



HAL
open science

High-order staggered schemes for compressible hydrodynamics. Weak consistency and numerical validation

Gautier Dakin, Bruno Despres, Stéphane Jaouen

► **To cite this version:**

Gautier Dakin, Bruno Despres, Stéphane Jaouen. High-order staggered schemes for compressible hydrodynamics. Weak consistency and numerical validation. *Journal of Computational Physics*, 2019, 376, pp.339-364. hal-01967344

HAL Id: hal-01967344

<https://hal.sorbonne-universite.fr/hal-01967344>

Submitted on 4 Jan 2019

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.

High-order staggered schemes for compressible hydrodynamics. Weak consistency and numerical validation

Gautier Dakin^a, Bruno Després^b, Stéphane Jaouen^a

^aCEA, DAM, DIF, F-91297 Arpaçon, France

^bUPMC, LJLL, Paris, France

Abstract

Staggered grids schemes, formulated in internal energy, are commonly used for CFD applications in industrial context. Here, we prove the consistency of a class of high-order Lagrange-Remap staggered schemes for solving the Euler equations in 1D and 2D. The main result of the paper is that using an *a posteriori* internal energy corrector, the Lagrangian schemes are proved to be conservative in mass, momentum and total energy and to be weakly consistent with the 1D Lagrangian formulation of the Euler equations. Extension in 2D is done using directional splitting methods and face-staggering. Numerical examples in both 1D and 2D illustrate the accuracy, the convergence and the robustness of the schemes.

Keywords: Finite Volume, internal energy corrector, weak consistency, Lagrange-remap, Euler equations, face-staggering, high-order accuracy

1. Introduction

Historically the first shock capturing scheme was formulated in internal energy and based on staggered grids [44]. Still, staggered grids schemes are routinely used for CFD applications in industrial context [41, 34, 16, 11, 18, 25, 5]: seminal references are [39, 6, 7, 49]. Two reasons can be argued which are that staggered schemes often need less degrees of freedom than collocated ones to obtain the same accuracy for acoustic propagation [1], and they naturally capture low Mach regimes [11] which is not so easy for standard collocated schemes (see [15] and references therein). However, solid mathematical foundations are scarcely available for such schemes, for many reasons. In this work, we concentrate on three reasons which are: firstly, the staggered structure is an obstacle for a natural use of Lax theorem [23] which establishes the weak consistency for collocated finite volume schemes for shock calculations; secondly, internal energy equation is not in divergence form which is mandatory for the Lax theorem; thirdly, the development of high-order accurate schemes, which is now a standard trend for industrial CFD codes, is even more challenging due to the first two points. In this paper, we aim at tackling these three issues.

The model problem are the Euler compressible hydrodynamics equations for the description of inviscid compressible flows in the absence of source terms

$$\partial_t \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ \rho e \end{pmatrix} + \nabla \cdot \begin{pmatrix} \rho \mathbf{u} \\ \rho \mathbf{u} \otimes \mathbf{u} + p \underline{\mathbf{I}} \\ (\rho e + p) \mathbf{u} \end{pmatrix} = \mathbf{0}. \quad (1)$$

The variables are the density ρ , the velocity field \mathbf{u} and the total energy e . It is convenient to use also kinetic energy $e_{\text{kin}} = \frac{1}{2} \|\mathbf{u}\|^2$, internal energy $\epsilon = e - e_{\text{kin}}$ and specific volume $\tau = \frac{1}{\rho}$. The system is closed with an equation of state which links pressure, internal energy and specific volume with $p = \text{EOS}(\tau, \epsilon)$. In this work we will rather use another hyperbolic system written in non-conservative form which is representative of the equations solved in

Email addresses: gautier.dakin@cea.fr (Gautier Dakin), despres@ann.jussieu.fr (Bruno Després), stephane.jaouen@cea.fr (Stéphane Jaouen)

staggered Lagrange+remap and Lagrange+ALE numerical methods. Without restriction, the equations are hereafter written in dimension one (extended to multi dimension with directional splitting)

$$\begin{cases} D_t \rho_0 \tau & - \partial_X u = 0, \\ D_t \rho_0 u & + \partial_X p = 0, \\ D_t \rho_0 \epsilon & + p \partial_X u = 0, \\ D_t \rho_0 e_{\text{kin}} & + u \partial_X p = 0, \end{cases} \quad (2)$$

where ρ_0 is the initial density and D_t is the material derivative. Equation (2) is non conservative. The terms $p \partial_X u$ and $u \partial_X p$ are well-defined for smooth enough functions, but are not in general in the sense of distributions. Discretization of non-conservative products has tremendous consequences for schemes solving eq. (2) [42, 31]: such discretizations are the key to build high-order accurate schemes that able to capture correctly shocks and discontinuities. An issue which is not discussed in our work is the entropy condition for system (1) and system (2). Although extremely important on the theoretical level [33, 12], the entropy condition is technically much too difficult to address in the context of high-order schemes analyzed in this work. Moreover, all our numerical experiments, and in particular the ones reported at the end of this work, show that numerical solutions are in accordance with the entropy principle. This fact leads us to think that the entropy principle is not a practical issue in our context.

Before describing our main results for the discretization of system (2), we give an overview of the literature on such topics. Staggered schemes originate from the vNR scheme [44], based on a node staggering which is *a priori* not conservative in total energy. Furthermore, without any artificial viscosity, the scheme is unable to correctly capture strong shocks. The main difficulty for schemes formulated in internal energy is that this is not any longer a conservation law. On a mathematical continuous level, the term appearing in the internal energy evolution is not defined in the sense of distributions, for velocity and pressure as bounded functions ($(u, p) \in L^\infty$). The use of artificial viscosity solves this problem by smoothing the pressure. With an appropriate definition of artificial viscosity, the internal energy evolution term becomes well-defined. The default of total energy conservation was highlighted in 1961 by Trulio and Trigger [42]. Indeed, for non-constant time-steps, the vNR scheme is not conservative in total energy. An implicit conservative version of the vNR scheme was proposed in [42], still formulated in internal energy. The spatial staggering of variables was kept but not the temporal one. Similarly, works done by Popov and Samarskii [31] developed a similar staggered scheme with implicitation in time. In the early 1970s, DeBar used a Lagrange-remap formalism for the Trulio–Trigger scheme [6, 7]. At the end of each Lagrangian phase, the variables were projected on the original grids. He identified a lack of conservation due to this procedure. In fact, the projection of momentum highly dissipates kinetic energy, and so leads to a dissipation of the reconstructed total energy. A correction in internal energy to recover global total energy conservation and ensure correct shock capturing was proposed. Later, and using the earlier works by DeBar, several multifluid Eulerian hydrocodes with interface reconstruction on 2D Cartesian grids [39] were developed, based on a face-staggering of variables. Those hydrocodes relied on the Trulio-Trigger implicit Lagrangian scheme, making use of a Lagrange-remap approach with Strang splitting. The splitting was made to consider first a 1D Lagrange-remap scheme in the x -direction, and then in the y -direction. This kind of directional splitting yields the advantage of an easy extension from one dimensional problems to multi-dimensional ones. Later, a strictly explicit predictor-corrector conservative version of the Trulio-Trigger scheme was reported by Woodward and Colella in [46]. This version was called the BBC scheme. It is a 2D Lagrange-remap scheme on staggered Cartesian grids based on a 1D Lagrange-remap setting with Strang dimensional splitting. The total energy conservation result has been credited to Noh [29]. The retained staggering of variables is the C-type one, based on Arakawa classification system. Caramana in 1998 [3] introduced the so-called compatible Lagrangian hydrodynamics for node-staggering schemes. The idea of compatible Lagrangian method is to discretize properly the internal energy evolution in order to automatically satisfy the conservation of total energy. In [2], the authors highlight the properties of such a discretization. Mainly, the emphasis is laid on accuracy, consistency and stability of the compatible Lagrangian scheme. Simultaneously, Youngs developed a staggered scheme in which the velocity components were based on the node of the grids [49, 41, 34]. The scheme, although formulated in internal energy, is conservative in total energy, using a similar internal energy corrector as DeBar during the remapping phase. Recently for unstructured grids, Herbin, Gallouet et al. [16, 11, 18] developed similar procedures to recover local conservation of total energy for the compressible Navier–Stokes and Euler equations for a face-staggering. Very recently, a paper by Llor et al. proposed a conservative, compatible and entropic version of the original vNR schemes

[25] staggered in both time and space whereas Dakin et al. proposed conservative and high-order accurate schemes using face-staggering in space and no staggering in time [5].

Our main theoretical results, developed for simplicity for the high-order schemes recently announced in [5], are twofolds. The theoretical results are described for system (2) which contains all the difficulties. Firstly, we prove in Proposition 1 that the high-order staggered schemes are weakly consistent in the sense of Lax for the barotropic case. Secondly, we prove in Theorem 1 the weak consistency for the remaining total energy equation: the technical difficulty is solved with a new high-order accurate internal energy corrector (called in [5] kinetic energy synchronization). In particular, we develop an original full treatment of internal energy correctors for general high-order methods. To our knowledge, these results are original with respect to the literature cited above. The remap or ALE phase yields no additional difficulty so is not described, although it is mandatory to transfer the good properties obtained for (2) to the full Eulerian system (1). The interest of our theoretical results for numerical purposes is assessed in a second part with extensive Lagrange+remap numerical tests in 1D and 2D for smooth flows and for shock/discontinuity regimes. The general behavior and the convergence of the internal energy corrector is shown for a high-order discretization of the Sod problem.

The paper is organized as follows. Section 2 is devoted to 1D Lagrange-remap staggered schemes and especially to the *a posteriori* internal energy corrector which enables both conservation of total energy and weak consistency. Section 3 details the extension of the 1D schemes to 2D problems via directional splitting methods. Last, section 4 provides numerical examples and results to illustrate the high-order accuracy of the schemes as well as their ability to handle shocks and discontinuities.

2. Explicit consistent and conservative high-order staggered schemes in 1D

The objective of this section is to present the machinery involved to build high-order schemes on face-staggered schemes and to study their properties.

2.1. Formulation of Runge–Kutta Lagrangian finite volume schemes

We detail the path to build high-order schemes using face-staggering discretization.

2.1.1. Semi-discrete formulation of the Lagrangian finite volume schemes

To get the semi-discrete formulation of the Lagrangian finite volume schemes, system (2) is integrated in time between t^n and t^{n+1} over a cell $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ for the thermodynamics variables $\rho_0\tau$ and $\rho_0\epsilon$ and over a cell $[x_i, x_{i+1}]$ for the ρ_0u and ρe_{kin} . It yields

$$\begin{cases} \Delta X(\overline{\rho_0\tau}_i^{n+1} - \overline{\rho_0\tau}_i^n) &= \int_{t^n}^{t^{n+1}} u_{i+\frac{1}{2}}(\theta) - u_{i-\frac{1}{2}}(\theta) d\theta, \\ \Delta X(\overline{\rho_0u}_{i+\frac{1}{2}}^{n+1} - \overline{\rho_0u}_{i+\frac{1}{2}}^n) &= \int_{t^n}^{t^{n+1}} p_{i+1}(\theta) - p_i(\theta) d\theta, \\ \Delta X(\overline{\rho_0\epsilon}_i^{n+1} - \overline{\rho_0\epsilon}_i^n) &= \int_{t^n}^{t^{n+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (p\partial_X u)(y, \theta) dy d\theta, \\ \Delta X(\overline{\rho_0e_{\text{kin}}}_{i+\frac{1}{2}}^{n+1} - \overline{\rho_0e_{\text{kin}}}_{i+\frac{1}{2}}^n) &= \int_{t^n}^{t^{n+1}} \int_{x_i}^{x_{i+1}} (u\partial_X p)(y, \theta) dy d\theta, \\ p_i &= EOS(\tau_i, \epsilon_i). \end{cases} \quad (3)$$

Notations $\overline{p\delta u}_i$ and $\overline{u\delta p}_{i+\frac{1}{2}}$ are introduced as $\overline{p\delta u}_i = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (p\partial_X u)(y, \theta) dy$ and $\overline{u\delta p}_{i+\frac{1}{2}} = \int_{x_i}^{x_{i+1}} (u\partial_X p)(y, \theta) dy$. Equation (3) rewrites

$$\begin{cases} \Delta X(\overline{\rho_0\tau}_i^{n+1} - \overline{\rho_0\tau}_i^n) &= \int_{t^n}^{t^{n+1}} u_{i+\frac{1}{2}}(\theta) - u_{i-\frac{1}{2}}(\theta) d\theta, \\ \Delta X(\overline{\rho_0u}_{i+\frac{1}{2}}^{n+1} - \overline{\rho_0u}_{i+\frac{1}{2}}^n) &= \int_{t^n}^{t^{n+1}} p_{i+1}(\theta) - p_i(\theta) d\theta, \\ \Delta X(\overline{\rho_0\epsilon}_i^{n+1} - \overline{\rho_0\epsilon}_i^n) &= \int_{t^n}^{t^{n+1}} \overline{p\delta u}_i(\theta) d\theta, \\ \Delta X(\overline{\rho_0e_{\text{kin}}}_{i+\frac{1}{2}}^{n+1} - \overline{\rho_0e_{\text{kin}}}_{i+\frac{1}{2}}^n) &= \int_{t^n}^{t^{n+1}} \overline{u\delta p}_{i+\frac{1}{2}}(\theta) d\theta, \\ p_i &= EOS(\tau_i, \epsilon_i). \end{cases} \quad (4)$$

Before performing any kind of time integration, one must first address the issue of computing with high-order accuracy the point-wise values of p , u , τ and ϵ .

2.1.2. Very high-order in space

To achieve high-order resolution, it is mandatory to compute with high-order accuracy the point-wise (resp. average) values from the average (resp. point-wise) ones. Tables 1 to 4 give the coefficients for centered symmetric polynomial reconstructions using eq. (5). Although other reconstructions may be used, centered and symmetric ones are retained here and are sufficient on uniform Cartesian grids.

$$\left\{ \begin{array}{l} \phi_{\xi(i)} = \sum C_k \bar{\phi}_{\xi(i)+k}, \\ \bar{\phi}_{\xi(i)} = \sum \widehat{C}_k \phi_{\xi(i)+k}, \\ \delta \phi_{\xi(i)} = \sum_{k \geq 0} d_k \left(\phi_{\xi(i)+k+\frac{1}{2}} - \phi_{\xi(i)-k-\frac{1}{2}} \right), \\ \phi_{\xi(i)} = \sum_{k \geq 0} r_k \left(\phi_{\xi(i)+k+\frac{1}{2}} + \phi_{\xi(i)-k-\frac{1}{2}} \right), \\ \phi_{\xi(i)} = \frac{(\rho_0 \phi)_{\xi(i)}}{(\rho_0)_{\xi(i)}}, \end{array} \right. \quad \text{with } \xi(i) = \begin{cases} i & \text{on primal grid,} \\ i + \frac{1}{2} & \text{on dual grid,} \end{cases} \quad (5)$$

The non-conservative terms $\overline{\psi \delta \phi}$ (with $(\psi, \phi) = (u, p)$ and $(\psi, \phi) = (p, u)$) of eq. (4) are computed by

1. Applying the δ operator defined in (5) to point-wise values of ϕ using coefficients in table 3.
2. Multiplying by point-wise values of ψ , then reconstructing average values using coefficients in table 2 and second equation of (5).

Order	C_0	$C_{\pm 1}$	$C_{\pm 2}$	$C_{\pm 3}$	$C_{\pm 4}$
2 nd	1	0	0	0	0
3 rd	$\frac{13}{12}$	$-\frac{1}{24}$	0	0	0
4 th and 5 th	$\frac{1067}{960}$	$-\frac{29}{480}$	$\frac{3}{640}$	0	0
6 th and 7 th	$\frac{30251}{26880}$	$-\frac{7621}{107520}$	$\frac{159}{17920}$	$-\frac{5}{7168}$	0
8 th and 9 th	$\frac{5851067}{5160960}$	$-\frac{100027}{1290240}$	$\frac{31471}{2580480}$	$-\frac{425}{258048}$	$\frac{35}{294912}$

Table 1: Coefficients for the finite volume computation of point-wise values from cell-average ones.

Order	\widehat{C}_0	$\widehat{C}_{\pm 1}$	$\widehat{C}_{\pm 2}$	$\widehat{C}_{\pm 3}$	$\widehat{C}_{\pm 4}$
2 nd	1	0	0	0	0
3 rd	$\frac{11}{12}$	$\frac{1}{24}$	0	0	0
4 th and 5 th	$\frac{863}{960}$	$\frac{77}{1440}$	$-\frac{17}{5760}$	0	0
6 th and 7 th	$\frac{215641}{241920}$	$\frac{6361}{107520}$	$-\frac{281}{53760}$	$\frac{367}{967680}$	0
8 th and 9 th	$\frac{41208059}{46448640}$	$\frac{3629953}{58060800}$	$-\frac{801973}{116121600}$	$\frac{49879}{58060800}$	$-\frac{27859}{464486400}$

Table 2: Coefficients for the finite volume computation of average values from point-wise ones.

2.1.3. Runge–Kutta based time discretization

We consider N th order *explicit* schemes with s sub-cycles with the following notations for Runge-Kutta sequences: α_m is the time step for the m th sub-cycle, $a_{m,l}$ the m, l term of the Butcher table and θ_l the l th reconstruction coefficient for the last step. It is given by table 5. The system (6) details one Runge-Kutta sub-cycle at time $t^{n+\alpha_m}$.

Order	d_0	d_1	d_2	d_3	d_4
2 nd	1	0	0	0	0
3 rd	$\frac{9}{8}$	$-\frac{1}{24}$	0	0	0
4 th and 5 th	$\frac{75}{64}$	$-\frac{25}{384}$	$\frac{3}{640}$	0	0
6 th and 7 th	$\frac{1225}{1024}$	$-\frac{245}{3072}$	$\frac{49}{5120}$	$-\frac{5}{7168}$	0
8 th and 9 th	$\frac{19845}{16384}$	$-\frac{735}{8192}$	$\frac{567}{40960}$	$-\frac{405}{229376}$	$\frac{35}{294912}$

Table 3: Coefficients for the δ operator.

Order	r_0	r_1	r_2	r_3	r_4
2 nd	$\frac{1}{2}$	0	0	0	0
3 rd	$\frac{9}{16}$	$-\frac{1}{16}$	0	0	0
4 th and 5 th	$\frac{75}{128}$	$-\frac{25}{256}$	$\frac{3}{256}$	0	0
6 th and 7 th	$\frac{1225}{2048}$	$-\frac{245}{2048}$	$\frac{49}{2048}$	$-\frac{5}{2048}$	0
8 th and 9 th	$\frac{19845}{32768}$	$-\frac{2205}{16384}$	$\frac{567}{16384}$	$-\frac{405}{65536}$	$\frac{35}{65536}$

Table 4: Coefficients for the interpolation of cell-centered values from staggered ones and *vice-versa*.

$$\left\{ \begin{array}{l} \overline{\rho_0 \tau_i^{n+\alpha_m}} = \overline{\rho_0 \tau_i^n} + \frac{\Delta t}{\Delta X} \sum_{l=0}^{m-1} a_{m,l} du_i^{n+\alpha_l}, \\ \overline{\rho_0 u_{i+\frac{1}{2}}^{n+\alpha_m}} = \overline{\rho_0 u_{i+\frac{1}{2}}^n} - \frac{\Delta t}{\Delta X} \sum_{l=0}^{m-1} a_{m,l} dp_{i+\frac{1}{2}}^{n+\alpha_l}, \\ \overline{\rho_0 \epsilon_i^{n+\alpha_m}} = \overline{\rho_0 \epsilon_i^n} - \frac{\Delta t}{\Delta X} \sum_{l=0}^{m-1} a_{m,l} p \delta u_i^{n+\alpha_l}, \\ p_i^{n+\alpha_m} = EOS(\tau_i^{n+\alpha_m}, \epsilon_i^{n+\alpha_m}). \end{array} \right. \quad (6)$$

Here, $d\phi$ is the difference between two consecutive point-wise values: $d\phi_i = \phi_{i+\frac{1}{2}} - \phi_{i-\frac{1}{2}}$ and $d\phi_{i+\frac{1}{2}} = \phi_{i+1} - \phi_i$. Note that in (6), meaning for intermediate Runge–Kutta time-step, there is no need to compute the evolution of the kinetic

α_1	$a_{1,0}$	0	0	0	...
α_2	$a_{2,0}$	$a_{2,1}$	0	0	...
\vdots	\vdots	\vdots	\ddots
α_s	$a_{s,0}$	$a_{s,s-1}$	0
1	θ_0	θ_1	...	θ_{s-1}	θ_s

Table 5: Example of Butcher table for explicit Runge–Kutta sequence with s sub-cycles.

energy, nor the position of the cells face $x_{i+\frac{1}{2}}$. System (7) details the final step at time t^{n+1} .

$$\left\{ \begin{array}{l} \overline{\rho_0 \tau_i}^{n+1} = \overline{\rho_0 \tau_i}^n + \frac{\Delta t}{\Delta X} \sum_{l=0}^s \theta_l d u_i^{n+\alpha_l}, \\ \overline{\rho_0 u_{i+\frac{1}{2}}}^{n+1} = \overline{\rho_0 u_{i+\frac{1}{2}}}^n - \frac{\Delta t}{\Delta X} \sum_{l=0}^s \theta_l d p_{i+\frac{1}{2}}^{n+\alpha_l}, \\ \overline{\rho_0 \epsilon_i}^{n+1} = \overline{\rho_0 \epsilon_i}^n - \frac{\Delta t}{\Delta X} \sum_{l=0}^s \theta_l p \delta u_i^{n+\alpha_l}, \\ \overline{\rho_0 e_{kin, i+\frac{1}{2}}}^{n+1} = \overline{\rho_0 e_{kin, i+\frac{1}{2}}}^n - \frac{\Delta t}{\Delta X} \sum_{l=0}^s \theta_l u \delta p_{i+\frac{1}{2}}^{n+\alpha_l}, \\ x_{i+\frac{1}{2}}^{n+1} = x_{i+\frac{1}{2}}^n + \Delta t \sum_{l=0}^s \theta_l u_{i+\frac{1}{2}}^{n+\alpha_l}, \\ p_i^{n+1} = EOS(\tau_i^{n+1}, \epsilon_i^{n+1}). \end{array} \right. \quad (7)$$

2.2. Properties of the staggered schemes (6)-(7)

The objective of this section is to study different properties of the schemes (6)-(7) and to give some preliminary results to the proof of the main theorem.

DEFINITION 1 (Consistency of a flux in a conservative finite volume approximation [12]). *Consider a finite volume scheme which writes under the form*

$$\mathbf{U}_i^{n+1} - \mathbf{U}_i^n + \frac{t^{n+1} - t^n}{\Delta x} \left[\mathbf{f}_{i+\frac{1}{2}}^* - \mathbf{f}_{i-\frac{1}{2}}^* \right] = \mathbf{0}. \quad (8)$$

Let Φ be a Lipschitz continuous function and let $\mathbf{f}_{i+\frac{1}{2}}^* = \Phi(\mathbf{U}_{i-p+1}^n, \dots, \mathbf{U}_{i+r}^n)$, $\forall i \in \mathbb{Z}$, $\forall n \in \mathbb{N}$ with $(r, p) \in \mathbb{N}^2$. The flux is said consistent if $\Phi(\mathbf{U}, \dots, \mathbf{U}) = \mathbf{f}(\mathbf{U})$.

DEFINITION 2 (Weak consistency, inspired from [8, 17, 18, 19]). *Consider a given numerical scheme. Assume that numerical solutions $\mathbf{U}_{\Delta x}$ are bounded in $(L^\infty)^N$ and BV^N . Assume that numerical solutions converge, as $\Delta x \rightarrow 0$, in $(L^1_{loc})^N$ toward a limit $\widehat{\mathbf{U}} \in (L^\infty)^N$. If we can prove that $\widehat{\mathbf{U}}$ is a weak solution, then the scheme is said weakly consistent.*

REMARK 1. *A practical and simple criterion for weak consistency is to show that the flux is consistent [23, 8, 12].*

The main difficulty in the analysis of schemes (6)-(7) is precisely that they do not have a natural divergence structure like (8). In particular, there is no natural flux for the internal and kinetic energies equations.

2.2.1. Conservation properties of the schemes

In this section, we focus on proving that the schemes preserve some quantities, among which mass and momentum. Two definitions of total energies are introduced in order to study the schemes properties concerning the conservation of total energy.

DEFINITION 3. *A first total energy, based on the kinetic energy reconstructed from the momentum, of the system at time $t = t^n$, denoted E^n , is defined as*

$$E^n = \Delta X \left(\sum_i \overline{\rho_0 \epsilon_i}^n + \sum_i \overline{\rho_0 u^2_{i+\frac{1}{2}}}^n \right). \quad (9)$$

DEFINITION 4. *A second total energy, based on the discretized kinetic energy, of the system at time $t = t^n$, denoted \mathcal{E}^n , is defined as*

$$\mathcal{E}^n = \Delta X \left(\sum_i \overline{\rho_0 \epsilon_i}^n + \sum_i \overline{\rho_0 e_{kin, i+\frac{1}{2}}}^n \right). \quad (10)$$

A desired feature is that the mass, momentum and the total energy E^n defined in definition 3 are conserved for periodic or wall boundary conditions. A first result is summarized in the following lemma.

LEMMA 1 (Mass and momentum conservation). *For all explicit Runge–Kutta sequences, all spatial reconstructions, the schemes (6)-(7) are conservative in mass and momentum.*

Proof. Conservation of mass writes $\sum_i \overline{\rho_0 \tau_i^{n+1}} - \overline{\rho_0 \tau_i^n} = \sum_i \frac{\Delta t}{\Delta X} \sum_l \theta_l (u_{i+\frac{1}{2}}^{n+\alpha_l} - u_{i-\frac{1}{2}}^{n+\alpha_l}) = \frac{\Delta t}{\Delta X} \sum_l \theta_l \sum_i (u_{i+\frac{1}{2}}^{n+\alpha_l} - u_{i-\frac{1}{2}}^{n+\alpha_l})$. Hence for wall or periodic boundary conditions, it yields $\sum_i \overline{\rho_0 \tau_i^{n+1}} - \overline{\rho_0 \tau_i^n} = 0$. The conservation for momentum is very similar. Indeed, $\sum_i \overline{\rho_0 u_{i+\frac{1}{2}}^{n+1}} - \overline{\rho_0 u_{i+\frac{1}{2}}^n} = \sum_i -\frac{\Delta t}{\Delta X} \sum_l \theta_l (p_{i+1}^{n+\alpha_l} - p_i^{n+\alpha_l}) = -\frac{\Delta t}{\Delta X} \sum_l \theta_l \sum_i (p_{i+1}^{n+\alpha_l} - p_i^{n+\alpha_l})$. Hence for periodic boundary conditions, it yields $\sum_i \overline{\rho_0 u_{i+\frac{1}{2}}^{n+1}} - \overline{\rho_0 u_{i+\frac{1}{2}}^n} = 0$. \square

A second result is that the schemes conserve the total energy \mathcal{E}^n defined in definition 4.

LEMMA 2 (\mathcal{E} global conservation). *For all explicit Runge-Kutta sequences, all spatial reconstructions, the schemes (6)-(7) formulated in internal energy conserve the quantity \mathcal{E} , meaning $\mathcal{E}^{n+1} - \mathcal{E}^n$.*

Proof.

$$\begin{aligned} \mathcal{E}^{n+1} - \mathcal{E}^n &= \sum_i (\overline{\rho_0 \epsilon_i^{n+1}} - \overline{\rho_0 \epsilon_i^n}) + \sum_i (\overline{\rho_0 e_{kin}^{n+1}} - \overline{\rho_0 e_{kin}^n}) \\ &= -\frac{\Delta t}{\Delta X} \sum_i \sum_{l=1}^s \theta_l (\overline{p \delta u_i^{n+\alpha_l}} + \overline{u \delta p_{i+\frac{1}{2}}^{n+\alpha_l}}) \\ &= -\frac{\Delta t}{\Delta X} \sum_i \sum_{l=1}^s \sum_k \sum_{k'} \theta_l \widehat{C}_k d_{k'} (p_{i+k}^{n+\alpha_l} u_{i+k+k'+\frac{1}{2}}^{n+\alpha_l} + u_{i+k+\frac{1}{2}}^{n+\alpha_l} p_{i+k+k'+1}^{n+\alpha_l} - p_{i+k}^{n+\alpha_l} u_{i+k-k'-\frac{1}{2}}^{n+\alpha_l} - u_{i+k+\frac{1}{2}}^{n+\alpha_l} p_{i+k-k'}^{n+\alpha_l}). \end{aligned}$$

Making the change of index $i \leftarrow i + k'$ in the first term and $i \leftarrow i + k' + 1$ in the second term of the RHS we get the result for wall (with non-trivial definitions of ghost-cell values) or periodic boundary conditions.

$$\mathcal{E}^{n+1} - \mathcal{E}^n = -\frac{\Delta t}{\Delta X} \sum_i \sum_{l=1}^s \sum_k \sum_{k'} \theta_l \widehat{C}_k d_{k'} (p_{i+k-k'}^{n+\alpha_l} u_{i+k+\frac{1}{2}}^{n+\alpha_l} + u_{i+k-k'-\frac{1}{2}}^{n+\alpha_l} p_{i+k}^{n+\alpha_l} - p_{i+k}^{n+\alpha_l} u_{i+k-k'-\frac{1}{2}}^{n+\alpha_l} - u_{i+k+\frac{1}{2}}^{n+\alpha_l} p_{i+k-k'}^{n+\alpha_l}) = 0.$$

\square

REMARK 2. *However, one observes numerically that the schemes (6)-(7) do not preserve the quantity E as defined in definition 3. More details about the non-conservation of E are given in the following.*

An internal energy corrector detailed in section 2.3 has been developed to tackle the conservation of E . The design of the internal energy corrector is based on the schemes conservation of \mathcal{E} .

2.2.2. Weak consistency of the barotropic version

In order to greatly simplify the proof of the main theorem, we lay first the stress on the weak consistency of the schemes barotropic version. Intermediate stages write

$$\begin{cases} \overline{\rho_0 \tau_i^{n+\alpha_m}} = \overline{\rho_0 \tau_i^n} + \frac{\Delta t}{\Delta X} \sum_{l=0}^{m-1} a_{m,l} (u_{i+\frac{1}{2}}^{n+\alpha_l} - u_{i-\frac{1}{2}}^{n+\alpha_l}), \\ \overline{\rho_0 u_{i+\frac{1}{2}}^{n+\alpha_m}} = \overline{\rho_0 u_{i+\frac{1}{2}}^n} - \frac{\Delta t}{\Delta X} \sum_{l=0}^{m-1} a_{m,l} (p_{i+1}^{n+\alpha_l} - p_i^{n+\alpha_l}), \\ p_i^{n+\alpha_m} = EOS(\tau_i^{n+\alpha_m}), \end{cases} \quad (11)$$

and the final stage writes

$$\begin{cases} \overline{\rho_0 \tau}_i^{n+1} = \overline{\rho_0 \tau}_i^n + \frac{\Delta t}{\Delta X} \sum_{l=0}^s \theta_l (u_{i+\frac{1}{2}}^{n+\alpha_l} - u_{i-\frac{1}{2}}^{n+\alpha_l}), \\ \overline{\rho_0 u}_{i+\frac{1}{2}}^{n+1} = \overline{\rho_0 u}_{i+\frac{1}{2}}^n - \frac{\Delta t}{\Delta X} \sum_{l=0}^s \theta_l (p_{i+1}^{n+\alpha_l} - p_i^{n+\alpha_l}), \\ x_{i+\frac{1}{2}}^{n+1} = x_{i+\frac{1}{2}}^n + \Delta t \sum_{l=0}^s \theta_l u_{i+\frac{1}{2}}^{n+\alpha_l}, \\ p_i^{n+1} = EOS(\tau_i^{n+1}). \end{cases} \quad (12)$$

PROPOSITION 1 (Weak consistency of the barotropic Lagrangian staggered schemes (11)-(12)). *For all explicit Runge–Kutta sequences, all consistent spatial reconstructions, the schemes (11)-(12) are weakly consistent.*

Proof. Here we use the fact that a scheme whose flux is consistent (definition 1) is weakly consistent (definition 2). This is why we have to verify that the scheme can be rewritten under the form (8). From equation (12), one defines the natural flux

$$f_{i+\frac{1}{2}}^{\star} = \sum_{l=0}^s \theta_l \begin{pmatrix} -u_{i+\frac{1}{2}}^{n+\alpha_l} \\ p_{i+1}^{n+\alpha_l} \end{pmatrix}. \quad (13)$$

Intermediate fluxes are defined from (11) as $f_{i+\frac{1}{2}}^{\alpha_m} = \sum_{l=0}^{m-1} a_{m,l} \begin{pmatrix} -u_{i+\frac{1}{2}}^{n+\alpha_l} \\ p_{i+1}^{n+\alpha_l} \end{pmatrix}$. The proof is done by induction on the intermediate time-steps. First one proves that the intermediate (resp. natural as defined in (13)) flux can be written as $\Phi^m(\mathbf{U}_{i-mr+1}, \dots, \mathbf{U}_{i+mr+1})$ (resp. $\Phi^{\star}(\mathbf{U}_{i-(s+1)r+1}, \dots, \mathbf{U}_{i+(s+1)r+1})$). Second, one proves that Φ^m (resp. Φ^{\star}) satisfies for constant state $(\rho_0 \tau, \rho_0, \rho_0 u, \widehat{\rho_0})^t$

$$\begin{cases} \Phi^m \left(\begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix} \right) = \alpha_m \begin{pmatrix} -u \\ p \end{pmatrix}, \\ \Phi^{\star} \left(\begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix} \right) = \begin{pmatrix} -u \\ p \end{pmatrix}. \end{cases}$$

We start the proof considering the first intermediate time-step. One has $f_{i+\frac{1}{2}}^{\alpha_1} = a_{1,0} \begin{pmatrix} -u_{i+\frac{1}{2}}^n \\ p_{i+1}^n \end{pmatrix}$, $\alpha_1 = a_{1,0}$ where

$$u_{i+\frac{1}{2}}^n = \frac{\sum_{k=-r}^r C_k \overline{\rho_0 u}_{i+\frac{1}{2}+k}^n}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+\frac{1}{2}+k}^n} \text{ and } p_i^n = p \left(\frac{\sum_{k=-r}^r C_k \overline{\rho_0 \tau}_{i+k}^n}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+k}^n} \right).$$

Hence, one can write $f_{i+\frac{1}{2}}^{\alpha_1}$ as a function Φ^1 with $f_{i+\frac{1}{2}}^{\alpha_1} = \Phi^1 \left(\begin{pmatrix} \overline{\rho_0 \tau}_{i-r+1}^n \\ \overline{\rho_0}_{i-r+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}-r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}-r} \end{pmatrix}, \dots, \begin{pmatrix} \overline{\rho_0 \tau}_{i+r+1}^n \\ \overline{\rho_0}_{i+r+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}+r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}+r} \end{pmatrix} \right)$. The function Φ^1 writes

$$\Phi^1 \left(\begin{pmatrix} \overline{\rho_0 \tau}_{i-r+1}^n \\ \overline{\rho_0}_{i-r+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}-r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}-r} \end{pmatrix}, \dots, \begin{pmatrix} \overline{\rho_0 \tau}_{i+r+1}^n \\ \overline{\rho_0}_{i+r+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}+r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}+r} \end{pmatrix} \right) = a_{1,0} \begin{pmatrix} \frac{\sum_{k=-r}^r C_k \overline{\rho_0} u_{i+\frac{1}{2}+k}^n}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+\frac{1}{2}+k}} \\ p \frac{\sum_{k=-r}^r C_k \overline{\rho_0} \tau_{i+k}^n}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+k}} \end{pmatrix}.$$

Hence, for constant state $(\rho_0 \tau, \rho_0, \rho_0 u, \widehat{\rho_0})^t$

$$\Phi^1 \left(\begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix} \right) = a_{1,0} \begin{pmatrix} \frac{\sum_{k=-r}^r C_k \rho_0 u}{\sum_{k=-r}^r C_k \widehat{\rho_0}} \\ p \frac{\sum_{k=-r}^r C_k \rho_0 \tau}{\sum_{k=-r}^r C_k \rho_0} \end{pmatrix},$$

using the fact that $\sum_k C_k = 1$, it leads to $\Phi^1 \left(\begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0 \tau \\ \rho_0 \\ \rho_0 u \\ \widehat{\rho_0} \end{pmatrix} \right) = a_{1,0} \begin{pmatrix} -\frac{\rho_0 u}{p \frac{\rho_0}{\widehat{\rho_0}}} \\ p(\tau) \end{pmatrix} = a_{1,0} \begin{pmatrix} -u \\ p(\tau) \end{pmatrix}$. In particular, still for constant states $(\rho_0 \tau, \rho_0, \rho_0 u, \widehat{\rho_0})^t$, one obtains that $\overline{\rho_0 \tau}_i^{n+\alpha_1} = \overline{\rho_0 \tau}_i^n = \rho_0 \tau$ and $\overline{\rho_0 u}_{i+\frac{1}{2}}^{n+\alpha_1} = \overline{\rho_0 u}_{i+\frac{1}{2}}^n = \rho_0 u$. Then by straightforward induction on the intermediate time-steps, any $f_{i+\frac{1}{2}}^{\alpha_m}$ writes as a function Φ^m as

$$f_{i+\frac{1}{2}}^{\alpha_m} = \Phi^m \left(\begin{pmatrix} \overline{\rho_0 \tau}_{i-mr+1}^n \\ \overline{\rho_0}_{i-mr+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}-mr}^n \\ \overline{\rho_0}_{i+\frac{1}{2}-mr} \end{pmatrix}, \dots, \begin{pmatrix} \overline{\rho_0 \tau}_{i+mr+1}^n \\ \overline{\rho_0}_{i+mr+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}+mr}^n \\ \overline{\rho_0}_{i+\frac{1}{2}+mr} \end{pmatrix} \right).$$

The function Φ^m writes

$$\Phi^m \left(\begin{pmatrix} \overline{\rho_0 \tau}_{i-mr+1}^n \\ \overline{\rho_0}_{i-mr+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}-mr}^n \\ \overline{\rho_0}_{i+\frac{1}{2}-mr} \end{pmatrix}, \dots, \begin{pmatrix} \overline{\rho_0 \tau}_{i+mr+1}^n \\ \overline{\rho_0}_{i+mr+1} \\ \overline{\rho_0 u}_{i+\frac{1}{2}+mr}^n \\ \overline{\rho_0}_{i+\frac{1}{2}+mr} \end{pmatrix} \right) = \sum_{l=0}^{m-1} a_{m,l} \begin{pmatrix} \frac{\sum_{k=-r}^r C_k \overline{\rho_0} u_{i+\frac{1}{2}+k}^{n+\alpha_l}}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+\frac{1}{2}+k}} \\ p \frac{\sum_{k=-r}^r C_k \overline{\rho_0} \tau_{i+k}^{n+\alpha_l}}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+k}} \end{pmatrix}.$$

Then for constant state $(\rho_0\tau, \rho_0, \rho_0u, \widehat{\rho_0})^t$ and by induction on the previous intermediate time-steps

$$\Phi^m \left(\begin{pmatrix} \rho_0\tau \\ \rho_0 \\ \rho_0u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0\tau \\ \rho_0 \\ \rho_0u \\ \widehat{\rho_0} \end{pmatrix} \right) = \sum_{l=0}^{m-1} a_{m,l} \left(p \left(\frac{-\rho_0u}{\rho_0\tau} \right) \right) = \sum_{l=0}^{m-1} a_{m,l} \left(p(\tau) \right) = \alpha_m \left(p(\tau) \right).$$

And in particular, still for constant states, one obtains that $\overline{\rho_0\tau}_i^{n+\alpha_m} = \overline{\rho_0\tau}_i^n = \rho_0\tau$ and $\overline{\rho_0u}_{i+\frac{1}{2}}^{n+\alpha_m} = \overline{\rho_0u}_{i+\frac{1}{2}}^n = \rho_0u$. Therefore, by induction, the natural flux $f_{i+\frac{1}{2}}^*$ writes as a vector values function Φ^* as

$$f_{i+\frac{1}{2}}^* = \Phi^* \left(\begin{pmatrix} \overline{\rho_0\tau}_{i-(s+1)r+1}^n \\ \overline{\rho_0}_{i-(s+1)r+1} \\ \overline{\rho_0u}_{i+\frac{1}{2}-(s+1)r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}-(s+1)r} \end{pmatrix}, \dots, \begin{pmatrix} \overline{\rho_0\tau}_{i+(s+1)r+1}^n \\ \overline{\rho_0}_{i+(s+1)r+1} \\ \overline{\rho_0u}_{i+\frac{1}{2}+(s+1)r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}+(s+1)r} \end{pmatrix} \right),$$

where Φ^* satisfies

$$\Phi^* \left(\begin{pmatrix} \overline{\rho_0\tau}_{i-(s+1)r+1}^n \\ \overline{\rho_0}_{i-(s+1)r+1} \\ \overline{\rho_0u}_{i+\frac{1}{2}-(s+1)r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}-(s+1)r} \end{pmatrix}, \dots, \begin{pmatrix} \overline{\rho_0\tau}_{i+(s+1)r+1}^n \\ \overline{\rho_0}_{i+(s+1)r+1} \\ \overline{\rho_0u}_{i+\frac{1}{2}+(s+1)r}^n \\ \overline{\rho_0}_{i+\frac{1}{2}+(s+1)r} \end{pmatrix} \right) = \sum_{l=0}^s \theta_l \left(\frac{\sum_{k=-r}^r C_k \overline{\rho_0u}_{i+\frac{1}{2}+k}^{n+\alpha_l}}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+\frac{1}{2}+k}} \right) \frac{1}{p \left(\frac{\sum_{k=-r}^r C_k \overline{\rho_0\tau}_{i+k}^{n+\alpha_l}}{\sum_{k=-r}^r C_k \overline{\rho_0}_{i+k}} \right)}.$$

Thus for constant states, $\Phi^* \left(\begin{pmatrix} \rho_0\tau \\ \rho_0 \\ \rho_0u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0\tau \\ \rho_0 \\ \rho_0u \\ \widehat{\rho_0} \end{pmatrix} \right) = \sum_{l=0}^s \theta_l \left(p \left(\frac{-\rho_0u}{\rho_0\tau} \right) \right) = \sum_{l=0}^s \theta_l \left(p(\tau) \right)$. Using the fact that $\sum_{l=0}^s \theta_l = 1$, it leads

to $\Phi^* \left(\begin{pmatrix} \rho_0\tau \\ \rho_0 \\ \rho_0u \\ \widehat{\rho_0} \end{pmatrix}, \dots, \begin{pmatrix} \rho_0\tau \\ \rho_0 \\ \rho_0u \\ \widehat{\rho_0} \end{pmatrix} \right) = \begin{pmatrix} -u \\ p(\tau) \end{pmatrix}$. Hence, the scheme is weakly consistent for the barotropic equations in the sense of definition 1. \square

2.3. A new local internal energy corrector

Compared to the barotropic schemes, an additional theoretical difficulty shows up with the energy equation for the hydrodynamics case (6)-(7). It is related to the fact that, even if mass and momentum are conserved (see lemma 1) and also that the total energy \mathcal{E} is preserved (see lemma 2), it is not the case for the total energy E . Experimentally, as shown on the left picture in figure 1 and stated in [42, 22]), we observe that the schemes (6)-(7) are unable to capture the shocks correctly, in the sense that the Rankine–Hugoniot jump relations are not recovered.

The idea is then to recouple E and \mathcal{E} using a correction of the internal energy at the end of the Lagrangian phase (6)-(7). The difference between the computed kinetic energy and the kinetic energy reconstructed from the velocity is reversed in the internal energy. This is very similar to what is done in works by Herbin, Latché and al. [17, 18, 19]. The main difference is that they perform the correction *a priori*, whereas here in our case the correction is applied *a posteriori*.

2.3.1. Internal energy corrector

The idea of the internal energy corrector is to force the conservation of the quantity E , using the already proved conservation of \mathcal{E} (see lemma 2). The difference between the computed kinetic energy and the kinetic energy reconstructed from the velocity is computed as follows. As the scheme is high-order accurate, the result is not so

straightforward. It follows the steps described hereafter. First the point-wise kinetic energy reconstructed from the velocity is defined by $(\frac{1}{2}\rho_0 u^2)_{i+\frac{1}{2}}^{n+1} = \frac{1}{2} \left(\sum_k C_k \overline{\rho_0 u^2}_{i+k+\frac{1}{2}}^{n+1} \right)^2 / \sum_k C_k \overline{\rho_0}^n_{i+k+\frac{1}{2}}$. Second it is averaged over a cell using the coefficients \widehat{C}_k presented in table 1, that is $(\frac{1}{2}\overline{\rho_0 u^2})_{i+\frac{1}{2}}^{n+1} = \sum_k \widehat{C}_k (\frac{1}{2}\rho_0 u^2)_{i+k+\frac{1}{2}}^{n+1}$. The difference between the two kinetic energies is $\Delta K_{i+\frac{1}{2}}^{n+1} = \overline{\rho_0 e_{\text{kin}}^{n+1}}_{i+\frac{1}{2}} - (\frac{1}{2}\overline{\rho_0 u^2})_{i+\frac{1}{2}}^{n+1}$. Third, linear interpolation is made to compute $\Delta K_i^{n+1} = \frac{1}{2}(\Delta K_{i+\frac{1}{2}}^{n+1} + \Delta K_{i-\frac{1}{2}}^{n+1})$. Last, the difference ΔK_i^{n+1} is added to the internal energy $\overline{\rho_0 \epsilon_i}^{n+1}$ whereas $\Delta K_{i+\frac{1}{2}}^{n+1}$ is subtracted to the kinetic ones. It writes as an *a posteriori* correction

$$\begin{cases} \overline{\rho_0 \epsilon_i}^{n+1, \star} &= \overline{\rho_0 \epsilon_i}^{n+1} + \Delta K_i^{n+1} \\ \overline{\rho_0 e_{\text{kin}}^{n+1, \star}}_{i+\frac{1}{2}} &= \overline{\rho_0 e_{\text{kin}}^{n+1}}_{i+\frac{1}{2}} - \Delta K_{i+\frac{1}{2}}^{n+1} = (\frac{1}{2}\overline{\rho_0 u^2})_{i+\frac{1}{2}}^{n+1} \end{cases} \quad (14)$$

The internal energy corrector can be applied at the end of each Runge–Kutta sub-cycle or only at the end of the time-step. Commonly, the internal energy corrector is performed only at the end of the time-step, to save CPU.

2.3.2. Properties of the internal energy corrector

LEMMA 3 (High-order accuracy of the internal energy corrector). *The internal energy corrector is high-order accurate in both time and space.*

Proof. Assume that the solution is smooth enough. Assume that the coefficients \widehat{C}_k and C_k yield N^{th} order of accuracy in space, and that the Lagrange phase is also of order N in both time and space. Then in particular, one has

$$\Delta K_{i+\frac{1}{2}}^{n+1} = \overline{\rho_0 e_{\text{kin}}^{n+1}}_{i+\frac{1}{2}} - (\frac{1}{2}\overline{\rho_0 u^2})_{i+\frac{1}{2}}^{n+1} = \mathcal{O}(\Delta X^N).$$

And then trivially, one gets that $\Delta K_i^{n+1} = \frac{1}{2}(\Delta K_{i+\frac{1}{2}}^{n+1} + \Delta K_{i-\frac{1}{2}}^{n+1}) = \mathcal{O}(\Delta X^N)$. As the Lagrange phase is assumed to be high-order accurate, one has that $\overline{\rho_0 \epsilon_i}^{n+1} = \overline{\rho_0 \epsilon(x_i, t^{n+1})} + \mathcal{O}(\Delta X^N)$. And then, one gets $\overline{\rho_0 \epsilon_i}^{n+1, \star} = \overline{\rho_0 \epsilon_i}^{n+1} + \Delta K_i = \overline{\rho_0 \epsilon(x_i, t^{n+1})} + \mathcal{O}(\Delta X^N)$, which concludes the proof, yielding high-order accuracy for the internal energy. \square

The following lemma gives conservation of total energy when applying the internal energy corrector.

LEMMA 4 (\mathcal{E} -conservation of the internal energy corrector). *The internal energy corrector satisfies $\mathcal{E}^{n+1, \star} = \mathcal{E}^{n+1}$.*

Proof. One has that $\mathcal{E}^{n+1, \star} - \mathcal{E}^{n+1} = \Delta X \left(\sum_i \Delta K_i - \sum_i \Delta K_{i+\frac{1}{2}} \right)$. Then, using $\Delta K_i = \frac{1}{2}(\Delta K_{i+\frac{1}{2}} + \Delta K_{i-\frac{1}{2}})$, it leads to $\mathcal{E}^{n+1, \star} - \mathcal{E}^{n+1} = \Delta X \left(\sum_i \frac{1}{2}(\Delta K_{i+\frac{1}{2}} + \Delta K_{i-\frac{1}{2}}) - \sum_i \Delta K_{i+\frac{1}{2}} \right)$. Performing change of discrete variables in the first summation, and assuming wall or periodic boundary conditions, it yields $\mathcal{E}^{n+1, \star} - \mathcal{E}^{n+1} = 0$. \square

LEMMA 5 (E -Conservation of the staggered schemes (6)-(7)-(14)). *The schemes (6)-(7) with the internal energy corrector (14) satisfy $E^{n+1, \star} = E^n$ (cf definition 3).*

Proof. We have $E^{n+1, \star} - E^n = \Delta X \sum_i (\overline{\rho_0 \epsilon_i}^{n+1, \star} - \overline{\rho_0 \epsilon_i}^n) + \Delta X \sum_i (\overline{\rho_0 u^2}_{i+\frac{1}{2}}^{n+1, \star} - \overline{\rho_0 u^2}_{i+\frac{1}{2}}^n)$. Introducing the term at time $t = t^{n+1}$, it becomes

$$\begin{aligned} E^{n+1, \star} - E^n &= \Delta X \sum_i (\overline{\rho_0 \epsilon_i}^{n+1, \star} - \overline{\rho_0 \epsilon_i}^{n+1} + \overline{\rho_0 \epsilon_i}^{n+1} - \overline{\rho_0 \epsilon_i}^n) \\ &\quad + \Delta X \sum_i (\overline{\rho_0 u^2}_{i+\frac{1}{2}}^{n+1, \star} - \overline{\rho_0 e_{\text{kin}, u}}_{i+\frac{1}{2}}^{n+1} + \overline{\rho_0 e_{\text{kin}, u}}_{i+\frac{1}{2}}^{n+1} - \overline{\rho_0 u^2}_{i+\frac{1}{2}}^n) \\ &= \Delta X \sum_i (\overline{\rho_0 \epsilon_i}^{n+1, \star} - \overline{\rho_0 \epsilon_i}^{n+1}) - \Delta X \sum_i (\overline{\rho_0 u^2}_{i+\frac{1}{2}}^{n+1, \star} - \overline{\rho_0 e_{\text{kin}, u}}_{i+\frac{1}{2}}^{n+1}) + \mathcal{E}^{n+1} - \mathcal{E}^n. \end{aligned}$$

Using the fact that $\overline{\rho_0 u_{i+\frac{1}{2}}^{n+1,\star}} = \overline{\rho_0 e_{\text{kin},u_{i+\frac{1}{2}}}^{n+1,\star}}$, it leads to $E^{n+1,\star} - E^n = \mathcal{E}^{n+1,\star} - \mathcal{E}^{n+1} + \mathcal{E}^{n+1} - \mathcal{E}^n$. Then, using lemma 2, one gets that $\mathcal{E}^{n+1} - \mathcal{E}^n = 0$ and using lemma 4, that $\mathcal{E}^{n+1,\star} - \mathcal{E}^{n+1} = 0$. Hence, $E^{n+1,\star} - E^n = 0$. Thus, applying the internal corrector gives conservation of the energy E between time $t = t^{n+1,\star}$ and time $t = t^n$. \square

2.4. Proof of the main theorem

We state what we consider as the main theoretical contribution of this work and detail its proof.

THEOREM 1 (Weak consistency of the staggered schemes (6)-(7)-(14)). *For all explicit Runge–Kutta sequences, for coefficients $C_k, \widehat{C}_k, d_k, r_k$ defined in tables 1 to 4, the schemes (6)-(7)-(14) are weakly consistent with the Euler equations in Lagrangian coordinates with the regularity assumptions of definition 2.*

REMARK 3. *The proof for weak consistency of the two first equations, specific volume and momentum conservations which show up in (6)-(7) is essentially similar to the one of proposition 1 for the barotropic case, so is not detailed. Instead we focus on the energy equation. However, due to the very intricate structure of the discrete energy equation, no explicit natural fluxes for total energy have been exhibited so far for any order in both space and time.*

For instance, let us consider first order Euler forward scheme with second order in space. Discrete total energy is defined as $\rho e_i = \rho \epsilon_i + \frac{1}{2}(\rho e_{\text{kin},u_{i+\frac{1}{2}}} + \rho e_{\text{kin},u_{i-\frac{1}{2}}})$. Then we get the following total energy evolution equation

$$\rho e_i^{n+1} - \rho e_i^n = -\frac{\Delta t}{\Delta X} \left(\frac{p_{i+1} + p_i}{2} u_{i+\frac{1}{2}} - \frac{p_i + p_{i-1}}{2} u_{i-\frac{1}{2}} \right), \quad (15)$$

which is trivially written under flux form. However generalization is not obvious. Generally, it means that the energy equation is not rewritten under conservative fluxes form. Hence the criterion of flux consistency of Lax–Wendroff is not directly applicable, this is why the proof is detailed hereafter in full length, starting directly from definition 2.

Proof. The assumptions presented in definition 2 for weak consistency are done. We first detail the proof for a forward Euler, second order in space scheme because it highlights the key elements of the method. The general case with a forward Euler and any order in space will be dealt with in a second stage. The most general case with any explicit Runge–Kutta sequences will not be detailed because it would add no new technical ideas and the notations are too heavy. For the sake of simplicity, in the following the time step $t^{n+1,\star}$ is denoted by t^{n+1} .

First stage. For a forward Euler, second order in space scheme, the internal and kinetic energies discrete equations write

$$\begin{cases} \rho_0 \epsilon_i^{n+1} - \rho_0 \epsilon_i^n &= -\frac{\Delta t}{\Delta X} p_i^n (u_{i+\frac{1}{2}}^n - u_{i-\frac{1}{2}}^n) + \Delta K_i^{n+1}, \\ \rho_0 e_{\text{kin},i+\frac{1}{2}}^{n+1} - \rho_0 e_{\text{kin},i+\frac{1}{2}}^n &= -\frac{\Delta t}{\Delta X} u_{i+\frac{1}{2}}^n (p_{i+1}^n - p_i^n) - \Delta K_{i+\frac{1}{2}}^{n+1}, \end{cases}$$

The idea is to take a test function $\phi \in \mathcal{C}_0^\infty$ with compact support. Denote $\phi_i^n = \phi(i\Delta X, t^n)$ and $\phi_{i+\frac{1}{2}}^n((i+\frac{1}{2})\Delta X, t^n)$. Multiply the first equation by $\Delta X \phi_i^{n+1}$ and the second by $\Delta X \phi_{i+\frac{1}{2}}^{n+1}$ then sum over the n and i and combine both. It leads to

$$\begin{aligned} & \sum_n \sum_i \Delta X \left[(\rho_0 \epsilon_i^{n+1} - \rho_0 \epsilon_i^n) \phi_i^{n+1} + (\rho_0 e_{\text{kin},i+\frac{1}{2}}^{n+1} - \rho_0 e_{\text{kin},i+\frac{1}{2}}^n) \phi_{i+\frac{1}{2}}^{n+1} \right] \\ & + \sum_n \sum_i \Delta t \left[p_i^n \phi_i^{n+1} (u_{i+\frac{1}{2}}^n - u_{i-\frac{1}{2}}^n) + u_{i+\frac{1}{2}}^n \phi_{i+\frac{1}{2}}^{n+1} (p_{i+1}^n - p_i^n) \right] \\ & - \sum_n \sum_i \Delta X \left[\Delta K_i^{n+1} \phi_i^{n+1} - \Delta K_{i+\frac{1}{2}}^{n+1} \phi_{i+\frac{1}{2}}^{n+1} \right] = 0. \end{aligned} \quad (16)$$

Denote $h = \max(\Delta X, \Delta t)$. Introducing the notation

$$\begin{cases} \mathcal{T}_1^h &= \sum_n \sum_i \Delta X \left[(\rho_0 \epsilon_i^{n+1} - \rho_0 \epsilon_i^n) \phi_i^{n+1} + (\rho_0 e_{\text{kin},i+\frac{1}{2}}^{n+1} - \rho_0 e_{\text{kin},i+\frac{1}{2}}^n) \phi_{i+\frac{1}{2}}^{n+1} \right], \\ \mathcal{T}_2^h &= -\sum_n \sum_i \Delta X \left[\Delta K_i^{n+1} \phi_i^{n+1} - \Delta K_{i+\frac{1}{2}}^{n+1} \phi_{i+\frac{1}{2}}^{n+1} \right], \\ \mathcal{T}_3^h &= \sum_n \sum_i \Delta t \left[p_i^n \phi_i^{n+1} (u_{i+\frac{1}{2}}^n - u_{i-\frac{1}{2}}^n) + u_{i+\frac{1}{2}}^n \phi_{i+\frac{1}{2}}^{n+1} (p_{i+1}^n - p_i^n) \right], \end{cases} \quad (17)$$

eq. (16) rewrites simply under the form $\mathcal{T}_1^h + \mathcal{T}_2^h + \mathcal{T}_3^h = 0$. We study separately the three terms.

- Terms \mathcal{T}_1^h is reordered into

$$\mathcal{T}_1^h = - \sum_n \Delta t \sum_i \Delta X \left[\rho_0 \epsilon_i^n \frac{\phi_i^{n+1} - \phi_i^n}{\Delta t} + \rho_0 e_{\text{kin}_{i+\frac{1}{2}}}^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+\frac{1}{2}}^n}{\Delta t} \right].$$

We will use the natural definition/notation for staircase functions with the indicator function (23)

$$\psi_h(x, t) = \sum_i \sum_n \chi_{]i, i+1[}^{n, n+1}[(t) \chi_{]x_{i-1/2}, x_{i+1/2}[}(x) \psi_i^n \text{ and } \psi_{h, \text{stag}}(x, t) = \sum_i \sum_n \chi_{]i, i+1[}^{n, n+1}[(t) \chi_{]x_i, x_{i+1}[}(x) \psi_{i+\frac{1}{2}}^n.$$

Then, using the internal energy corrector, it yields that $\rho_0 e_{\text{kin}_{i+\frac{1}{2}}}^n = \frac{1}{2}(\rho_0 u^2)_{i+\frac{1}{2}}^n$ and so

$$\mathcal{T}_1^h = - \int_0^T \int_{\Omega} (\rho_0 \epsilon)_h \partial_t \phi_h dx dt - \int_0^T \int_{\Omega} \left(\frac{1}{2} \rho_0 u^2 \right)_{h, \text{stag}} \partial_t \phi_{h, \text{stag}} dx dt + \int_{\Omega} (\rho_0 \epsilon)_h^0 \phi_h^0 dx + \int_{\Omega} \left(\frac{1}{2} \rho_0 u^2 \right)_{h, \text{stag}}^0 \phi_{h, \text{stag}}^0 dx.$$

Using the convergence hypothesis of definition 2 and the regularity of the test function ϕ , one can pass to the limit as ΔX and Δt tend to 0. It leads to

$$\lim_{h \rightarrow 0} \mathcal{T}_1^h = - \int_0^T \int_{\Omega} \widehat{\rho_0 \epsilon} \partial_t \phi dx dt - \int_0^T \int_{\Omega} \frac{1}{2} \widehat{\rho_0 u^2} \partial_t \phi dx dt + \int_{\Omega} \widehat{\rho_0 \epsilon}(x, 0) \phi(x, 0) dx + \int_{\Omega} \frac{1}{2} \widehat{\rho_0 u^2}(x, 0) \phi(x, 0) dx.$$

Using $\rho_0 e = \rho_0 \epsilon + \frac{1}{2} \rho_0 u^2$, one gets $\lim_{h \rightarrow 0} \mathcal{T}_1^h = - \int_0^T \int_{\Omega} \widehat{\rho_0 e} \partial_t \phi dx dt + \int_{\Omega} \widehat{\rho_0 e}(x, 0) \phi(x, 0) dx$.

- Now, focus on \mathcal{T}_2^h . It satisfies $\mathcal{T}_2^h = - \sum_n \sum_i \Delta X \left[\Delta K_i^{n+1} \phi_i^{n+1} - \Delta K_{i+\frac{1}{2}}^{n+1} \phi_{i+\frac{1}{2}}^{n+1} \right]$. Using $\Delta K_i^{n+1} = \frac{1}{2}(\Delta K_{i+\frac{1}{2}}^{n+1} + \Delta K_{i-\frac{1}{2}}^{n+1})$,

it gives after reordering that $\mathcal{T}_2^h = - \sum_n \sum_i \Delta X \Delta K_{i+\frac{1}{2}}^{n+1} \left(\frac{\phi_{i+1}^{n+1} + \phi_i^{n+1}}{2} - \phi_{i+\frac{1}{2}}^{n+1} \right)$. Using the boundedness in L^∞ of $\Delta K_{i+\frac{1}{2}}^{n+1}$ and regularity of ϕ , it leads to $|\mathcal{T}_2^h| \leq C_\phi \Delta X \|(\Delta K)_h\|_{L^\infty}$, which gives immediately $\lim_{h \rightarrow 0} |\mathcal{T}_2^h| = 0$.

- The term \mathcal{T}_3^h satisfies $\mathcal{T}_3^h = \sum_n \sum_i \Delta t \left[p_i^n \phi_i^{n+1} (u_{i+\frac{1}{2}}^n - u_{i-\frac{1}{2}}^n) + u_{i+\frac{1}{2}}^n \phi_{i+\frac{1}{2}}^{n+1} (p_{i+1}^n - p_i^n) \right]$. Reordering the terms, it yields

$$\begin{aligned} \mathcal{T}_3^h &= \sum_n \Delta t \sum_i \left[p_i^n u_{i+\frac{1}{2}}^n \phi_i^{n+1} - p_{i+1}^n u_{i+\frac{1}{2}}^n \phi_{i+1}^{n+1} + u_{i+\frac{1}{2}}^n \phi_{i+\frac{1}{2}}^{n+1} (p_{i+1}^n - p_i^n) \right] \\ &= \sum_n \Delta t \sum_i \left[u_{i+\frac{1}{2}}^n p_i^n (\phi_i^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}) + u_{i+\frac{1}{2}}^n p_{i+1}^n (\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+1}^{n+1}) \right] \\ &= - \sum_n \Delta t \Delta X \sum_i u_{i+\frac{1}{2}}^n \left[\frac{1}{2} p_i^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_i^{n+1}}{\Delta X} + \frac{1}{2} p_{i+1}^n \frac{\phi_{i+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X} \right] \\ &= - \sum_n \Delta t \Delta X \sum_i u_{i+\frac{1}{2}}^n \left(\left[\frac{p_i^n + p_{i+1}^n}{4} \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_i^{n+1}}{\Delta X} + \frac{p_i^n + p_{i+1}^n}{4} \frac{\phi_{i+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X} \right] + \left[\frac{p_i^n - p_{i+1}^n}{4} \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_i^{n+1}}{\Delta X} + \frac{p_{i+1}^n - p_i^n}{4} \frac{\phi_{i+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X} \right] \right). \end{aligned}$$

The previous expression is decomposed into two terms denoted $\mathcal{T}_{3,1}^h$ and $\mathcal{T}_{3,2}^h$ with

$$\begin{cases} \mathcal{T}_{3,1}^h &= - \sum_n \Delta t \Delta X \sum_i u_{i+\frac{1}{2}}^n \left[\frac{p_i^n - p_{i+1}^n}{4} \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_i^{n+1}}{\Delta X} + \frac{p_{i+1}^n - p_i^n}{4} \frac{\phi_{i+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X} \right] \\ \mathcal{T}_{3,2}^h &= - \sum_n \Delta t \Delta X \sum_i u_{i+\frac{1}{2}}^n \left[\frac{p_i^n + p_{i+1}^n}{4} \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_i^{n+1}}{\Delta X} + \frac{p_i^n + p_{i+1}^n}{4} \frac{\phi_{i+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X} \right] \end{cases}$$

The $\mathcal{T}_{3,1}^h$ is dealt using the assumptions that p_h is in BV , u_h is bounded in L^∞ , and ϕ is in \mathcal{C}^1 which gives $|\mathcal{T}_{3,1}^h| \leq \Delta X C_\phi \|u_h\|_{L^\infty} \|p_h\|_{BV}$. Hence, passing to the limit, it leads to $\lim_{h \rightarrow 0} |\mathcal{T}_{3,1}^h| = 0$.

On the other hand, one easily notices that $\mathcal{T}_{3,2}^h$ rewrites as

$$\mathcal{T}_{3,2}^h = - \sum_n \Delta t \Delta X \sum_i u_{i+\frac{1}{2}}^n \frac{p_i^n + p_{i+1}^n}{2} \left[\frac{1}{2} \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_i^{n+1}}{\frac{\Delta X}{2}} + \frac{1}{2} \frac{\phi_{i+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\frac{\Delta X}{2}} \right] = - \int_0^T \int_\Omega (pu)_h \partial_X \phi_h dx dt$$

Using the regularity of ϕ it leads, passing to the limit, to $\lim_{h \rightarrow 0} \mathcal{T}_{3,2}^h = - \int_0^T \int_\Omega \widetilde{p} u \partial_X \phi dx dt$. Reassembling all the terms, it yields that

$$\lim_{h \rightarrow 0} \mathcal{T}_1^h + \mathcal{T}_2^h + \mathcal{T}_3^h = - \int_0^T \int_\Omega \widetilde{\rho}_0 e \partial_t \phi dx dt - \int_0^T \int_\Omega \widetilde{p} u \partial_X \phi dx dt + \int_\Omega \widetilde{\rho}_0 e(x, 0) \phi(x, 0) dx.$$

And, hence, we get $\int_0^T \int_\Omega \widetilde{\rho}_0 e \partial_t \phi dx dt + \int_0^T \int_\Omega \widetilde{p} u \partial_X \phi dx dt = \int_\Omega \widetilde{\rho}_0 e(x, 0) \phi(x, 0) dx$. Previous equation gives weak consistency for the second order in space, forward Euler staggered scheme with internal energy corrector.

Second stage. Now, the problem of high-order in space is tackled. There is no difficulty for the terms \mathcal{T}_1^h and \mathcal{T}_2^h , but this is not the case for \mathcal{T}_3^h , where the desired result is not obvious. For the sake of simplicity here, we consider that $\widehat{C}_0 = 1$, $\widehat{C}_k = 0, \forall |k| > 0$. The results does not change, provided $\sum_k \widehat{C}_k = 1$ but it greatly alleviates the algebra of

the proof. One has that $\mathcal{T}_3^h = - \sum_n \Delta t \sum_i \sum_{k \geq 0} d_k \left[p_i^n \phi_i^{n+1} (u_{i+k+\frac{1}{2}}^n - u_{i-k-\frac{1}{2}}^n) + u_{i+\frac{1}{2}}^n \phi_{i+\frac{1}{2}}^{n+1} (p_{j+k+1}^n - p_{j-k}^n) \right]$. Reordering the terms, so that only $u_{i+\frac{1}{2}}^n$ shows up, leads to

$$\mathcal{T}_3^h = - \sum_n \Delta t \sum_i u_{i+\frac{1}{2}}^n \sum_{k \geq 0} d_k \left[p_{i-k}^n (\phi_{i-k}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}) + p_{i+k+1} (\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+k+1}^{n+1}) \right].$$

Highlighting the space derivatives of ϕ gives

$$\mathcal{T}_3^h = + \sum_n \Delta t \sum_i u_{i+\frac{1}{2}}^n \sum_{k \geq 0} \left(k + \frac{1}{2} \right) d_k \left[p_{i-k}^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i-k}^{n+1}}{\Delta X (k + \frac{1}{2})} + p_{i+k+1} \frac{\phi_{i+k+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X (k + \frac{1}{2})} \right].$$

Noticing that $(k + \frac{1}{2}) d_k = r_k, k \geq 0$, it yields

$$\mathcal{T}_3^h = + \sum_n \Delta t \sum_i u_{i+\frac{1}{2}}^n \sum_{k \geq 0} r_k \left[p_{i-k}^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i-k}^{n+1}}{\Delta X (k + \frac{1}{2})} + p_{i+k+1} \frac{\phi_{i+k+1}^{n+1} - \phi_{i+\frac{1}{2}}^{n+1}}{\Delta X (k + \frac{1}{2})} \right].$$

As previously for the case of second order accuracy in space, the conclusion is reached using the assumption that p_h is in BV , u_h is bounded in L^∞ , and ϕ is in \mathcal{C}^1 . One has $\lim_{h \rightarrow 0} \mathcal{T}_{3,2}^h = - \int_0^T \int_\Omega \widetilde{p} u \partial_X \phi dx dt$. And, hence, it leads to

$\int_0^T \int_\Omega \widetilde{\rho}_0 e \partial_t \phi dx dt + \int_0^T \int_\Omega \widetilde{p} u \partial_X \phi dx dt = \int_\Omega \widetilde{\rho}_0 e(x, 0) \phi(x, 0) dx$. Previous equation gives weak consistency for forward Euler staggered scheme with internal energy corrector. Using Runge–Kutta sequences adds only more technical difficulty in the algebra, but does not alter the weak consistency result. Idem for the use of the coefficients \widehat{C}_k . The key point for consistency is to use the same coefficients d_k and \widehat{C}_k for both the internal and kinetic energies equations. \square

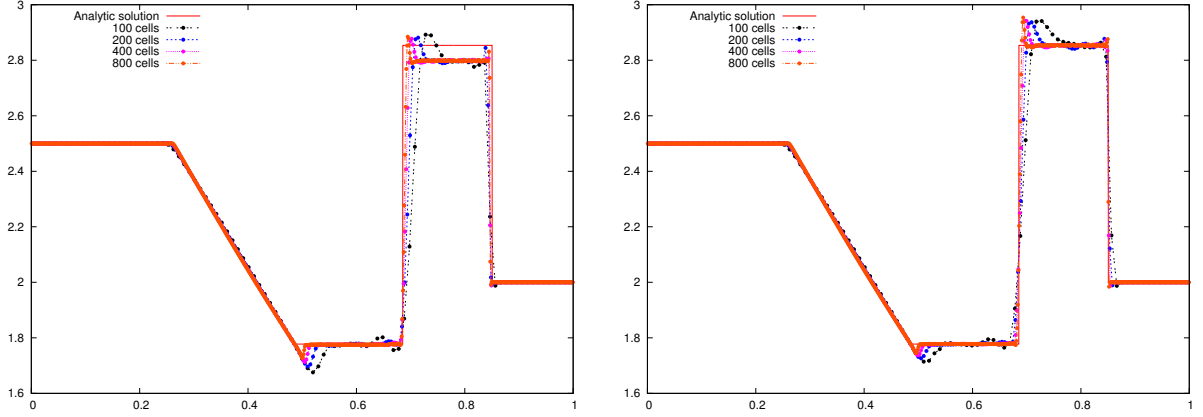


Figure 1: Illustration of the importance of the internal energy corrector. Without the internal energy corrector (left), the scheme does not converge toward the weak solution for the Sod shock tube. Using the internal energy corrector (right), although oscillatory, jump relations are recovered for small mesh sizes.

REMARK 4. Without internal energy corrector, for a forward Euler second order in space scheme, the first term writes

$$\begin{aligned} \widehat{\mathcal{T}}_1^h &= - \sum_n \Delta t \sum_i \Delta X \left[\rho_0 \epsilon_i^n \frac{\phi_i^{n+1} - \phi_i^n}{\Delta t} + \rho_0 e_{kin, i+\frac{1}{2}}^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+\frac{1}{2}}^n}{\Delta t} \right] \\ &= - \sum_n \Delta t \sum_i \Delta X \left(\left[\rho_0 \epsilon_i^n \frac{\phi_i^{n+1} - \phi_i^n}{\Delta t} + \left(\frac{1}{2} \rho_0 u^2 \right)_{i+\frac{1}{2}}^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+\frac{1}{2}}^n}{\Delta t} \right] - \sum_n \Delta t \sum_i \Delta X \left[\left(\rho_0 e_{kin, i+\frac{1}{2}}^n - \left(\frac{1}{2} \rho_0 u^2 \right)_{i+\frac{1}{2}}^n \right) \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+\frac{1}{2}}^n}{\Delta t} \right] \right) \end{aligned}$$

Define $\mathcal{T}_{1,1}^h$ and $\mathcal{T}_{1,2}^h$ by

$$\begin{cases} \mathcal{T}_{1,1}^h &= - \sum_n \Delta t \sum_i \Delta X \left[\rho_0 \epsilon_i^n \frac{\phi_i^{n+1} - \phi_i^n}{\Delta t} + \left(\frac{1}{2} \rho_0 u^2 \right)_{i+\frac{1}{2}}^n \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+\frac{1}{2}}^n}{\Delta t} \right], \\ \mathcal{T}_{1,2}^h &= - \sum_n \Delta t \sum_i \Delta X \left[\left(\rho_0 e_{kin, i+\frac{1}{2}}^n - \left(\frac{1}{2} \rho_0 u^2 \right)_{i+\frac{1}{2}}^n \right) \frac{\phi_{i+\frac{1}{2}}^{n+1} - \phi_{i+\frac{1}{2}}^n}{\Delta t} \right]. \end{cases}$$

The term $\mathcal{T}_{1,1}^h$ has been dealt with as it is equal to the term \mathcal{T}_1^h of the proof. Now, consider the term $\mathcal{T}_{1,2}^h$. Then under regularity hypothesis on the test function, one obtains $|\mathcal{T}_{1,2}^h| \leq C_\phi \|\rho_0 e_{kin} - (\frac{1}{2} \rho_0 u^2)\|_{L^1([0;T] \times \Omega)}$. Experimentally, one observes that without internal energy corrector, $\|\rho_0 e_{kin} - (\frac{1}{2} \rho_0 u^2)\|_{L^1([0;T] \times \Omega)}$ does not tend to 0 as ΔX and Δt tend to 0. As an example, for the Sod shock tube [36], we perform the computation with and without the internal energy corrector. Results are displayed in fig. 1. On the left picture, we show that the scheme does not converge toward the analytical solution without the internal energy corrector. On the right one, we show that adding the internal energy corrector, the profile obtained in internal energy is much more satisfactory. In table 6, we present the values of $\|\rho_0 e_{kin} - (\frac{1}{2} \rho_0 u^2)\|_{L^1([0;T] \times \Omega)}$, for the scheme without internal energy corrector, to assess that it does not tend to 0 experimentally. Further studies on the Sod shock tube are presented in the sequel.

REMARK 5. One notices that the default of conservation in E is strongly linked to the quantity $\|\rho_0 e_{kin} - (\frac{1}{2} \rho_0 u^2)\|_{L^1([0;T] \times \Omega)}$. It tends to attest in the present case that for the presented schemes the key for weak consistency (and correct shock capturing) is to ensure conservation of E .

2.5. The remapping stage

The remapping is the algorithm designed to project the Lagrangian quantities on the original Cartesian grids, so that one gets a Cartesian Euler scheme. A possible choice is to project the reconstructed total energy. Here, for robustness purposes, we would rather remap independently both the internal and the kinetic energies.

ΔX	$\frac{1}{100}$	$\frac{1}{200}$	$\frac{1}{400}$	$\frac{1}{800}$	$\frac{1}{1600}$
$\ \rho_0 e_{\text{kin}} - (\frac{1}{2}\rho_0 u^2)\ _{l^1([0:T]\times\Omega)}$	6.7e-3	4.7e-3	3.7e-3	3.3e-3	3.1e-3

Table 6: Illustration of the interest and importance of the internal energy corrector. Without the internal energy corrector, the term $\|\rho_0 e_{\text{kin}} - (\frac{1}{2}\rho_0 u^2)\|_{l^1([0:T]\times\Omega)}$ does not tend to 0 as ΔX and Δt tends to zero.

The quantities to be remapped at the end of the Lagrangian phase (6)-(7)-(14) are $\rho_0, \rho_0 \epsilon$ on the primal grid $\{x_{i+\frac{1}{2}}\}$ and $\rho_0, \rho_0 u, \rho_0 e_{\text{kin}}$ on the dual one $\{x_i\}$. The projection briefly explained hereafter is equal to the one proposed in [10, 45] but adapted here to the staggered grids as done in [5].

2.5.1. Lagrange polynomials based conservative projection

At the end of the Lagrangian phase, the primal deformed grid $\{x_{i+\frac{1}{2}}^{n+1}\}$ is known. In order to project the staggered variables $\rho_0, \rho_0 u, \rho_0 e_{\text{kin}}$, one must first deduce the deformation of the dual grid $\{x_i^{n+1}\}$. This is done by using the coefficients r_k presented in table 4, and using $x_i^{n+1} = \sum_k r_k (x_{i+k+\frac{1}{2}}^{n+1} + x_{i-k-\frac{1}{2}}^{n+1})$, which leads to locations of cell centers at high-order accuracy provided $\{x_{i+\frac{1}{2}}^{n+1}\}$ is also known at high-order accuracy. Take $\phi \in L^\infty$, one has $\Delta X \overline{(\rho_0 \phi)}_{\xi(i)}^{n+1} = \int_{X_{\xi(i)-\frac{1}{2}}}^{X_{\xi(i)+\frac{1}{2}}} (\rho_0 \phi)(Y, t^{n+1}) dX + \mathcal{O}(\Delta X^N)$. Using the change of variables $(x, t) \rightarrow (X, t)$, it yields

$$\Delta X \overline{(\rho_0 \phi)}_{\xi(i)}^{n+1} = \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \mathcal{O}(\Delta X^N). \quad (18)$$

On the other hand, one has the following identity $\Delta X \overline{(\rho \phi)}_{\xi(i)}^{n+1} = \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \mathcal{O}(\Delta X^N)$. Using the integral linearity, it gives

$$\Delta X \overline{(\rho \phi)}_{\xi(i)}^{n+1} = \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \mathcal{O}(\Delta X^N).$$

Plugging eq. (18) into the previous equation, it yields

$$\Delta X \overline{(\rho \phi)}_{\xi(i)}^{n+1} = \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \Delta X \overline{(\rho_0 \phi)}_{\xi(i)}^{n+1} + \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy + \mathcal{O}(\Delta X^N),$$

which written under conservative form, dropping the $\mathcal{O}(\Delta X^N)$, gives

$$\overline{(\rho \phi)}_{\xi(i)}^{n+1} = \overline{(\rho_0 \phi)}_{\xi(i)}^{n+1} - \left[\frac{x_{\xi(i)+\frac{1}{2}}^{n+1} - x_{\xi(i)-\frac{1}{2}}^{n+1}}{\Delta X} (\rho_0 \phi)_{\xi(i)+\frac{1}{2}}^* - \frac{x_{\xi(i)-\frac{1}{2}}^{n+1} - x_{\xi(i)-\frac{3}{2}}^{n+1}}{\Delta X} (\rho_0 \phi)_{\xi(i)-\frac{1}{2}}^* \right], \quad (19)$$

where $(\rho_0 \phi)_{\xi(i)+\frac{1}{2}}^*$ satisfies $(\rho_0 \phi)_{\xi(i)+\frac{1}{2}}^* = \frac{1}{x_{\xi(i)+\frac{1}{2}}^{n+1} - x_{\xi(i)-\frac{1}{2}}^{n+1}} \int_{x_{\xi(i)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy$. One easily notices that $(\rho_0 \phi)_{\xi(i)+\frac{1}{2}}^*$ can be written as $(\rho_0 \phi)_{\xi(i)+\frac{1}{2}}^* = \frac{1}{x_{\xi(i)+\frac{1}{2}}^{n+1} - x_{\xi(i)-\frac{1}{2}}^{n+1}} \left(\int_{x_{\xi(i^*)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)+\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy - \int_{x_{\xi(i^*)-\frac{1}{2}}^{n+1}}^{x_{\xi(i)-\frac{1}{2}}^{n+1}} (\rho \phi)(y, t^{n+1}) dy \right)$ with i^* an integer still to be determined to ensure both accuracy and stability. Then introducing the function $H_{\xi(i^*)}^{\rho \phi}(x) = \int_{x_{\xi(i^*)-\frac{1}{2}}^{n+1}}^x (\rho \phi)(y, t^{n+1}) dy$, one gets

$$(\rho_0 \phi)_{\xi(i)+\frac{1}{2}}^* = \frac{1}{x_{\xi(i)+\frac{1}{2}}^{n+1} - x_{\xi(i)-\frac{1}{2}}^{n+1}} \left(H_{\xi(i^*)}^{\rho \phi}(x_{\xi(i)+\frac{1}{2}}^{n+1}) - H_{\xi(i^*)}^{\rho \phi}(x_{\xi(i)-\frac{1}{2}}^{n+1}) \right). \quad (20)$$

Here, upwinded centered Lagrange polynomials are used to interpolate value of $H_{\xi(i^*)}^{\rho\phi}$. The upwinding is done in function of sign of $x_{\xi(i+\frac{1}{2})}^{n+1} - x_{\xi(i+\frac{1}{2})}$. It yields natural value for i^* as a function of the upwinding and the order of the scheme N . In practice, one has $i^* = i - 1 - \lfloor \frac{N}{2} \rfloor$ if $x_{\xi(i+\frac{1}{2})}^{n+1} > x_{\xi(i+\frac{1}{2})}$ and $i^* = i - \lfloor \frac{N-1}{2} \rfloor$ otherwise.

2.5.2. Properties of the remap step

LEMMA 6. *The remap step (19) is conservative in mass ($\phi = 1$), momentum ($\phi = u$), internal ($\phi = \epsilon$) and kinetic ($\phi = e_{kin}$) energies. It conserves by summation the total energy \mathcal{E} defined in definition 4.*

Proof. The proof is straightforward. Indeed due to the conservative form depicted in eq. (19), the projection is conservative in mass, momentum, internal and kinetic energies. Thus, as \mathcal{E} is the sum of both internal and kinetic energies, it is also conserved. \square

As pointed out previously, the conservation of E is a desired feature. The dissipation of total energy during the remap phase is mentioned in the early literature. Indeed, as pointed out by DeBar [6, 7] "kinetic energy disappears in the momentum advection process, and must be compensated for in the internal energy if total energy conservation is to be maintained". It was also formulated similarly later by Youngs [49, 41]. Using the conservation of \mathcal{E} , the internal energy corrector eq. (14) is applied at the end of the remapping stage. It thus yields straightforwardly conservation of both \mathcal{E} and E . Hence, three algorithms are available.

1. Lagrange phase \rightarrow Internal energy corrector,
2. Lagrange phase \rightarrow Internal energy corrector \rightarrow Remap phase \rightarrow Internal energy corrector,
3. Lagrange phase \rightarrow Remap phase \rightarrow Internal energy corrector,

The first algorithm is used to solve the Euler equations in Lagrangian coordinates, whereas the other two are used for the standard Euler equations. The numerical examples shown at the end of this paper are obtained with the third algorithm. A possible modification of the projection is to use monotonicity limiters in order to ensure the monotonic behaviour of the projection. In practice, one may apply the monotonicity preserving limiters [38] for more robustness during the remap phase.

3. 2D extension with directional splitting

As presented in [10, 45, 5], the extension to the multidimensional case is realized using directional splitting. The Euler system in 2D writes

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) + \partial_y(\rho v) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) + \partial_y(\rho uv) = 0, \\ \partial_t(\rho v) + \partial_x(\rho uv) + \partial_y(\rho v^2 + p) = 0, \\ \partial_t(\rho e) + \partial_x(\rho ue + pu) + \partial_y(\rho ve + pv) = 0. \end{cases} \quad (21)$$

3.1. Derivation of the subsystems using the operator splitting technique

The main idea is to split system presented in eq. (21) according to the x - and y -direction. It writes

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0 \\ \partial_t(\rho v) + \partial_x(\rho uv) = 0 \\ \partial_t(\rho e) + \partial_x(\rho ue + pu) = 0 \end{cases} \quad \begin{cases} \partial_t \rho + \partial_y(\rho v) = 0 \\ \partial_t(\rho u) + \partial_y(\rho uv) = 0 \\ \partial_t(\rho v) + \partial_y(\rho v^2 + p) = 0 \\ \partial_t(\rho e) + \partial_y(\rho ve + pv) = 0 \end{cases} \quad (22)$$

Splitting techniques relies on solving alternatively first and second system of eq. (22) with weighted time-steps in order to reach high-order accuracy. The theory of operator splitting and especially of high-order splitting sequences are extensively detailed by McLachlan in [27, 26, 28] and very high-order splitting methods are described by Yoshida in [48]. Using directionnal splitting methods, each subsystems of eq. (22) is solved using the 1D schemes proposed in this paper. However, slight modifications must be first performed. Indeed, as one wishes for global conservation of mass, momentum and total energy, use of values averaged in both directions is required, using rectangle control volumes.

3.2. Modifications of the 1D schemes for the 2D finite volume case

The first important point to mention is the special distribution of variables on the staggered grids in both 2D and 3D. The extension of the internal energy corrector proposed for the 1D schemes is straightforward for multidimensional case.

3.2.1. nD distribution of variables on the modified Arakawa C-type grids

The distribution of variables on the modified Arakawa C-type grids is very similar to the one for the 1D case. The x -velocity u is staggered along the x -direction as well as the density and the kinetic energy related to the x -velocity u . It will be denoted in the following by $e_{\text{kin},u}$. Then similarly, the y -velocity is staggered along the y -direction as well as the density and the kinetic energy $e_{\text{kin},v}$ related to the y -velocity v . If one wishes to extend the schemes to the 3D case, then the z -velocity denoted w should be staggered along the z -direction along with the density and the kinetic energy $e_{\text{kin},w}$. Distribution of variables is depicted on fig. 2. Then, for such a choice of variables, the total energy is the sum of the internal energy and the kinetic energies in each direction. This a key ingredient to yield conservation as will be shown hereafter.

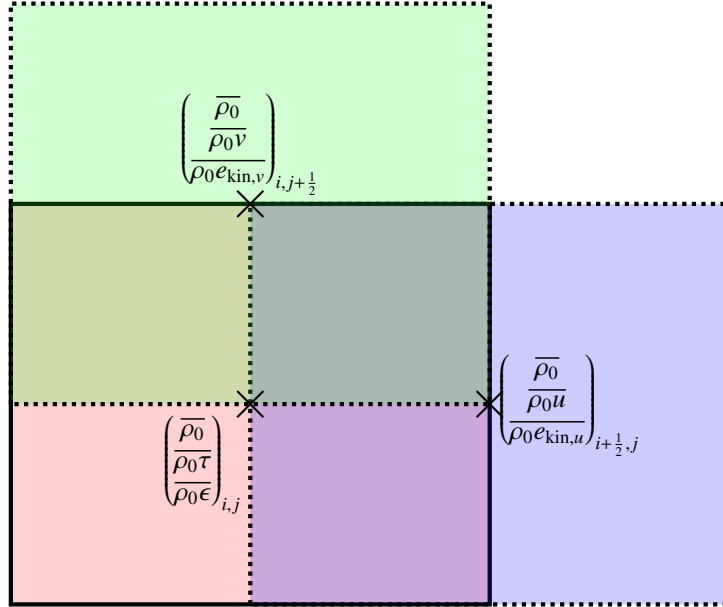


Figure 2: Staggered finite volume space discretization on Cartesian grids

3.2.2. Derivation of a procedure to apply the 1D schemes in one direction using the 2D finite volume formalism

The aim here is to apply with slight modifications the 1D schemes for two dimensions problem using directional splitting method. For two dimension problems, the degrees of freedom are the 2D-average values inside a cell. Thus it is mandatory at the beginning of a sweep, to deduce from 2D average values the 1D average values along one direction. The procedure originates from [10, 45] and is extended here to staggered grids. A sweep along the x -direction proceeds as follows:

1. Interpolate the 2D average values $\bar{\bar{U}}$ along the y -direction to get 1D average values \bar{U} of the variables according to eq. (5). It writes for cell-centered variables $\bar{U}_{i,j}^n = \sum_k C_k \bar{\bar{U}}_{i,j+k}^n$. This way, we only get 1D-cell-average values along the x -direction. This is exactly the values needed to use the 1D scheme.
2. Compute the 1D Lagrange evolution terms using \bar{U} . Note that the velocity in the y -direction and its related kinetic energy do not change. The Lagrange evolution step gives values of the deformed grid $\{x_{i+\frac{1}{2},j}\}$. Interpolation gives value for the $\{x_{i,j}\}$ and $\{x_{i+\frac{1}{2},j+\frac{1}{2}}\}$ grids. The first grid is used to compute remap fluxes of the centered

variables $(\rho_0, \rho_0\tau, \rho_0\epsilon)$, the second for the variables $(\rho_0, \rho_0u, \rho_0e_{\text{kin},u})$ staggered along the x -direction, and the third one for the variables $(\rho_0, \rho_0v, \rho_0e_{\text{kin},v})$ staggered along the y -direction.

3. Denote by $\Delta\mathbf{U}$ the evolution terms (see fig. 3). Reconstruct the average values of $\Delta\mathbf{U}$ in the y -direction using eq. (5) denoted $\overline{\Delta\mathbf{U}}$. It writes for cell-centered variables $\overline{\Delta\mathbf{U}}_{i,j}^n = \sum_k \widehat{C}_k \Delta\mathbf{U}_{i,j+k}^n$.
4. Apply the reconstructed 2D Lagrange-remap terms $\overline{\Delta\mathbf{U}}$ on the 2D-cell-average values. It leads for cell-centered variables to $\overline{\mathbf{U}}_{i,j}^{n+1} = \overline{\mathbf{U}}_{i,j}^n + \overline{\Delta\mathbf{U}}_{i,j}^n$.

The procedure is summarized in fig. 3.

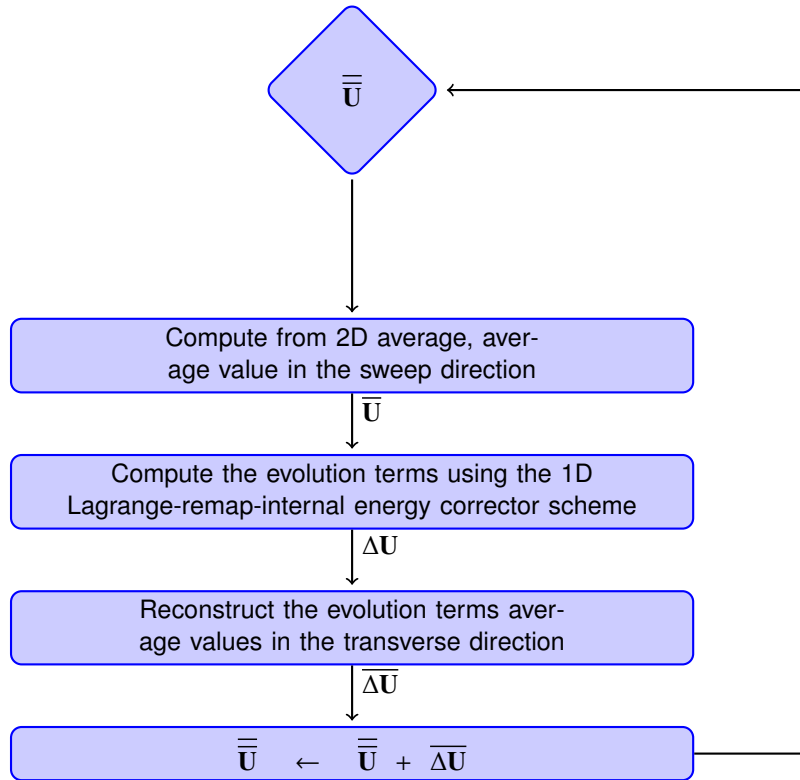


Figure 3: Flow chart of the 2D scheme

3.2.3. Properties of the 2D schemes

LEMMA 7. *The 2D schemes (6)-(7)-(14)-(19) are conservative in mass, momentum and total energies \mathcal{E} and E .*

Proof. With the proposed face-staggering of variables, the 2D schemes satisfy lemmas 2, 5 and 6 direction by direction and so are globally conservative in mass, momentum and total energy for any dimensional splitting sequences. \square

REMARK 6. *Extension to the 3D case is straightforward.*

4. Numerical results

The goal of this section is to validate the theoretical approach as well as to assess numerically both the robustness and the accuracy of the designed schemes.

The reference solutions are computed using the GAD scheme presented in [20]. It is a cell-centered (or collocated) scheme which is naturally consistent with the Euler equations. In the following, the function χ denotes for the indicator function. In particular, one has

$$\chi_F(x) = \begin{cases} 1 & \text{for } x \in F, \\ 0 & \text{otherwise.} \end{cases} \quad (23)$$

The third order Runge–Kutta sequence used for the following numericals results is the SSPRK3 [13, 14]. The fourth order Runge–Kutta sequence is the $\frac{3}{8}$ -Kutta sequence [21]. The fifth order Runge–Kutta sequence is the Dormand–Prince sequence [9]. Last, the sixth, seventh and eighth order Runge–Kutta sequence are the robust Verner sequences available in [43].

4.1. Numerical validation in 1D

The numerical test-suite for validation contains among others three smooth test-problems which are the Cook–Cabot breaking wave test-case proposed in 2004 [4], a slight modification of the breaking wave using a non-convex equation of state and last an acoustic propagation which highlights the advantages concerning staggered grids schemes over cell-centered ones concerning the propagation of waves. Then, four shock test-problems are shown to illustrate the correct capture of shocks, among which the Sod test-case, the Woodward–Colella double blast wave and the Noh compression. The idea is to validate the schemes on a very large variety of test-cases to assess both accuracy and robustness. This is the real difficulty of the proposed test-suite. Recall that for all shock problems, additional artificial viscosities or hyperviscosities are never used. The dissipation induced by the time and space discretization is enough for the proposed test-suite.

4.1.1. Cook–Cabot breaking wave test-case [4]

The Cook–Cabot test-case is designed to assess numerically the order of accuracy of the schemes as the variables profiles are smooth until a given time T_{shock} where a discontinuity occurs. The breaking wave [4] initial data are set as follows:

$$\begin{cases} \rho = \rho_0(1 + \alpha \sin(2\pi x)), \\ p = p_0 \frac{\rho^\gamma}{\rho_0^\gamma}, \\ c = c_0 \frac{\rho}{\rho_0^{(\gamma-1)/2}}, \\ u = \frac{2}{\gamma-1}(c_0 - c), \end{cases} \quad \text{for } -0.5 \leq x \leq 0.5 \quad (24)$$

with the constants defined as $\rho_0 = 10^{-3}$, $p_0 = 10^6$, $\gamma = \frac{5}{3}$, $c_0 = \sqrt{\gamma \frac{p}{\rho}}$ and $\alpha = 0.1$. Finally $T_{\text{shock}} = \frac{1}{(\gamma+1)\pi\alpha c_0}$. The fluid is supposed to be a perfect gas. "For this set of initial conditions, two of the three characteristics are initially constant, with the third satisfying a Burgers-like equation" [4]. The exact solution until T_{shock} is the initial profile advected with velocity $u - c$. The momentum error in l^1 -norm as well as the experimental order of convergence are displayed in table 7. Expected orders of convergence are almost reached. For very high-order methods, the machine precision is already reached with 200 cells.

N_x	STAG-3		STAG-4		STAG-5		STAG-6		STAG-7		STAG-8	
50	9.3e-5	.	6.4e-6	.	5.3e-7	.	1.0e-7	.	3.1e-8	.	5.6e-9	.
100	1.2e-5	2.91	4.3e-7	3.89	2.0e-8	4.68	2.1e-9	5.64	2.6e-10	6.88	5.1e-11	6.79
200	1.6e-6	2.95	3.0e-8	3.86	7.7e-10	4.73	4.1e-11	5.69	2.8e-12	6.59	5.4e-13	6.56
400	2.0e-7	2.98	2.0e-9	3.93	2.6e-11	4.87	1.2e-12	5.1	8.2e-13	★	8.6e-13	★
800	2.6e-8	2.99	1.2e-10	3.96	1.8e-12	3.87	1.4e-12	★	1.7e-12	★	1.7e-12	★
1600	3.2e-9	2.99	8.7e-12	3.85	3.6e-12	★	1.5e-12	★	3.0e-12	★	2.8e-12	★
3200	4.0e-10	3.00	6.2e-12	★	3.8e-12	★	2.2e-12	★	3.3e-12	★	3.1e-12	★

Table 7: l^1 -error in momentum and experimental order of convergence for the Lagrange-remap staggered scheme taken on the Cook–Cabot breaking wave test problem [4], until $t = 0.9T_{\text{shock}}$. ★ indicates machine precision reached.

4.1.2. Sod test-case [36]

The Sod shock tube [36] is a simple Riemann problem for the Euler equations. This test-case proves useful to determine the ability of the scheme to handle shocks and especially the capacity to recover correct discrete Rankine-Hugoniot relations on the shock using the proposed internal energy corrector. Initially, a left state and a right state trigger a rarefaction, contact discontinuity and shock. The domain is $[0 : 1]$ and the initial data are

$$\begin{cases} \rho_0(x) = 1.0\chi_{\{x<0.5\}} + 0.125\chi_{\{x>0.5\}}, \\ p_0(x) = 1.0\chi_{\{x<0.5\}} + 0.1\chi_{\{x>0.5\}}, \\ u_0(x) = 0, \\ \gamma = 1.4. \end{cases} \quad (25)$$

Wall boundary conditions are imposed at $x = 0$ and at $x = 1$. In fig. 4, profiles of density and internal energy are depicted with the analytic solution for a mesh containing 100 cells. In table 8, convergence results on density in norm l^1 are proposed. Although oscillatory due to the absence of artificial viscosities, convergence in the l^1 -norm is achieved.

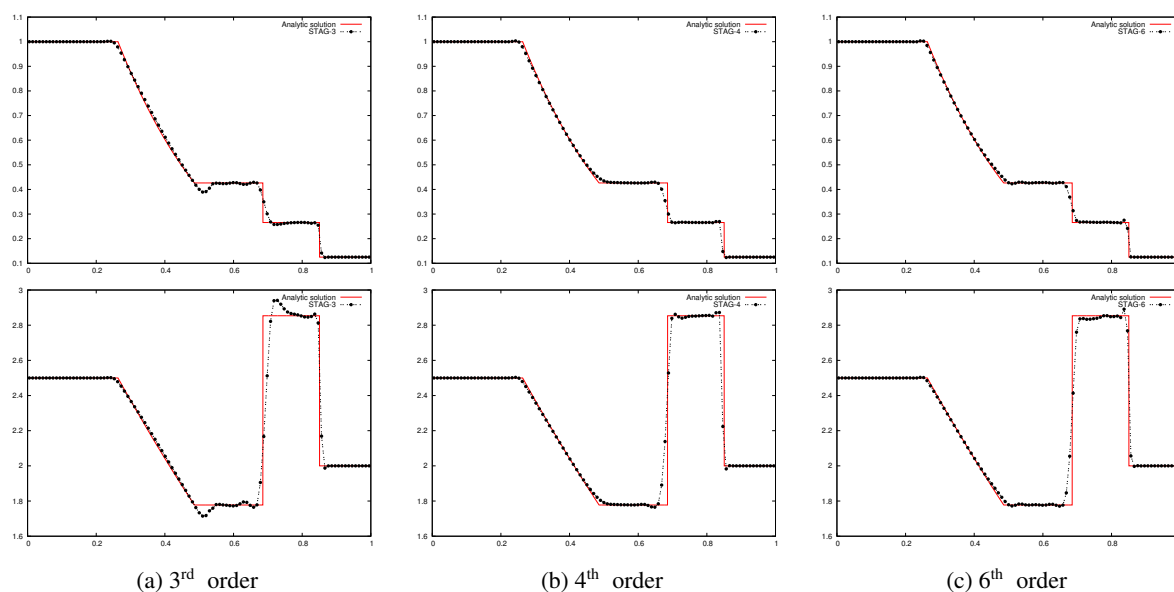


Figure 4: Density (top) and internal energy (bottom) profiles on $[0 : 1]$ for the Sod test-case problem [36] at time $t = 0.2$, CFL=0.7, 100 cells, monotonicity limiters used during the remap phase, no artificial viscosities during the Lagrangian phase, for the 3rd, 4th and 6th order staggered schemes.

N_x	STAG-3	STAG-4	STAG-5	STAG-6	STAG-7	STAG-8
50	1.16e-2	1.00e-2	9.70e-3	1.03e-2	1.02e-2	8.69e-3
200	3.47e-3	2.57e-3	2.50e-3	3.08e-3	5.64e-3	2.53e-3
800	8.65e-4	7.82e-4	7.51e-4	7.09e-4	6.95e-4	6.59e-4
3200	2.82e-4	2.38e-4	2.17e-4	2.24e-4	2.17e-4	2.02e-4
12800	1.02e-4	6.86e-5	7.02e-5	9.43e-5	8.67e-5	6.02e-5
25600	6.20e-5	3.80e-5	3.72e-5	7.07e-5	6.20e-5	4.94e-5

Table 8: l^1 -error in density for the Lagrange-remap staggered scheme taken on the Sod test problem [36], until $t = 0.2$.

4.1.3. Noh test-case [30]

We use the Noh test-case to assess the strong robustness of the schemes without any artificial viscosity or hyperviscosity. This test-case [30] is a compression with a complete conversion of kinetic energy into internal energy. The

domain is fixed at $[0 : 1]$. A continuous incoming flux of gas at $x = 1$ is entering the computational domain with a constant speed and compress the gas located around $x = 0$. We consider an incoming constant state of gas at $x = 1$ and a wall boundary at $x = 0$. The initial data are $(\rho_0, u_0, p_0, \gamma) = (1, -1, 10^{-8}, \frac{5}{3})$. The analytical solution writes

$$\begin{cases} \rho(x, t) &= 4.0\chi_{\{x < \frac{t}{3}\}} + 1.0\chi_{\{x > \frac{t}{3}\}}, \\ u &= -1.0\chi_{\{x > \frac{t}{3}\}}, \\ p &= \frac{4}{3}\chi_{\{x < \frac{t}{3}\}} + 10^{-8}\chi_{\{x > \frac{t}{3}\}}, \end{cases} \quad (26)$$

which gives an infinite shock intensity. This is a real difficulty for most schemes as highlighted in [30]. In fig. 5, profiles of density and pressure are depicted with the analytic solution for a mesh containing 400 cells over $[0 : 1]$. Zoom is made on $[0 : 0.25]$. The higher the order, the more oscillatory the profile is. This is due to the high-order approximations done in the scheme. The important point is that even without artificial viscosity, these high order schemes are robust enough to handle such a difficult test-case. If one is interested to get rid of the remaining oscillations, a possibility is to add artificial viscosity to smear them out.

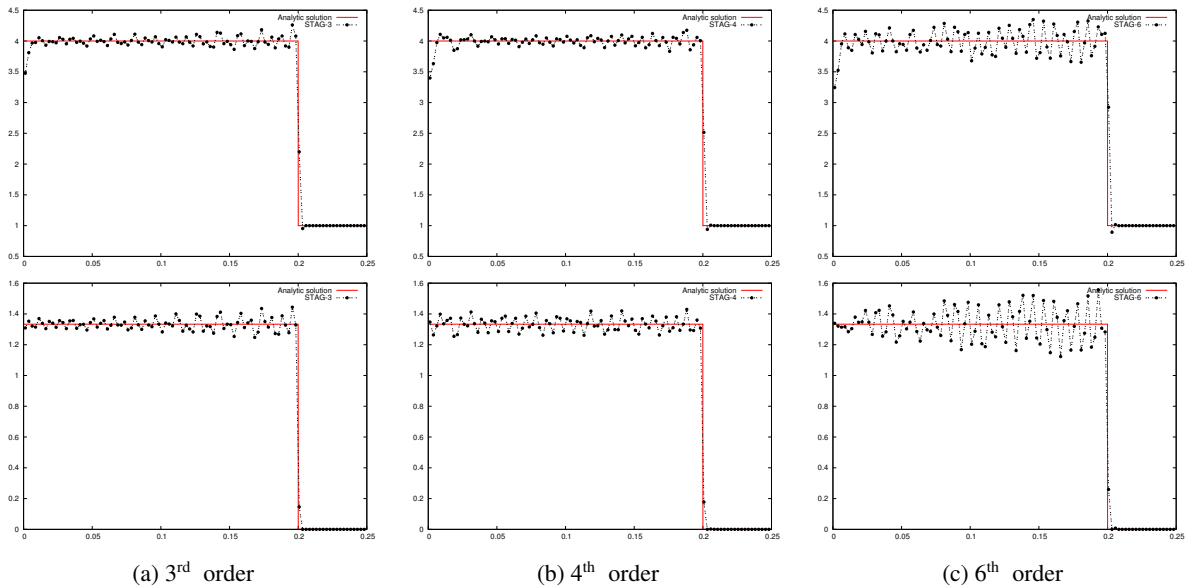


Figure 5: Density (top) and pressure (bottom) profiles on $[0 : 0.25]$ for the Noh test-case problem [30] at time $t = 0.6$, CFL=0.7, 400 cells, monotonicity limiters used during the remap phase, no artificial viscosities during the Lagrangian phase, for the 3rd, 4th and 6th order staggered schemes.

4.1.4. Shu-Osher test-case [35]

The Shu-Osher test-case [35] is a Mach 3 shock wave interacting with a sinusoidal density field. Computations till $t = 1.8$ with CFL=0.7 are reported in fig. 6. This test-case highlights the interest of high-order accuracy even on a shock problem, and especially the restitution of the density profile with high-order accurate schemes. Reference solution is obtained using the GAD scheme with CFL=0.5 and 50000 cells. Initial data, on a $[-5 : 5]$ domain, are

$$\begin{cases} \rho_0(x) &= \frac{27}{7}\chi_{\{x < -4\}} + (1 + \frac{\sin(5x)}{5})\chi_{\{x > -4\}}, \\ p_0(x) &= \frac{31}{3}\chi_{\{x < -4\}} + 1\chi_{\{x > -4\}}, \\ u_0(x) &= \frac{4\sqrt{35}}{9}\chi_{\{x < -4\}}, \\ \gamma &= 1.4. \end{cases} \quad (27)$$

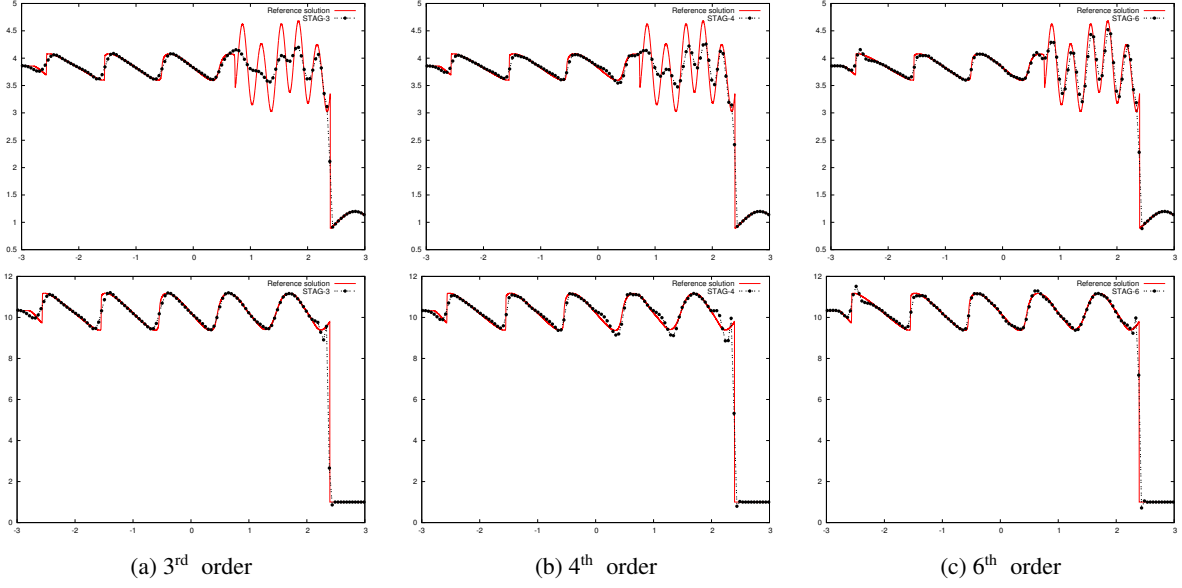


Figure 6: Density (top) and pressure (bottom) profiles on $[-3 : 3]$ for the Shu-Osher test-case problem [35] at time $t = 1.8$, CFL=0.7, 200 cells, monotonicity limiters used during the remap phase, no artificial viscosities during the Lagrangian phase, for the 3rd, 4th and 6th order staggered schemes.

4.1.5. Interacting blast-waves test-case [46]

The interacting blast-waves test-case was proposed in [46]. It is a three states shock tube. The left blast propagates to the right and the right one to the left till interaction between both. This test-case highlights the robustness of the schemes. The domain is set to $[0 : 1]$. Wall boundary conditions are imposed at $x = 0$ and $x = 1$. The initial data are $\rho_0 = 1$, $u_0 = 0$, $\gamma = 1.4$ and $p_0(x) = 1000\chi_{\{x < 0.1\}} + 0.01\chi_{\{0.1 < x < 0.9\}} + 100\chi_{\{0.9 < x\}}$. Density and pressure profiles are shown in fig. 7. Reference solution is obtained using the GAD scheme with CFL=0.5 and 50000 cells.

4.2. Numerical validation in 2D

4.2.1. Isentropic vortex advection [47]

We assess high-order accuracy on the 2D vortex test [47] whose initial data are given by (with $r^2 = x^2 + y^2$)

$$\begin{cases} \rho_0(x, y) &= \left(1 - \frac{(\gamma - 1)\beta^2}{8\gamma\pi^2} e^{1-r^2}\right)^{\frac{1}{\gamma-1}}, \\ \vec{u}_0(x, y) &= (2, 1)^t + \frac{\beta}{2\pi} e^{\frac{1-r^2}{2}} \cdot (-y, x)^t, \\ p_0(x, y) &= \rho_0(x, y)^\gamma, \\ \gamma &= 1.4 \end{cases} \quad (28)$$

with $\gamma = 1.4$ and $\beta = 5$. Computations are performed till $t = 20$ with a CFL number of 0.9 on the computational domain $\Omega = [-10, 10]^2$. Periodic boundary conditions are imposed. The l^1 -error as well as experimental order of convergence are presented in table 9. Expected orders of accuracy of the schemes are reached.

4.2.2. Sedov test-case [32]

With the Sedov test-case, we assess the robustness of the staggered schemes as well as the ability to reconstitute correct cylindrical symmetry. Let $r_{\text{Sedov}} = \frac{1}{\sqrt{2}} \sqrt{\Delta X^2 + \Delta Y^2}$. Initial data are $\rho_0 = 1$, $u_0 = v_0 = 0$, $\gamma = 1.4$ and $p_0(x, y) = \frac{(\gamma - 1)\epsilon_{\text{Sedov}}}{\pi r_{\text{Sedov}}^2} \chi_{\{x^2 + y^2 < r_{\text{Sedov}}^2\}} + 10^{-14} \chi_{\{x^2 + y^2 > r_{\text{Sedov}}^2\}}$ with $\epsilon_{\text{Sedov}} = 0.851072$. A scatter plot is realized to display

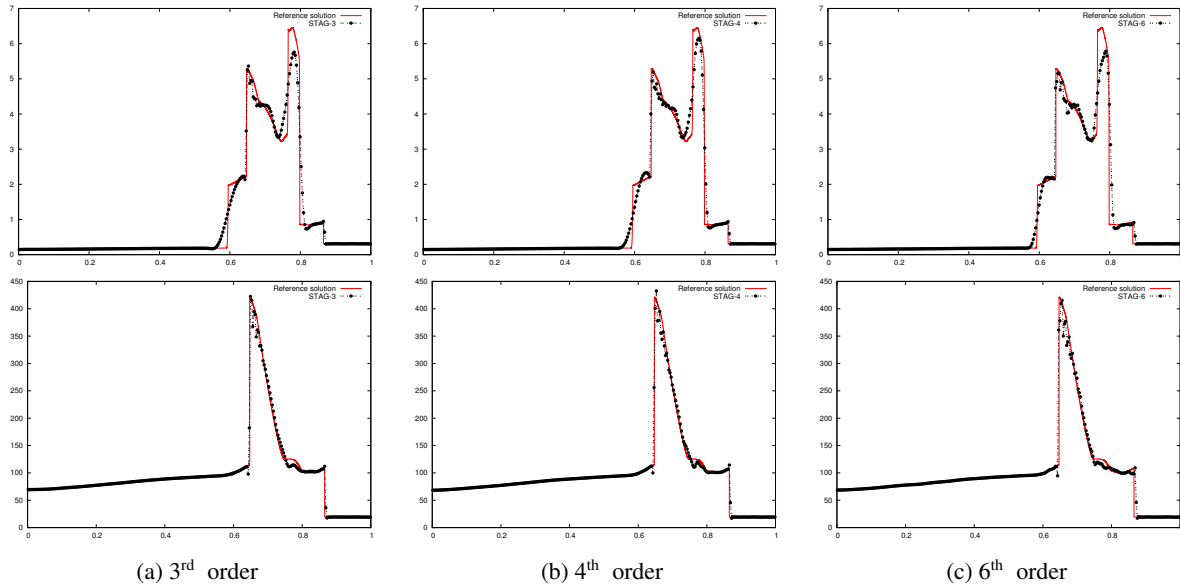


Figure 7: Density (top) and pressure (bottom) profiles on $[0 : 1]$ for the Woodward test-case problem [46] at time $t = 0.038$, $CFL=0.7$, 300 cells, monotonicity limiters used during the remap phase, no artificial viscosities during the Lagrangian phase, for the 3rd, 4th and 6th order staggered schemes.

N_x	STAG-3		STAG-4		STAG-5		STAG-6		STAG-7		STAG-8	
50	3.3e-1	.	1.5e-1	.	2.6e-1	.	1.7e-1	.	1.5e-1	.	1.1e-1	.
100	9.5e-2	1.79	1.9e-2	3.01	4.9e-2	2.41	8.9e-3	4.27	1.2e-2	3.70	2.0e-3	5.83
200	1.6e-2	2.54	1.0e-3	4.19	1.9e-3	4.68	6.5e-5	7.10	8.0e-5	7.20	5.2e-6	8.59
400	2.2e-3	2.89	6.1e-5	4.06	6.1e-5	4.96	7.2e-7	6.48	6.3e-7	7.00	1.6e-8	8.37
800	2.8e-4	2.97	3.9e-6	3.99	1.9e-6	4.98	9.9e-9	6.18	5.0e-9	6.97	1.1e-10	7.17
1600	3.5e-5	2.99	2.4e-7	3.99	5.98e-8	4.99	1.5e-10	6.02	3.9e-11	6.99	3.4e-12	★

Table 9: l^1 -error in density and experimental order of convergence for the Lagrange-remap staggered scheme taken on the isentropic vortex advection test problem [47], until $t = 20$, $CFL=0.9$. ★ indicates machine precision reached.

profiles of density along each radius in fig. 8 using 100 cells in each direction. Even without the use of artificial viscosities, the density profile is quite smooth for each scheme. The higher the order of the staggered schemes, the better the maximum of density near the shock is recovered. The shock position is in good agreement with the analytic solution for the three staggered schemes.

4.2.3. Rayleigh–Taylor instability [37, 40, 24]

Our last test-case is representative of multi dimensional CFD calculations for a complex problem in a low Mach regime with both gravity and Navier–Stokes viscous stress tensor. It is a Rayleigh–Taylor instability. For such problems, collocated schemes struggle to reconstitute correct hydrostatic profiles, and accuracy is damped. On the contrary, staggered schemes are not prone to such difficulties. Our initial data for a single perturbation mode are

$$\begin{cases} \rho_0(x, y) &= 2\chi_{\{y>0\}} + 1\chi_{\{y<0\}}, \\ u_0(x, y) &= 0, \\ v_0(x, y) &= 0.25a(1 + \cos(4\pi x))(1 + \cos(3\pi y))\chi_{\{|y| < 1/6\}}, \\ p_0(x, y) &= K_0 + \rho_0(x, y)gy, \end{cases} \quad (29)$$

where $g = -0.1$, $K_0 = 2.5$, $a = 10^{-2}$. In order to highlight the role of viscosity, computations are run first with the Euler schemes and then with the Compressible Navier–Stokes (CNS) schemes. The viscous parameters are either

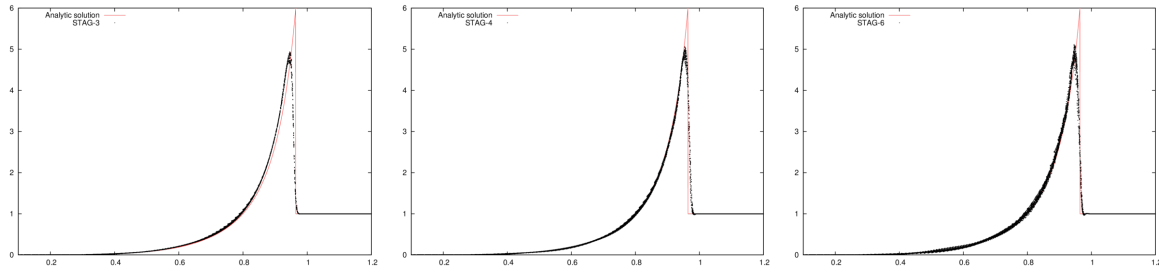


Figure 8: Scatter plot of density profiles for the Sedov blast-wave test-case using the third, fourth and sixth order staggered schemes (CFL=0.7) at $t = 1.0$; 100 cells in each direction.

$\mu = \lambda = 0$ (Euler) or $\mu = 10^{-4}$ and $\lambda = -\frac{2}{3}\mu$ (CNS). Periodic boundary conditions are set on the left and right boundaries, whereas wall boundary conditions are imposed on the top and bottom boundaries. The computation domain is set to $[-0.25 : 0.25] \times [-0.75 : 0.75]$. Results are depicted in fig. 9. Without viscous stress tensor, the higher the order, the more modes develop. Without dissipation, Euler schemes are unable to recover correctly the Rayleigh–Taylor expected profiles, and do not seem to converge as already explained in [24]. On the contrary, using even a small coefficient of viscosity prevents such modes from developing, and seems to help a lot the numerical code to reach numerical convergence. This last test problem also highlights that it is relatively easy to incorporate additional physics to the schemes that have been presented and justified in this work. Here we still consider explicit time integration, and that the diagonal part of the viscous tensor is discretized at the cell centers, whereas the non-diagonal ones is at the cell corners.

5. Conclusion

Using an *a posteriori* internal energy corrector, the proposed staggered schemes are proved to be both conservative in total energy but also weakly consistent with the Euler equations (see theorem 1). They are able to capture properly shocks and discontinuities. Ongoing work includes the use of implicit Runge–Kutta sequences or Rosenbrock ones for time-integration instead of explicit Runge–Kutta sequences. Implicitation is particularly of great interest for low Mach regimes where the Lagrangian CFL condition is the most restrictive, compared to the remap one. The conservation and consistency for such sequences are still to investigate.

References

- [1] Arakawa, A., Lamb, V. R., 1977. Computational design of the basic dynamical processes of the ucla general circulation model. *Methods in computational physics* 17, 173–265.
- [2] Bauer, A. L., Burton, D. E., Caramana, E., Loubère, R., Shashkov, M. J., Whalen, P., 2006. The internal consistency, stability, and accuracy of the discrete, compatible formulation of lagrangian hydrodynamics. *Journal of Computational Physics* 218 (2), 572–593.
- [3] Caramana, E., Burton, D., Shashkov, M., Whalen, P., 1998. The construction of compatible hydrodynamics algorithms utilizing conservation of total energy. *Journal of Computational Physics* 146 (1), 227–262.
- [4] Cook, A. W., Cabot, W. H., 2004. A high-wavenumber viscosity for high-resolution numerical methods. *Journal of Computational Physics* 195 (2), 594–601.
- [5] Dakin, G., Jourden, H., 2016. High-order accurate lagrange-remap hydrodynamic schemes on staggered cartesian grids. *Comptes Rendus Mathématique*.
- [6] DeBar, R. B., 1974. Fundamentals of the KRAKEN code. Tech. rep., Lawrence Livermore National Lab., CA (USA).
- [7] DeBar, R. B., 1974. Method in two-d eulerian hydrodynamics. Tech. rep., Lawrence Livermore National Lab., CA (USA).
- [8] Després, B., 2010. Weak consistency of the cell-centered lagrangian glace scheme on general meshes in any dimension. *Computer Methods in Applied Mechanics and Engineering* 199 (41), 2669–2679.
- [9] Dormand, J. R., Prince, P. J., 1980. A family of embedded runge-kutta formulae. *Journal of computational and applied mathematics* 6 (1), 19–26.
- [10] Duboc, F., Enaux, C., Jaouen, S., Jourden, H., Wolff, M., 2010. High-order dimensionally split Lagrange-remap schemes for compressible hydrodynamics. *C. R. Acad. Sci. Paris, Ser. I* 348, 105–110.
- [11] Gallouët, T., Herbin, R., Latché, J.-C., 2010. Kinetic energy control in explicit finite volume discretizations of the incompressible and compressible navier-stokes equations. *International Journal on Finite Volumes* 7 (2), 1–6.
- [12] Godlewski, E., Raviart, P.-A., 2013. Numerical approximation of hyperbolic systems of conservation laws. Vol. 118. Springer Science & Business Media.

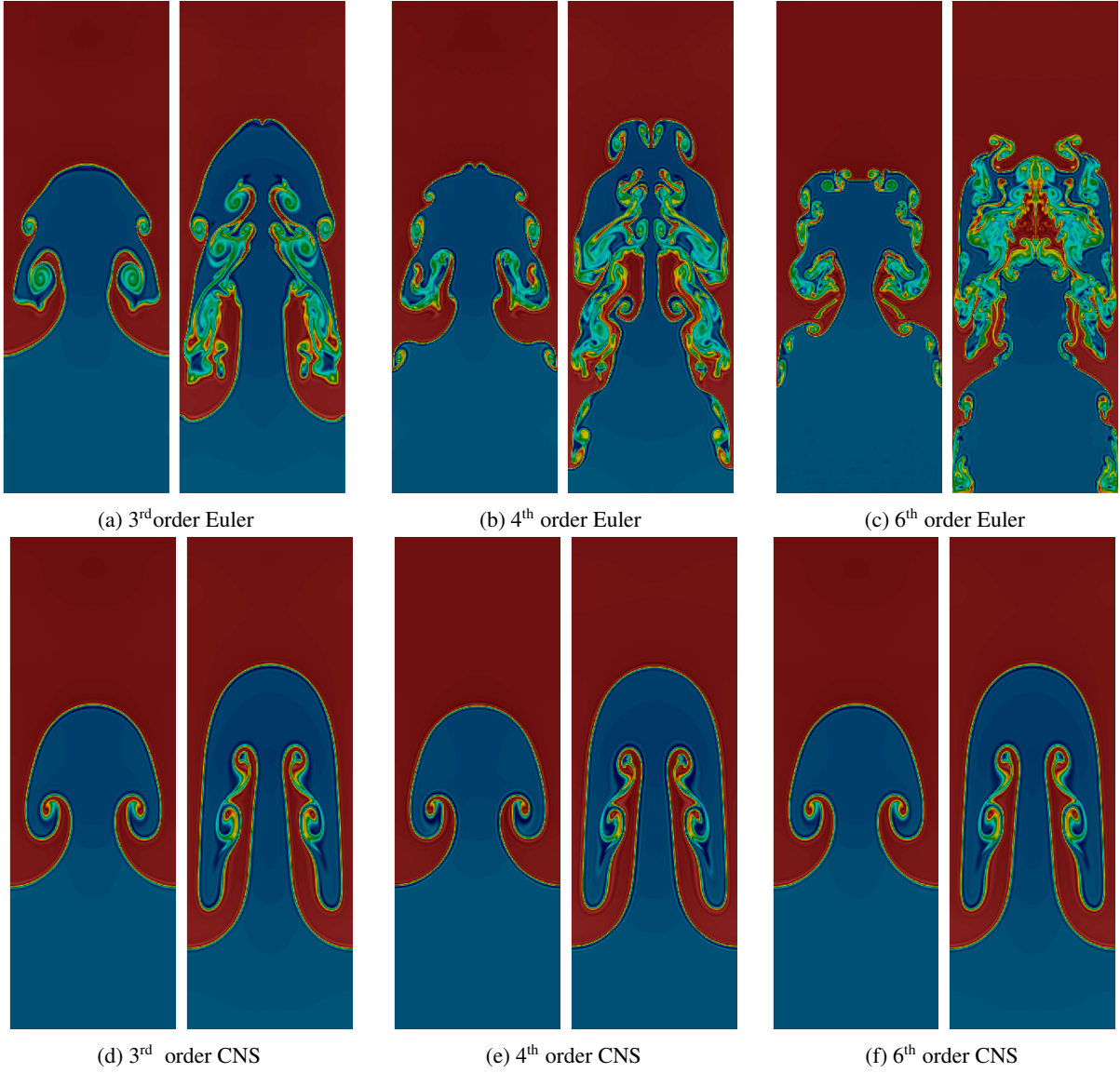


Figure 9: Density profiles on the Rayleigh–Taylor mono-mode instability for the Euler equations (top) and for the Compressible Navier–Stokes (CNS) equations with $\mu = 10^{-4}$ and $\lambda = -\frac{2}{3}\mu$ (bottom) using third, fourth and sixth order schemes, at time $t = 9.5$ (left) and $t = 12.75$ (right) with 200 cells in the x -direction and 600 in the y -direction.

- [13] Gottlieb, S., Shu, C.-W., 1998. Total variation diminishing runge-kutta schemes. *Mathematics of computation of the American Mathematical Society* 67 (221), 73–85.
- [14] Gottlieb, S., Shu, C.-W., Tadmor, E., 2001. Strong stability-preserving high-order time discretization methods. *SIAM review* 43 (1), 89–112.
- [15] Guillard, H., Nkonga, B., 2017. On the behaviour of upwind schemes in the low mach number limit: A review. In: *Handbook of Numerical Analysis*. Vol. 18. Elsevier, pp. 203–231.
- [16] Herbin, R., Latché, J.-C., 2010. Kinetic energy control in the mac discretization of the compressible navier-stokes equations. *International Journal of Finites Volumes* 7.
- [17] Herbin, R., Latché, J.-C., Nguyen, T. T., 2013. Consistent explicit staggered schemes for compressible flows part ii: the euler equation.
- [18] Herbin, R., Latché, J.-C., Nguyen, T. T., 2013. Explicit staggered schemes for the compressible euler equations. In: *ESAIM: Proceedings*. Vol. 40. EDP Sciences, pp. 83–102.
- [19] Herbin, R., Latché, J.-C., Nguyen, T. T., 2017. Consistent segregated staggered schemes with explicit steps for the isentropic and full euler equations.
- [20] Heuzé, O., Jaouen, S., Jourden, H., 2009. Dissipative issue of high-order shock capturing schemes with non-convex equations of state. *J. Comput. Physics* 228, 833–860.
- [21] Kutta, W., 1901. Beitrag zur näherungsweise integration totaler differentialgleichungen.
- [22] Lax, P. D., 1973. Hyperbolic systems of conservation laws and the mathematical theory of shock waves. Vol. 11. SIAM.
- [23] Lax, P. D., Wendroff, B., 1960. Systems of conservation laws. *Communications on Pure and Applied mathematics* 13 (2), 217–237.
- [24] Liska, R., Wendroff, B., 2003. Comparison of several difference schemes on 1d and 2d test problems for the euler equations. *SIAM Journal on Scientific Computing* 25 (3), 995–1017.
- [25] Llor, A., Claisse, A., Fochesato, C., 2016. Energy preservation and entropy in lagrangian space-and time-staggered hydrodynamic schemes. *Journal of Computational Physics* 309, 324–349.
- [26] McLachlan, R. I., 1995. On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM Journal on Scientific Computing* 16 (1), 151–168.
- [27] McLachlan, R. I., Atela, P., 1992. The accuracy of symplectic integrators. *Nonlinearity* 5 (2), 541.
- [28] McLachlan, R. I., Quispel, G. R. W., 2002. Splitting methods. *Acta Numerica* 11, 341–434.
- [29] Noh, W., 1976. Numerical methods in hydrodynamic calculations. Lawrence Livermore Laboratory Report Technical Report UCRL-52112.
- [30] Noh, W. F., 1987. Errors for calculations of strong shocks using an artificial viscosity and an artificial heat flux. *Journal of Computational Physics* 72 (1), 78–120.
- [31] Popov, Y. P., Samarskii, A. A., 1969. Completely conservative difference schemes. *Zhurnal Vyssei Matematiki Matematicheskoi Fiziki* 9 4, 953–958.
- [32] Sedov, L. I., 1946. Propagation of strong shock waves. *Journal of Applied Mathematics and Mechanics* 10, 241–250.
- [33] Serre, D., 1999. *Systems of Conservation Laws 1: Hyperbolicity, entropies, shock waves*. Cambridge University Press.
- [34] Shanmuganathan, S., Youngs, D., Griffond, J., Thornber, B., Williams, R., 2014. Accuracy of high-order density-based compressible methods in low mach vortical flows. *International Journal for Numerical Methods in Fluids* 74 (5), 335–358.
- [35] Shu, C.-W., Osher, S., 1989. Efficient Implementation of Essentially Non-oscillatory Shock-capturing Schemes, II. *J. Comput. Phys.* 83, 32–78.
- [36] Sod, G. A., 1978. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *J. Comput. Physics* 27, 1–31.
- [37] Strutt, J., Rayleigh, L., 1883. Investigation of the character of the equilibrium of an incompressible heavy fluid of variable density. *Proc. London Math. Soc* 14 (1), 8.
- [38] Suresh, A., Huynh, H., 1997. Accurate monotonicity-preserving schemes with Runge-Kutta stepping. *J. Comput. Physics* 136, 83–99.
- [39] Sutcliffe, W. G., 1974. BBC hydrodynamics. Lawrence Livermore Laboratory Report Technical Report UCID-17013.
- [40] Taylor, G., 1950. The instability of liquid surfaces when accelerated in a direction perpendicular to their planes. i. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 201. The Royal Society, pp. 192–196.
- [41] Thornber, B., Drikakis, D., Youngs, D., 2008. Large-eddy simulation of multi-component compressible turbulent flows using high resolution methods. *Computers & Fluids* 37 (7), 867–876.
- [42] Trulio, J. G., Trigger, K. R., 1961. Numerical solution of the one dimensional Lagrangian hydrodynamics equations. Lawrence Radiation Laboratory Report Technical Report UCRL-6267.
- [43] Verner, J., 2013. Jim Verner’s Refuge for Runge-Kutta Pairs. <http://people.math.sfu.ca/jverner/>.
- [44] von Neumann, J., Richtmyer, R. D., 1950. A method for numerical calculation of hydrodynamic shocks. *J. Appl. Phys.* 21, 232–237.
- [45] Wolff, M., 2011. Mathematical and numerical analysis of the resistive magnetohydrodynamics system with self-generated magnetic field terms. Ph.D. thesis, Université de Strasbourg.
- [46] Woodward, P., Colella, P., 1984. The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Physics* 54, 115–173.
- [47] Yee, H., Sandham, N., Djomehri, M., 1999. Low dissipative high-order shock-capturing methods using characteristics-based filters. *Journal of Computational Physics* 150, 199–238.
- [48] Yoshida, H., 1990. Construction of higher order symplectic integrators. *Phys. Letters A* 150, 262–267.
- [49] Youngs, D. L., 2007. “The Lagrangian Remap Method”. *Implicit Large Eddy Simulation: computing turbulent flow dynamics* Cambridge University Press.