it reads

$$\begin{cases} \partial_t(\rho\varphi) + \frac{1}{1-\alpha}\partial_x(u\rho\varphi) = 0, & \varphi \in \{1, u, E, \Pi\} \\ \partial_t(\rho\mathcal{T}) + \frac{1}{1-\alpha}\partial_x u = 0, \end{cases} \quad (40)$$

and will be called the advection system.

The pressure system (39) is hyperbolic and involves the characteristic velocities  $\{\pm\frac{1}{\alpha}a/\rho, 0, 0, 0\}$  that are all associated with linearly degenerate fields. The advection system (40) is only weakly hyperbolic as its Jacobian matrix admits  $(1 - \alpha)u$  as multiple eigenvalues but is not diagonalizable. Nevertheless, let us underline that the algorithms we will consider for approximating the solutions of (40) will verify a local maximum principle under a CFL condition so that stability will be ensured for the advection step (see section 5.2).

Before continuing, let us comment on equations (39) and (36). The factors  $\alpha$  and  $1 - \alpha$  that appear in the fluxes and source terms of these equations correspond to the case  $N = 2$  of the flux splitting stability argument presented at the beginning of section 4.

Then, although the trivial equation stationary (37) is now split into two non-stationary parts within (39) and (39), the overall scheme will indeed guarantee that  $(\rho\mathcal{T})_j^n = 1$  for  $j \in \mathbb{Z}$  and  $n \in \mathbb{N}$ .

## 5.2. The convex combination

We propose the following discretization strategy:

1. Compute  $\mathbf{U}_j^P$  as the update of the initial state  $\mathbf{U}_j^n$  by approximating the solution of (39):

$$\begin{cases} \rho_j^P = \rho_j^n, \\ (\rho u)_j^P = (\rho u)_j^n - \frac{1}{\alpha} \frac{\Delta t}{\Delta x} (\Pi_{j+1/2}^{*,\theta} - \Pi_{j-1/2}^{*,\theta}) - \frac{1}{\alpha} \Delta t \{\rho \partial_x \phi\}_j^n, \\ (\rho E)_j^P = (\rho E)_j^n - \frac{1}{\alpha} \frac{\Delta t}{\Delta x} (\Pi_{j+1/2}^{*,\theta} u_{j+1/2}^* - \Pi_{j-1/2}^{*,\theta} u_{j-1/2}^*) - \frac{1}{\alpha} \Delta t \{\rho u \partial_x \phi\}_j^n, \\ (\rho \Pi)_j^P = (\rho \Pi)_j^n - \frac{1}{\alpha} \frac{\Delta t}{\Delta x} (a_{j+1/2}^2 u_{j+1/2}^* - a_{j-1/2}^2 u_{j-1/2}^*), \\ (\rho \mathcal{T})_j^P = 1 + \frac{1}{\alpha} \frac{\Delta t}{\Delta x} (u_{j+1/2}^* - u_{j-1/2}^*). \end{cases} \quad (41)$$

2. Compute  $\mathbf{U}_j^A$  as the update of the initial state  $\mathbf{U}_j^n$  by approximating the solution of (40): for  $\varphi \in \{1, u, E, \Pi\}$

$$\begin{cases} (\rho\varphi)_j^A = (\rho\varphi)_j^n - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} (u_{j+1/2}^* (\rho\varphi)_{j+1/2}^n - u_{j-1/2}^* (\rho\varphi)_{j-1/2}^n), \\ (\rho\mathcal{T})_j^A = (\rho\mathcal{T})_j^n - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} (u_{j+1/2}^* - u_{j-1/2}^*). \end{cases} \quad (42)$$

3. Evaluate  $\mathbf{U}_j^{n+1}$  as the convex combination of  $\mathbf{U}_j^P$  and  $\mathbf{U}_j^A$ :

$$\mathbf{U}_j^{n+1} = \alpha \mathbf{U}_j^P + (1 - \alpha) \mathbf{U}_j^A \quad (43)$$

It can be verified that the update in (43) is equivalent to the FSLP scheme in (31), for any value of  $\alpha \in ]0, 1[$ . This means that the flux of the FSLP scheme can be expressed as an arbitrary convex combination of the fluxes involved in the update. As explained at the beginning of section 4, this interpretation allows us to choose  $\alpha$  optimally in order to obtain the least restrictive CFL condition, given by (33).

### 5.3. Stability of the pressure step

In this section, we prove the stability of the pressure step. We chose to move all the derivations in the appendix as the arguments we use are already present in [16] in the proof of the stability of the acoustic step (12) of the OSLP method (19). We introduce the CFL condition associated with the pressure step.

$$\frac{1}{\alpha} \frac{\Delta t}{\Delta x} \max_{j \in \mathbb{Z}} \left( \max \left( 1/\rho_j^n, 1/\rho_{j+1}^n \right) a_{j+1/2} \right) \leq \frac{1}{2}, \quad (44)$$

It is identical to the acoustic CFL (20) but  $1/\alpha$  times as restrictive.

**Proposition 5.1.** *Suppose that  $a$  is chosen large enough so that (17) is verified and that both  $\mathcal{T}_L^* > 0$ ,  $\mathcal{T}_R^* > 0$  from (B.11e) are positive. Suppose also that the low-Mach correction  $\theta$  is chosen large enough so that (C.7) is valid. Under the CFL condition (44) we have that:*

1. *the density and the internal energy verify  $\rho_j^P > 0$  and  $e_j^P > 0$ , for all  $j$ ,*
2. *the discretization (41) satisfies the entropy inequality*

$$\rho_j^P s^{EOS}(\mathcal{T}_j^P, e_j^P) - \rho_j^n s(1/\rho_j^n, e_j^n) + \frac{1}{\alpha} \frac{\Delta t}{\Delta x} (q_{j+1/2}^n - q_{j-1/2}^n) \geq 0, \quad (45)$$

with  $q_{j+1/2}^n = q_\Delta(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$ , where  $q_\Delta$  is a flux function consistent with 0 as  $\Delta t, \Delta x \rightarrow 0$ .

*Proof.* The positivity of the internal energy and the entropy inequality of (45) are direct consequences of the approximate Riemann solver properties of proposition Appendix C.2 and the consistency in the integral sense [61].  $\square$

The condition (C.7) is identical to the OSLP low-Mach stability condition (22) but with the surrogate density  $\mathcal{T}$  instead of  $1/\rho$ . In practice, conditions for stability for the pressure update are strictly the same as for the acoustic step (12) from [32] apart from the factor  $1/\alpha$  in the CFL condition.

### 5.4. Stability of the advection step

We introduce the CFL condition associated with the advection step.

$$\frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} \max_{j \in \mathbb{Z}} \left( \left( u_{j-\frac{1}{2}}^* \right)^+ - \left( u_{j+\frac{1}{2}}^* \right)^- \right) < 1, \quad (46)$$

It is identical to the transport CFL (21) but  $1/1-\alpha$  times as restrictive.

**Proposition 5.2.** *Under the CFL condition (46), the discretization (42) of the advection subsystem verifies the following properties.*

1.  $\mathbf{U}_j^A$  is a positive linear combination of  $\mathbf{U}_{j-1}^n$ ,  $\mathbf{U}_j^n$  and  $\mathbf{U}_{j+1}^n$ .
2. Under the CFL conditions (46):  $b_j^A$  is a convex combination of  $b_{j-1}^n$ ,  $b_j^n$  and  $b_{j+1}^n$  for  $b \in \{u, E, \mathcal{T}\}$ .
3. Under the CFL conditions (46): if  $e_j^n > 0$  for all  $j \in \mathbb{Z}$  then  $e_j^A > 0$  for all  $j \in \mathbb{Z}$ ,
4. The discretization (42) satisfies the entropy inequality

$$\rho_j^A s^{EOS}(\mathcal{T}_j^A, e_j^A) - \rho_j^n s_j^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} \left( u_{j+1/2}^* \rho_{j+1/2}^n s_{j+1/2}^n - u_{j-1/2}^* \rho_{j-1/2}^n s_{j-1/2}^n \right) \geq 0 \quad (47)$$

*Proof.* The advection scheme (42) can be recast into

$$\mathbf{U}_j^A = -\frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,-} \mathbf{U}_{j+1}^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,+} \mathbf{U}_{j-1}^n + \left[ 1 - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} (u_{j+1/2}^{*,+} - u_{j-1/2}^{*,-}) \right] \mathbf{U}_j^n, \quad (48)$$

which proves 1. One can also write

$$\left(\frac{\mathbf{U}}{\rho}\right)_j^A = \lambda_j^{(+1)}\left(\frac{\mathbf{U}}{\rho}\right)_{j+1}^n + \lambda_j^{(0)}\left(\frac{\mathbf{U}}{\rho}\right)_j^n + \lambda_j^{(-1)}\left(\frac{\mathbf{U}}{\rho}\right)_{j-1}^n, \quad (49)$$

with

$$\lambda_j^{(+1)} = -\frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,-} \left(\frac{\rho_{j+1}^n}{\rho_j^A}\right), \quad \lambda_j^{(0)} = \left[1 - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} (u_{j+1/2}^{*,+} - u_{j-1/2}^{*,-})\right] \left(\frac{\rho_j^n}{\rho_j^A}\right), \quad \lambda_j^{(-1)} = \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,+} \left(\frac{\rho_{j-1}^n}{\rho_j^A}\right). \quad (50)$$

By (48) we have that

$$\rho_j^A = -\frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,-} \rho_{j+1}^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,+} \rho_{j-1}^n + \left[1 - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} (u_{j+1/2}^{*,+} - u_{j-1/2}^{*,-})\right] \rho_j^n, \quad (51)$$

so that  $\lambda_j^{(+1)} + \lambda_j^{(0)} + \lambda_j^{(-1)} = 1$ , which proves that  $b_j^A$  is a convex combination of  $b_{j-1}^n$ ,  $b_j^n$  and  $b_{j+1}^n$  for  $b \in \{u, E\}$ . Let us now consider the case of  $\mathcal{T}^A$ . By (42), we have that

$$\mathcal{T}_j^A = \mathcal{T}_j^n \frac{\rho_j^n}{\rho_j^A} - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,+} \frac{1}{\rho_j^A} - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,-} \frac{1}{\rho_j^A} + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,+} \frac{1}{\rho_j^A} + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,-} \frac{1}{\rho_j^A}. \quad (52)$$

However, since we chose  $\rho_i^n \mathcal{T}_i^n = 1$  for all  $i \in \mathbb{Z}$ , We can write that

$$\mathcal{T}_j^A = \mathcal{T}_j^n \frac{\rho_j^n}{\rho_j^A} - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,+} \frac{\rho_j^n}{\rho_j^A} \mathcal{T}_j^n - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,-} \frac{\rho_{j+1}^n}{\rho_j^A} \mathcal{T}_{j+1}^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,+} \frac{\rho_{j-1}^n}{\rho_j^A} \mathcal{T}_{j-1}^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,-} \frac{\rho_j^n}{\rho_j^A} \mathcal{T}_j^n \quad (53)$$

$$= \lambda_j^{(+1)} \mathcal{T}_{j+1}^n + \lambda_j^{(-1)} \mathcal{T}_{j-1}^n + \lambda_j^{(0)} \mathcal{T}_j^n. \quad (54)$$

Consequently  $\mathcal{T}_j^A$  is also a convex combination of  $\mathcal{T}_{j-1}^n$ ,  $\mathcal{T}_j^n$  and  $\mathcal{T}_{j+1}^n$ , which proves 2. For statement 3, we consider the concave function  $K$  introduced in the proof of lemma Appendix A.1, and we have that  $e_j^A = K(u_j^A, E_j^A)$ . Thanks to statement 2. we can thus write that.

$$e_j^A = K\left(\sum_{k=0,\pm 1} \lambda_j^{(k)} u_{j+k}^n, \sum_{k=0,\pm 1} \lambda_j^{(k)} E_{j+k}^n\right) \geq \sum_{k=0,\pm 1} \lambda_j^{(k)} K(u_{j+k}^n, E_{j+k}^n) = \sum_{k=0,\pm 1} \lambda_j^{(k)} e_{j+k}^n > 0, \quad (55)$$

which proves statement 3.

Now using the lemma Appendix A.1, we have that

$$s(\mathcal{T}_j^A, e_j^A) = \mathcal{U}(\mathcal{T}_j^A, u_j^A, E_j^A) = \mathcal{U}\left(\sum_{k=0,\pm 1} \lambda_j^{(k)} \mathcal{T}_{j+k}^n, \sum_{k=0,\pm 1} \lambda_j^{(k)} u_{j+k}^n, \sum_{k=0,\pm 1} \lambda_j^{(k)} E_{j+k}^n\right) \quad (56)$$

$$\geq \sum_{k=0,\pm 1} \lambda_j^{(k)} \mathcal{U}(\mathcal{T}_{j+k}^n, u_{j+k}^n, E_{j+k}^n) = \sum_{k=0,\pm 1} \lambda_j^{(k)} s(\mathcal{T}_{j+k}^n, e_{j+k}^n). \quad (57)$$

This inequality also reads

$$s(\mathcal{T}_j^A, e_j^A) \geq s_j^n \frac{\rho_j^n}{\rho_j^A} - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,+} \frac{\rho_j^n}{\rho_j^A} s_j^n - \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j+1/2}^{*,-} \frac{\rho_{j+1}^n}{\rho_j^A} s_{j+1}^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,+} \frac{\rho_{j-1}^n}{\rho_j^A} s_{j-1}^n + \frac{1}{1-\alpha} \frac{\Delta t}{\Delta x} u_{j-1/2}^{*,-} \frac{\rho_j^n}{\rho_j^A} s_j^n. \quad (58)$$

If the CFL condition (46) is met then  $\rho_j^A \geq 0$  and by multiplying (58) by  $\rho_j^A$  we get (47).  $\square$

### 5.5. Stability of the FSLP method

**Proposition 5.3.** *If the following conditions are met*

1. *The CFL condition (33) is met,*
2. *the parameter  $a$  is large enough so that (17) is verified and  $\mathcal{T}_L^* > 0$ ,  $\mathcal{T}_R^* > 0$  in (B.11e) for all  $j \in \mathbb{Z}$ ,*
3. *both density and internal energies are positive, i.e.  $\rho_j^n > 0$  and  $e_j^n > 0$  for all  $j \in \mathbb{Z}$ ,*
4. *the parameter  $\theta$  is large enough so that (C.7) is valid at each interface,*

*then the flux-splitting update (31)*

*(b) preserves positivity for both density and internal energy i.e.  $\rho_j^{n+1} > 0$  and  $e_j^{n+1} > 0$  for all  $j \in \mathbb{Z}$ ,*

*(c) is endowed with the following entropy inequality:*

$$\rho_j^{n+1} s(1/\rho_j^{n+1}, e_j^{n+1}) - \rho_j^n s(1/\rho_j^n, e_j^n) + \frac{\Delta t}{\Delta x} (\mathcal{Q}_{j+1/2} - \mathcal{Q}_{j-1/2}) \geq 0, \quad (59)$$

*with  $\mathcal{Q}_{j+1/2} = u_{j+1/2}^* \rho_{j+1/2}^n s_{j+1/2}^n + q_{j+1/2}^n$ .*

*Proof.* (a) Let us start by ensuring that the CFL conditions (44), (46) are satisfied so that the advection and pressure steps are stable. By choosing  $\alpha$  so that  $\alpha c_j = (1 - \alpha)v_j = v_j + c_j$  where  $v_j = \left( (u_{j-\frac{1}{2}}^*)^+ - (u_{j+\frac{1}{2}}^*)^- \right)$  and  $c_j = 2 \max \left[ \max(1/\rho_{j-1}^n, 1/\rho_j^n) a_{j-1/2}, \max(1/\rho_j^n, 1/\rho_{j+1}^n) a_{j+1/2} \right]$ , it is straightforward that (44), (46) are equivalent and corresponds to (33). This choice of  $\alpha$  seems local as it depends on the characteristic speed of each cell considered. However, it can be chosen globally as the minimizer of  $f(\alpha) = \max(\alpha c_j, (1 - \alpha)v_j) = c_i + v_i$  where  $i$  is the index of the cell with the largest speed sum of the simulation domain.

Thanks to the propositions 5.2 and 5.1 we have  $\rho^A > 0$ ,  $\rho^P > 0$ ,  $e^A > 0$  and  $e^P > 0$  thus, the positivity is straightforward for the density as  $\rho_j^{n+1} = (\rho_j^A + \rho_j^P)/2 > 0$ . For the internal energy, we consider the function  $\Lambda((\rho, \rho u, \rho E)) = (\rho E) - \frac{(\rho u)^2}{2\rho}$ , that is proven to be concave in Appendix A.1. We have that:

$$\begin{aligned} (\rho e)_j^{n+1} &= (\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} = \Lambda(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}) \\ &= \Lambda((1 - \alpha)\rho_j^A + \alpha\rho_j^P, (1 - \alpha)(\rho u)_j^A + \alpha(\rho u)_j^P, (1 - \alpha)(\rho E)_j^A + \alpha(\rho E)_j^P) \\ &\geq (1 - \alpha)\Lambda(\rho_j^A, (\rho u)_j^A, (\rho E)_j^A) + \alpha\Lambda(\rho_j^P, (\rho u)_j^P, (\rho E)_j^P) = (1 - \alpha)(\rho e)_j^A + \alpha(\rho e)_j^P > 0. \end{aligned} \quad (60)$$

by concavity.

For (b): propositions 5.2 and 5.1 ensure that both entropy inequalities (45) and (47) are satisfied. We then use the concavity of the function  $\eta(\rho, \rho\mathcal{T}, \rho u, \rho E) = \rho s\left(\frac{\rho\mathcal{T}}{\rho}, \frac{\rho E}{\rho} - \frac{(\rho u)^2}{2\rho^2}\right)$  that is proven in Appendix A.1 and the fact that  $(\rho\mathcal{T})_j^{n+1} = \alpha(\rho\mathcal{T})_j^P + (1 - \alpha)(\rho\mathcal{T})_j^A = 1$ . Noting  $\alpha^P = \alpha$  and  $\alpha^A = 1 - \alpha$ , we have:

$$\begin{aligned} \rho_j^{n+1} s(1/\rho_j^{n+1}, e_j^{n+1}) &= \rho_j^{n+1} s\left(\frac{1}{\rho_j^{n+1}}, \frac{(\rho E)_j^{n+1}}{\rho_j^{n+1}} - \frac{1}{2}\left(\frac{(\rho u)_j^{n+1}}{\rho_j^{n+1}}\right)^2\right) = \eta(\rho_j^{n+1}, 1, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}) \\ &= \eta(\rho_j^{n+1}, (\rho\mathcal{T})_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}) = \eta\left(\sum_{k=A,P} \alpha^k \rho_j^k, \sum_{k=A,P} \alpha^k (\rho\mathcal{T})_j^k, \sum_{k=A,P} \alpha^k (\rho u)_j^k, \sum_{k=A,P} \alpha^k (\rho E)_j^k\right). \end{aligned} \quad (61)$$

$$\rho_j^{n+1} s(1/\rho_j^{n+1}, e_j^{n+1}) \geq \sum_{k=A,P} \alpha^k \eta(\rho_j^k, (\rho\mathcal{T})_j^k, (\rho u)_j^k, (\rho E)_j^k) = \sum_{k=A,P} \alpha^k \rho_j^k s(\mathcal{T}_j^k, e_j^k). \quad (62)$$

by concavity. Using (45) and (47), we get:

$$\rho_j^{n+1} s(1/\rho_j^{n+1}, e_j^{n+1}) \geq \rho_j^n s(1/\rho_j^n, e_j^n) - \frac{\Delta t}{\Delta x} (q_{j+1/2}^n - q_{j-1/2}^n) - \frac{\Delta t}{\Delta x} (u_{j+1/2}^* \rho_{j+1/2}^n s_{j+1/2}^n - u_{j-1/2}^* \rho_{j-1/2}^n s_{j-1/2}^n), \quad (63)$$

which proves (b).  $\square$

## 6. Low Mach behavior, extension to multi-dimensional and higher order of accuracy

In this section, we briefly address the behavior of the scheme in the low Mach regime and propose simple means to extend the FSLP method to multi-dimensional problems and improve its accuracy with higher-order techniques.

### 6.1. Low Mach behavior

Many simulation cases involve flows in which the material velocity is relatively low compared to the sound velocity. A common way to characterize this situation is to consider the numbers  $L, t_0, \rho_0, u_0, p_0, u_0 = p_0 \rho_0, c_0 = \sqrt{p_0/\rho_0}$  and  $(\partial_x \phi)_0$  that are the characteristic magnitudes for length, time, density, velocity, pressure, sound velocity, and  $\partial_x \phi$ , respectively. We then introduce the following non-dimensional variables:  $\tilde{x} = x/L, \tilde{t} = t/t_0, \tilde{\rho} = \rho/\rho_0, \tilde{u} = u/u_0, \tilde{e} = e/e_0, \tilde{p} = p/p_0, \widetilde{(\partial_x \phi)} = \partial_x \phi / (\partial_x \phi)_0$ , and we define the Mach number  $\text{Ma}$  and the Froude number  $\text{Fr}$  by  $\text{Ma} = u_0/c_0$  and  $\text{Fr} = u_0/\sqrt{L(\partial_x \phi)_0}$ . Following [38, 97], we consider a particular flow regime such that  $\text{Ma} = \text{Fr}$  so that the system (1) takes the following non-dimensional form

$$\partial_{\tilde{t}} \tilde{\rho} + \partial_{\tilde{x}} (\tilde{\rho} \tilde{u}) = 0, \quad \partial_{\tilde{t}} (\tilde{\rho} \tilde{u}) + \partial_{\tilde{x}} (\tilde{\rho} \tilde{u}^2) + \frac{1}{\text{Ma}^2} (\partial_{\tilde{x}} \tilde{p} + \tilde{\rho} \widetilde{(\partial_x \phi)}) = 0, \quad \partial_{\tilde{t}} (\tilde{\rho} \tilde{E}) + \partial_{\tilde{x}} (\tilde{\rho} \tilde{E} \tilde{u} + \tilde{p} \tilde{u}) = -\tilde{\rho} \tilde{u} \widetilde{(\partial_x \phi)}. \quad (64a)$$

Thanks to system (64), one can see that in the limit  $\text{Ma} \rightarrow 0$ , a singularity may appear in the momentum equation. Supposing now that  $\text{Ma} \ll 1$ , this suggests to distinguish two cases similarly as in [32]: in the first case the term  $\partial_{\tilde{x}} \tilde{p} + \tilde{\rho} \widetilde{(\partial_x \phi)}$  will always remain of magnitude  $O(\text{Ma}^2)$ , so that  $\tilde{\rho}, \tilde{u}$  and  $\tilde{E}$  will also remain of order  $O(\text{Ma}^0)$ . In this case, we will say that the system is in the low Mach regime. In the second case, the term  $\partial_{\tilde{x}} \tilde{p} + \tilde{\rho} \widetilde{(\partial_x \phi)}$  will not remain of magnitude  $O(\text{Ma}^2)$  in such way that  $\tilde{\rho} \tilde{u}$ , may experience large variations from  $O(\text{Ma}^2)$  to  $O(\text{Ma}^0)$ , yielding significant growth of  $\text{Ma}$  and thus a change in the Mach regime. These variations characterize all-regime flows with respect to the Mach number. Let us remark that the finer definition of well-prepared initial conditions used in [38] verifies the looser notion of low Mach regime considered in this work.

As it was mentioned earlier, the behavior of the Euler equations in the low Mach regime and adapted simulation strategies raise issues that have been intensively investigated for many years and are still very actively studied (see [25, 26, 43, 27, 44, 29, 31, 28, 36, 37, 30, 32, 33, 34, 35, 34, 38, 39, 40, 41, 42, 45, 49, 46, 47, 48] and the references therein). In this work, we propose transposing the low-Mach error analysis of the OSLP method presented in [32] to the FSLP scheme. This task is straightforward, although it requires lengthy and tedious calculations. Therefore, for the sake of brevity, we only recall the main points of this approach. We consider a non-dimensional expression of the FSLP solver for a one-dimensional problem and evaluate the truncation error obtained with a smooth solution of (64) that satisfies the low Mach regime hypothesis.  $\partial_{\tilde{x}} \tilde{p} + \tilde{\rho} \widetilde{(\partial_x \phi)} = O(\text{Ma}^2)$ . Similarly to the OSLP scheme, the magnitudes of the resulting truncation error estimates are uniform with respect to  $\text{Ma}$  except for the momentum equation that features an error term of order  $O(\theta \Delta x / \text{Ma})$ . Consequently, choosing  $\theta = O(\text{Ma})$  when  $\text{Ma} \ll 1$  will help the scheme preserve a uniform truncation error with respect to  $\text{Ma}$ . A well-known consequence of this choice is that in regions where  $\text{Ma} \ll 1$ , the non-centered part of the pressure term  $\Pi_{j+1/2}^{*,\theta}$  will be moderated.

The numerical tests proposed in sections 8.3, 8.6 and 8.7 show that this simple correction work similarly for both FSLP OSLP methods: in the low Mach regime, both schemes provide accurate results. Nevertheless, we need to emphasize that the modification of the scheme induced by  $\theta$  is not flawless and should be considered with care. Spurious oscillations may occur [58, 59] and the inequality (C.7) that ensures the entropy property of the scheme may not be verified in the limit  $\text{Ma} \rightarrow 0$ .

Let us finally highlight that as in [32, 1] the present approach is rather pragmatic and does not provide reliable analysis and explanation for the low Mach issues. Indeed, we do not study the delicate question of the asymptotic regime  $\text{Ma} \rightarrow 0$  [36, 37, 34, 38, 39, 40, 41, 42], we neither address the strong time step limitation due to the CFL

conditions (44) when  $\text{Ma} \ll 1$  that can be circumvented by using Implicit-Explicit strategies [32, 45, 46, 47, 48]. Moreover, the present lines are derived within a one-dimensional setting that does not allow fully expressing issues related to low Mach flows.

## 6.2. Extension to higher order

The FSLP algorithm can be implemented thanks to a simple single-step evaluation of numerical fluxes. This enables the use of classical high-order enhancements that are available in the literature for finite volume methods such as MUSCL-Hancock [89, 90, 91, 92, 98, 53], (W)ENO [93, 94] or MOOD [95, 96]. For the sake of simplicity, in this paper, we will only show numerical results with the MUSCL method for which the positivity can be proven under a half CFL condition. Let us consider a linear reconstruction of the primitive variables  $\mathbf{V} = (\rho, u, p)$  in each cells

$$\tilde{\mathbf{V}}_j^n(x) = \mathbf{V}_j^n + (x - x_j)\mathbf{p}_j^n$$

where the slopes  $\mathbf{p}_j^n = \mathbf{p}^n(\mathbf{V}_{j-1}^n, \mathbf{V}_j^n, \mathbf{V}_{j+1}^n)$  are obtained using a standard slope limiter such as the minmod function [99]. Let us introduce the function  $H: \mathbf{V} \mapsto \mathbf{U}$  that converts a state's conservative representation into its corresponding set of primitive variables. The reconstruction provides a second-order evaluation of the conserved quantities at each interface with

$$\mathbf{U}_{j+1/2,-}^{n,HO} = H(\tilde{\mathbf{V}}_j^n(x_{j+1/2})), \quad \mathbf{U}_{j-1/2,+}^{n,HO} = H(\tilde{\mathbf{V}}_{j-1}^n(x_{j-1/2})), \quad (65)$$

that we use to evaluate the FSLP flux function (34) at each interface by setting:

$$\mathbf{U}_j^{n+1} - \mathbf{U}_j^n + \frac{\Delta t}{\Delta x} (\mathbf{F}^{\text{FSLP}}(\mathbf{U}_{j+1/2,-}^{n,HO}, \mathbf{U}_{j+1/2,+}^{n,HO}) - \mathbf{F}^{\text{FSLP}}(\mathbf{U}_{j-1/2,-}^{n,HO}, \mathbf{U}_{j-1/2,+}^{n,HO})) = \Delta t \mathbf{S}_j(\mathbf{U}_{j+1/2,-}^{n,HO}, \mathbf{U}_{j+1/2,+}^{n,HO}, \mathbf{U}_{j-1/2,-}^{n,HO}, \mathbf{U}_{j-1/2,+}^{n,HO}). \quad (66)$$

The gravity source term can also be computed with the same formula as in the first-order method by replacing cell-averaged values with the high-precision face-centered values:

$$\mathbf{S}_j(\mathbf{U}_{j+1/2,-}^{n,HO}, \mathbf{U}_{j+1/2,+}^{n,HO}, \mathbf{U}_{j-1/2,-}^{n,HO}, \mathbf{U}_{j-1/2,+}^{n,HO}) = \begin{pmatrix} 0 \\ \{\rho \partial_x \phi\}_j^{n,HO} \\ \{\rho u \partial_x \phi\}_j^{n,HO} \end{pmatrix} \quad (67)$$

$$\begin{cases} \{\rho \partial_x \phi\}_j^{n,HO} = \frac{(\rho \partial_x \phi)_{j+1/2}^{HO} + (\rho \partial_x \phi)_{j-1/2}^{HO}}{2} \\ \{\rho u \partial_x \phi\}_j^{n,HO} = \frac{u_{j+1/2}^{*,HO} (\rho \partial_x \phi)_{j+1/2}^{HO} + u_{j-1/2}^{*,HO} (\rho \partial_x \phi)_{j-1/2}^{HO}}{2} \\ (\rho \partial_x \phi)_{j+1/2}^{HO} = \frac{\rho_{j+1/2,-}^{n,HO} + \rho_{j+1/2,+}^{n,HO}}{2} (\partial_x \phi)_{j+1/2}^{HO} \end{cases}$$

where  $(\partial_x \phi)_{j+1/2}^{HO}$  is a second-order accurate evaluation of the derivative of the gravitational potential at the interface  $x_{j+1/2}$ . Note that if the potential is known explicitly, it can be computed exactly at the interfaces. The extension of the well-balanced property is not straightforward and beyond the scope of this paper. Second-order well-balanced methods can be found in [97]. The second-order extension (66) of the FSLP scheme is positive for density and internal energy as long as it is ensured that:

$$\frac{\Delta t}{\Delta x} \max_{j \in \mathbb{Z}} \left( 2 \max(1/\rho_j^n, 1/\rho_{j+1}^n) a_{j+1/2} + \left(u_{j-\frac{1}{2}}^*\right)^+ - \left(u_{j+\frac{1}{2}}^*\right)^- \right) < \frac{1}{2} \quad (68)$$

The stability of the second-order method under the conditions above is a direct consequence of the stability of the first-order method. For the second-order extension in time, one can use either the SSP-RK2 method [100, 101] or a classical Hancock update [92].

### 6.3. Multidimensional extension

Before going any further, let us introduce the notations for our 2D space discretization: we consider two strictly increasing sequences  $(x_{i+1/2})_{i \in \mathbb{Z}}$  and  $(y_{j+1/2})_{j \in \mathbb{Z}}$  and divide the real plane into cells where the  $ij^{\text{th}}$  cell is the interval  $(x_{i-1/2}, x_{i+1/2}) \times (y_{j-1/2}, y_{j+1/2})$ . The space steps of the  $ij^{\text{th}}$  cell are  $\Delta x = x_{i+1/2} - x_{i-1/2} > 0$  and  $\Delta y_j = y_{j+1/2} - y_{j-1/2} > 0$ . We consider a discrete initial data  $\mathbf{U}_{ij}^0$  defined by  $\mathbf{U}_{ij}^0 = \frac{1}{\Delta x_i \Delta y_j} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} \mathbf{U}^0(x, y) dx dy$ , for  $(i, j) \in \mathbb{Z}^2$ . Let us introduce the Euler equations of gas dynamics in two dimensions of space:

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) + \partial_y \mathbf{G}(\mathbf{U}) = \mathbf{S}(\mathbf{U}, \phi), \quad \text{for } (x, y) \in \mathbb{R}^2, t > 0, \quad (69)$$

with  $\mathbf{U} = (\rho, \rho u, \rho v, \rho E)^T$ ,  $\mathbf{F}(\mathbf{U}) = (\rho u, \rho u^2 + p, \rho u v, \rho u E + p u)^T$ ,  $\mathbf{G}(\mathbf{U}) = (\rho v, \rho v u, \rho v^2 + p, \rho v E + p v)^T$ , and  $\mathbf{S}(\mathbf{U}, \phi) = -\rho \partial_x \phi(0, 1, 0, u)^T - \rho \partial_y \phi(0, 0, 1, v)^T$  where  $v$  is the velocity in the  $y$  direction and  $\phi$  is smooth enough so that we can consider that  $\partial_x \phi$ ,  $\partial_y \phi$  are also regular and bounded. We take advantage of the rotational invariance of the 2D Euler system and discretize the fluxes direction by direction:

$$\left\{ \begin{array}{l} \rho_{i,j}^{n+1} = \rho_{i,j}^n + \frac{\Delta t}{\Delta x} (u_{i+1/2,j}^* \rho_{i+1/2,j}^n - u_{i-1/2,j}^* \rho_{i-1/2,j}^n) + \frac{\Delta t}{\Delta y} (v_{i,j+1/2}^* \rho_{i,j+1/2}^n - v_{i,j-1/2}^* \rho_{i,j-1/2}^n), \\ (\rho u)_{i,j}^{n+1} = (\rho u)_{i,j}^n + \frac{\Delta t}{\Delta x} (u_{i+1/2,j}^* (\rho u)_{i+1/2,j}^n + \Pi_{i+1/2,j}^{\theta^x,*} - u_{i-1/2,j}^* (\rho u)_{i-1/2,j}^n - \Pi_{i-1/2,j}^{\theta^x,*}) \\ \quad + \frac{\Delta t}{\Delta y} (v_{i,j+1/2}^* (\rho u)_{i,j+1/2}^n - v_{i,j-1/2}^* (\rho u)_{i,j-1/2}^n) - \{\rho \partial_x \phi\}_{i,j}^n, \\ (\rho v)_{i,j}^{n+1} = (\rho v)_{i,j}^n + \frac{\Delta t}{\Delta x} (u_{i+1/2,j}^* (\rho v)_{i+1/2,j}^n - u_{i-1/2,j}^* (\rho v)_{i-1/2,j}^n) \\ \quad + \frac{\Delta t}{\Delta y} (v_{i,j+1/2}^* (\rho v)_{i,j+1/2}^n + \Pi_{i,j+1/2}^{\theta^y,*} - v_{i,j-1/2}^* (\rho v)_{i,j-1/2}^n - \Pi_{i,j-1/2}^{\theta^y,*}) - \{\rho \partial_y \phi\}_{i,j}^n, \\ (\rho E)_{i,j}^{n+1} = (\rho E)_{i,j}^n + \frac{\Delta t}{\Delta x} (u_{i+1/2,j}^* (\rho E)_{i+1/2,j}^n + \Pi_{i+1/2,j}^{\theta^x,*} u_{i+1/2,j}^* - u_{i-1/2,j}^* (\rho E)_{i-1/2,j}^n - \Pi_{i-1/2,j}^{\theta^x,*} u_{i-1/2,j}^*) \\ \quad + \frac{\Delta t}{\Delta y} (v_{i,j+1/2}^* (\rho E)_{i,j+1/2}^n + \Pi_{i,j+1/2}^{\theta^y,*} v_{i,j+1/2}^* - v_{i,j-1/2}^* (\rho E)_{i,j-1/2}^n - \Pi_{i,j-1/2}^{\theta^y,*} v_{i,j-1/2}^*) - \{(\rho u \partial_x + \rho v \partial_y) \phi\}_{i,j}^n. \end{array} \right. \quad (70)$$

with

$$\left\{ \begin{array}{l} u_{i+1/2,j}^* = \frac{u_{i+1,j}^n + u_{i,j}^n}{2} - \frac{1}{2a_{i+1/2,j}} (p_{i+1,j}^n - p_{i,j}^n + \frac{\rho_{i+1,j}^n + \rho_{i,j}^n}{2} (\phi_{i+1,j} - \phi_{i,j})), \\ v_{i+1/2,j}^* = \frac{v_{i,j+1}^n + v_{i,j}^n}{2} - \frac{1}{2a_{i,j+1/2}} (p_{i,j+1}^n - p_{i,j}^n + \frac{\rho_{i,j+1}^n + \rho_{i,j}^n}{2} (\phi_{i,j+1} - \phi_{i,j})), \\ \Pi_{i+1/2,j}^{\theta^x,*} = \frac{p_{i+1,j}^n + p_{i,j}^n}{2} - \theta_{i+1/2,j}^x \frac{a_{i+1/2,j}}{2} (u_{i+1,j}^n - u_{i,j}^n), \\ \Pi_{i,j+1/2}^{\theta^y,*} = \frac{p_{i,j+1}^n + p_{i,j}^n}{2} - \theta_{i,j+1/2}^y \frac{a_{i,j+1/2}}{2} (u_{i,j+1}^n - u_{i,j}^n), \end{array} \right. \quad (71)$$

as well as the source terms discretization:

$$\left\{ \begin{array}{l} \{\rho \partial_x \phi\}_{i,j}^n = \frac{(\rho \partial_x \phi)_{i+1/2,j} + (\rho \partial_x \phi)_{i-1/2,j}}{2}, \\ \{\rho \partial_y \phi\}_{i,j}^n = \frac{(\rho \partial_y \phi)_{i,j+1/2} + (\rho \partial_y \phi)_{i,j-1/2}}{2}, \\ \{\rho u \partial_x \phi\}_{i,j}^n = \frac{u_{i+1/2,j}^* (\rho \partial_x \phi)_{i+1/2,j} + u_{i-1/2,j}^* (\rho \partial_x \phi)_{i-1/2,j}}{2}, \\ \{\rho u \partial_y \phi\}_{i,j}^n = \frac{v_{i,j+1/2}^* (\rho \partial_y \phi)_{i,j+1/2} + v_{i,j-1/2}^* (\rho \partial_y \phi)_{i,j-1/2}}{2}, \\ (\rho \partial_x \phi)_{i+1/2,j} = \frac{\rho_{j+1}^n + \rho_{j,j}^n}{2} \frac{\phi_{i+1,j} - \phi_{i,j}}{\Delta x}, \\ (\rho \partial_y \phi)_{i,j+1/2} = \frac{\rho_{i,j+1}^n + \rho_{i,j}^n}{2} \frac{\phi_{i,j+1} - \phi_{i,j}}{\Delta y}. \end{array} \right. \quad (72)$$

## 7. flux-splitting as a relaxation approximation

The goal of this section is to highlight the connection between the FSLP flux-splitting approach and a relaxation approximation. In the previous sections, we concluded that the FSLP approach could be expressed as an averaging



procedure (43) where  $U_j^P$  and  $U_j^A$  are defined as approximate solutions of two systems (39) and (40) that respectively only account for the pressure and the advection effects. We propose to translate that three-step process thanks to a relaxation approximation. Suppose that  $\alpha \in (0, 1)$  is a constant and let  $\nu$  be a positive parameter, we consider the system

$$\partial_t \begin{bmatrix} \rho^P \\ \rho^P u^P \\ \rho^P E^P \\ \rho^P \Pi^P \\ \rho^P \mathcal{T}^P \\ \phi \end{bmatrix} + \frac{1}{\alpha} \partial_x \begin{bmatrix} 0 \\ \Pi^P \\ \Pi^P u^P \\ a^2 u^P \\ -u^P \\ 0 \end{bmatrix} + \frac{1}{\alpha} \begin{bmatrix} 0 \\ \rho^P \\ \rho^P u^P \\ 0 \\ 0 \\ 0 \end{bmatrix} \partial_x \phi = \nu \begin{bmatrix} \alpha \rho^P + (1-\alpha)\rho^A - \rho^P \\ \alpha \rho^P u^P + (1-\alpha)\rho^A u^A - \rho^P u^P \\ \alpha \rho^P E^P + (1-\alpha)\rho^A E^A - \rho^P E^P \\ p^{\text{EOS}}(1/\rho^P, e^P) - \Pi^P \\ 1 - \rho^P \mathcal{T}^P \\ 0 \end{bmatrix}, \quad (73a_v)$$

$$\partial_t \begin{bmatrix} \rho^A \\ \rho^A u^A \\ \rho^A E^A \\ \rho^A \Pi^A \\ \rho^A \mathcal{T}^A \end{bmatrix} + \left( \frac{1}{1-\alpha} \right) \partial_x \begin{bmatrix} \rho^A u^P \\ \rho^A u^A u^P \\ \rho^A E^A u^P \\ \rho^A \Pi^A u^P \\ u^P \end{bmatrix} = \nu \begin{bmatrix} \alpha \rho^P + (1-\alpha)\rho^A - \rho^A \\ \alpha \rho^P u^P + (1-\alpha)\rho^A u^A - \rho^A u^A \\ \alpha \rho^P E^P + (1-\alpha)\rho^A E^A - \rho^A E^A \\ p^{\text{EOS}}(1/\rho^A, e^A) - \Pi^A \\ 1 - \rho^A \mathcal{T}^A \end{bmatrix}. \quad (73b_v)$$

The system (73<sub>v</sub>) features a pair of duplicate conservative variables ( $U^P, U^A$ ) and 4 other variables:  $\Pi^P, \Pi^A, \mathcal{T}^A$  and  $\mathcal{T}^P$ . The variables  $\Pi^P$  and  $\Pi^A$  are surrogate for the thermodynamical pressure, while  $\mathcal{T}^A$  and  $\mathcal{T}^P$  play the role of a pseudo-specific volume. It is possible to view (73<sub>v</sub>) as a Suliciu relaxation approximation with a separation of the acoustic and transport operators. Indeed, (73<sub>v</sub>) implies that

$$\partial_t \begin{bmatrix} \alpha \rho^P + (1-\alpha)\rho^A \\ \alpha \rho^P u^P + (1-\alpha)\rho^A u^A \\ \alpha \rho^P E^P + (1-\alpha)\rho^A E^A \end{bmatrix} + \partial_x \begin{bmatrix} \rho^A u^P \\ \rho^A u^A u^P + \Pi^P \\ \rho^A E^A u^P + \Pi^P u^P \end{bmatrix} = \begin{bmatrix} 0 \\ -\rho^P \\ -\rho^P u^P \end{bmatrix} \partial_x \phi \quad (74a)$$

$$\partial_t [\alpha \rho^P \Pi^P + (1-\alpha)\rho^A \Pi^A] + \partial_x (\rho^A \Pi^A u^P + a^2 u^P) = \nu \left[ \alpha p^{\text{EOS}}\left(\frac{1}{\rho^P}, e^P\right) + (1-\alpha)p^{\text{EOS}}\left(\frac{1}{\rho^A}, e^A\right) - \alpha \Pi^P - (1-\alpha)\Pi^A \right], \quad (74b)$$

$$\partial_t [\alpha \rho^P \mathcal{T}^P + (1-\alpha)\rho^A \mathcal{T}^A] = \nu [1 - \alpha \rho^P \mathcal{T}^P - (1-\alpha)\rho^A \mathcal{T}^A]. \quad (74c)$$

Taking the limit  $\nu \rightarrow +\infty$  formally enforces that  $U^P = U^A = U = (\rho, \rho u, \rho E)^T$  and  $\Pi^A = \Pi^P = p^{\text{EOS}}(1/\rho, e)$ , so that (74a) enables to retrieve the Euler system (1). This suggests that we can use the relaxation system (73<sub>v</sub>) as an approximation of (1) in the limit  $\nu \rightarrow +\infty$ . The equation (74b) plays here a similar role as the surrogate pressure equation in the classic Suliciu approximation [60, 62, 63]. The sole purpose of equation (74c) is to ensure that  $\alpha \rho^P \mathcal{T}^P + (1-\alpha)\rho^A \mathcal{T}^A = 1$  in the regime  $\nu \rightarrow \infty$ . In our discretization strategy, we classically mimic the  $\nu \rightarrow \infty$  regime for  $t \in [t^n, t^{n+1})$ , by enforcing  $(U^P, \Pi^P, \mathcal{T}^A, U^A, \Pi^A, \mathcal{T}^A)(t = t^n) = (U, p^{\text{EOS}}(1/\rho, e), 1/\rho, U, p^{\text{EOS}}(1/\rho, e), 1/\rho)(t = t^n)$  and by solving the relaxation off-equilibrium system (73<sub>v=0</sub>). The properties of the off-equilibrium system (73<sub>v=0</sub>) are briefly summarized in the following proposition whose proof is given in Appendix D.

**Proposition 7.1.** *The system (73<sub>v=0</sub>) is hyperbolic with a set of characteristic velocities given by:  $\frac{u^P}{1-\alpha}$  (with an algebraic multiplicity 4), 0 (with an algebraic multiplicity 5) and  $\pm \frac{a}{\alpha \rho^P}$ . Moreover, (73<sub>v=0</sub>) only involves linearly degenerate fields.*

The relaxation formulation (73<sub>v</sub>) sheds some more light on the similarities between the flux-splitting we propose here and the acoustic/transport operator splitting strategy presented in [1]. Indeed, the source term and pressure effects can be treated separately from the advection terms. The difference is that although the operators are separated, they are re-distributed within a larger single system instead of two separate systems.

By discretizing the pressure and advection parts of (9) identically than in section 5.2, we re-obtain the same update formula (31), which yields the FSLP scheme (34). Finally, let us mention that it is possible to build an alternate flux-splitting method for the system (1) by seeking the solution of the Riemann problem for (73<sub>v=0</sub>). This option is not studied in the present work.

## 8. Numerical experiments

In this section, we consider that the fluid is a perfect gas with the EOS  $p = (\gamma - 1)\rho e$  and that the potential  $\phi$  takes the form  $\phi(x, y) = -gy$  for tests that involve the source term.

We will present numerical experiments with the FSLP method and the HLLC Riemann solver [92] using first and second-order discretizations. The second-order accuracy is achieved using a MUSCL-Hancock strategy [92] for both the HLLC and FSLP solvers. Let us mention that the slope reconstruction is performed on the primitive variables with a minmod slope limiter [54, 53, 92]. For the OSLP method. Noting  $(a/\rho)_{j+1/2} = \max(1/\rho_j^n, 1/\rho_{j+1}^n)a_{j+1/2}$ , the time steps  $\Delta t$  is computed as follows:

$$\Delta t = C^{\text{CFL}} \Delta x \frac{1}{\max_{j \in \mathbb{Z}} \left[ \max \left\{ 2 \max \left[ (a/\rho)_{j-1/2}, (a/\rho)_{j+1/2} \right], \left( \left( u_{j-1/2}^* \right)^+ - \left( u_{j+1/2}^* \right)^- \right) \right\} \right]}, \quad (75)$$

For the FSLP method, it is computed as follows:

$$\Delta t = C^{\text{CFL}} \Delta x \frac{1}{\max_{j \in \mathbb{Z}} \left[ 2 \max \left[ (a/\rho)_{j-1/2}, (a/\rho)_{j+1/2} \right] + \left( \left( u_{j-1/2}^* \right)^+ - \left( u_{j+1/2}^* \right)^- \right) \right]}, \quad (76)$$

where the parameter  $C^{\text{CFL}}$  is given by the table 1 so that the CFL conditions (20) and(21) for the OSLP method, (33) for the first-order FSLP method and (68) for the second-order FSLP method are all checked. For the HLLC solver, the standard CFL from [102] is used.

Numerical scheme	first-order	second-order
OSLP	1.0	N.A.
FSLP	1.0	1/2
HLLC	1.0	1/2

Table 1: Values for  $C^{\text{CFL}}$  used in the simulations.

The parameter  $\theta$  related to the low Mach correction is defined at each interface  $(i + 1/2, j)$  and  $(i, j + 1/2)$  by

$$\theta_{i+1/2,j}^x = \max(u_{i,j}/c_i, u_{i+1,j}/c_{i+1,j}), \quad \theta_{i,j+1/2}^y = \max(v_{i,j}/c_i, u_{i,j+1}/c_{i,j+1}). \quad (77)$$

### 8.1. Sod shock tube test case

We consider here the classical Sod shock tube test case [103, 92]: we set  $\gamma = 1.4$  and the initial conditions are:

$$(\rho, u, p)(x, t = 0) = \begin{cases} (1, 0, 1) & \text{if } x < 0.5, \\ (0.125, 0, 0.1) & \text{if } x > 0.5 \end{cases}$$

The goal of this test is to study the ability of our solver to handle different wave types. The initial discontinuity generates three waves: a leftward going rarefaction, a contact discontinuity, and a shock that both travel towards the right of the computational domain. Figure 1 shows the profile obtained at  $t = 0.2s$  with five different solvers: OSLP, FSLP/HLLC for the first and second-order methods. At first order, the HLLC solver provides the sharpest resolution of the shock and contact discontinuity. The differences between the FSLP and OSLP methods are hardly visible. None of the schemes suffers from spurious oscillations and both the position and the amplitude of the waves match the exact solution.

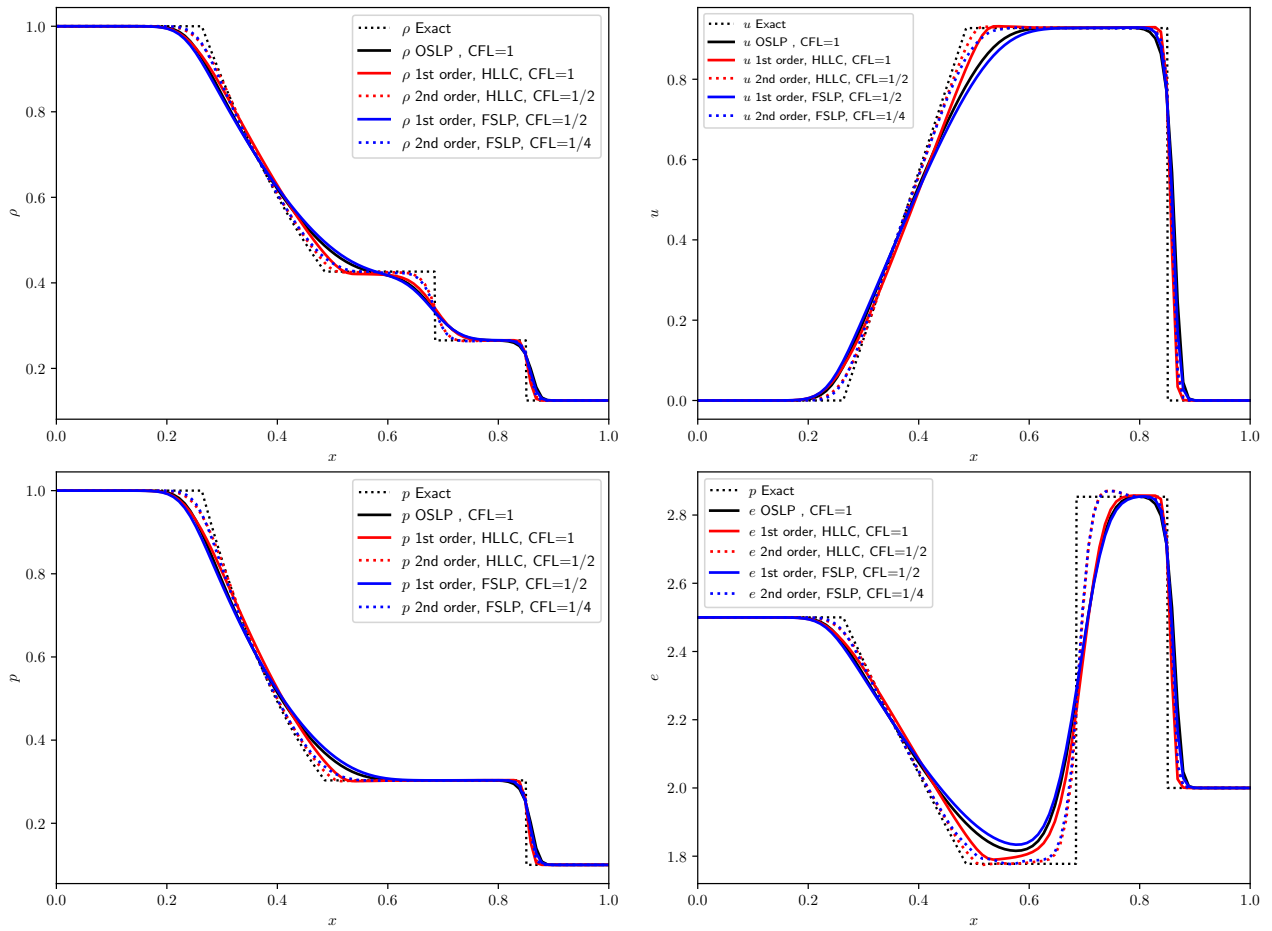


Figure 1: Sod shock tube test case. Profile at  $t = 0.2s$  of the density (top left), velocity (top right), pressure (bottom left), and specific internal energy (bottom right). The results are obtained with the OSLP method, first and second-order FSLP method, first and second-order HLLC scheme, and the exact solution on a 100-cell grid.

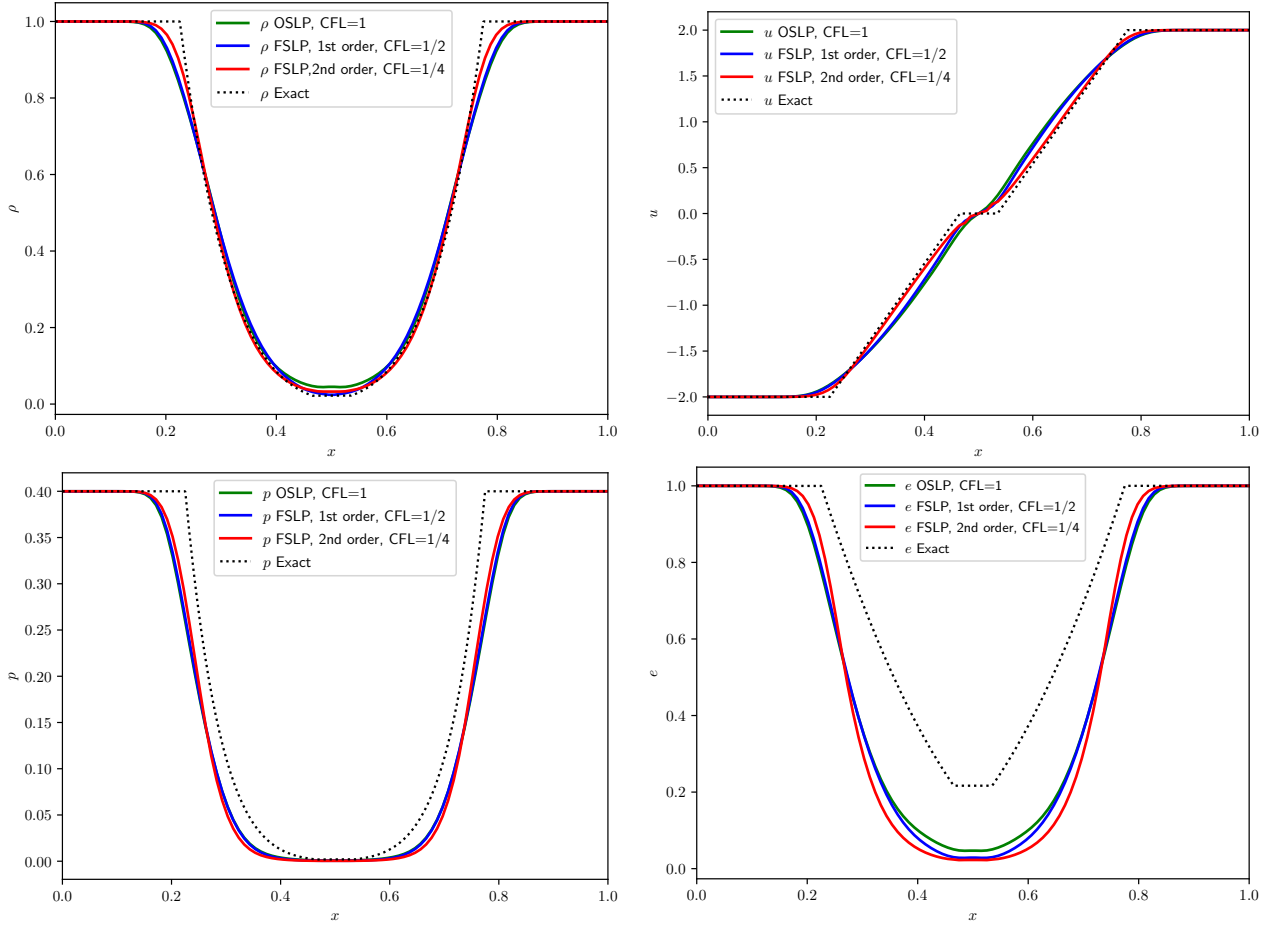


Figure 2: Two-rarefaction test case. Profile at  $t = 0.1s$  of the density (top left), velocity (top right), pressure (bottom left), and specific internal energy (bottom right). The results are obtained with the OSLP method, the first and second-order FSLP method, and the exact solution on a 100-cell grid.

### 8.2. Two-rarefaction test case

We now consider the two-rarefaction test proposed by Einfeldt [104, 92] for a perfect gas with  $\gamma = 1.4$ . The initial conditions are

$$(\rho, u, p)(x, t = 0) = \begin{cases} (1, -2, 0.4), & \text{if } x < 0.5, \\ (1, 2, 0.4), & \text{if } x > 0.5. \end{cases}$$

The resulting wave pattern features two rarefaction waves that split from the position  $x = 0.5$ , traveling towards each end of the computational domain. As a result, a near vacuum region presenting low densities and pressures appears in the middle of the domain.

Figure 2 shows that all methods are robust enough to preserve positivity for mass, pressure, and energy so that they are able to reach the end of the simulation. Moreover, none of the numerical schemes exhibit entropy-related issues like the apparition of nonphysical shocks within the wave pattern.

### 8.3. The Gresho Vortex

The Gresho vortex [105] involves a stationary vortex that can be parameterized by the maximum value of the Mach number  $M_a$  across the computational domain. Therefore this test is very useful for studying the performance of numerical schemes in the low-Mach regime. We consider a perfect gas with  $\gamma = 1.4$ . Using polar coordinates  $(r, \theta)$ ,

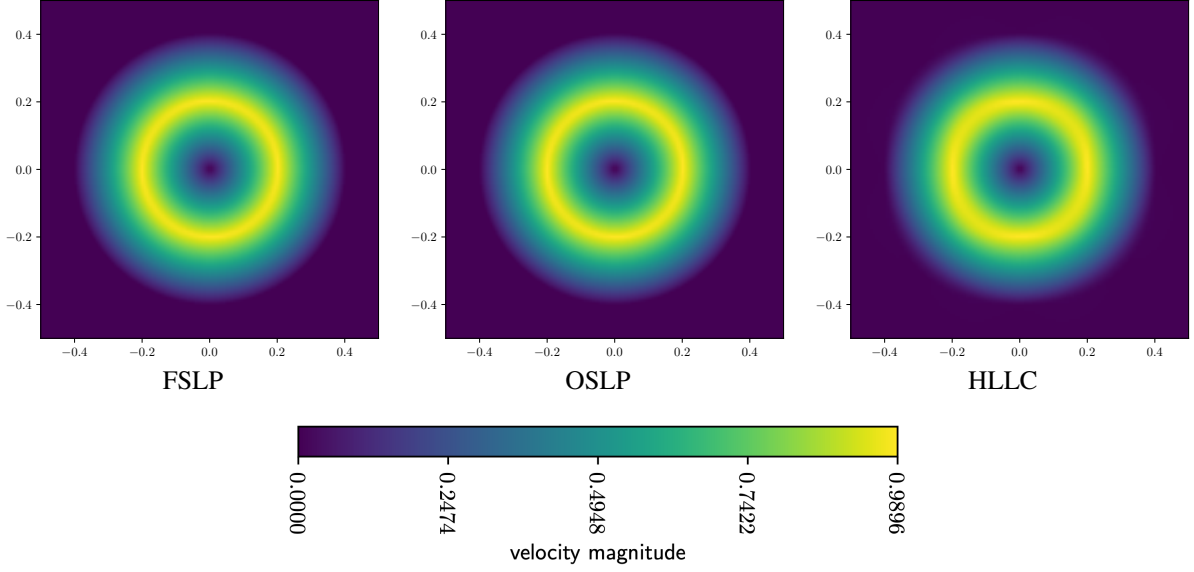


Figure 3: Comparison of the final velocity magnitude map for the Gresho vortex test case with  $Ma = 10^{-1}$  obtained with the FSLP, OSLP, and HLLC solvers on a  $128 \times 128$ -cell grid at  $t = 0.1s$ .

the initial conditions read:

$$\rho(r, \theta, t = 0) = 1, \quad (78a)$$

$$(u_r, u_\theta)(r, \theta, t = 0) = \begin{cases} (0, 5r) & \text{if } 0 \leq r < 0.2, \\ (0, 2 - 5r) & \text{if } 0.2 \leq r < 0.4, \\ (0, 0) & \text{if } 0.4 \leq r, \end{cases} \quad (78b)$$

$$p(r, \theta, t = 0) = \begin{cases} p_0 + 12.5r^2 & \text{if } 0 \leq r < 0.2, \\ p_0 + 12.5r^2 + 4 - 20r + 4 \ln(5r) & \text{if } 0.2 \leq r < 0.4, \\ p_0 - 2 + 4 \ln 2 & \text{if } 0.4 \leq r, \end{cases} \quad (78c)$$

where  $p_0 = \frac{1}{\gamma M_a^2}$ . For the simulations, we will use three different values for the reference Mach number:  $Ma \in \{10^{-1}, 10^{-3}, 10^{-5}\}$ . We will compare the distributions of the velocity magnitude obtained at  $t = 10^{-2}s$  with the initial conditions.

Figures 3, 4, 5 give us the final velocity magnitude map for the Gresho vortex obtained with different solvers and Mach numbers. For  $Ma = 10^{-1}$ , we can see in figure 3 that on all three simulations, the initial velocity ring is preserved. Figure 4 displays the results for  $Ma = 10^{-3}$ : one can see that the FSLP and OSLP methods can both preserve the velocity ring thanks to the low Mach correction while the HLLC methods fail to do so. The same behavior is observed for  $Ma = 10^{-5}$  (see figure 5). In order to measure the effect of the solver's numerical diffusion, we evaluate the ratio  $e_{\text{kin}}/e_{\text{kin}}^0$  of the kinetic energy obtained at the final instant and the initial instant with

$$e_{\text{kin}} = \sum_j \frac{1}{2} \rho_j^n (u_j^n)^2 + (v_j^n)^2 \Delta x^2, \quad e_{\text{kin}}^0 = \sum_j \frac{1}{2} \rho_j^0 ((u_j^0)^2 + (v_j^0)^2) \Delta x^2. \quad (79)$$

The results are displayed in table 2. They show that both FSLP and OSLP solvers better preserve the kinetic energy than the HLLC method in the low Mach regime.

#### 8.4. Two-dimensional Riemann problems

We now intend to study the ability of the FSLP method to capture more complex wave patterns in a two-dimensional setting, including shocks and rarefaction waves. To that end, we consider the popular 2D Riemann prob-

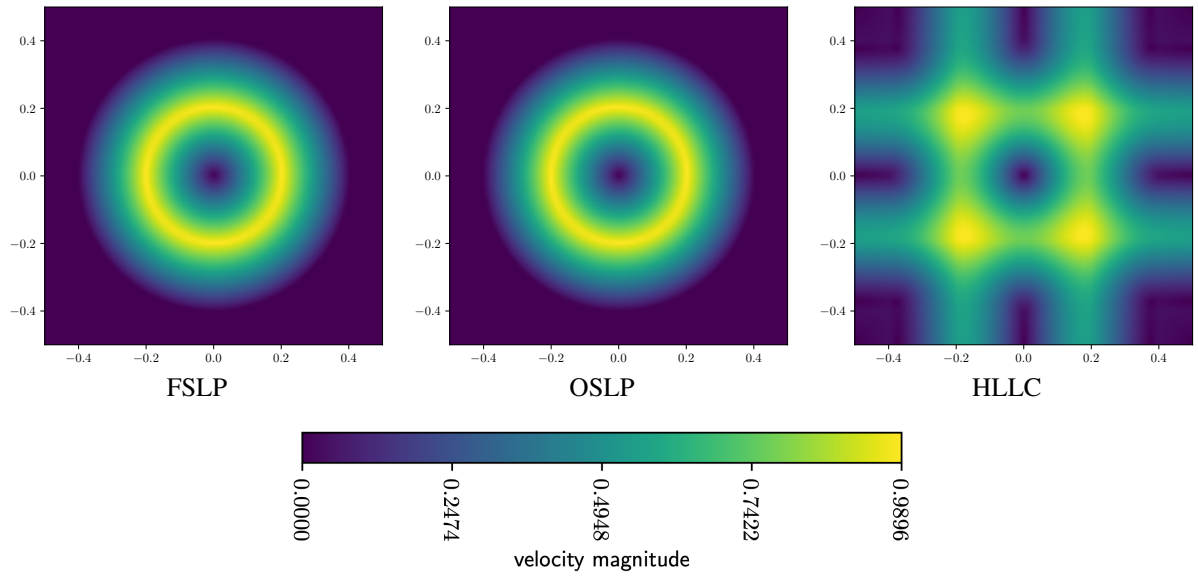


Figure 4: Comparison of the final velocity magnitude map for the Gresho vortex test case with  $Ma = 10^{-3}$  obtained with the FSLP, OSLP, and HLLC solvers on a  $128 \times 128$ -cell grid at  $t = 0.1s$ .

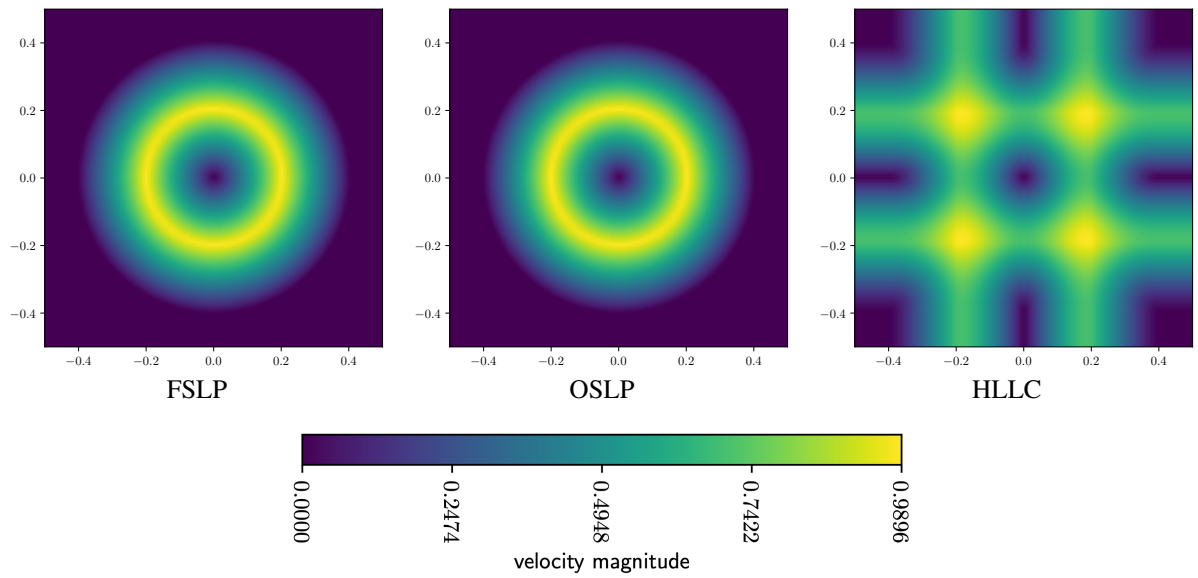


Figure 5: Comparison of the final velocity magnitude map for the Gresho vortex test case with  $Ma = 10^{-5}$  obtained with the FSLP, OSLP, and HLLC solvers on a  $128 \times 128$ -cell grid at  $t = 0.1s$ .

Table 2: Gresho vortex test case: evaluation of the kinetic energy in the computational domain for different values of the Mach number  $Ma$ .

	$Ma = 10^{-1}$	$Ma = 10^{-3}$	$Ma = 10^{-5}$
$e_{kin}/e_{kin}^0$ (at $t = 10^{-2}$ ) — OSLP scheme	0.9966	0.9966	0.9966
$e_{kin}/e_{kin}^0$ (at $t = 10^{-2}$ ) — FSLP scheme	0.9966	0.9966	0.9966
$e_{kin}/e_{kin}^0$ (at $t = 10^{-2}$ ) — HLLC scheme	0.9762	0.5262	0.5167

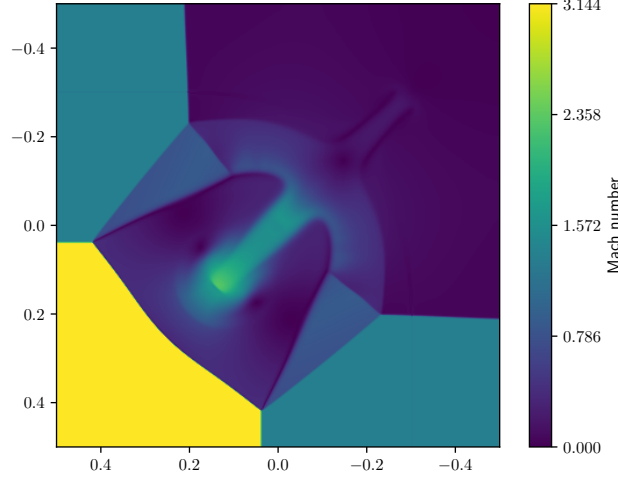


Figure 6: 2D Riemann problem. Mapping of the Mach number as  $t = 0.8s$ . The reference simulation is obtained with a second-order HLLC method on a  $384 \times 384$ -cell mesh.

s

lem of the literature referred to as Configuration 3 in [106]. The computational domain is the rectangle  $[0, 1] \times [0, 1]$ , with the initial conditions

$$(\rho, u, v, p)(x, y, t = 0) = \begin{cases} (0.138, 1.206, 1.206, 0.029) & \text{if } x < 0.8, y < 0.8 \quad (\text{bottom left}) \\ (0.5323, 0.0, 1.206, 0.3) & \text{if } x > 0.8, y < 0.8 \quad (\text{bottom right}) \\ (0.5323, 1.206, 0.0, 0.3) & \text{if } x < 0.8, y > 0.8 \quad (\text{top left}) \\ (1.5, 0.0, 0.0, 1.5) & \text{if } x > 0.8, y > 0.8 \quad (\text{top right}). \end{cases} \quad (80)$$

We impose homogeneous Neumann conditions at the boundaries. We compute a reference solution thanks to a second-order HLLC method on a  $384 \times 384$ -grid. The waves at play produce a jet that propagates along the diagonal  $x = y$  creating an important low Mach region in the center and the top right part of the domain (see figure 6).

Figure 7 shows a mapping of the density obtained with the OSLP (first-order), the FSLP (first and second-order), and the HLLC (first and second-order) schemes using a  $128 \times 128$ -cell mesh. One can see that the overall wave pattern is rendered successfully by all numerical schemes. The results of the FSLP, OSLP and HLLC schemes for first-order methods are similar. second-order methods all better succeed in capturing the shape of the jet as depicted in figure 8. Nevertheless, we can note that spurious oscillations appear in the simulation performed with the second-order FSLP scheme with low-Mach correction. These spurious waves propagate along the  $x$  and  $y$  axes in the top right part of the domain. We believe they are caused by the lack of numerical dissipation around the low-Mach shocks due to the combination of the low-Mach correction and the second-order reconstruction. A more careful choice of  $\theta$  than (77) is required to ensure the discrete entropy inequality (see (22)). Improving the second-order discretization for the FSLP scheme would, for example, require proposing a better choice for  $\theta$  but such a task is beyond the scope of the present work.

### 8.5. Hydrostatic equilibrium test

In order to challenge the well-balanced ability of the FSLP scheme, we consider the atmosphere at rest test (see, for example [1]). It involves a fluid column of a perfect gas in a rectangular  $[0, 2] \times [0, 1]$  domain. For this test, the

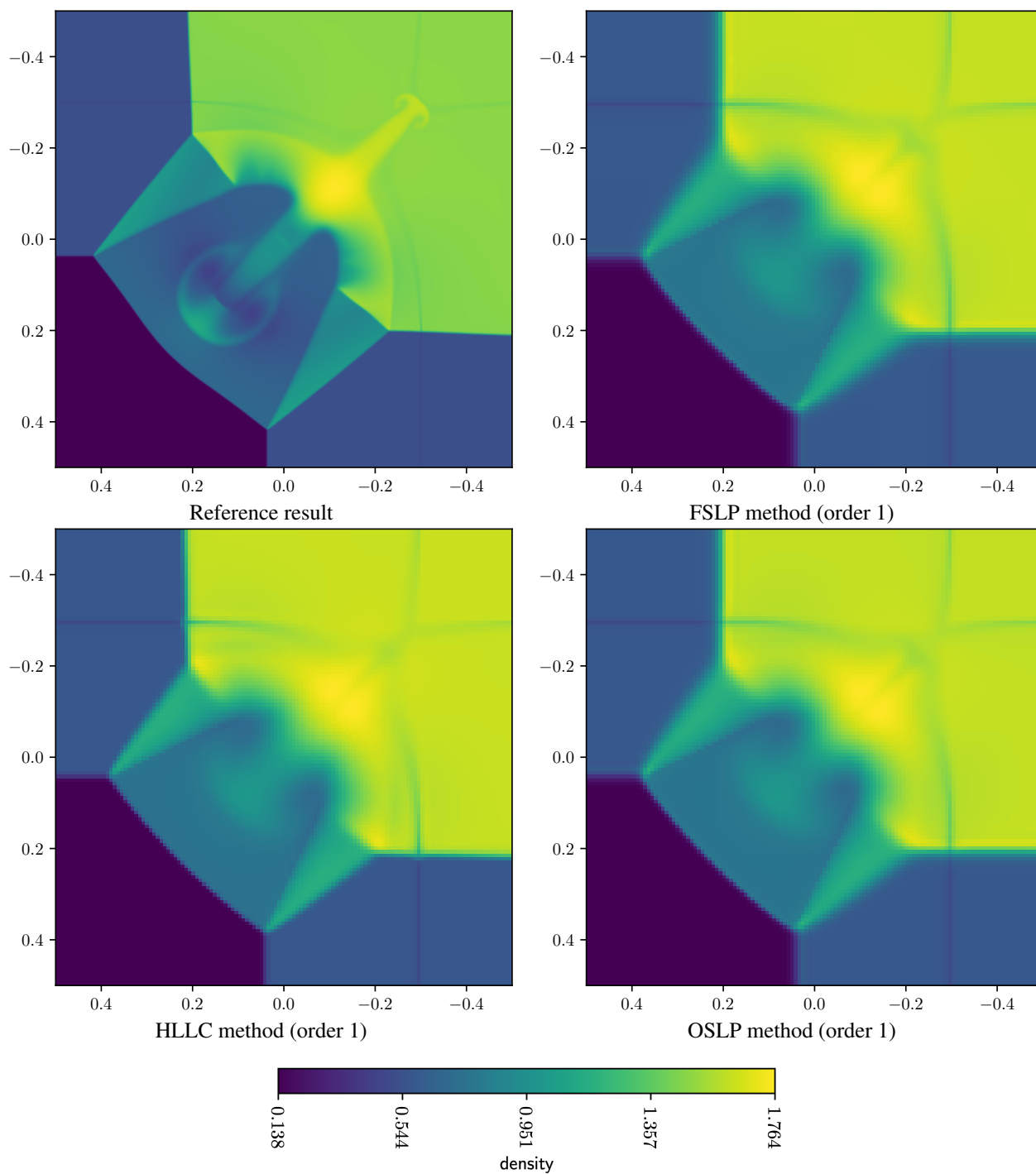


Figure 7: 2D Riemann problem. Mapping of the density number as  $t = 0.8s$ . The reference simulation is obtained with a second-order HLLC method on a  $384 \times 384$ -cell mesh. The other simulations are performed on a  $128 \times 128$ -cell grid with the first-order FSLP method (top right), the HLLC first-order method (bottom left), and the first-order OSLP method (bottom right).



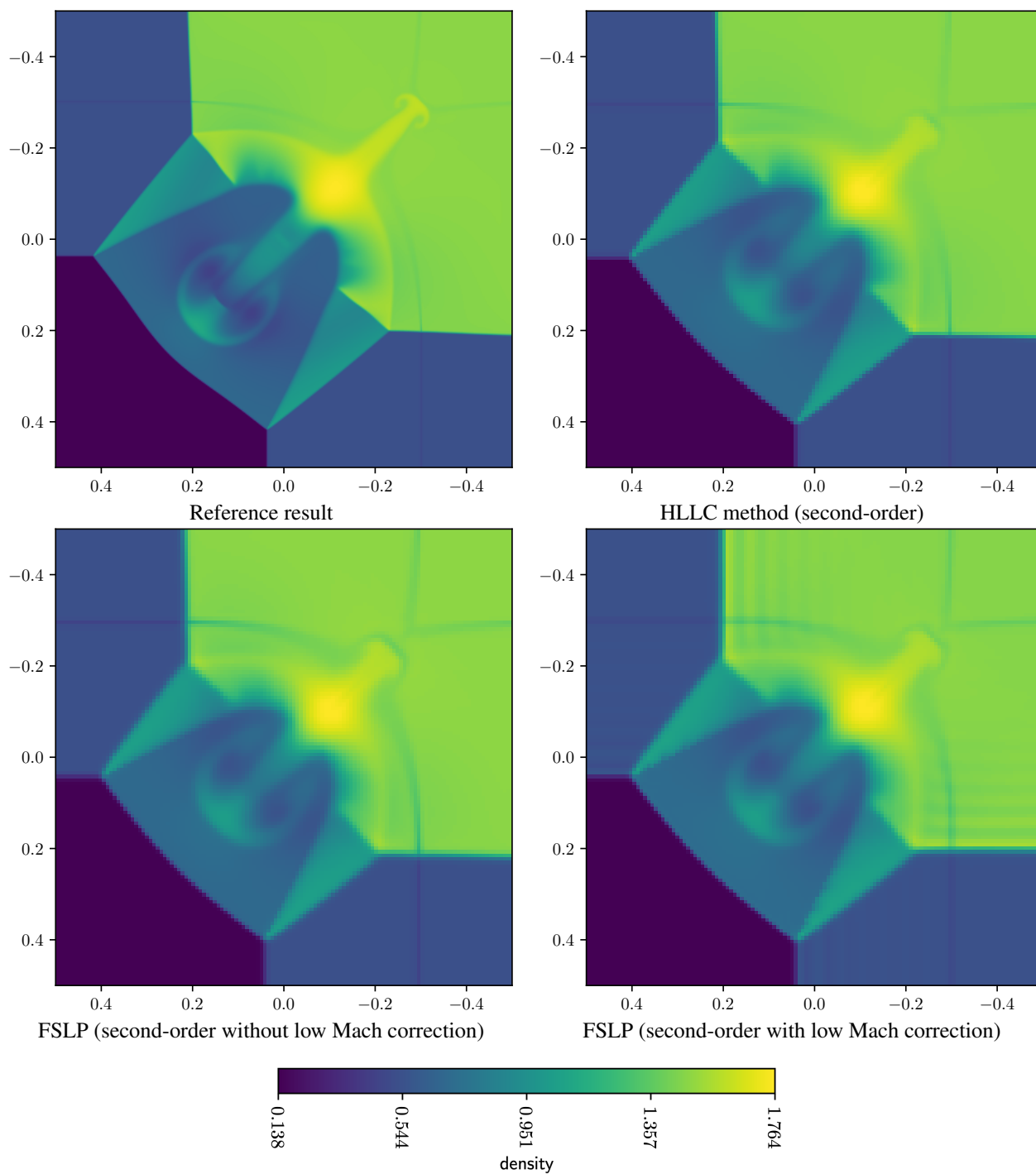


Figure 8: 2D Riemann problem. Mapping of the density number as  $t = 0.8s$ . The reference simulation is obtained with a second-order HLLC method on a  $384 \times 384$ -cell mesh. The other simulations are performed on a  $128 \times 128$ -cell grid with the second-order HLLC method (top right), the FSLP second-order method without low-Mach correction (bottom left), the second-order FSLP method with low-Mach correction (bottom right).

gravity acceleration is set to  $g = -1$  so that  $\phi(x, y) = -y$ . For the EOS of the fluid, we set  $\gamma = 5/3$  and  $c_v = 1$ , where  $c_v$  is the heat capacity at constant volume so that the temperature  $T$  of the gas is given by  $e = c_v T$ . We consider periodic boundary conditions for the left and right sides of the domain. At the top and bottom of the domain, wall boundaries are imposed for the normal velocity, while the temperature is linearly extrapolated. The initial condition is built by imposing a linear temperature profile as follows

$$T(x, y = 0, t = 0) = 3.78565, \quad \nabla T(x, y, t = 0) = (0, -1.2)^T, \quad (81a)$$

$$\rho(x, y = 0, t = 0) = 1, \quad \nabla(c_v(\gamma - 1)\rho T)(x, y, t = 0) = (0, \rho g)^T. \quad (81b)$$

The computational domain is discretized over a  $100 \times 50$  on which we let the solver evolve the profile for  $t \in [0, 100s]$ . Table 3 displays the value of the  $\max_{i,j} |v_{i,j}^n|$  at  $t = 100s$  and shows that both the OSLP and the FSLP first-order methods preserve the velocity magnitude at zero-machine precision.

Table 3: Hydrostatic equilibrium test. Measure of the velocity magnitude at  $t = 100s$

Solver	OSLP	FSLP
Average speed	$1.342 \times 10^{-14}$	$2.056 \times 10^{-14}$

It is important to mention that a direct second-order extension of the well-balanced method, as presented in section 6.2 will fail to preserve the hydrostatic equilibrium. This question of designing a well-balanced high-order method has been successfully investigated in the literature [19, 22, 24]. Adapting these techniques to the FSLP scheme is beyond the scope of this paper.

### 8.6. Rayleigh-Taylor instability

We now consider the Rayleigh-Taylor test performed in [1]: the computational domain is  $[-1/4, 1/4] \times [-3/4, 3/4]$  and the fluid is a perfect gas with  $\gamma = 5/3$ . At  $t = 0$  a dense layer of fluid lies on top of a lighter layer so that the configuration is unstable. The gravity acceleration is  $g = -0.1$  thus  $\phi(x, y) = -0.1 \times y$ . The initial conditions are given by

$$\rho(x, y, t = 0) = \begin{cases} 1 & \text{for } y < 0, \\ 2 & \text{for } y \geq 0, \end{cases} \quad (82a)$$

$$p(x, y, t = 0) = -\rho\phi, \quad (82b)$$

$$(u, v)(x, y, t = 0) = \left(0, \frac{C}{4} (1 + \cos(4\pi x)) (1 + \cos(3\pi y))\right). \quad (82c)$$

The initial velocity (82c) imposes a single-mode perturbation of magnitude  $C = 0.01$  that will break the hydrostatic equilibrium.

This test allows measuring and comparing the effect of the numerical diffusion of each method as it tends to limit the development of high-frequency modes in the instability. Figure 10 and 9 respectively show the density and Mach number of a reference second-order HLLC simulation obtained with a  $200 \times 600$ -cell mesh. We observe a sharp transition between both fluid layers, and the interface presents lateral arms with secondary rolls.

Figure 10 shows simulations ran with both the FSLP solver and the HLLC solver on a coarse  $50 \times 150$ -cell mesh obtained with first and second-order methods. The HLLC method presents an important amount of numerical diffusion: it only shows a single mode growth, and no lateral arm is created. On the other hand, The FSLP method with low-Mach correction can produce the arms that appear on the reference HLLC simulation. It shows that our new method can better capture high-frequency flow features with much lower resolution than the classic HLLC solver, similar to OSLP. This is due to the low-Mach nature of this test: as displayed in figure 9 one can indeed see that  $M_a \in [0, 0.165]$ . Therefore the low-Mach correction at play in the FSLP solver has an important effect on the result. Note, however how this correction does not fix the important amount of numerical diffusion that appears at the interface between both layers with the FSLP solver. At second-order, the HLLC solution shown in figure 10 does present lateral arms, similar to the first-order FSLP method. The second-order FSLP method presents many secondary

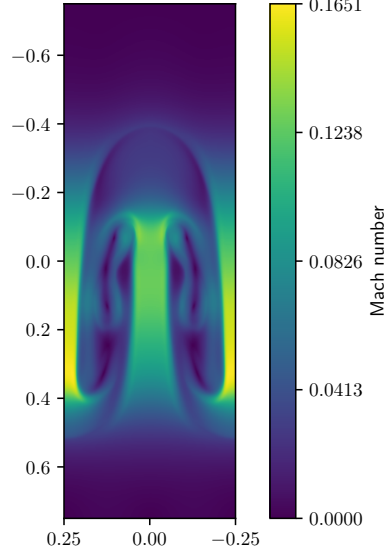


Figure 9: Rayleigh-Taylor instability: mapping of the Mach number profile for reference solution obtained with a second-order HLLC method on a  $200 \times 300$ -cell mesh at  $t = 12.4s$ .

rolls both on the front of the main mode and on the lateral arms. This agreement with the reference solution displayed in figure 10 shows the higher accuracy of the second-order FSLP method. Finally, let us mention that the results obtained with the OSLP in figure 10 resemble the first-order FSLP simulation of figure 10.

### 8.7. The stationary vortex in a gravitational field

The stationary vortex in a gravity field test [97] is a modified version of the Gresho vortex [105] where a gravitational field and a background hydrostatic equilibrium state are added. It allows testing the low-Mach properties of numerical methods. We consider the sub-case of the setup proposed in [97] with  $Fr = Ma, RT = 1/Ma$ , and an adiabatic index  $\gamma = 5/3$ . The potential and the initial conditions are given by:

$$\Phi(r) = \begin{cases} 12.5r^2 & \text{if } r \leq 0.2 \\ 0.5 - \ln(0.2) + \ln(r) & \text{if } 0.2 < r \leq 0.4 \\ \ln(2) - 0.5 \frac{r_c}{r_c - 0.4} + 2.5 \frac{r_c}{r_c - 0.4} r - 1.25 \frac{1}{r_c - 0.4} r^2 & \text{if } 0.4 < r \leq r_c \\ \ln(2) - 0.5 \frac{r_c}{r_c - 0.4} + 1.25 \frac{r_c^2}{r_c - 0.4} & \text{if } r > r_c. \end{cases}$$

with  $r_c = 0.5$ . The density is given by:

$$\rho = \exp(-Ma^2\Phi), \quad (83)$$

The radial velocity is null, and the tangential velocity is given by

$$u_\theta(r) = \frac{1}{u_r} \begin{cases} 5r & \text{if } r \leq 0.2 \\ 2 - 5r & \text{if } 0.2 < r \leq 0.4 \\ 0 & \text{if } r > 0.4 \end{cases} \quad (84)$$

The pressure is  $p = \rho/Ma^2 + p_2$  with:

$$p_2(r) = \frac{1}{u_r^2} \begin{cases} p_{21}(r) & \text{if } r \leq 0.2 \\ p_{21}(0.2) + p_{22}(r) & \text{if } 0.2 < r \leq 0.4 \\ p_{21}(0.2) + p_{22}(0.4) & \text{if } r > 0.4 \end{cases} \quad (85)$$

where  $u_r = 0.4\pi$  and

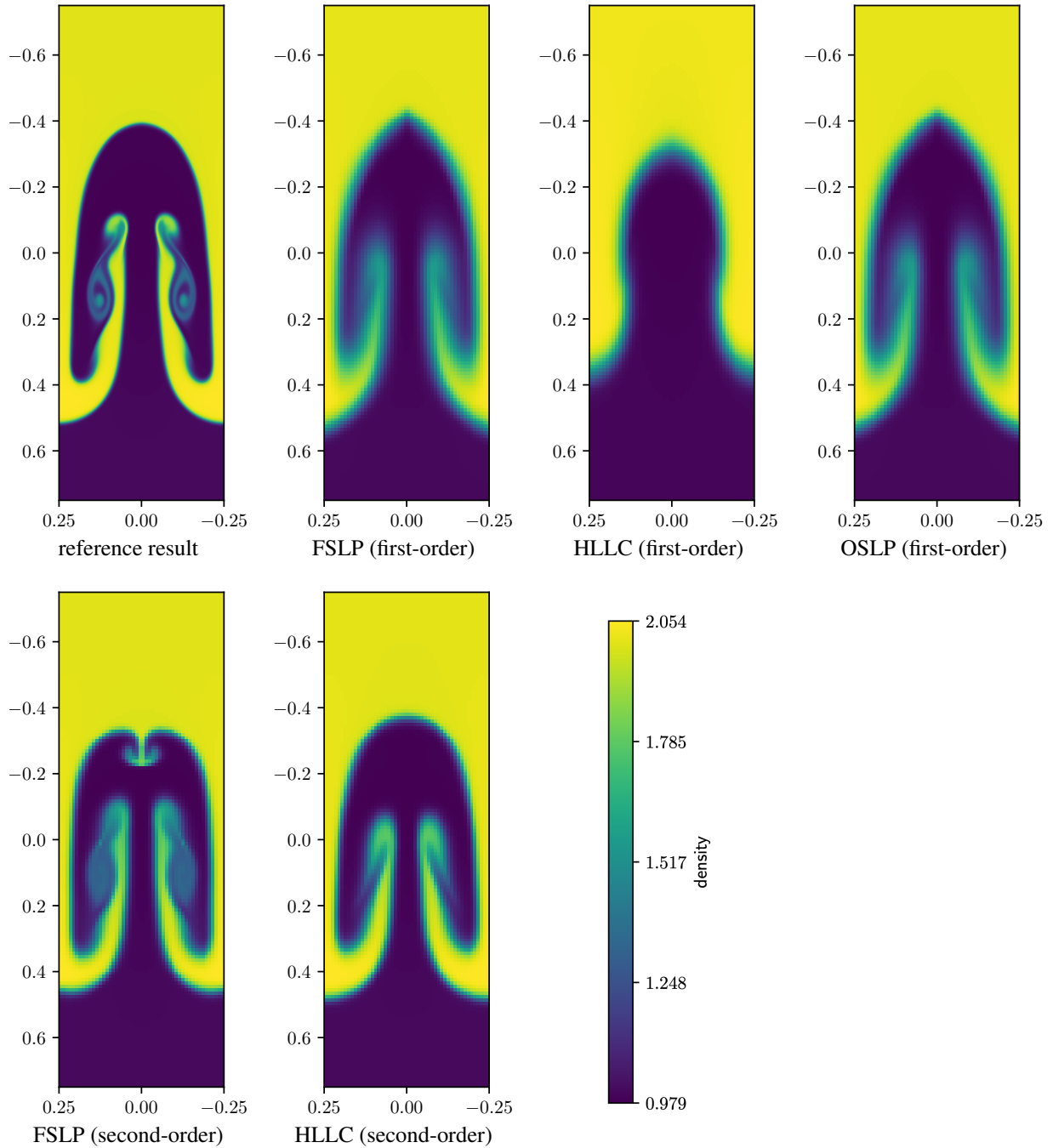


Figure 10: Rayleigh-Taylor instability: mapping of the density profile for: the reference results obtained with the second-order HLLC method on a  $200 \times 600$ -cell grid, the FSLP (first and second-order), the HLLC (first and second-order), and the OSLP methods on a  $50 \times 150$ -cell mesh at  $t = 12.4s$ .

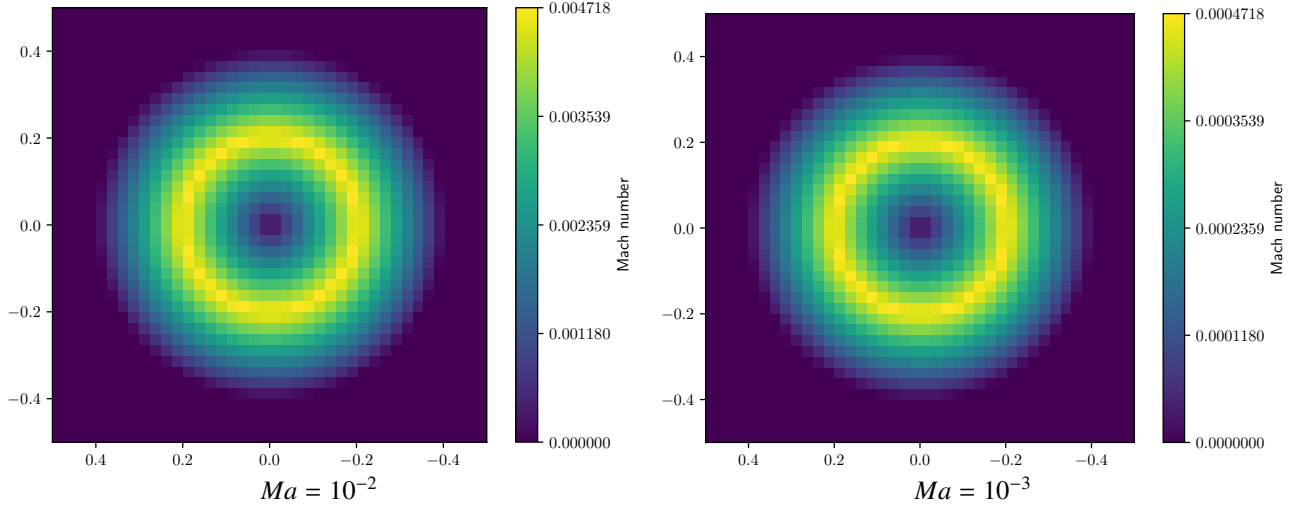


Figure 11: Comparison of the initial Mach number distribution for the Gresho vortex test case with  $Ma = 10^{-2}$  and  $Ma = 10^{-3}$  obtained with the FSLP method with resolutions  $40^2$ .

$$\begin{aligned}
 p_{21}(r) &= \left(1 - \exp\left(-12.5Ma^2r^2\right)\right) \\
 p_{22}(r) &= \frac{1}{(1 - Ma^2)(1 - 0.5Ma^2)} \exp\left((-0.5 + \ln(0.2))Ma^2\right) \\
 &\quad \left(r^{-Ma^2} \left(Ma^4(r(10 - 12.5r) - 2) - 4Ma^4(\gamma - 1)^2 + Ma^2(r(12.5r - 20) + 6)\right) \right. \\
 &\quad \left. + \exp\left(-\ln(0.2)Ma^2\right)(4 - 2.5Ma^2 + 0.5Ma^4)\right).
 \end{aligned} \tag{86}$$

We consider the domain  $[0, 1]^2$  and define the radius from the center  $r = (x - 0.5)^2 + (y - 0.5)^2$ . The initial Mach number distribution is shown for two configurations corresponding to  $Ma = 10^{-2}$  and  $Ma = 10^{-3}$  in figure 11. We let the vortex evolve until  $t = 1$  s corresponds to a full revolution and display the final Mach number distribution with different resolutions in figures 3, 13. We also give the final to initial kinetic energy ratio in table 4. It is clear from the figures and the table that the numerical diffusion is indeed roughly independent of the Mach regime.

Table 4: Vortex in a gravitational potential test case: evaluation of the ratio kinetic energies  $e_{kin}/e_{kin}^0$  at  $t = 1$  s in the computational domain for different values of the Mach number  $Ma$ .

	$Ma = 10^{-2}$	$Ma = 10^{-3}$
$40^2$	0.5723	0.5727
$80^2$	0.7261	0.7258
$160^2$	0.8386	0.8388

## Reproducing the numerical experiments and figures

All the simulations shown in this paper were performed with the open source code ARK<sup>2</sup>-MHD, which can be found at [https://gitlab.erc-atmo.eu/remi.bourgeois/ark-2-mhd/-/tree/test\\_case\\_unsplit\\_paper\\_%232](https://gitlab.erc-atmo.eu/remi.bourgeois/ark-2-mhd/-/tree/test_case_unsplit_paper_%232). All parameter files and plotting scripts can be found in the folder `/test_case_unsplit_paper`.

## 9. Conclusion

We have presented the recasting of an operator splitting Lagrange-Projection solver for gas dynamics into a corresponding flux-splitting finite-volume method. This FSLP method is obtained thanks to a simple modification: it only

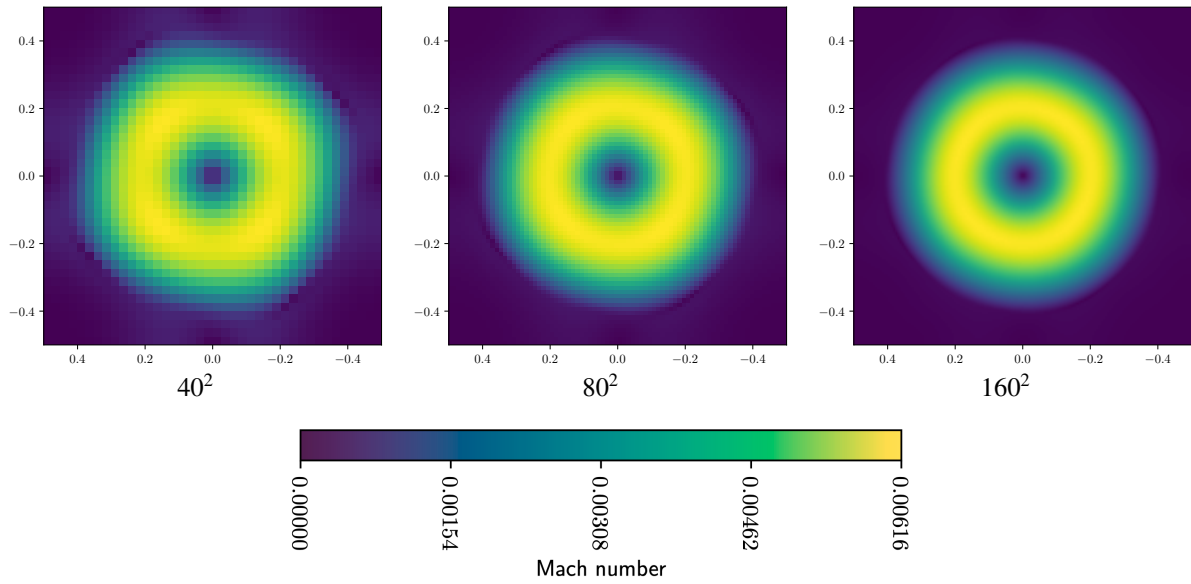


Figure 12: Comparison of the Mach number distribution for the Gresho vortex test case with  $Ma = 10^{-2}$  obtained with the FSLP method with resolutions  $40^2$ ,  $80^2$ ,  $160^2$  at  $t = 1s$ .

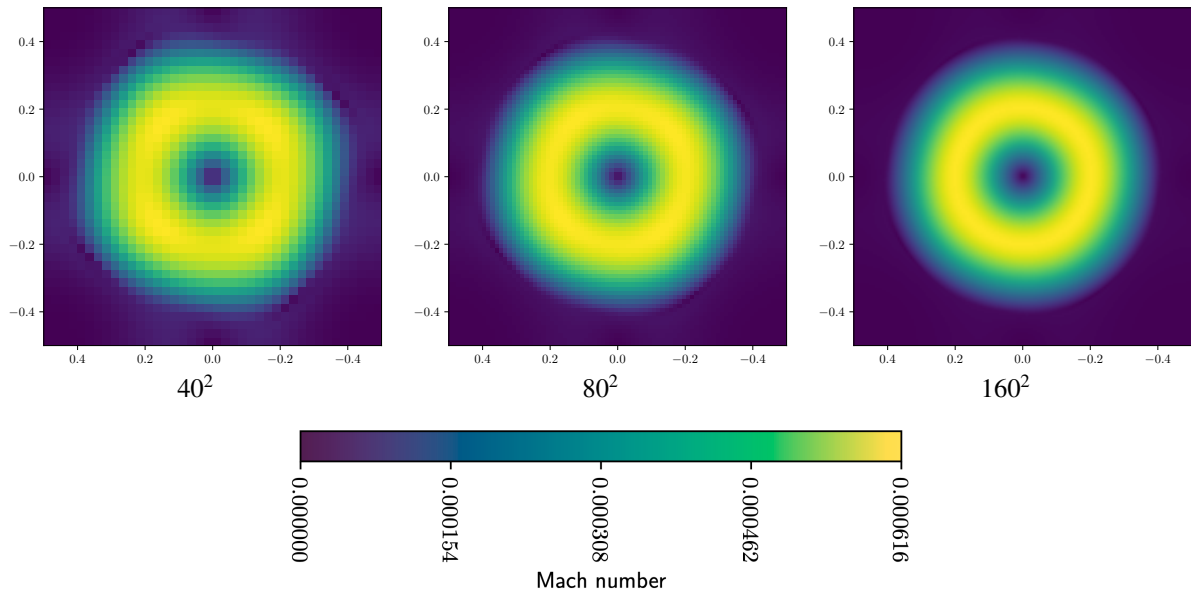


Figure 13: Comparison of the Mach number distribution for the Gresho vortex test case with  $Ma = 10^{-3}$  obtained with the FSLP method with resolutions  $40^2$ ,  $80^2$ ,  $160^2$  at  $t = 1s$ .

differs from the OSLP method in the states used to compute the transport step. The method relies on a flux evaluation that separates pressure-related terms from the advection terms in the spirit of [70, 82, 77, 83]. Two different interpretations of this flux-splitting scheme were proposed to understand better and analyze the resulting method. First, we showed that the FSLP discretization could be written as a convex combination of two updated states resulting from approximating two subsystems that respectively account for pressure and advection effects. This approach allowed us to derive the stability properties of the proposed algorithm. Second, we discussed the interpretation of the FSLP method as the result of the discretization of a larger relaxation system that accounts separately for pressure and advection terms within a single step. We showed that the FSLP method is more computationally efficient than the OSLP method. As a flux-based solver, the resulting FSLP method was straightforwardly extended to multiple dimensions of space and to a high order of accuracy thanks to a standard MUSCL method.

The initial OSLP solver has several interesting numerical advantages: a well-balanced treatment of the source term and a low Mach fix that provides a uniform truncation error with respect to the Mach number. Both properties were preserved through the recasting process. The robustness and accuracy of our new flux-splitting method were tested against a set of benchmark problems, including one and two-dimensional problems, high and low Mach flows with first and second-order discretizations. The results further confirm the numerical stability of our approach.

In the future, we plan to perform a similar recasting by considering an Implicit-Explicit OSLP solver to prevent the severe CFL limitations imposed by the sound velocity in the low Mach regime. The methods can also be extended to several other flow models like two-phase flow models, magneto-hydrodynamics and the M1 model for radiative transfer.

## Acknowledgment

Pascal Tremblin acknowledges support by the European Research Council under Grant Agreement ATMO 757858. Rémi Bourgeois thanks Teddy Pichard for the very fruitful discussions about the convex combination interpretation of the FSLP method.

## References

- [1] T. Padiou, P. Tremblin, E. Audit, P. Kestener, and S. Kokh. A high-performance and portable all-Mach regime flow solver code with well-balanced gravity. Application to compressible convection. *The Astrophysical Journal*, 875(2):p. 128, 2019.
- [2] L. Gosse and A.-Y. Le Roux. A well-balanced scheme designed for inhomogeneous scalar conservation laws. *C. R. Math. Acad. Sci. Paris*, 323(5):pp. 543–546, 1996.
- [3] J. M. Greenberg and A.-Y. Leroux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.*, 33(1):pp. pp. 1–16, 1996.
- [4] R. J. LeVeque. Balancing source terms and flux gradients in high-resolution Godunov methods: The quasi-steady wave-propagation algorithm. *J. of Comp. Phys.*, 146(1):pp. 346–365, 1998.
- [5] L. Gosse. A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms. *Computers & Mathematics with Applications*, 39(9-10):pp. 135–159, 2000.
- [6] L. Gosse and G. Toscani. Asymptotic-preserving and well-balanced schemes for radiative transfer and the Rosseland approximation. *Numer. Math.*, 2:pp. 223–250, 2004.
- [7] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM Journal on Scientific Computing*, 25(6):pp. 2050–2065, 2004.
- [8] M. Lukáčová-Medvid'ová, S. Noelle, and M. Kraft. Well-balanced finite volume evolution Galerkin methods for the shallow water equations. *J. of Comp. Phys.*, 221(1):pp. 122–147, 2007.
- [9] S. Noelle, Y. Xing, and C.-W. Shu. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. of Comp. Phys.*, 226(1):pp. 29–58, 2007.
- [10] M. J. Castro Díaz, T. Chacón Rebollo, E. D. Fernández-Nieto, and C. Parés. On well-balanced finite volume methods for nonconservative nonhomogeneous hyperbolic systems. *SIAM Journal on Scientific Computing*, 29(3):pp. 1093–1126, 2007.
- [11] M. Pelanti, F. Bouchut, and A. Mangeney. A Roe-type scheme for two-phase shallow granular flows over variable topography. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(5):pp. 851–885, 2008.
- [12] L. Gosse. *Computing qualitatively correct approximations of balance laws: exponential-fit, well-balanced and asymptotic-preserving*. Number volume 2 in SIMAI Springer series. Springer, Milan, 2013.
- [13] R. Käppeli and S. Mishra. Well-balanced schemes for the Euler equations with gravitation. *J. of Comp. Phys.*, 259:pp. 199–219, 2014.
- [14] P. Chandrashekar and C. Klingenberg. A second order well-balanced finite volume scheme for Euler equations with gravity. *SIAM Journal on Scientific Computing*, 37:pp. B382–B402, 2015.
- [15] V. Desveaux, M. Zenk, C. Berthon, and C. Klingenberg. A well-balanced scheme to capture non-explicit steady states in the Euler equations with gravity. *International Journal for Numerical Methods in Fluids*, 81(2):pp. 04–127, 2016.

- [16] C. Chalons, P. Kestener, S. Kokh, and M. Stauffert. A large time-step and well-balanced Lagrange-Projection type scheme for the shallow-water equations. *Comm. in Math. Sci.*
- [17] R. Touma, U. Koley, and C. Klingenberg. Well-balanced unstaggered central schemes for the Euler equations with gravitation. *SIAM Journal on Scientific Computing*, 38(5):pp. B773–B807, 2016.
- [18] V. Michel-Dansac, C. Berthon, S. Clain, and F. Foucher. A well-balanced scheme for the shallow-water equations with topography. *Computers & Mathematics with Applications*, 72(3):pp. 568–593, 2016.
- [19] M. J. Castro, T. Morales de Luna, and C. Parés. Well-balanced schemes and path-conservative numerical methods. In *Handbook of Numerical Analysis*, pages pp. 131–175. Elsevier, 2017.
- [20] A. Chertock, S. Cui, A. Kurganov, S. N. Özcan, and E. Tadmor. Well-balanced schemes for the Euler equations with gravitation: Conservative formulation using global fluxes. *J. of Comp. Phys.*, 358:pp. 36–52, 2018.
- [21] M. J. Castro and C. Parés. Well-balanced high-order finite volume methods for systems of balance laws. *Journal of Scientific Computing*, 82(2), 2020.
- [22] T. Morales de Luna, M. J. Castro Díaz, and C. Chalons. High-order fully well-balanced Lagrange-projection scheme for shallow water. *Communications in Mathematical Sciences*, 18(3):pp. 781–807, 2020.
- [23] J. P. Berberich, P. Chandrashekar, and C. Klingenberg. High order well-balanced finite volume methods for multi-dimensional systems of hyperbolic balance laws. *Computers and Fluids*, 219:pp. 104858, 2021.
- [24] A. Del Grosso and C. Chalons. Second-order well-balanced Lagrange-projection schemes for blood flow equations. *Calcolo*, 58(4), 2021.
- [25] E. Turkel. Preconditioned methods for solving the incompressible and low speed compressible equations. *J. Comp. Phys.*, 72(2):pp. 277–298, 1987.
- [26] H. Guillard and C. Viozat. On the behaviour of upwind schemes in the low Mach number limit. *Computers & Fluids*, 28(1):pp. 63–86, 1999.
- [27] H. Guillard and A. Murrone. On the behavior of upwind schemes in the low Mach number limit: II. Godunov type schemes. *Computers and Fluids*, 33(4):pp. 655–675, 2004.
- [28] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number. *J. Comp. Phys.*, 229(4):pp. 978–1016, 2010.
- [29] F. Rieper and G. Bader. The influence of cell geometry on the accuracy of upwind schemes in the low Mach number regime. *J. of Comp. Phys.*, 228(8):pp. 2918–2933, 2009.
- [30] S. Dellacherie, P. Omnes, and F. Rieper. The influence of cell geometry on the Godunov scheme applied to the linear wave equation. *J. of Comp. Phys.*, 229(14):pp. 5315–5338, 2010.
- [31] F. Dauvergne, J.-M. Ghidaglia, F. Pascal, and J.-M. Rovarch. Renormalization of the numerical diffusion for an upwind finite volume method. application to the simulation of Kelvin-Helmholtz instability. *Finite volumes for complex applications. V. Proceedings of the 5th International Symposium, Aussois, June 2008, R. Eymard and J.-M. Hérard editors*, pages pp. 321–328, 2008.
- [32] C. Chalons, M. Girardin, and S. Kokh. An all-regime Lagrange-projection-like scheme for the gas dynamics equations on unstructured meshes. *Comm. in Comp. Phys.*, 20(1):pp. 188–233, July 2016.
- [33] S. Dellacherie, J. Jung, P. Omnes, and P.-A. Raviart. Construction of modified Godunov-type schemes accurate at any Mach number for the compressible Euler system. *Math. Models Methods Appl. Sci.*, 26(13):pp. 2525–2615, 2016.
- [34] H. Zakerzadeh. On the Mach-uniformity of the Lagrange-projection scheme. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2016.
- [35] W. Barsukow. Truly multi-dimensional all-speed schemes for the Euler equations on Cartesian grids. *J. of Comp. Phys.*, 435:p. 27, 2021. Id/No 110216.
- [36] P. Degond and M. Tang. All speed scheme for the low Mach number limit of the isentropic Euler equations. *Comm. in Comp. Phys.*, 10:pp. 1–31, 2011.
- [37] F. Cordier, P. Degond, and A. Kumbaro. An Asymptotic-Preserving all-speed scheme for the Euler and Navier-Stokes equations. *J. Comp. Phys.*, 231(17):pp. 5685–5704, 2012.
- [38] G. Bispen, M. Lukáčová-Medvid'ová, and L. Yelash. Asymptotic preserving IMEX finite volume schemes for low Mach number Euler equations with gravitation. *J. of Comp. Phys.*, 335:pp. 222–248, 2017.
- [39] C. Berthon, C. Klingenberg, and M. Zenk. An all Mach number relaxation upwind scheme. *The SMAI journal of computational mathematics*, 6:pp. 1–31, 2020.
- [40] G. Dimarco, R. Loubère, and M.-H. Vignal. Study of a new asymptotic preserving scheme for the Euler system in the low Mach number limit. *SIAM Journal on Scientific Computing*, 39(5):pp. a2099–a2128, 2017.
- [41] S. Boscarino, G. Russo, and L. Scandurra. All Mach number second order semi-implicit scheme for the Euler equations of gas dynamics. *Journal of Scientific Computing*, 77(2):pp. 850–884, 2018.
- [42] F. Bouchut, C. Chalons, and S. Guisset. An entropy satisfying two-speed relaxation system for the barotropic Euler equations: application to the numerical approximation of low Mach number flows. *Numerische Mathematik*, 145(1):pp. 35–76, 2020.
- [43] H. Paillere, C. Viozat, A. Kumbaro, and I. Toumi. Comparison of low Mach number models for natural convection problems. *Heat and Mass Transfer*, 36(6):pp. 567–573, 2000.
- [44] A. Beccantini, E. Studer, S. Gounand, J.-P. Magnaud, T. Kloczko, C. Corre, and S. Kudriakov. Numerical simulations of a transient injection flow at low Mach number regime. *International Journal for Numerical Methods in Engineering*, 76(5):pp. 662–696, 2008.
- [45] G. Dimarco, R. Loubère, V. Michel-Dansac, and M.-H. Vignal. Second-order implicit-explicit total variation diminishing schemes for the Euler system in the low Mach regime. *J. of Comp. Phys.*, 372:pp. 178–201, 2018.
- [46] W. Boscheri, G. Dimarco, R. Loubère, M. Tavelli, and M.-H. Vignal. A second order all Mach number IMEX finite volume solver for the three dimensional Euler equations. *J. of Comp. Phys.*, 415:p. 30, 2020. Id/No 109486.
- [47] F. Bouchut, E. Franck, and L. Navoret. A low cost semi-implicit low-Mach relaxation scheme for the full Euler equations. *Journal of Scientific Computing*, 83(1):p. 24, 2020.
- [48] J. Zeifang, J. Schütz, K. Kaiser, A. Beck, M. Lukáčová-Medvid'ová, and S. Noelle. A novel full-euler low mach number IMEX splitting.



- Comm. in Comp. Phys.*, 27(1):pp. 292–320, 2020.
- [49] P. Bruel, S. Delmas, J. Jung, and V. Perrier. A low Mach correction able to deal with low Mach acoustics. *J. of Comp. Phys.*, 378:pp. 723–759, 2019.
- [50] H. Weyl. Shock waves in arbitrary fluids. *Comm. Pure Appl. Math.*, 2(2-3):pp. 103–122, 1949.
- [51] H.B. Callen. *Thermodynamics and an introduction to thermostatistics*. John Wiley & sons, 1985.
- [52] J. Smoller. *Shock waves and reaction diffusion equations*. Springer Verlag, 1983.
- [53] R.J. LeVeque. *Finite-Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [54] E. Godlewski and P.-A. Raviart. *Hyperbolic Systems of Conservation Laws*. Ellipse, 1990.
- [55] D. Serre. *Systems of Conservation Laws*, volume 1. Cambridge University Press, 1999.
- [56] S. Del Pino and H. Jourden. Arbitrary high-order schemes for the linear advection and wave equations: application to hydrodynamics and aeroacoustics. *Comptes Rendus Mathematique*, 342(6):pp. 441–446, 2006.
- [57] F. Duboc, C. Enaux, S. Jaouen, H. Jourden, and M. Wolff. High-order dimensionally split Lagrange-remap schemes for compressible hydrodynamics. *Comptes Rendus Mathematique*, 348(1-2):pp. 105–110, 2010.
- [58] S. Dellacherie. Checkerboard modes and wave equation. In *Algorithmy 2009. 18th conference on scientific computing, Vysoké Tatry – Podbanské, Slovakia, March 15–20, 2009. Proceedings of contributed papers and posters.*, pages pp. 71–80. Bratislava: Slovak University of Technology, Faculty of Civil Engineering, Department of Mathematics and Descriptive Geometry, 2009.
- [59] J. Jung and V. Perrier. Steady low Mach number flows: identification of the spurious mode and filtering method. *J. of Comp. Phys.*, 468:p. 27, 2022. Id/No 111462.
- [60] I. Suliciu. On the thermodynamics of rate-type fluids and phase transitions. I. rate-type fluids. *International Journal of Engineering Science*, 36(9):pp. 921–947, 1998.
- [61] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser, Basel Berlin, 2004.
- [62] C. Chalons and J.-F. Coulombel. Relaxation approximation of the Euler equations. *J. Math. Anal. Appl.*, 348(2):pp. 872–893, 2008.
- [63] F. Coquel, E. Godlewski, and N. Seguin. Relaxation of fluid systems. *Mathematical Models and Methods in Applied Sciences*, 22(08):pp. 1250014, 2012.
- [64] F. Liu, I. Jennions, and A. Jameson. Computation of turbomachinery flow by a convective-upwind-split-pressure (CUSP) scheme. In *36th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV,U.S.A., 1998.
- [65] D. Darracq, S. Champagneux, and A. Corjon. Time-accurate fluid-structure coupling for turbulent flows. In C.-H. Bruneau, editor, *Sixteenth International Conference on Numerical Methods in Fluid Dynamics*, volume 515, pages pp. 31–36. Springer Berlin Heidelberg, 1998.
- [66] S. Evje and K. K. Fjelde. Hybrid flux-splitting schemes for a two-phase flow model. *J. of Comp. Phys.*, 175(2):pp. 674–701, 2002.
- [67] H. Paillère, C. Corre, and J.R. García Cascales. On the extension of the AUSM+ scheme to compressible two-fluid models. *Computers & Fluids*, 32(6):pp. 891–916, 2003.
- [68] J.R. García-Cascales and H. Paillère. Application of AUSM schemes to multi-dimensional compressible two-phase flow problems. *Nuclear Engineering and Design*, 236(12):pp. 1225–1239, 2006.
- [69] J. L. Steger and R.F Warming. Flux vector splitting of the inviscid gas dynamic equations with application to finite-difference methods. *J. of Comp. Phys.*, 40(2):pp. 263–293, 1981.
- [70] G.-C. Zha and E. Bilgen. Numerical solutions of Euler equations by using a new flux vector splitting scheme. *International Journal for Numerical Methods in Fluids*, 17(2):pp. 115–144, 1993.
- [71] M.-S. Liou and C. J. Steffen. A new flux splitting scheme. *J. of Comp. Phys.*, 107(1):pp. 23–39, 1993.
- [72] A. Jameson. Analysis and design of numerical scheme for gas dynamics, 2: artificial diffusion and discrete shock structure. *International Journal of Computational Fluid Dynamics*, 5(1-2):pp. 1–38, 1995.
- [73] M.-S. Liou. Recent progress and applications of AUSM+. In *Sixteenth International Conference on Numerical Methods in Fluid Dynamics*, pages pp. 302–307. Springer Berlin Heidelberg, 1998.
- [74] M.-S. Liou. A sequel to AUSM, Part II: AUSM+-up for all speeds. *J. of Comp. Phys.*, 214(1):pp. 137–170, 2006.
- [75] M.-S. Liou. A Sequel to AUSM: AUSM+. *J. of Comp. Phys.*, 129(2):pp. 364–382, 1996.
- [76] F. Bouchut. Entropy satisfying flux vector splittings and kinetic BGK models. *Numerische Mathematik*, 94(4):pp. 623–672, 2003.
- [77] E. F. Toro and M. E. Vázquez-Cendón. Flux splitting schemes for the Euler equations. *Computers and Fluids*, 70:pp. 1–12, 2012.
- [78] B. Dubroca. Solveur de Roe positivement conservatif. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 329(9):pp. 827–832, 1999.
- [79] G. Gallice. Schémas de type Godunov entropiques et positifs préservant les discontinuités de contact. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 331(2):pp. 149–152, 2000.
- [80] G. Gallice. Positive and entropy stable Godunov-type schemes for gas dynamics and MHD equations in Lagrangian or Eulerian coordinates. *Numerische Mathematik*, 94(4):pp. 673–713, 2003.
- [81] A. Chan, G. Gallice, R. Loubère, and P.-H. Maire. Positivity preserving and entropy consistent approximate Riemann solvers dedicated to the high-order MOOD-based finite volume discretization of Lagrangian and Eulerian gas dynamics. *Computers & Fluids*, 229:p. 105056, 2021.
- [82] S. M. Deshpande, N. Balakrishnan, and S. V. Raghurama Rao. PVU and wave-particle splitting schemes for Euler equations of gas dynamics. *Sadhana*, 19(6):pp. 1027–1054, 1994.
- [83] K. Borah, G. Natarajan, and A. K. Dass. A novel second-order flux splitting for ideal magnetohydrodynamics. *J. of Comp. Phys.*, 313:pp. 159–180, 2016.
- [84] R. Baraille, G. Bourdin, F. Dubois, and A.Y. Roux. Une version à pas fractionnaires du schéma de Godunov pour l'hydrodynamique. *C. R. Acad. Sci. Paris*, 314:pp. 147–152, 1992.
- [85] T. Buffard and J.M. Hérard. A conservative fractional step method to solve non-isentropic Euler equations. *Computer Methods in Applied Mechanics and Engineering*, 144(3-4):pp. 199–225, 1997.
- [86] C. Chalons, F. Coquel, S. Kokh, and N. Spillane. Large time-step numerical scheme for the seven-equation model of compressible two-

- phase flows. In J. Fořt, J. Fürst, J. Halama, R. Herbin, and F. Hubert, editors, *Finite Volumes for Complex Applications VI: Problems & Perspectives*, volume 4, pages pp. 225–233. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [87] F. Coquel, J.-M. Hérard, and K. Saleh. A splitting method for the isentropic Baer-Nunziato two-phase flow model. *ESAIM: Proceedings*, 38:pp. 241–256, 2012.
- [88] C. Chalons, M. Girardin, and S. Kokh. An all-regime Lagrange-projection like scheme for 2D homogeneous models for two-phase flows on unstructured meshes. *J. of Comp. Phys.*, 335, 2017.
- [89] B. van Leer. Towards the ultimate conservative difference scheme. III: Upstream- centered finite-difference schemes for ideal compressible flow. *J. of Comp. Phys.*, 23:pp. 263–275, 1977.
- [90] B. van Leer. Towards the ultimate conservative difference scheme. IV: A new approach to numerical convection. *J. of Comp. Phys.*, 23:pp. 276–299, 1977.
- [91] B. van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov’s method. *J. of Comp. Phys.*, 32:pp. 101–136, 1979.
- [92] E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics - A Practical Introduction*. Springer, 3rd edition, 2009.
- [93] X.-D. Liu, S. Osher, and T. Chan. Weighted essentially non-oscillatory schemes. *J. of Comp. Phys.*, 115(1):pp. 200–212, 1994.
- [94] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. *J. of Comp. Phys.*, 126(1):pp. 202–228, 1996.
- [95] Steven Diot, Raphaël Loubère, and Stéphane Clain. The multidimensional optimal order detection method in the three-dimensional case: very high-order finite volume method for hyperbolic systems. *International Journal for Numerical Methods in Fluids*, 73(4):pp. 362–392, 2013.
- [96] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for systems of conservation laws—multi-dimensional optimal order detection (mood). *J. of Comp. Phys.*, 230(10):pp. 4028–4050, 2011.
- [97] A. Thomann, G. Puppo, and C. Klingenberg. An all speed second order well-balanced IMEX relaxation scheme for the Euler equations with gravity. *J. of Comp. Phys.*, 420:p. 109723, 2020.
- [98] E. Godlewski and P.-A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Springer New York, 2021.
- [99] H. C. Yee. A class of high-resolution explicit and implicit shock-capturing methods. 1989.
- [100] R. J. Spiteri and S. J. Ruuth. A new class of optimal high-order strong-stability-preserving time discretization methods. *SIAM Journal on Numerical Analysis*, 40(2):pp. 469–491, 2002.
- [101] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Mathematics of computation*, 67(221):pp. 73–85, 1998.
- [102] E. F. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock waves*, 4(1):pp. 25–34, 1994.
- [103] G. A. Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *J. of Comp. Phys.*, 27(1):pp. 1–31, 1978.
- [104] B. Einfeldt, C. D. Munz, P. L. Roe, and B. Sjögreen. On Godunov-type methods near low densities. *J. of Comp. Phys.*, 92(2):pp. 273–295, 1991.
- [105] P. M. Gresho and S. T. Chan. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. Part 2: Implementation. *Int. J. Numer. Meth. Fluids*, 11(5):pp. 621–659, 1990.
- [106] R. Liska and B. Wendroff. Comparison of several difference schemes on 1D and 2D test problems for the Euler equations. *SIAM Journal on Scientific Computing*, 25(3):pp. 995–1017, 2003.

## Appendix A. A few classic convexity properties

We recall hereafter a few classic convexity/concavity properties related to admissible states, entropy, and energy of our flow model that can be found in the literature (see for example [98]). We propose short self-contained proofs of these properties for the sake of completeness.

**Lemma Appendix A.1.** *We have the following properties.*

(a) *The function  $\Lambda : U = (\rho, \rho u, \rho E) \in [0, +\infty) \times \mathbb{R} \times [0, +\infty) \mapsto \Lambda(U) = (\rho E) - \frac{(\rho u)^2}{2\rho}$  is concave.*

(b) *The set  $\Omega$  defined by (2) is convex.*

(c) *The function  $\mathcal{U} : (\mathcal{T}, u, E) \mapsto s(\mathcal{T}, E - \frac{u^2}{2})$  is strictly concave.*

(d) *The function  $\eta : (\rho, \rho \mathcal{T}, \rho u, \rho E) \mapsto \rho s\left(\frac{\rho \mathcal{T}}{\rho}, \frac{\rho E}{\rho} - \frac{(\rho u)^2}{2\rho^2}\right)$  is strictly concave.*

*Proof.* Let  $\theta_1 = 1 - \theta_2 \in [0, 1]$  and  $U_k \in [0, +\infty) \times \mathbb{R} \times [0, +\infty)$  for  $k = 1, 2$ , we have

$$\Lambda\left(\sum_{k=1,2} \theta_k U_k\right) - \sum_{k=1,2} \theta_k \Lambda(U_k) = \sum_{k=1,2} \theta_k (\rho E)_k + \frac{\theta_1 \theta_2}{\sum_{k=1,2} \theta_k \rho_k} \left[ (\rho u)_1 \sqrt{\frac{\rho_2}{\rho_1}} - (\rho u)_2 \sqrt{\frac{\rho_1}{\rho_2}} \right]^2 \geq 0, \quad (\text{A.1})$$

which proves (a). For (b), consider again  $\theta_1 = 1 - \theta_2 \in [0, 1]$  and  $U_k \in \Omega$ ,  $k = 1, 2$ . If we note  $U = \sum_{k=1,2} \theta_k U_k$ , then  $U \in \Omega$ . Indeed, we have that  $\sum_{k=1,2} \theta_k \rho_k \geq 0$ , and as  $\rho e = \Lambda(U) \geq \sum_{k=1,2} \theta_k \Lambda(U_k) = \sum_{k=1,2} \theta_k \rho_k e_k \geq 0$ , where  $e_k = E_k - (u_k^2)/2$ . This implies that  $e \geq 0$ .

For (c): The function  $K : (u, E) \mapsto E - u^2/2$  is strictly concave and we have  $\mathcal{U}(\mathcal{T}, u, E) = s(\mathcal{T}, K(u, E))$ . Consider  $\lambda \in [0, 1]$  and let us note  $\lambda = \lambda_1$  and  $\lambda_2 = 1 - \lambda$ . We have that

$$\mathcal{U}\left(\sum_{k=1,2} \lambda_k \mathcal{T}_k, \sum_{k=1,2} \lambda_k u_k, \sum_{k=1,2} \lambda_k E_k\right) - \sum_{k=1,2} \lambda_k \mathcal{U}(\mathcal{T}_k, u_k, E_k) \quad (\text{A.2})$$

$$= s\left(\sum_{k=1,2} \mathcal{T}_k, K\left(\sum_{k=1,2} \lambda_k u_k, \sum_{k=1,2} \lambda_k E_k\right)\right) - \sum_{k=1,2} \lambda_k s(\mathcal{T}_k, K(u_k, E_k)) \quad (\text{A.3})$$

$$= s\left(\sum_{k=1,2} \mathcal{T}_k, K\left(\sum_{k=1,2} \lambda_k u_k, \sum_{k=1,2} \lambda_k E_k\right)\right) - s\left(\sum_{k=1,2} \mathcal{T}_k, \sum_{k=1,2} \lambda_k K(u_k, E_k)\right) \\ + s\left(\sum_{k=1,2} \mathcal{T}_k, \sum_{k=1,2} \lambda_k K(u_k, E_k)\right) - \sum_{k=1,2} \lambda_k s(\mathcal{T}_k, K(u_k, E_k)). \quad (\text{A.4})$$

As  $K$  is concave, we get that

$$K\left(\sum_{k=1,2} \lambda_k u_k, \sum_{k=1,2} \lambda_k E_k\right) \geq \sum_{k=1,2} \lambda_k K(u_k, E_k). \quad (\text{A.5})$$

By (3) we know that  $e' \mapsto s^{\text{EOS}}(\overline{\mathcal{T}}, e')$  is increasing so that

$$s\left(\sum_{k=1,2} \mathcal{T}_k, K\left(\sum_{k=1,2} \lambda_k u_k, \sum_{k=1,2} \lambda_k E_k\right)\right) - s\left(\sum_{k=1,2} \mathcal{T}_k, \sum_{k=1,2} \lambda_k K(u_k, E_k)\right) \geq 0. \quad (\text{A.6})$$

We also know that  $s$  is concave therefore

$$s\left(\sum_{k=1,2} \mathcal{T}_k, \sum_{k=1,2} \lambda_k K(u_k, E_k)\right) - \sum_{k=1,2} \lambda_k s(\mathcal{T}_k, K(u_k, E_k)) \geq 0. \quad (\text{A.7})$$

By replacing (A.6) and (A.7) into (A.4) we obtain that

$$\mathcal{U}\left(\sum_{k=1,2} \lambda_k \mathcal{T}_k, \sum_{k=1,2} \lambda_k u_k, \sum_{k=1,2} \lambda_k E_k\right) \geq \sum_{k=1,2} \lambda_k \mathcal{U}(\mathcal{T}_k, u_k, E_k). \quad (\text{A.8})$$

for (d): If we note again  $U = (\rho, \rho u, \rho E)$ , by (6), we have

$$\eta(\rho, \rho \mathcal{T}, \rho u, \rho E) = \rho s\left(\frac{\rho \mathcal{T}}{\rho}, \frac{(\rho E)}{\rho} - \frac{(\rho u)^2}{2\rho^2}\right) = S\left(\rho, \rho \mathcal{T}, \rho E - \frac{(\rho u)^2}{2\rho}\right) = S(\rho, \rho \mathcal{T}, \Lambda(U)). \quad (\text{A.9})$$

Now we consider  $U_k = (\rho_k, \rho_k u_k, \rho_k E_k) \in \Omega$  and  $\mathcal{T}_k \geq 0$ ,  $k = 1, 2$ , we have

$$\eta\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \sum_{k=1,2} \theta_k \rho_k u_k, \sum_{k=1,2} \theta_k \rho_k E_k\right) - \sum_{k=1,2} \theta_k \eta(\rho_k, \rho_k \mathcal{T}_k, \rho_k u_k, \rho_k E_k) \\ = S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \Lambda\left(\sum_{k=1,2} \theta_k U_k\right)\right) - \sum_{k=1,2} \theta_k S(\rho_k, \rho_k \mathcal{T}_k, \Lambda(U_k)) \\ = S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \Lambda\left(\sum_{k=1,2} \theta_k U_k\right)\right) - S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \sum_{k=1,2} \theta_k \Lambda(U_k)\right) \\ + S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \sum_{k=1,2} \theta_k \Lambda(U_k)\right) - \sum_{k=1,2} \theta_k S(\rho_k, \rho_k \mathcal{T}_k, \Lambda(U_k)) \quad (\text{A.10})$$

As  $\Lambda$  is concave, we have  $\Lambda\left(\sum_{k=1,2} \theta_k U_k\right) \geq \sum_{k=1,2} \theta_k \Lambda(U_k)$  and as  $\mathcal{E}' \mapsto S(\bar{\rho}, \bar{\mathcal{T}}, \mathcal{E}')$  is increasing, we have

$$S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \Lambda\left(\sum_{k=1,2} \theta_k U_k\right)\right) - S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \sum_{k=1,2} \theta_k \Lambda(U_k)\right) \geq 0. \quad (\text{A.11})$$

Using the fact that  $S$  is concave, we also get

$$S\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \sum_{k=1,2} \theta_k \Lambda(U_k)\right) - \sum_{k=1,2} \theta_k S(\rho_k, \rho_k \mathcal{T}_k, \Lambda(U_k)) \geq 0. \quad (\text{A.12})$$

Injecting (A.11) and (A.12) into (A.10) provides

$$\eta\left(\sum_{k=1,2} \theta_k \rho_k, \sum_{k=1,2} \theta_k \rho_k \mathcal{T}_k, \sum_{k=1,2} \theta_k \rho_k u_k, \sum_{k=1,2} \theta_k \rho_k E_k\right) \geq \sum_{k=1,2} \theta_k \eta(\rho_k, \rho_k \mathcal{T}_k, \rho_k u_k, \rho_k E_k). \quad (\text{A.13})$$

□

## Appendix B. Approximate Riemann solver for the pressure subsystem

In this section, we present the derivation of an approximate Riemann solver for the pressure subsystem (39), following the lines of [32, 1]. We express (39) in the following compact form:

$$\partial_t \mathbf{U} + 2\partial_x \mathbf{P}(\mathbf{U}) = \mathbf{S}(\mathbf{U}), \quad \partial_t \Pi + \partial_x (2a^2 u) = 0, \quad \partial_t (\rho \mathcal{T}) - 2\partial_x u^P = 0, \quad \partial_t \phi = 0. \quad (\text{B.1})$$

where  $\mathbf{P}(\mathbf{U})^T = (0, \Pi, \Pi u)$ . Let  $\Delta x_L > 0$ ,  $\Delta x_R > 0$ , we consider  $\bar{x} \in \mathbb{R}$  and the following piecewise initial data

$$(\mathbf{U}, \Pi, \mathcal{T}, \phi)(x, t = 0) = \begin{cases} (\mathbf{U}_L, \Pi_L, \mathcal{T}_L, \phi_L) & \text{if } x \leq \bar{x}, \\ (\mathbf{U}_R, \Pi_R, \mathcal{T}_R, \phi_R) & \text{if } x > \bar{x}, \end{cases} \quad (\text{B.2})$$

that verifies the equilibrium relations:

$$(\mathbf{U}_k, \Pi_k, \mathcal{T}_k, \phi_k) = \left[ (\rho_k, \rho_k u_k, E_k)^T, p^{\text{EOS}}\left(\frac{1}{\rho_k}, e_k\right), \frac{1}{\rho_k}, \phi_k \right], \quad k = L, R, \quad (\text{B.3})$$

with  $\phi_L = \frac{1}{\Delta x_L} \int_{-\Delta x_L}^0 \phi(\bar{x} + x) dx$  and  $\phi_R = \frac{1}{\Delta x_R} \int_0^{\Delta x_R} \phi(\bar{x} + x) dx$ . We seek a self-similar function  $(\mathbf{U}_{\text{RP}}, \Pi_{\text{RP}}, \mathcal{T}_{\text{RP}}, \phi_{\text{RP}})$  composed of four constant states separated by three discontinuities as follows:

$$(\mathbf{U}_{\text{RP}}, \Pi_{\text{RP}}, \mathcal{T}_{\text{RP}}, \phi_{\text{RP}})\left(\frac{x - \bar{x}}{t}; \mathbf{U}_L, \Pi_L, \mathcal{T}_L, \phi_L, \mathbf{U}_R, \Pi_R, \mathcal{T}_R, \phi_R\right) = \begin{cases} (\mathbf{U}_L, \Pi_L, \mathcal{T}_L, \phi_L), & \text{if } \frac{x - \bar{x}}{t} \leq -\frac{2a}{\rho_L}, \\ (\mathbf{U}_L^*, \Pi_L^*, \mathcal{T}_L^*, \phi_L), & \text{if } -\frac{2a}{\rho_L} < \frac{x - \bar{x}}{t} \leq 0, \\ (\mathbf{U}_R^*, \Pi_R^*, \mathcal{T}_R^*, \phi_R), & \text{if } 0 < \frac{x - \bar{x}}{t} \leq \frac{2a}{\rho_R}, \\ (\mathbf{U}_R, \Pi_R, \mathcal{T}_R, \phi_R), & \text{if } \frac{2a}{\rho_R} < \frac{x - \bar{x}}{t}, \end{cases} \quad (\text{B.4})$$

where the intermediate states  $\mathbf{U}_k^*$ ,  $\Pi_k^*$  and  $\mathcal{T}_k^*$  are required to satisfy the four following properties.

1. The approximate Riemann solver should be consistent in the integral sense with the pressure subsystem (B.1): for  $\Delta t$  such that  $\frac{2a}{\min(\rho_L, \rho_R)} \Delta t < \frac{1}{2} \min(\Delta x_L, \Delta x_R)$ , we have

$$\begin{bmatrix} 2\mathbf{P}(\mathbf{U}_R) - 2\mathbf{P}(\mathbf{U}_L) \\ 2a^2(u_R^* - u_L^*) \\ -(2u_R^* - 2u_L^*) \end{bmatrix} = -\frac{2a}{\rho_L} \begin{bmatrix} \mathbf{U}_L^* - \mathbf{U}_L \\ (\rho \Pi)_L^* - (\rho \Pi)_L \end{bmatrix} + \frac{2a}{\rho_R} \begin{bmatrix} \mathbf{U}_R - \mathbf{U}_R^* \\ (\rho \Pi)_R - (\rho \Pi)_R^* \end{bmatrix} + (\Delta x_L + \Delta x_R) \{\mathbf{S}\}, \quad (\text{B.5})$$

with  $\{\mathbf{S}\}$  a function that is a consistent approximation of  $\mathbf{S}$ , that is to say:

$$\lim_{\substack{\Phi_L, \Phi_R \rightarrow \phi(\bar{x}) \\ \Delta x_L, \Delta x_R \rightarrow 0 \\ (\mathbf{U}_R, \Pi_R), (\mathbf{U}_L, \Pi_L) \rightarrow (\bar{\mathbf{U}}, p^{\text{EOS}}(\bar{\rho}, \bar{e}))}} \{\mathbf{S}\} = \mathbf{S}(\bar{\mathbf{U}}, \phi)(x = \bar{x}). \quad (\text{B.6})$$

2. In the case  $\phi_L = \phi_R$ , it should be degenerate to an approximate Riemann for the homogeneous problem obtained with (B.1) when  $\mathbf{S} = \mathbf{0}$ .
3. If  $\mathbf{U}_L$  and  $\mathbf{U}_R$  satisfy the following discrete version of the hydrostatic condition (8):

$$\Pi_R - \Pi_L = -\frac{\rho_L + \rho_R}{2}(\phi_R - \phi_L), \quad u_L = u_R = 0, \quad (\text{B.7})$$

then  $(\mathbf{U}_L^*, \Pi_L^*) = (\mathbf{U}_L, \Pi_L)$  and  $(\mathbf{U}_R^*, \Pi_R^*) = (\mathbf{U}_R, \Pi_R)$ .

Let us build the states  $(\mathbf{U}_R^*, \Pi_R^*)$  and  $(\mathbf{U}_L^*, \Pi_L^*)$  so that they verify the above properties. We note

$$\Pi_R^* - \Pi_L^* + \mathcal{M} = 0. \quad (\text{B.8})$$

First, we impose that  $\rho_L^*$  and  $\rho_R^*$  are consistent with the exact solution of (B.1) by setting  $\rho_L^* = \rho_L$  and  $\rho_R^* = \rho_R$ . Then we also require that the Rankine-Hugoniot jump conditions obtained in the case  $\mathbf{S} = \mathbf{0}$  are valid across the waves of velocity  $-2a/\rho_L$  and  $+2a/\rho_R$

$$\frac{2a}{\rho_L} \begin{bmatrix} \mathbf{U}_L^* - \mathbf{U}_L \\ (\rho\Pi)_L^* - (\rho\Pi)_L \\ (\rho\mathcal{T})_L^* - (\rho\mathcal{T})_L \end{bmatrix} + \begin{bmatrix} 2\mathbf{P}(\mathbf{U}_L^*) - 2\mathbf{P}(\mathbf{U}_L) \\ 2a^2u_L^* - 2a^2u_L \\ -2u_L^* + 2u_L \end{bmatrix} = 0, \quad -\frac{2a}{\rho_R} \begin{bmatrix} \mathbf{U}_R - \mathbf{U}_R^* \\ (\rho\Pi)_R - (\rho\Pi)_R^* \\ (\rho\mathcal{T})_R - (\rho\mathcal{T})_R^* \end{bmatrix} + \begin{bmatrix} 2\mathbf{P}(\mathbf{U}_R) - 2\mathbf{P}(\mathbf{U}_R^*) \\ 2a^2u_R - 2a^2u_R^* \\ -2u_R + 2u_R^* \end{bmatrix} = 0 \quad (\text{B.9})$$

Finally, we postulate that the velocity is continuous across the stationary wave by setting

$$u_L^* = u_R^* = u^*, \quad (\text{B.10})$$

and we also impose that  $(\Pi u)_k^* = \Pi_k^* u_k^* = \Pi_k^* u^*$ ,  $k = L, R$ . Then, relations (B.5), (B.9), (B.8) yield

$$\rho_L^* = \rho_L, \quad \rho_R^* = \rho_R, \quad (\text{B.11a})$$

$$E_L^* = E_L - \frac{1}{a} \left( (\Pi^* + \frac{\mathcal{M}}{2}) u^* - \Pi_L u_L \right), \quad E_R^* = E_R + \frac{1}{a} \left( (\Pi^* - \frac{\mathcal{M}}{2}) u^* - \Pi_R u_R \right), \quad (\text{B.11b})$$

$$u^* = u_R^* = u_L^* = \frac{u_R + u_L}{2} - \frac{1}{2a} (\Pi_R - \Pi_L) - \frac{\mathcal{M}}{2a}, \quad \Pi^* = \frac{\Pi_R + \Pi_L}{2} - \frac{a}{2} (u_R - u_L), \quad (\text{B.11c})$$

$$\Pi_L^* = \Pi^* + \frac{\mathcal{M}}{2}, \quad \Pi_R^* = \Pi^* - \frac{\mathcal{M}}{2}, \quad (\text{B.11d})$$

$$\mathcal{T}_L^* = \frac{1}{\rho_L} + \frac{1}{a} (u^* - u_L), \quad \mathcal{T}_R^* = \frac{1}{\rho_R} - \frac{1}{a} (u^* - u_R), \quad (\text{B.11e})$$

where the jump  $\mathcal{M}$  can be identified as

$$\mathcal{M} = \frac{\Delta x_L + \Delta x_R}{2} \{\rho \partial_x \phi\}, \quad \mathcal{M} u^* = \frac{\Delta x_L + \Delta x_R}{2} \{\rho u \partial_x \phi\}. \quad (\text{B.12})$$

At this point, the functions  $\{\rho \partial_x \phi\}$  and  $\{\rho u \partial_x \phi\}$  are still yet to be specified. Let us consider the constraint 3: if it is satisfied then for a state that verifies (B.7) the jumps  $\mathcal{M}$  and  $\mathcal{M} u^*$  necessarily take the value  $\mathcal{M} = -(\Pi_R - \Pi_L)$  and  $\mathcal{M} u^* = 0$ . A simple choice that fulfills this requirement is

$$\{\rho \partial_x \phi\} = (\rho_L + \rho_R) \frac{\phi_R - \phi_L}{\Delta x_L + \Delta x_R}, \quad \{\rho u \partial_x \phi\} = (\rho_L + \rho_R) u^* \frac{\phi_R - \phi_L}{\Delta x_L + \Delta x_R}. \quad (\text{B.13a})$$

Relations (B.11) and (B.13a) give a complete definition of the approximate Riemann solver (B.4). This solver yields a definition for the conservative numerical flux  $\mathbf{P}_\Delta(\mathbf{U}_L, \Pi_L, \phi_L, \mathbf{U}_R, \Pi_R, \phi_R)$  and a source term discretization (located at the interface)  $\mathbf{S}_\Delta(\mathbf{U}_L, \Pi_L, \phi_L, \mathbf{U}_R, \Pi_R, \phi_R)$  thanks to the consistency in the integral sense. We get

$$\mathbf{P}_\Delta(\mathbf{U}_L, \Pi_L, \phi_L, \mathbf{U}_R, \Pi_R, \phi_R) = \frac{\mathbf{P}(\mathbf{U}_R, \Pi_R) + \mathbf{P}(\mathbf{U}_L, \Pi_L)}{2} - \frac{a}{2\rho_L} (\mathbf{U}_L^* - \mathbf{U}_L) - \frac{a}{2\rho_R} (\mathbf{U}_R - \mathbf{U}_R^*), \quad (\text{B.14a})$$

$$\mathbf{S}_\Delta(\mathbf{U}_L, \Pi_L, \phi_L, \mathbf{U}_R, \Pi_R, \phi_R) = [0, -\{\rho \partial_x \phi\}, -\{\rho u \partial_x \phi\}]^T, \quad (\text{B.14b})$$

so that for two neighbouring states  $(\mathbf{U}_j^n, \Pi_j^n, \phi_j)$  and  $(\mathbf{U}_{j+1}^n, \Pi_{j+1}^n, \phi_{j+1})$  across the cell interface  $j + 1/2$  that separates the cell  $j$  and the cell  $j + 1$ , the numerical conservative flux  $(0, \Pi_{j+1/2}^*, \Pi_{j+1/2}^* u_{j+1/2}^*)$  is defined by

$$(0, \Pi_{j+1/2}^*, \Pi_{j+1/2}^* u_{j+1/2}^*) = \mathbf{P}_\Delta(\mathbf{U}_j^n, \Pi_j^n, \phi_j, \mathbf{U}_{j+1}^n, \Pi_{j+1}^n, \phi_{j+1}), \quad (\text{B.15})$$

and the discrete source term  $\mathbf{S}_j$  within the cell  $j$  is given by

$$\mathbf{S}_j = \frac{\Delta x_{j+1/2}}{2\Delta x_j} \mathbf{S}_{j+1/2} + \frac{\Delta x_{j-1/2}}{2\Delta x_j} \mathbf{S}_{j-1/2}, \quad \mathbf{S}_{j+1/2} = \mathbf{S}_\Delta(\mathbf{U}_j^n, \Pi_j^n, \phi_j, \mathbf{U}_{j+1}^n, \Pi_{j+1}^n, \phi_{j+1}). \quad (\text{B.16})$$

Let us now give some properties of the approximate Riemann solver. Let us note  $e_k^* = E_k^* - (u_k^*)^2/2$ , the following lemma is a direct consequence of (B.9) that exhibits a reminiscent property associated with the Riemann invariants associated of the system (B.1) when  $\mathbf{S} = \mathbf{0}$ .

**Lemma Appendix B.1.**

$$e_k^* - \frac{(\Pi_k^*)^2}{2a^2} = e_k - \frac{(\Pi_k)^2}{2a^2}, \quad \mathcal{T}_k^* + \frac{\Pi_k^*}{a} = \mathcal{T}_k + \frac{\Pi_k}{a}, \quad k = L, R. \quad (\text{B.17})$$

The following positivity result is a direct consequence of (B.11e).

**Proposition Appendix B.1.**

1. If  $a$  is chosen large enough then  $\mathcal{T}_L^* > 0$  and  $\mathcal{T}_R^* > 0$ .
2.  $\mathcal{T}_L^* > 0$  and  $\mathcal{T}_R^* > 0$  is equivalent to  $u_L - a\mathcal{T}_L = u_L - a/\rho_L < u^* < u_R + a\mathcal{T}_R = u_R + a/\rho_R$ .

Following the lines of [16], we first prove two preliminary stability-related results. The differences from Lemma 1 of [16] is that the Riemann states we are dealing with here depend on the  $\mathcal{M}$  terms and that the specific volume we use is  $\mathcal{T}$  instead of  $1/\rho$  (that are different in the sub-system framework). However, the proof turns out to be almost identical.

**Proposition Appendix B.2.** Consider the intermediate states defined by (B.11).

and noting  $s_k = s^{\text{EOS}}(\mathcal{T}_k, s_k)$ , we have

$$e_k^* - e^{\text{EOS}}(\mathcal{T}_k^*, s_k) - \frac{(p^{\text{EOS}}(\mathcal{T}_k^*, s_k) - \Pi_k^*)^2}{2a^2} \geq 0, \quad (\text{B.18})$$

with  $e_k^* = E_k^* - \frac{u_k^{*2}}{2}$ .

*Proof.* We only describe the case  $k = R$ . Consider the function:

$$\begin{aligned} \chi(\mathcal{T}) = e^{\text{EOS}}(\mathcal{T}, s_R) - \frac{p^{\text{EOS}}(\mathcal{T}, s_R)^2}{2a^2} - e^{\text{EOS}}(\mathcal{T}_R^*, s_R) + \frac{p^{\text{EOS}}(\mathcal{T}_R^*, s_R)^2}{2a^2} \\ + p^{\text{EOS}}(\mathcal{T}_R^*, s_R) \left( \mathcal{T} + \frac{p^{\text{EOS}}(\mathcal{T}, s_R)}{a^2} - \mathcal{T}_R^* - \frac{p^{\text{EOS}}(\mathcal{T}_R^*, s_R)}{a^2} \right). \end{aligned} \quad (\text{B.19})$$

One can check that  $\chi'(\mathcal{T}) = (p^{\text{EOS}}(\mathcal{T}_R^*, s_R) - p^{\text{EOS}}(\mathcal{T}, s_R)) (1 - \rho^2 c^2(\mathcal{T}, s_R)/a^2)$ . We have  $\partial_{\mathcal{T}} p < 0$  from 3, we also assume that  $a$  is large enough. We have two different cases:

$$\begin{aligned} \mathcal{T}_R^* < \mathcal{T} < \mathcal{T}_R &\implies \chi'(\mathcal{T}) > 0 &\implies \chi(\mathcal{T}_R^*) < \chi(\mathcal{T}) < \chi(\mathcal{T}_R). \\ \mathcal{T}_R^* > \mathcal{T} > \mathcal{T}_R &\implies \chi'(\mathcal{T}) < 0 &\implies \chi(\mathcal{T}_R^*) < \chi(\mathcal{T}) < \chi(\mathcal{T}_R). \end{aligned} \quad (\text{B.20})$$

As  $\chi(\mathcal{T}_R^*) = 0$ , we have  $\chi(\mathcal{T}_R) > 0$ , in both cases. Accounting for (B.17), we get

$$\begin{aligned}
0 < \chi(\mathcal{T}_R) &= e^{\text{EOS}}(\mathcal{T}_R, s_R) - \frac{p^{\text{EOS}}(\mathcal{T}_R, s_R)^2}{2a^2} - e^{\text{EOS}}(\mathcal{T}_R^*, s_R) + \frac{p^{\text{EOS}}(\mathcal{T}_R^*, s_R)^2}{2a^2} \\
&\quad + p^{\text{EOS}}(\mathcal{T}_R^*, s_R) \left( \mathcal{T}_R + \frac{p^{\text{EOS}}(\mathcal{T}_R, s_R)}{a^2} - \mathcal{T}_R^* - \frac{p^{\text{EOS}}(\mathcal{T}_R^*, s_R)}{a^2} \right) \\
&= e_R^* - \frac{(\Pi_R^*)^2}{2a^2} - e^{\text{EOS}}(\mathcal{T}_R^*, s_R) + \frac{p^{\text{EOS}}(\mathcal{T}_R^*, s_R)^2}{2a^2} + p^{\text{EOS}}(\mathcal{T}_R^*, s_R) \left( \frac{\Pi_R^*}{a^2} - \frac{p^{\text{EOS}}(\mathcal{T}_R^*, s_R)}{a^2} \right) \\
&= e_R^* - e^{\text{EOS}}(\mathcal{T}_R^*, s_R) - \frac{(p^{\text{EOS}}(\mathcal{T}_R^*, s_R) - \Pi_R^*)^2}{2a^2}. \tag{B.21}
\end{aligned}$$

Similar lines can be used for  $k = L$ .  $\square$

We present a result concerning the behavior of the numerical scheme in the low Mach regime defined in section 6.1: we consider a one-dimensional smooth solution of the pressure subsystem (B.1) such that  $\partial_{\tilde{x}} \tilde{p} + \tilde{\rho}(\partial_{\tilde{x}} \phi) = O(\text{Ma}^2)$ . Then, we proceed as in [32] by evaluating the truncation error (in the sense of the Finite Difference) obtained by substituting these low Mach flow parameters into the finite volume update formula derived from the fluxes (B.15). We obtain the following results.

**Proposition Appendix B.3.** *In the low Mach regime, the rescaled discretization of the pressure system is consistent with*

$$\partial_{\tilde{t}} \tilde{p} = 0, \quad \partial_{\tilde{t}}(\tilde{\rho} \tilde{u}) + \frac{1}{\text{Ma}^2}(\partial_{\tilde{x}} \tilde{p} + \tilde{\rho}(\partial_{\tilde{x}} \phi)) = O(\Delta \tilde{t}) + O\left(\frac{\Delta \tilde{x}}{\text{Ma}}\right), \quad \partial_{\tilde{t}}(\tilde{\rho} \tilde{E}) + \partial_{\tilde{x}}(\tilde{\rho} \tilde{u}) = O(\Delta \tilde{t}) + O(\text{Ma} \Delta \tilde{x}). \tag{B.22}$$

If one performs a similar evaluation for the full FSLP scheme, one can see that the truncation error term  $O\left(\frac{\Delta \tilde{x}}{\text{Ma}}\right)$  that appears in the momentum equation of (B.22) will be the only error term whose magnitude is not uniform with respect to Ma. Similarly as in [28, 32, 1, 33], this truncation error term can be traced back to the non-centered part of  $\Pi_{j+1/2}^*$ . To tackle this issue, we adopt the modification used in [32, 1] by replacing  $\Pi_{j+1/2}^*$  with

$$\Pi_{j+1/2}^{*\theta} = \frac{1}{2}(\Pi_j^n + \Pi_{j+1}^n) - \theta_{j+1/2} \frac{a_{j+1/2}}{2}(u_{j+1}^n - u_j^n), \tag{B.23}$$

where  $\theta_{j+1/2} \in [0, 1]$ . This results in the update relation (12) that is a finite approximation of (B.1) with the flux definition (15). We will see in Appendix C how this resulting modified flux can still be associated with an Approximate Riemann solver.

### Appendix C. All-regime approximate Riemann solver for the pressure subsystem

Following similar lines as in [32]: although the modified pressure scheme (15) is defined as a flux scheme, it is possible to find an approximate Riemann solver  $(\mathbf{U}_{\text{RP}}^\theta, \Pi_{\text{RP}}^\theta, \mathcal{T}_{\text{RP}}^\theta)$  that enables to retrieve the numerical flux  $\mathbf{P}_{j+1/2}^\theta = (0, \Pi_{j+1/2}^{*\theta}, \Pi_{j+1/2}^{*\theta} u_{j+1/2}^*)$ . We suppose that  $(\mathbf{U}_{\text{RP}}^\theta, \Pi_{\text{RP}}^\theta, \mathcal{T}_{\text{RP}}^\theta)$  has the same structure as  $(\mathbf{U}_{\text{RP}}, \Pi_{\text{RP}}, \mathcal{T}_{\text{RP}})$ , we consider

$$(\mathbf{U}_{\text{RP}}^\theta, \Pi_{\text{RP}}^\theta, \mathcal{T}_{\text{RP}}^\theta, \phi_{\text{RP}}) \left( \frac{x - \tilde{x}}{t}; U_L, \Pi_L, \mathcal{T}_L, \phi_L, U_R, \Pi_R, \mathcal{T}_R, \phi_R \right) = \begin{cases} (U_L, \Pi_L, \mathcal{T}_L, \Phi_L), & \text{if } \frac{x - \tilde{x}}{t} \leq -\frac{a}{\rho_L}, \\ (U_L^{*\theta}, \Pi_L^{*\theta}, \mathcal{T}_L^{*\theta}, \Phi_L), & \text{if } -\frac{a}{\rho_L} < \frac{x - \tilde{x}}{t} \leq 0, \\ (U_R^{*\theta}, \Pi_R^{*\theta}, \mathcal{T}_R^{*\theta}, \Phi_R), & \text{if } 0 < \frac{x - \tilde{x}}{t} \leq \frac{a}{\rho_R}, \\ (U_R, \Pi_R, \mathcal{T}_R, \Phi_R), & \text{if } \frac{a}{\rho_R} < \frac{x - \tilde{x}}{t}, \end{cases} \tag{C.1}$$

where  $\Pi_k$ ,  $\mathcal{T}_k$  and  $\Phi_k$  verify (B.3),  $k = L, R$ . The states  $(\mathbf{U}_k^{*,\theta}, \Pi_k^{*,\theta}, \mathcal{T}_k^{*,\theta})$ ,  $k = L, R$  are yet to be defined. First, we impose that  $(\mathbf{U}_{RP}, \Pi_{RP}, \mathcal{T}_{RP})$  verifies the consistency in the integral sense

$$\begin{bmatrix} \frac{1}{\alpha} \mathbf{P}(\mathbf{U}_R) - \frac{1}{\alpha} \mathbf{P}(\mathbf{U}_L) \\ \frac{1}{\alpha} a^2 (u_R - u_L) \\ -(\frac{1}{\alpha} u_R - \frac{1}{\alpha} u_L) \end{bmatrix} = -\frac{a}{\alpha \rho_L} \begin{bmatrix} \mathbf{U}_L^{*,\theta} - \mathbf{U}_L \\ (\rho \Pi)_L^{*,\theta} - (\rho \Pi)_L \\ (\rho \mathcal{T})_L^{*,\theta} - (\rho \mathcal{T})_L \end{bmatrix} + \frac{a}{\alpha \rho_R} \begin{bmatrix} \mathbf{U}_R - \mathbf{U}_R^{*,\theta} \\ (\rho \Pi)_R - (\rho \Pi)_R^{*,\theta} \\ (\rho \mathcal{T})_R - (\rho \mathcal{T})_R^{*,\theta} \end{bmatrix} + \frac{\Delta x_L + \Delta x_R}{2} \begin{bmatrix} \frac{1}{\alpha} \{\mathbf{S}\} \\ 0 \\ 0 \end{bmatrix}. \quad (\text{C.2})$$

We then enforce that the numerical flux resulting from (C.2) is  $\mathbf{P}_\Delta^\theta$ , which boils down to require that

$$\begin{bmatrix} \frac{1}{\alpha} \mathbf{P}_\Delta^\theta \\ \frac{1}{\alpha} a^2 u_\Delta^\theta \\ -\frac{1}{\alpha} u_\Delta^\theta \end{bmatrix} (\mathbf{U}_L, \Pi_L, \phi_L, \mathbf{U}_R, \Pi_R, \phi_R) = \begin{bmatrix} \mathbf{P}(\mathbf{U}_R, \Pi_R) + \mathbf{P}(\mathbf{U}_L, \Pi_L) \\ a^2 u_R + a^2 u_L \\ -(u_R + u_L) \end{bmatrix} - \frac{a}{\rho_L} \begin{bmatrix} \mathbf{U}_L^\theta - \mathbf{U}_L \\ (\rho \Pi)_L^\theta - (\rho \Pi)_L \\ (\rho \mathcal{T})_L^\theta - (\rho \mathcal{T})_L \end{bmatrix} - \frac{a}{\rho_R} \begin{bmatrix} \mathbf{U}_R - \mathbf{U}_R^\theta \\ (\rho \Pi)_R - (\rho \Pi)_R^\theta \\ (\rho \mathcal{T})_R - (\rho \mathcal{T})_R^\theta \end{bmatrix}. \quad (\text{C.3})$$

Choosing  $\rho_k^{*,\theta} = \rho_k$ ,  $k = L, R$ , relation (C.2) and (C.3) provide a linear system with respect to  $u_k^{*,\theta}$ ,  $\Pi_k^{*,\theta}$ ,  $\mathcal{T}_k^{*,\theta}$  and  $E_k^{*,\theta}$ ,  $k = 1, 2$  whose solution is

$$\rho_L^* = \rho_L, \quad \rho_R^* = \rho_R, \quad (\text{C.4a})$$

$$E_L^{*,\theta} = E_L^* - (1 - \theta) \frac{u_R - u_L}{2} u^*, \quad E_R^{*,\theta} = E_R^* + (1 - \theta) \frac{u_R - u_L}{2} u^*, \quad (\text{C.4b})$$

$$u_L^{*,\theta} = u^* - (1 - \theta) \frac{u_R - u_L}{2}, \quad u_R^{*,\theta} = u^* + (1 - \theta) \frac{u_R - u_L}{2}, \quad (\text{C.4c})$$

$$\Pi_L^{*,\theta} = \Pi_L^*, \quad \Pi_R^{*,\theta} = \Pi_R^*, \quad (\text{C.4d})$$

$$\mathcal{T}_L^{*,\theta} = \mathcal{T}_L^*, \quad \mathcal{T}_R^{*,\theta} = \mathcal{T}_R^*. \quad (\text{C.4e})$$

We now turn to positivity-preserving related properties. Let us note  $e_k^{*,\theta} = E_k^{*,\theta} - u_k^{*,\theta 2}/2$ , we have the following result.

**Proposition Appendix C.1.** *Assuming again that  $a$  is large enough, we have*

$$e_k^{*,\theta} - e^{EOS}(\mathcal{T}_k^{*,\theta}, s_k) - \frac{(p^{EOS}(\mathcal{T}_k^{*,\theta}, s_k) - \Pi_k^{*,\theta})^2}{2a^2} + \frac{(1 - \theta)^2 (u_R - u_L)^2}{8} \geq 0. \quad (\text{C.5})$$

*Proof.* Let us consider the case  $k = R$ , by (C.4) we get

$$\begin{aligned} e_R^{*,\theta} - e_R^* &= E_R^{*,\theta} - E_R^* - \frac{1}{2} (u_R^{*,\theta 2} - u_R^{*2}) \\ &= (1 - \theta) \frac{u_R - u_L}{2} u^* - \frac{1}{2} \left( (u^*)^2 + u^* (1 - \theta) (u_R - u_L) + (1 - \theta)^2 \frac{(u_R - u_L)^2}{4} - (u^*)^2 \right) \\ &= -\frac{1}{8} (1 - \theta)^2 (u_R - u_L)^2. \end{aligned} \quad (\text{C.6})$$

Using (B.18), we obtain

$$\begin{aligned} e_R^{*,\theta} - e^{EOS}(\mathcal{T}_R^{*,\theta}, s_R) &= e_R^{*,\theta} - e_R^* + e_R^* - e^{EOS}(\mathcal{T}_R^{*,\theta}, s_R) = -\frac{1}{8} (1 - \theta)^2 (u_R - u_L)^2 + e_R^* - e^{EOS}(\mathcal{T}_R^{*,\theta}, s_R) \\ &\geq -\frac{1}{8} (1 - \theta)^2 (u_R - u_L)^2 + \frac{(p^{EOS}(\mathcal{T}_k^{*,\theta}, s_k) - \Pi_k^{*,\theta})^2}{2a^2}. \end{aligned}$$

Similar lines can be used for the case  $k = L$ . □

The relation (C.5) highlights the role of the inequality

$$\frac{1}{2a^2} (p^{EOS}(\mathcal{T}_k^{*,\theta}, s_k) - \Pi_k^*)^2 - \frac{(1 - \theta)^2 (u_R - u_L)^2}{8} \geq 0, \quad k = L, R \quad (\text{C.7})$$

in obtaining stability properties for the modified scheme. We have the following proposition.



**Proposition Appendix C.2.** Let us note:  $s_k^{*,\theta} = s^{\text{EOS}}(\mathcal{T}_k^{*,\theta}, e_k^{*,\theta})$ , if (C.7) is satisfied, then

- the modified approximate Riemann solver (C.1) preserves the positivity of the internal energy, that is to say:  $e_k^{*,\theta} > 0$ ,  $k = R, L$ ,
- the modified approximate Riemann solver (C.1) verifies  $s_k^{*,\theta} \geq s_k$ ,  $k = R, L$ ,
- the modified approximate Riemann solver (C.1) is entropy satisfying in the sense that

$$-a(s_L^{*,\theta} - s_L) + a(s_R - s_R^{*,\theta}) \geq 0. \quad (\text{C.8})$$

*Proof.* If (Appendix C.2) is satisfied, then  $e_k^{*,\theta} \geq e^{\text{EOS}}(\mathcal{T}_k^{*,\theta}, s_k)$ . By the assumption on the EOS, we have that  $e_k^{*,\theta} > 0$ . Now, considering a fixed  $\bar{\mathcal{T}} > 0$ , by (3) we know that  $e' \mapsto s^{\text{EOS}}(\bar{\mathcal{T}}, e')$  is increasing, thus we deduce that  $s^{\text{EOS}}(\mathcal{T}_k^{*,\theta}, e^{\text{EOS}}(\mathcal{T}_k^{*,\theta}, s_k^{*,\theta})) = s_k^{*,\theta} \geq s^{\text{EOS}}(\mathcal{T}_k^{*,\theta}, e^{\text{EOS}}(\mathcal{T}_k^{*,\theta}, s_k)) = s_k$ ,  $k = L, R$ . This implies (C.8).  $\square$

#### Appendix D. Eigenstructure of the off-equilibrium ( $73_{v=0}$ )

We propose in this section to study the eigenstructure of the relaxation system ( $73_{v=0}$ ). Let us first express the acoustic part of ( $73_{v=0}$ ) using a change of variables: accounting for  $e^P = E^P - (u^P)^2/2$ , the evolution equations for  $E^P$ , for  $\Pi^P$  and  $\mathcal{T}^P$  in ( $73_{a_{v=0}}$ ) yield

$$\partial_t(\rho^P e^P) + 2\Pi^P \partial_x u^P = 0, \quad 2\partial_x u^P = \partial_t(\rho^P \Pi^P / a^2). \quad (\text{D.1})$$

We thus obtain the stationary equations

$$\partial_t \left[ e^P - \frac{(\Pi^P)^2}{2a^2} \right] = 0, \quad \partial_t \left[ \mathcal{T}^P + \frac{\Pi^P}{a^2} \right] = 0. \quad (\text{D.2})$$

So now the acoustic subsystem ( $73_{a_{v=0}}$ ) takes the simple form

$$\partial_t \phi = 0, \quad \partial_t \rho^P = 0, \quad \partial_t \left[ e^P - \frac{(\Pi^P)^2}{2a^2} \right] = 0, \quad (\text{D.3a})$$

$$\partial_t(\rho^P u^P) + 2\partial_x \Pi^P + 2\rho^P \partial_x \phi^P = 0, \quad \partial_t(\rho^P \Pi^P) + 2a^2 \partial_x u^P = 0, \quad \partial_t \left[ \mathcal{T}^P + \frac{\Pi^P}{a^2} \right] = 0. \quad (\text{D.3b})$$

We now turn to the advection part of ( $73_{v=0}$ ): the subsystem ( $73_{b_{v=0}}$ ) takes the simple form

$$\partial_t \rho^A + \partial_x(2\rho^A u^P) = 0, \quad \partial_t \left[ \rho^A \mathcal{T}^A - \frac{\rho^P \Pi^P}{a^2} \right] = 0, \quad \partial_t b^A + 2u^P \partial_x b^A = 0, \quad b^A \in \{u^A, E^A, \Pi^A\}. \quad (\text{D.4})$$

Therefore if we set

$$\mathbf{W}^T = \left[ u^P, \Pi^P, \rho^P, \phi, e^P - \frac{(\Pi^P)^2}{2a^2}, \mathcal{T}^P + \frac{\Pi^P}{a^2}, \rho^A \mathcal{T}^A - \frac{\rho^P \Pi^P}{a^2}, u^A, \Pi^A, E^A, \rho^A \right], \quad (\text{D.5})$$

we can see that ( $73_{v=0}$ ) can be recast into the following quasilinear system

$$\partial_t \mathbf{W} + \mathbf{M}(\mathbf{W}) \partial_x \mathbf{W} = 0, \quad \mathbf{M}(\mathbf{W}) = \begin{bmatrix} 0 & \frac{2}{\rho^P} & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{2a^2}{\rho^P} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2u^P & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2u^P & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2u^P & 0 \\ 2\rho^A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2u^P \end{bmatrix}. \quad (\text{D.6})$$

It is then straightforward to see that the eigenvalues of  $M(W)$  are  $2u^P$  (with an algebraic multiplicity 4), 0 (with an algebraic multiplicity 5) and  $\pm 2a/\rho^P$ .

The eigenvectors  $(\mathbf{r}_0^{(k)})_{k=1,\dots,3}$ ,  $(\mathbf{r}_{u^P}^{(k)})_{k=1,\dots,4}$  and  $\mathbf{r}_\pm$  that are respectively associated with 0,  $2u^P$  and  $\pm 2a/\rho^P$  are

$$\mathbf{r}_0^{(1)} = [0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0]^T, \quad \mathbf{r}_0^{(2)} = [0, -\rho^P, 0, 1, 0, 0, 0, 0, 0, 0, 0]^T, \quad (\text{D.7a})$$

$$\mathbf{r}_0^{(3)} = [0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0]^T, \quad \mathbf{r}_0^{(4)} = [0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0]^T \quad (\text{D.7b})$$

$$\mathbf{r}_0^{(5)} = [0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0]^T, \quad (\text{D.7c})$$

$$\mathbf{r}_{u^P}^{(1)} = [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0]^T, \quad \mathbf{r}_{u^P}^{(2)} = [0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0]^T, \quad (\text{D.7d})$$

$$\mathbf{r}_{u^P}^{(3)} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0]^T, \quad \mathbf{r}_{u^P}^{(4)} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1]^T, \quad (\text{D.7e})$$

$$\mathbf{r}_+ = \left[ 1, a, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{\rho^A \rho^P}{\rho^P u^P - a} \right]^T, \quad \mathbf{r}_- = \left[ 1, -a, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{\rho^A \rho^P}{\rho^P u^P + a} \right]^T, \quad (\text{D.7f})$$

so that (D.6) is hyperbolic and only involves linearly degenerate fields.