An induced limitation in the reconstruction step for Euler equations of compressible gas dynamics in arbitrary dimension

Philippe Hoch

Comptes Rendus
Mathématique
Draft

An induced limitation in the reconstruction step for Euler equations of compressible gas dynamics in arbitrary dimension

Une limitation induite pour l'étape de reconstruction dans les équations d'Euler en dimension quelconque

Philippe Hoch [a]

[a] CEA-DAM, DIF, 91297, Arpajon Cedex, France
E-mail: philippe.hoch@cea.fr

Abstract. We are interested in the limitation process for the reconstruction of quantities related to Euler's equations of compressible gas dynamics for a general pressure law of type $P(\rho,\epsilon)$ (density, specific internal energy). For example, for perfect gas laws, we recall the constraints $\rho > 0$ and $\epsilon > 0$, and that the velocity $\mathbf{U}$ is *a priori* not bounded in the continuous problem. Nevertheless it is in $L^2(\Omega,\rho)$ as a consequence of the relation on the energies $E = \epsilon + \frac{1}{2}|\mathbf{U}|^2$ in $L^1(\Omega,\rho)$ (due to global conservation of total energy ρE). We show a similar result for conservative reconstruction in any space dimension and for an arbitrary reconstruction order. The use of the Leibniz formula on the specific variables ϵ, $\mathbf{U}$ and $\frac{1}{2}|\mathbf{U}|^2$ allows to obtain also such a discrete **induced** control of reconstructed velocity thanks to control of reconstructed density and energies. We build a direct limitation on the weight variable ρ and also especially on the specific variable ϵ. In particular, the latter makes it possible to limit, in an **induced** way, the velocity $\mathbf{U}$. The limited reconstruction of the conservative variables is deduced from the assembly of these different limitation processes. We illustrate in dimension $d = 1$ and $d = 2$ on some test cases, our reconstructions of orders 2 and 3.

Résumé. On s'intéresse au processus de limitation pour la reconstruction des quantitées liées aux équations d'Euler de la dynamique des gaz compressibles pour une loi de pression générale de type $P(\rho,\epsilon)$ (densité, énergie interne massique). Par exemple, pour la loi des gaz parfait, les contraintes sont $\rho > 0$ et $\epsilon > 0$, et la vitesse $\mathbf{U}$ n'est *a priori* pas bornée dans le problème continu. Elle est néanmoins dans $L^2(\Omega,\rho)$ comme conséquence de la relation sur les énergies $E = \epsilon + \frac{1}{2}|\mathbf{U}|^2$ dans $L^1(\Omega,\rho)$ (par conservation globale de l'énergie totale ρE). On montre un principe similaire dans le cadre d'une reconstruction conservative en dimension d'espace quelconque et pour un ordre de reconstruction lui aussi arbitraire. L'utilisation de la formule de Leibniz sur les variables massiques ϵ, $\mathbf{U}$ et $\frac{1}{2}|\mathbf{U}|^2$ permet en effet d'obtenir en discret un contrôle **induit** de la vitesse reconstruite grâce au contrôle des reconstructions de la densité et des énergies. Nous construisons une limitation directe sur la variable poids ρ et aussi surtout sur la variable massique ϵ. En particulier, cette dernière permet de limiter, de manière **induite**, la vitesse $\mathbf{U}$. La reconstruction limitée des variables conservatives se déduit de l'assemblage de ces différents processus de limitation. Nous illustrons sur des cas tests, en dimension $d = 1$ et $d = 2$, les reconstructions d'ordres 2 et 3 ainsi obtenues.

2020 Mathematics Subject Classification. 00X99.

1. Introduction/Framework

We are interested in a novel limitation process for finite volume schemes. We propose to limit some variable in order to fulfill some principles (physical or invariant domain validity) in such a way that it induce the limitation of some other variables. Here, we apply this strategy in the case of compressible Euler system of gas dynamics. In this case, we focus on the construction of the velocity limitation induced by the limitation of specific internal energy (and density). These equations model the conservation of mass, momentum and total energy, and write:

$$\partial_t \mathscr{U}(t,\mathbf{x}) + \nabla \cdot (\mathscr{F}(\mathscr{U}(t,\mathbf{x}))) = \mathbf{0} \tag{1}$$

where $\mathbf{x} = (x_1, ..., x_d)$ and:

$$\begin{aligned} &\mathscr{U} = (\rho, \rho\boldsymbol{U}, \rho E), \qquad \mathscr{F}(\mathscr{U}) = (\rho\boldsymbol{U}, \rho\boldsymbol{U}\otimes\boldsymbol{U} + PI_d, \boldsymbol{U}(\rho E + P)) \\ &\boldsymbol{U} \text{ the velocity,} \qquad E \text{ the specific total energy with} \\ &E = \epsilon + \frac{1}{2}|\boldsymbol{U}|^2, \text{ and } \epsilon \text{ the specific internal energy.} \end{aligned} \tag{2}$$

In the following, vectors are noted in bold. Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in $\mathbb{R}^d$, their dot product is denoted by $(\mathbf{a},\mathbf{b})$ and tensorial product by $\mathbf{a}\otimes\mathbf{b}$. For our purpose the pressure P is considered as $P(\rho,\epsilon)$, where $\rho \in I_{density}$ and $\epsilon \in I_{energy}$ (I_Q is an open interval of $\mathbb{R}$, for exemple for perfect gas EOS (3): $I_{density} = I_{energy} =]0, +\infty[$):

$$P = (\gamma - 1)\rho\epsilon, \qquad (\gamma > 1 : \text{adiabatic constant}). \tag{3}$$

Note however that this work is valid for general pressure laws $P = P(\rho,\epsilon)$ (as far as (1) is hyperbolic). Consider a two point finite volume discretization of (1), let's define some notations.

- Ω_j is a cell, $|\Omega_j|$ his volume. $\partial\Omega_j$ denote the cell boundary, and f is a $d-1$ dimensional hypersurface (f is an edge for $d=2$, a face for $d=3$, etc, $f \subset \partial\Omega_j$), $|f|$ the associated $d-1$ measure, and the cell index k will denote the neighboring cell sharing f with j.
- r is a node, $k_i(r)$ design the set of cells containing the node r.
- in dimension $d=2$, $r+1/2$ will also denotes an edge (see also Figuree 1).

The so called finite volume unknown is the cell average $\mathscr{U}_j(t) := \frac{1}{|\Omega_j|}\int_{\Omega_j} \mathscr{U}(t,\mathbf{x})dx$, we refer e.g. to [8, 28]:

$$\frac{d}{dt}\mathscr{U}_j(t) + \frac{1}{|\Omega_j|}\sum_{f\in\partial\Omega_j}\int_f \mathscr{F}(\mathscr{U}(t,\mathbf{x})).\mathbf{N}ds = 0, \qquad \text{let } \mathscr{V}_j \text{ an approximation of } \mathscr{U}_j, \tag{4}$$

$$\frac{d}{dt}\mathscr{V}_j(t) + \frac{1}{|\Omega_j|}\sum_{f\in\partial\Omega_j} |f| G(\mathscr{V}_j, \mathscr{V}_k, \tilde{\mathbf{N}}_f) = 0. \tag{5}$$

$\tilde{\mathbf{N}}_f$ is an exact or averaged unit normal to f. The function G in (5) is a numerical flux that is locally conservative ($G(a,b,\mathbf{N}) + G(b,a,-\mathbf{N}) = 0$) and consistent ($G(a,a,\mathbf{N}) = \mathscr{F}(a).\mathbf{N}$), for example: Roe, Rusanov, HLL, Relaxation, VFFC...

We consider a high-order spatial reconstruction, represented by a (high degree polynomial) function $\mathscr{R}_j(\mathbf{x},\mathscr{V})$. We also need to deal with a high-order quadrature flux formula $(\omega_l^f, \mathbf{x}_l^f)_{l=1}^s$ (e.g. Gauss Legendre), so that we consider:

$$\frac{d}{dt}\mathscr{V}_j(t) + \frac{1}{|\Omega_j|}\sum_{f\in\partial\Omega_j}\sum_l \omega_l^f |f| G(\mathscr{R}_j(\mathbf{x}_l^f,\mathscr{V}), \mathscr{R}_k(\mathbf{x}_l^f,\mathscr{V}), \tilde{\mathbf{N}}_f(\mathbf{x}_l^f)) = 0. \tag{6}$$

In case of dimension $d=2$ (then from (7) to (10)), let us also consider the extension of numerical fluxes to nodes then to composite finite volume schemes (see [15]). In the latter, the numerical fluxes are defined at all co-dimension object (from 1 to d) on quadrature points of $\partial\Omega_j$.These are either localised at nodes G_j^r, or at some interior edge points $G_j^{r+1/2}$. We recall that these nodal and

composite numerical fluxes $G_j^r, G_j^{r+1/2}$ are locally conservative and consistent (and that the edge part $r+1/2$ may be any of the classical numerical edge flux of (6)). The associated local normals at these points are defined by vectors $\mathbf{C}_j^r = \frac{1}{2}(\boldsymbol{x}_{r+1} - \boldsymbol{x}_{r-1})^\perp$, $\mathbf{C}_j^{r+1/2} = (\boldsymbol{x}_{r+1} - \boldsymbol{x}_r)^\perp$ and for edge invariant quadrature formula:

$$\frac{d}{dt}\mathcal{V}_j(t) + \frac{1}{|\Omega_j|}\left((1-\theta)\sum_{r\in\partial\Omega_j} G_j^r(\mathcal{V}_{k_1(r)},...,\mathcal{V}_{k_m(r)})\,\mathbf{C}_j^r + \theta\sum_{r+1/2\in\partial\Omega_j} G_j^{r+1/2}(\mathcal{V}_j,\mathcal{V}_k)\,\mathbf{C}_j^{r+1/2}\right) = 0. \quad (7)$$

Here again, we consider extension to higher order in both quadrature and polynomial reconstruction of (7)

Hypothesis 1.1. *We assume that there exists $\{q_i,\theta_i\}_{i=1}^s$ some interior quadrature parameters and weights ($\theta_i \geq 0$) such that:*

$$1 - \sum_{i=1}^{s}\theta_i \geq 0. \quad (8)$$

Let $q_0 = 0$ and $q_{s+1} = 1$ be the extremity parameters, and the associated extremity weights:

$$\theta_0 = \theta_{s+1} = \frac{1}{2}(1 - \sum_{i=1}^{s}\theta_i). \quad (9)$$

Let us now consider, the arbitrary order composite (in terms of geometric objects of non-unique co-dimension fluxes type) scheme [15] (here, in dimension $d = 2$):

$$\begin{aligned}\frac{d}{dt}\mathcal{V}_j(t) + \frac{1}{|\Omega_j|}\Bigg((1-\sum_{i=1}^{s}\theta_i)\sum_{r\in\partial\Omega_j} G_j^r(\mathscr{R}_{k_1(r)}(\mathbf{x}_r,\mathcal{V}),...,\mathscr{R}_{k_m(r)}(\mathbf{x}_r,\mathcal{V}))\,\mathbf{C}_j^r + \\ \sum_{i=1}^{s}\theta_i\sum_{e\in\partial\Omega_j} G_j^{r+1/2}(\mathscr{R}_j(\mathbf{x}^e(q_i),\mathcal{V}),\mathscr{R}_k(\mathbf{x}^e(q_i),\mathcal{V}))\,\mathbf{C}_j^{r+1/2}\Bigg) = 0.\end{aligned} \quad (10)$$

- Taking $s = 1$ in (8) (9) gives $q_1 = \frac{1}{2}$, $\theta_1 = \theta \in [0,1]$ so recovering (7) (we may reach the third/fourth order with $\theta = \frac{2}{3}$).
- Taking $(\theta_i, q_i) = (\omega_i, q_i), \forall i = 1,...,s$ (Gauss Legendre) in (8) (9) so that $\sum_{i=1}^{s}\omega_i = 1$ and then by (9) $\theta_0 = \theta_{s+1} = 0$ so recovering arbitrary high-order pure edges schemes (6) (see e.g. [7]).

A major numerical problem is the preservation of the set of admissible states. We recall that for usual equations of state, these are nothing but the positivity of density and temperature which, in mainy cases, corresponds to positivity of specific internal energy ϵ.

One of the main difficulty is due to the fact that such invariant domain are expressed using the definition of the specific internal energy (or temperature), which is a nonlinear function of conservative variables (density, momentum, and total energy). Nevertheless, these latter are of prime importance to approximate correct weak (and entropic) solutions. Therefore, a limitation process of the specific energy leads to complex modifications of the conservative variables. In the literature, the question of the "good" variables to limit is not really addressed and this is the purpose of our contribution. Up to our knowledge, for d-dimensional Euler gas-dynamic equations, no previous works propose a direct algebraic procedure based on non-linear reconstructions (non polynomial) of specific quantities (using Leibniz formula) of arbitrary order.

The litterature on limitation processes for hyperbolic conservation laws is very rich and is still an active research domain. For finite volume methods it may be either included in the reconstruction (of averages unknowns $\mathcal{V}_j$) process or directly applied to the reconstructed flux G (with blending technics). Here, we adopt the philosophy of the former strategy : limitation of high-order reconstruction. It can be divided into two classes: the a-priori and a-posteriori processes.

The latter includes APITALI [14,16], MOOD [4] processes. These methods act by systematicaly testing the numerical solution under some criteria with an a-posteriori limitation. The process

first selects "bad" cells, then the main action is to decrease for this set some high-order parts (in compact or in a hierarchical way). Due to the existence of a low order (generaly first order) scheme fullfiling by construction the given criteria, the iterative loop always converge by choosing appropriate decreasing sequences of degree's reconstruction (see [3]).

The first class includes all the developpment based on the seminal constructive work of Harten based on *a priori* TVD limitation criteria in dimension $d = 1$. Unfortunately, due to a rather negative result of [9] implying that a TVD scheme is at most first order in dimension $d \geq 2$, many efforts and contributions have been done on relaxed stability constraints. Among all the works, one may cite UNO/ENO/WENO methods [1, 12, 13, 20], maximum principle when the continuous PDE satisfies such a property [6, 24, 29], and invariant domain preservation [10, 22, 30].

The paper is organised as follow, we present an arbitrary high-order reconstruction for volumic (density), but also and more important of the specific (primary) quantities (**U** and E). The latter is based on arbitrary order multi-variate Leibniz formula. The second section is devoted to a non-linear reconstruction of internal energy, we show that this reconstruction depends on both high-order velocity terms and high-order internal energy terms. The high-order terms is a made of three contributions, each of them possessing different properties of homogeneity. By imposing equality between the different contributions, we obtain, as a by-product, a natural dependence between limiters, thus inducing those of velocity reconstruction. The third section is devoted to the construction of the so called induced limitation process: a scalar limiter for internal energy is proposed, for which a limiter is automaticaly deduced (and not specifically designed!) for the velocity. In the last section, we asses this induced limitation strategy on some discontinuous solution in dimension $d = 1$ and $d = 2$.

1.1. *Obtaining arbitrary order for volumic and specific variables*

We deal with arbitrary order reconstruction for a pure volumic field Q (e.g. density). If $n+1$ is the order (e.g. same order as the quadrature formula in (6) or in hypothesis 1.1 for (10)), we write:

$$P(\mathbf{x}, Q) = \overline{Q}_j + (\nabla Q)_j(\mathbf{x}-\mathbf{x}_j) + \frac{1}{2}((\nabla^2 Q)_j(\mathbf{x}-\mathbf{x}_j),(\mathbf{x}-\mathbf{x}_j)) + ... + \frac{1}{n!}(\nabla^n Q)_j \overbrace{(\mathbf{x}-\mathbf{x}_j,...,\mathbf{x}-\mathbf{x}_j)}^{n\ times}. \quad (11)$$

Here $\nabla = {}^t(\partial_{x_1},...,\partial_{x_d})$ denotes the gradient operator, $(\nabla^n Q)_j$ is a n-multilinear form acting on $\mathbb{R}^{n\times d}$, and $\mathbf{x}_j = \frac{1}{|\Omega_j|}\int\limits_{\Omega_j} \mathbf{x} d\mathbf{x}$ is the centroïd of cell Ω_j (or a point inside Ω_j).

Each $(\nabla^l Q)_j$ must be an approximation of $(\nabla^l Q)(\mathbf{x}_j)$, and $\overline{Q}_j$ is given by

$$\overline{Q}_j = \begin{cases} Q_j - \sum\limits_{l=1}^{n} c_j^l(Q), & \text{if conservation is mandatory,} \\ Q_j, & \text{otherwise,} \end{cases} \quad (12)$$

with $Q_j = \frac{1}{|\Omega_j|}\int\limits_{\Omega_j} Q d\mathbf{x}$ and

$$\forall\, l = 1,...,n \qquad c_j^l(Q) := \frac{1}{l!}\frac{1}{|\Omega_j|}\int_{\Omega_j} (\nabla^l Q)_j(\mathbf{x}-\mathbf{x}_j,...,\mathbf{x}-\mathbf{x}_j) d\mathbf{x}. \quad (13)$$

We recall that high-order derivatives in (11) may be represented by the tensorial form using canonical basis $e_i \in \mathbb{R}^d = (0,...0,\underbrace{1}_{i^{th}},0...,0)$ and with the Einstein sum convention reads as:

$$(\nabla^n Q) = (\partial_{x_{i_1}}...\partial_{x_{i_n}} Q)\ (e_{i_1}\otimes ... \otimes e_{i_n}) \qquad \text{or} \quad (14)$$

$$(\nabla^n Q)(\mathbf{a},...,\mathbf{a}) = \sum_{i_1=1}^{d} ... \sum_{i_n=1}^{d} (\partial_{x_{i_1}}...\partial_{x_{i_n}} Q)\ a_{i_1}...a_{i_n} \qquad \text{with } \mathbf{a} = \mathbf{x}-\mathbf{x}_j. \quad (15)$$

Notice that $(\partial_{x_{i_1}}...\partial_{x_{i_n}}Q)\prod_{j=1}^{n}a_{i_j}$ can be written as $\prod_{i=1}^{d}(\partial_{x_i}^{m_i}Q)\prod_{i=1}^{d}a_i^{m_i}$ with $m_i=\#\{j;i_j=i\}$. Using the multi-index notation, let $\mathbf{m}$ be a list of d natural numbers $\mathbf{m}=(m_1,m_2,...,m_d)$, $|\mathbf{m}|=\sum_{i=1}^{d}m_i$. Using

$$\partial^{\mathbf{m}} \text{ corresponds to } \overbrace{\partial_{x_1}...\partial_{x_1}}^{m_1}\overbrace{\partial_{x_2}...\partial_{x_2}}^{m_2}\cdots\overbrace{\partial_{x_d}...\partial_{x_d}}^{m_d}, \tag{16}$$

we can rewrite (15)

$$(\nabla^n Q)(\mathbf{a},...,\mathbf{a})=\sum_{|m|=n}\binom{n}{\mathbf{m}}(\partial^{\mathbf{m}}Q)\prod_{i=1}^{d}a_i^{m_i},\qquad \text{with } \binom{n}{\mathbf{m}}=\frac{n!}{m_1!...m_d!}=\frac{n!}{\prod_{i=1}^{d}m_i!}. \tag{17}$$

Remark 1. We also may use other multivariate calculus by using the following notation:

$$\forall n\geq 1\quad (\nabla^n Q)(\mathbf{a},...,\mathbf{a})=[\mathbf{a},\nabla]^n Q=\left(\sum_{i=1}^{d}(a_i\partial_{x_i})\right)^n Q,\ \text{where } \partial_{x_k}a_l=\delta_{kl}\ \text{(Kronecker symbol)}. \tag{18}$$

Here the right hand side of (18) represents a non symetric linear application ($[A,B]\neq[B,A]$) of the vector direction $\mathbf{a}$ with the vector gradient ∇ to the power n, the resulting differential operator being applied to Q (see the two examples below). Using the multinomial Newton formula, the operator $[\mathbf{a},\nabla]^n$ (18) writes

$$\left(\sum_{i=1}^{d}(a_i\partial_{x_i})\right)^n=\sum_{|\mathbf{m}|=n}\binom{n}{\mathbf{m}}\prod_{i=1}^{d}(a_i\partial_{x_i})^{m_i}, \tag{19}$$

and we finally get for $n\geq 1$ (see also (17)):

$$(\nabla^n Q)(\mathbf{a},...,\mathbf{a})=\sum_{|\mathbf{m}|=n}\binom{n}{\mathbf{m}}\prod_{i=1}^{d}a_i^{m_i}\partial^{\mathbf{m}}Q. \tag{20}$$

Example 2. For example, if $n=2$, and $d=2$, the right hand side of (17) (or (20)) applied to Q gives

$$(\nabla^2 Q)(\mathbf{a},\mathbf{a})=a_1^2\partial_{x_1}^2Q+2a_1a_2\partial_{x_1}\partial_{x_2}Q+a_2^2\partial_{x_2}^2Q.$$

which is nothing but the tensorial form (with Hessian) in (11) (with $\mathbf{a}=\mathbf{x}-\mathbf{x}_j$).

Example 3. Taking $n=3$, and $d=2$, the right hand side of (17) (or (20)) applied to Q gives:

$$(\nabla^3 Q)(\mathbf{a},\mathbf{a},\mathbf{a})=a_1^3\partial_{x_1}^3Q+3a_1^2a_2\partial_{x_1}^2\partial_{x_2}Q+3a_1a_2^2\partial_{x_1}\partial_{x_2}^2Q+a_2^3\partial_{x_2}^3Q.$$

We now focus on nonlinear strategies for the reconstruction of specific variables using Leibniz formula.

1.1.1. *Using Leibniz rule for specific quantities*

For any specific quantity S, whenever $P(\mathbf{x},\rho)>0$, we define its nonlinear reconstruction by the **rational function**:

$$R(\mathbf{x},S):=\frac{P(\mathbf{x},\rho S)}{P(\mathbf{x},\rho)}. \tag{21}$$

Where the reconstruction of $P(\mathbf{x},\rho S)$ is given by (11) with $Q=\rho S$. Instead of considering the volumic variable ρS, the idea is to deal with the variations of each variable ρ and S independently (Leibniz formula) and then gather all the information (see the subsections 1.1.2 and Remark 5 below). Assume that the fields ρ and S are regular, we can use the Leibniz rule at any differentiation

order of the product ρS so that: Leibniz multi-variate approach using the identity (18)(19)(20), replacing Q by the product ρS:

$$(\nabla)^n(\rho S)(\mathbf{a},..,\mathbf{a}) = \sum_{|\mathbf{m}|=n} \binom{n}{\mathbf{m}} \prod_{i=1}^{d} a_i^{m_i} \partial^m(\rho S). \tag{22}$$

Noting $\mathbf{0_d} = (0,...,0)$, we can use the multi-index Leibniz formula:

$$\partial^{\mathbf{m}}(\rho S) = \sum_{\mathbf{0_d} \le \mathbf{k} \le \mathbf{m}} \binom{\mathbf{m}}{\mathbf{k}} \partial^{\mathbf{k}}\rho \; \partial^{\mathbf{m}-\mathbf{k}} S \qquad \text{with } \binom{\mathbf{m}}{\mathbf{k}} = \prod_{i=1}^{d} \binom{m_i}{k_i}. \tag{23}$$

In (23), we recall some conventions for multi-indices comparison operators:

Definition 4. *Let* $\mathbf{k}$ *and* $\mathbf{m}$ *be two multi-indices, we have the classical component-wise definition:*

$$\mathbf{k} \le \mathbf{m} \qquad \textit{means that} \qquad \forall i \; k_i \le m_i.$$

We define the $<$ *operator*

$$\mathbf{k} < \mathbf{m} \qquad \textit{means that} \qquad \forall i \; k_i \le m_i \;\; \textit{and} \;\; \exists i \; k_i < m_i. \tag{24}$$

Now, for (23), we finally end up with

$$\forall \mathbf{a} \in \mathbb{R}^d, \qquad \nabla^n(\rho S)(\mathbf{a},...,\mathbf{a}) = \sum_{|\mathbf{m}|=n} \binom{n}{\mathbf{m}} \prod_{i=1}^{d} a_i^{m_i} \sum_{\mathbf{0_d} \le \mathbf{k} \le \mathbf{m}} \binom{\mathbf{m}}{\mathbf{k}} \partial^{\mathbf{k}}\rho \; \partial^{\mathbf{m}-\mathbf{k}} S. \tag{25}$$

For example if $n = 2$ and $d = 2$, then (25) writes:

$$1)\ m_1 = 2\ m_2 = 0: \qquad \binom{2}{(2\,0)} a_1^2 \left[\binom{2}{0} \overbrace{\partial^0_{x_1} \rho \, \partial^{2-0}_{x_1} S}^{k_1=0} + \binom{2}{1} \overbrace{\partial^1_{x_1} \rho \, \partial^{2-1}_{x_1} S}^{k_1=1} + \binom{2}{2} \overbrace{\partial^2_{x_1} \rho \, \partial^{2-2}_{x_1} S}^{k_1=2} \right]$$

$$2)\ m_1 = 1\ m_2 = 1: \begin{cases} \binom{2}{(1\,1)} a_1 a_2 \left[\binom{1}{0}\binom{1}{0} \overbrace{\partial^0_{x_1}\partial^0_{x_2} \rho \, \partial^{1-0}_{x_1}\partial^{1-0}_{x_2} S}^{k_1=0,k_2=0} + \binom{1}{0}\binom{1}{1} \overbrace{\partial^0_{x_1}\partial^1_{x_2} \rho \, \partial^{1-0}_{x_1}\partial^{1-1}_{x_2} S}^{k_1=0,k_2=1} + \right. \\ \left. \binom{1}{1}\binom{1}{0} \overbrace{\partial^1_{x_1}\partial^0_{x_2} \rho \, \partial^{1-1}_{x_1}\partial^{1-0}_{x_2} S}^{k_1=1,k_2=0} + \binom{1}{1}\binom{1}{1} \overbrace{\partial^1_{x_1}\partial^1_{x_2} \rho \, \partial^{1-1}_{x_1}\partial^{1-1}_{x_2} S}^{k_1=1,k_2=1} \right] \end{cases}$$

$$3)\ m_1 = 0\ m_2 = 2: \qquad \binom{2}{(0\,2)} a_2^2 \left[\binom{2}{0} \overbrace{\partial^0_{x_2} \rho \, \partial^{2-0}_{x_2} S}^{k_2=0} + \binom{2}{1} \overbrace{\partial^1_{x_2} \rho \, \partial^{2-1}_{x_2} S}^{k_2=1} + \binom{2}{2} \overbrace{\partial^2_{x_2} \rho \, \partial^{2-2}_{x_2} S}^{k_2=2} \right]$$

after summation, this gives the more common tensorial form:

$$(\nabla^2(\rho S)\mathbf{a},\mathbf{a}) = \left((\rho \nabla^2 S + (\nabla\rho \otimes \nabla S) + (\nabla S \otimes \nabla\rho) + S\nabla^2\rho)\mathbf{a},\mathbf{a}\right). \tag{26}$$

1.1.2. *Computing arbitrary high-order specific quantities*

In view of formula (25) (with $k_i = 0$ terms), we need to compute at least

$$\prod_{i=1}^{d} \partial^{m_i}_{x_i} S, \qquad \forall l = 1,...,n \;\; \forall \mathbf{m};\ |\mathbf{m}| = l \tag{27}$$

which is nothing but all $\nabla^l S,\ l \le n$. Hence, we choose a numerical method to compute high-order terms in each cell Ω_j: $(\nabla S)_j$, ... $(\nabla^n S)_j$ (each must be an approximation of $(\nabla^i S)(\mathbf{x}_j)$). Note also that by symmetry $k_i = m_i$ we need to compute all the high-order term for ρ: $\nabla^l \rho,\ l \le n$ using (20).

Remark 5. We use formula (20) for the volumic quantity ρ to obtain $\nabla^n \rho$ ($n \geq 1$) (replacing S by ρ in (27)). For the specific quantity S, the high-order derivatives (27) are gathered with those of ρ in (25) to deduce $P(\mathbf{x}, \rho S)$ in (11).

2. Obtaining homogeneity properties for non-linear reconstruction on U and ϵ

2.1. *Arbitrary order (limitation-free)*

In case of arbitrary order reconstruction of conservative quantities, we have defined in all cell j

$$P(\mathbf{x}, \rho) = \overline{\rho}_j + \sum_{l=1}^{n} \frac{1}{l!} (\nabla^l \rho)_j (\mathbf{x}-\mathbf{x}_j, ..., \mathbf{x}-\mathbf{x}_j), \tag{28}$$

$$\boldsymbol{P}(\mathbf{x}, \rho\mathbf{U}) = \overline{\rho\mathbf{U}}_j + \sum_{l=1}^{n} \frac{1}{l!} (\nabla^l (\rho\mathbf{U}))_j (\mathbf{x}-\mathbf{x}_j, ..., \mathbf{x}-\mathbf{x}_j), \tag{29}$$

$$P(\mathbf{x}, \rho E) = \overline{\rho E}_j + \sum_{l=1}^{n} \frac{1}{l!} (\nabla^l (\rho E))_j (\mathbf{x}-\mathbf{x}_j, ..., \mathbf{x}-\mathbf{x}_j). \tag{30}$$

Using (25), (29) and (30) may be written as

$$\boldsymbol{P}(\mathbf{x}, \rho\mathbf{U}) = \overline{\rho\mathbf{U}}_j + \sum_{l=1}^{n} \frac{1}{l!} \sum_{|\mathbf{m}|=l} \binom{l}{\mathbf{m}} \prod_{i=1}^{d} (x_i - x_{i,j})^{m_i} \sum_{\mathbf{0_d} \leq \mathbf{k} \leq \mathbf{m}} \binom{\mathbf{m}}{\mathbf{k}} (\partial^{\mathbf{k}} \rho)_j \, (\partial^{\mathbf{m}-\mathbf{k}} \mathbf{U})_j \tag{31}$$

$$P(\mathbf{x}, \rho E) = \overline{\rho E}_j + \sum_{l=1}^{n} \frac{1}{l!} \sum_{|\mathbf{m}|=l} \binom{l}{\mathbf{m}} \prod_{i=1}^{d} (x_i - x_{i,j})^{m_i} \sum_{\mathbf{0_d} \leq \mathbf{k} \leq \mathbf{m}} \binom{\mathbf{m}}{\mathbf{k}} (\partial^{\mathbf{k}} \rho)_j \, (\partial^{\mathbf{m}-\mathbf{k}} E)_j \tag{32}$$

Notation 6. The Leibniz rule in (31) or (32) for the volumic variable ρS (S = **U** or E) shows a separate dependancy on both ρ and the specific variable S. Hence, we will use the following notation

$$\begin{aligned} P(\mathbf{x}, \rho S) &= P(\mathbf{x}, [\rho]_0^n, [S]_0^n) \\ &\text{where } [.]_i^k \text{ denotes partial derivatives of global order between } i \text{ and } k. \end{aligned} \tag{33}$$

We may write now the primary specific variables **U** and E as rational functions. As usual, isolating the constant value in cell from variable function, we have

Definition 7. *Let the reconstructed density be given by (28) and assume that it does not vanish in Ω_j. Then the rational reconstruction of* **U** *and E given by (21) rewrites as*

$$\begin{cases} \boldsymbol{R}(\mathbf{x}, \mathbf{U}) := \frac{\boldsymbol{P}(\mathbf{x}, \rho\mathbf{U})}{P(\mathbf{x}, \rho)}, \\ \boldsymbol{R}(\mathbf{x}, \mathbf{U}) = \mathbf{U}_j + \boldsymbol{HR}(\mathbf{x}, \mathbf{U}), \qquad \left(\boldsymbol{HR}(\mathbf{x}, \mathbf{U}) = \frac{\boldsymbol{P}(\mathbf{x}, \rho\mathbf{U}) - \mathbf{U}_j P(\mathbf{x}, \rho)}{P(\mathbf{x}, \rho)} = \frac{\boldsymbol{HP}(\mathbf{x}, \rho\mathbf{U})}{P(\mathbf{x}, \rho)} \right), \end{cases} \tag{34}$$

and

$$\begin{cases} R(\mathbf{x}, E) := \frac{P(\mathbf{x}, \rho E)}{P(\mathbf{x}, \rho)}, \\ R(\mathbf{x}, E) = E_j + HR(\mathbf{x}, E), \qquad \left(HR(\mathbf{x}, E) = \frac{P(\mathbf{x}, \rho E) - E_j P(\mathbf{x}, \rho)}{P(\mathbf{x}, \rho)} = \frac{HP(\mathbf{x}, \rho E)}{P(\mathbf{x}, \rho)} \right), \end{cases} \tag{35}$$

where each of the $HR(\mathbf{x}, S)$ terms in (34) and (35) ($S = E$ or $S = \mathbf{U}$) may be considered as a rational high-order correction with respect to the constant term.

Definition 8. *Specific internal energy reconstruction*
The specific internal energy reconstruction is defined as:

$$R(\mathbf{x}, \epsilon) := R(\mathbf{x}, E) - \frac{1}{2} |\boldsymbol{R}(\mathbf{x}, \mathbf{U})|^2. \tag{36}$$

It is then obtained by using the reconstruction of the -primary- specific variables **U** *(34) and E (35), and it may be therefore considered as a -secondary- specific variable.*

Lemma 9. *The high-order part $HR(\mathbf{x},.)$ in (34) and (35) has the following properties:*

(1) *it reads as a rational function (see notation in (33)):*

$$HR(\mathbf{x},S) := \frac{HP(\mathbf{x},[\rho]_0^{n-1},[S]_1^n)}{P(\mathbf{x},\rho)} \tag{37}$$

(2) *the numerator in (37) is linear with respect to all S derivatives:*

$$\forall \lambda \in I\!R \qquad HP(\mathbf{x},[\rho]_0^{n-1},[\lambda S]_1^n) = \lambda HP(\mathbf{x},[\rho]_0^{n-1},[S]_1^n). \tag{38}$$

Proof. First, using (13),

$$\begin{aligned} c_j^l(\rho S) &= \frac{1}{l!}\frac{1}{|\Omega_j|}\int_{\Omega_j}(\nabla^l(\rho S))_j(\mathbf{x}-\mathbf{x}_j,...,\mathbf{x}-\mathbf{x}_j)d\mathbf{x}, \\ &= \frac{1}{l!}\frac{1}{|\Omega_j|}\int_{\Omega_j}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^d(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}\le\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}S)_j d\mathbf{x} \end{aligned}$$

We recall that:

$$\sum_{\mathbf{0_d}\le\mathbf{k}\le\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)\ (\partial^{\mathbf{m}-\mathbf{k}}S) = \sum_{0\le k_1\le m_1}...\sum_{0\le k_d\le m_d}\prod_i\binom{m_i}{k_i}(\prod_i\partial_{x_i}^{k_i})\rho\ (\prod_i\partial_{x_i}^{m_i-k_i})S. \tag{39}$$

In order to clarify manipulation of multi-index sums, we use Definition 4 (24). Now, isolating the constant term for S ($k_i = m_i$, $\forall i$) in (39), we have:

$$\sum_{\mathbf{0_d}\le\mathbf{k}\le\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}\partial^{\mathbf{k}}\rho\ \partial^{\mathbf{m}-\mathbf{k}}S = S\partial^{\mathbf{m}}\rho + \sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}\partial^{\mathbf{k}}\rho\ \partial^{\mathbf{m}-\mathbf{k}}S, \tag{40}$$

hence

$$c_j^l(\rho S) = S_j c_j^l(\rho) + \delta_j^l(\rho S),$$

where we have noted

$$\delta_j^l(\rho S) := \frac{1}{l!}\frac{1}{|\Omega_j|}\int_{\Omega_j}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^d(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}S)_j d\mathbf{x}. \tag{41}$$

We thus have

$$\overline{\rho S}_j = S_j\overline{\rho}_j - \sum_{l=1}^n\delta_j^l(\rho S). \tag{42}$$

Since

$$P(\mathbf{x},\rho S) = \overline{\rho S}_j + \sum_{l=1}^n\frac{1}{l!}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^d(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}S)_j,$$

we can isolate the constant term for S (see (40)). Thus, using (42), we have

$$\begin{aligned} P(\mathbf{x},\rho S) &= S_j\left(\overline{\rho}_j + \sum_{l=1}^n\frac{1}{l!}(\nabla^l\rho)_j(\mathbf{x}-\mathbf{x}_j,..,\mathbf{x}-\mathbf{x}_j)\right) - \sum_{l=1}^n\delta_j^l(\rho S) \\ &\quad + \sum_{l=1}^n\frac{1}{l!}\sum_{|m|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^d(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}S)_j \\ P(\mathbf{x},\rho S) &:= S_j P(\mathbf{x},\rho) + HP(\mathbf{x},[\rho]_0^{n-1},[S]_1^n), \end{aligned} \tag{43}$$

where (see also (41)):

$$HP(\mathbf{x},[\rho]_0^{n-1},[S]_1^n) := -\sum_{l=1}^n\delta_j^l(\rho S) + \sum_{l=1}^n\frac{1}{l!}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^d(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}S)_j \tag{44}$$

□

Notice that by keeping explicit dependency of different energies ϵ and $\frac{1}{2}|\mathbf{U}|^2$ in total specific energy E, with $E_j = \epsilon_j + \frac{1}{2}|\mathbf{U}_j|^2$, we have:

$$\begin{aligned} HR(\mathbf{x},E) &= \frac{P(\mathbf{x},\rho\epsilon) + P(\mathbf{x},\rho\frac{1}{2}|\mathbf{U}|^2) - (\epsilon_j + \frac{1}{2}|\mathbf{U}_j|^2)P(\mathbf{x},\rho)}{P(\mathbf{x},\rho)} \\ &= \frac{\big(P(\mathbf{x},\rho\epsilon) - \epsilon_j P(\mathbf{x},\rho)\big) + \big(P(\mathbf{x},\rho\frac{1}{2}|\mathbf{U}|^2) - \frac{1}{2}|\mathbf{U}_j|^2 P(\mathbf{x},\rho)\big)}{P(\mathbf{x},\rho)} \\ HR(\mathbf{x},E) &:= HR^{(E)}(\mathbf{x},\epsilon) + HR^{(E)}(\mathbf{x},\frac{1}{2}|\mathbf{U}|^2). \end{aligned} \tag{45}$$

this may define another way of computing the reconstructed specific total energy:

$$\begin{aligned} R^{sum}(\mathbf{x},E) &:= R^{(E)}(\mathbf{x},\epsilon) + R^{(E)}(\mathbf{x},\frac{1}{2}|\mathbf{U}|^2), \\ &= \epsilon_j + HR^{(E)}(\mathbf{x},\epsilon) + \frac{1}{2}|\mathbf{U}_j|^2 + HR^{(E)}(\mathbf{x},\frac{1}{2}|\mathbf{U}|^2) = E_j + HR^{(E)}(\mathbf{x},\epsilon) + HR^{(E)}(\mathbf{x},\frac{1}{2}|\mathbf{U}|^2) \end{aligned} \tag{46}$$

As we will see thereafter, the core of the approach is to express the high-order part of the kinetic energy in relation (45) as a function of high order terms of velocity already computed in $\boldsymbol{R}(\mathbf{x},\mathbf{U})$. Indeed, the variable ϵ needs a specific attention: its definition makes use of the kinetic energy, so that the interplay between ϵ and $\mathbf{U}$ is not trivial.

2.2. *Computing high-order kinetic energy terms with multi-variate Leibniz formula*

We recall the two specific total energy reconstruction (36) and (46):

$$\begin{cases} R^{def}(\mathbf{x},E) &= R(\mathbf{x},\epsilon) + \frac{1}{2}|R(\mathbf{x},\mathbf{U})|^2, \\ R^{sum}(\mathbf{x},E) &= R^{(E)}(\mathbf{x},\epsilon) + R^{(E)}(\mathbf{x},\frac{1}{2}|\mathbf{U}|^2). \end{cases} \tag{47}$$

Note that imposing $R^{def}(\mathbf{x},E) = R^{sum}(\mathbf{x},E)$ in (47) gives an equation for $R(\mathbf{x},\epsilon)$. Hence, owing to (35), (34), the relation (36) read

$$\begin{aligned} R(\mathbf{x},\epsilon) &= E_j + HR(\mathbf{x},E) - \frac{1}{2}|\mathbf{U}_j + \boldsymbol{HR}(\mathbf{x},\mathbf{U})|^2, \\ &= E_j + HR(\mathbf{x},E) - \frac{1}{2}\big(|\mathbf{U}_j|^2 + 2(\boldsymbol{HR}(\mathbf{x},\mathbf{U}),\mathbf{U}_j) + |\boldsymbol{HR}(\mathbf{x},\mathbf{U})|^2\big), \\ &= E_j - \frac{1}{2}|\mathbf{U}_j|^2 + HR(\mathbf{x},E) - (\boldsymbol{HR}(\mathbf{x},\mathbf{U}),\mathbf{U}_j) - \frac{1}{2}|\boldsymbol{HR}(\mathbf{x},\mathbf{U})|^2. \end{aligned}$$

Using (45), we obtain:

$$R(\mathbf{x},\epsilon) = \epsilon_j + HR^{(E)}(\mathbf{x},\epsilon) + \underbrace{\Big(HR^{(E)}(\mathbf{x},\frac{1}{2}|\mathbf{U}|^2) - (\boldsymbol{HR}(\mathbf{x},\mathbf{U}),\mathbf{U}_j)\Big)}_{:=D(\mathbf{x},\mathbf{U})} - \frac{1}{2}|\boldsymbol{HR}(\mathbf{x},\mathbf{U})|^2. \tag{48}$$

Proposition 10. *The scalar valued rational function $D(\mathbf{x},\mathbf{U})$ in (48) has the following properties:*

(1) *it writes (see notation in (33))*

$$D(\mathbf{x},\mathbf{U}) := \frac{D(\mathbf{x},[\rho]_0^{n-1},[\mathbf{U}]_1^{n-1})}{P(\mathbf{x},\rho)} \tag{49}$$

(2) *the numerator of (49) is a second-order homogeneous polynomial w.r.t the last variable:*

$$\text{order 2 homogeneous: } \forall\lambda\in\mathbb{R} \qquad D(\mathbf{x},[\rho]_0^{n-1},[\lambda\mathbf{U}]_1^{n-1}) = \lambda^2 D(\mathbf{x},[\rho]_0^{n-1},[\mathbf{U}]_1^{n-1}) \tag{50}$$

Proof. By (45) and (37), (34) (same denominator), $D(\mathbf{x},\mathbf{U})$ reads as a rational function

$$D(\mathbf{x},\mathbf{U}) = \frac{HP(\mathbf{x},\rho\frac{1}{2}|\mathbf{U}|^2) - (\boldsymbol{HP}(\mathbf{x},\rho\mathbf{U}),\mathbf{U}_j)}{P(\mathbf{x},\rho)} \tag{51}$$

$$= \frac{HP(\mathbf{x},[\rho]_0^{n-1},[\frac{1}{2}|\mathbf{U}|^2]_1^n) - \left(\boldsymbol{HP}(\mathbf{x},[\rho]_0^{n-1},[\mathbf{U}]_1^n),\mathbf{U}_j\right)}{P(\mathbf{x},\rho)} \tag{52}$$

Using (44) for the kinetic energy ($S=\frac{1}{2}|\mathbf{U}|^2$), the first term of the numerator in (52) reads as

$$HP(\mathbf{x},[\rho]_0^{n-1},[\frac{1}{2}|\mathbf{U}|^2]_1^n) = \tag{53}$$
$$\left(-\sum_{l=1}^{n}\delta_j^l(\rho\frac{1}{2}|\mathbf{U}|^2) + \sum_{l=1}^{n}\frac{1}{l!}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^{d}(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}(\frac{1}{2}|\mathbf{U}|^2)_j\right),$$

and the second term reads as:

$$(\boldsymbol{HP}(\mathbf{x},[\rho]_0^{n-1},[\mathbf{U}]_1^n),\mathbf{U}_j) = \tag{54}$$
$$\left(-\sum_{l=1}^{n}\delta_j^l(\rho\mathbf{U}) + \sum_{l=1}^{n}\frac{1}{l!}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^{d}(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\ (\partial^{\mathbf{m}-\mathbf{k}}\mathbf{U})_j,\mathbf{U}_j\right).$$

Now, subtracting (54) to (53), we must give an expression of

$$(\partial^{\mathbf{m}-\mathbf{k}}(\frac{1}{2}|\mathbf{U}|^2))_j - ((\partial^{\mathbf{m}-\mathbf{k}}\mathbf{U})_j,\mathbf{U}_j) \tag{55}$$

The second term in (55) writes (omitting cell index j):

$$(\partial^{\mathbf{m}-\mathbf{k}}\mathbf{U},\mathbf{U}) = \sum_{s=1}^{d} u_s\partial^{\mathbf{m}-\mathbf{k}}u_s. \tag{56}$$

and the first reads as:

$$\partial^{\mathbf{m}-\mathbf{k}}(\frac{1}{2}|\mathbf{U}|^2) = \frac{1}{2}\sum_{s=1}^{d}\partial^{\mathbf{m}-\mathbf{k}}(u_su_s),$$
$$= \frac{1}{2}\sum_{s=1}^{d}\sum_{\mathbf{0_d}\le\mathbf{t}\le\mathbf{m}-\mathbf{k}}\binom{\mathbf{m}-\mathbf{k}}{\mathbf{t}}\partial^{\mathbf{t}}u_s\,\partial^{\mathbf{m}-\mathbf{k}-\mathbf{t}}u_s. \tag{57}$$

Isolating the constant term with respect to u_s, ie considering the two extremal terms in the sum over $\mathbf{t}$, that is:

$$\forall i\ t_i = 0 \text{ and } \forall i\ t_i = m_i - k_i,$$

and using the previous Definition 4 (see (24)) also for the two bounds $\mathbf{t}=\mathbf{0}_d$ and $\mathbf{t}=\mathbf{m}-\mathbf{k}$, we have:

$$\sum_{\mathbf{0_d}\le\mathbf{t}\le\mathbf{m}-\mathbf{k}}\binom{\mathbf{m}-\mathbf{k}}{\mathbf{t}}\partial^{\mathbf{t}}u_s\,\partial^{\mathbf{m}-\mathbf{k}-\mathbf{t}}u_s = (u_s)\,\partial^{\mathbf{m}-\mathbf{k}}u_s + \partial^{\mathbf{m}-\mathbf{k}}u_s\,(u_s) + \sum_{\mathbf{0_d}<\mathbf{t}<\mathbf{m}-\mathbf{k}}\binom{\mathbf{m}-\mathbf{k}}{\mathbf{t}}\partial^{\mathbf{t}}u_s\,\partial^{\mathbf{m}-\mathbf{k}-\mathbf{t}}u_s \tag{58}$$

so that using (58), (57) and (56) in (55), we finally end up with (49):

$$D(\mathbf{x},[\rho]_0^{n-1},[\mathbf{U}]_1^{n-1}) = HP\big(\mathbf{x},[\rho]_0^{n-1},[\frac{1}{2}\big|\mathbf{U}\big|^2]_1^{n-1}\big), \tag{59}$$

with

$$HP\big(\mathbf{x},[\rho]_0^{n-1},[\tfrac{1}{2}\big|\mathbf{U}\big|^2]_1^{n-1}\big) = -\sum_{l=1}^{n}\left(\delta_j^l(\rho(\tfrac{1}{2}|\mathbf{U}|^2)) - (\delta_j^l(\rho\mathbf{U}),\mathbf{U}_j)\right) +$$
$$\sum_{l=1}^{n}\tfrac{1}{l!}\sum_{|\mathbf{m}|=l}\binom{l}{\mathbf{m}}\prod_{i=1}^{d}(x_i-x_{i,j})^{m_i}\sum_{\mathbf{0_d}\le\mathbf{k}<\mathbf{m}}\binom{\mathbf{m}}{\mathbf{k}}(\partial^{\mathbf{k}}\rho)_j\left(\tfrac{1}{2}\sum_{s=1}^{d}\sum_{\mathbf{0_d}<\mathbf{t}<\mathbf{m}-\mathbf{k}}\binom{\mathbf{m}-\mathbf{k}}{\mathbf{t}}(\partial^{\mathbf{t}}u_s)_j\ (\partial^{\mathbf{m}-\mathbf{k}-\mathbf{t}}u_s)_j\right). \tag{60}$$

In the last internal sum in (60), due to the strict inequality bounds $\mathbf{0_d}<\mathbf{t}<\mathbf{m}-\mathbf{k}$ (see (24)), we have:

- Lower bound: $\exists i^-$ such that $t_{i^-}>0$ and
- Upper bound: $\exists i^+$ such that $(m_{i^+}-k_{i^+}-t_{i^+})>0$,

considering now $u_s \to \lambda u_s$, is nothing but multiplying $\partial^{\mathbf{t}} u_s$ and $\partial^{\mathbf{m-k-t}} u_s$ by λ, which gives the expected 2-homogeneity property. □

Hence, we have obtain in (59) and (60) high-order terms for kinetic energy as a result of those of velocity (already computed!), note however that the highest order is $n-1$ (and not n).

2.3. *A limited reconstruction of internal energy*

With the previous results (48) (59), we can establish the following unlimited reconstruction of specific internal energy (here, let just suppose that $P(\mathbf{x},\rho) \neq 0$):

$$R(\mathbf{x},\epsilon) = \epsilon_j + \frac{HP(\mathbf{x},[\rho]_0^{n-1},[\epsilon]_1^n)}{P(\mathbf{x},\rho)} + \frac{HP(\mathbf{x},[\rho]_0^{n-1},[\frac{1}{2}|\mathbf{U}|^2]_1^{n-1})}{P(\mathbf{x},\rho)} - \frac{1}{2}\left|\frac{\boldsymbol{HP}(\mathbf{x},[\rho]_0^{n-1},[\mathbf{U}]_1^n)}{P(\mathbf{x},\rho)}\right|^2 \tag{61}$$

The high-order part of specific internal energy reconstruction is build with three high-order contributions. The first comes from ϵ data itself, the second term is made from kinetic energy data all derivatives are obtain by those coming from velocity thanks to Leibniz formula (see (59) (60)). Finally, the third term comes from the high-order velocity reconstruction and the algebraic relation (36).
Let λ_ρ, λ_ϵ, $\lambda_{\mathbf{U}}$ be function parameters (called limiters) for ρ, ϵ and $\mathbf{U}$. We begin by introducing the limited version of the reconstructed volumic density (at least verifying positivity at quadrature points, cf (68) or (71)):

$$P^\lambda(\mathbf{x},\rho) \text{ and } \text{ noting } [\rho^\lambda]_i^k \text{ the resulting limited version of } [\rho]_i^k \tag{62}$$

Now, collecting all the properties linked to high-order part of all the specific variable of the previous subsections, we can rewrite a function parameterized version of specific internal energy ϵ of (48) (see also (34)(35)(36)(37)):

$$R^\lambda(\mathbf{x},\epsilon) = \epsilon_j + \lambda_\epsilon \frac{HP(\mathbf{x},[\rho^\lambda]_0^{n-1},[\epsilon]_1^n)}{P^\lambda(\mathbf{x},\rho)} \tag{63a}$$

$$+(\lambda_{\mathbf{U}})^2 \frac{HP(\mathbf{x},[\rho^\lambda]_0^{n-1},[\frac{1}{2}|\mathbf{U}|^2]_1^{n-1})}{P^\lambda(\mathbf{x},\rho)} \tag{63b}$$

$$-\frac{(\lambda_{\mathbf{U}})^2}{2}\left|\frac{\boldsymbol{HP}(\mathbf{x},[\rho^\lambda]_0^{n-1},[\mathbf{U}]_1^n)}{P^\lambda(\mathbf{x},\rho)}\right|^2 \tag{63c}$$

The high-order part in (63a) and (63c) use linear dependancy (37) (38) and the one in (63b) is derived from the 2-homogeneous property (49) (50) (59).

3. Construction of induced limitation for velocity reconstruction

In view of physical constraints for the continuous system $\rho \in I_{density}$, $\epsilon \in I_{energy}$, we may decide to apply a **direct** limitation only for ρ, ϵ for arbitrary order reconstruction (not only second or third order like in [15]). We could limit independently λ_ϵ and $\lambda_{\mathbf{U}}$ (both beeing non-negative scalars), but it is interesting to notice that choosing in (63a)-(63c),

$$\lambda_{\mathbf{U}} = \sqrt{\lambda_\epsilon} \tag{64}$$

we obtain an induced limiter for the velocity (built from that of internal specific energy). Using (63a)-(63c) and (64), we obtain a *direct* limited reconstruction for the internal energy ϵ:

$$R^\lambda(\mathbf{x},\epsilon) = \epsilon_j + \lambda_\epsilon\left(\frac{HP(\mathbf{x},[\rho^\lambda]_0^{n-1},[\epsilon]_1^n)}{P^\lambda(\mathbf{x},\rho)} + \frac{HP(\mathbf{x},[\rho^\lambda]_0^{n-1},[\frac{1}{2}|\mathbf{U}|^2]_1^{n-1})}{P^\lambda(\mathbf{x},\rho)} - \frac{1}{2}\left|\frac{\boldsymbol{HP}(\mathbf{x},[\rho^\lambda]_0^{n-1},[\mathbf{U}]_1^n)}{P^\lambda(\mathbf{x},\rho)}\right|^2\right) \tag{65}$$

and an *induced* limited velocity reconstruction for $\mathbf{U}$ using (34) (38) and (64):

$$\boldsymbol{R}^{\lambda}(\mathbf{x},\mathbf{U}) = \mathbf{U}_j + \lambda_{\mathbf{U}} \frac{\boldsymbol{HP}(\mathbf{x},[\rho^{\lambda}]_0^{n-1},[\mathbf{U}]_1^n)}{P^{\lambda}(\mathbf{x},\rho)}, \qquad \text{with } \lambda_{\mathbf{U}} = \sqrt{\lambda_{\epsilon}}. \tag{66}$$

Now, we have obtained what we claimed at the begining: the velocity reconstruction (66) is limited thanks to the limitation of density and the specific internal energy reconstruction (65) which are here, the variables of interest. Note also, that in view of (65) and (64), the reconstructed total specific energy E have a same limitation factor from his specific internal and kinetic energies. This result seems to be optimal in terms of limiter for E which balances the limitation of the two sub energies ϵ and $\frac{1}{2}|\mathbf{U}|^2$.

Consequences.

- Each of (65) and (66) may be written as $R^{\lambda}(\mathbf{x},Q) = Q_j + \lambda_j \hat{HR}(\mathbf{x},Q)$, where $\hat{HR}(\mathbf{x},Q)$ is the "high-order" extension (polynomial for volumic variables and rational fraction for specific quantities). Taking $\lambda_j \equiv 1$ gives the unlimited version (and for conservative variables (31) (32)) while $\lambda_j \equiv 0$ reduces to first order $R^{\lambda}(\mathbf{x},Q) = Q_j$ (and then $\boldsymbol{P}(\mathbf{x},\rho\mathbf{U}) = \rho_j \mathbf{U}_j$ and $P(\mathbf{x},\rho E) = \rho_j E_j$).
- Classical scalar limiters can be used, Dukowicz, Barth-Jespersen [2], MLP [23] or min-mod. They all give $R^{\lambda}(\mathbf{x}_l,Q) \geq 0$ for all quadrature points $\mathbf{x}_l$ of $\partial\Omega_j$. Some of them also impose that

$$\min_{j'\in V(j)} Q_{j'} \leq R^{\lambda}(\mathbf{x}_l,Q) \leq \max_{j'\in V(j)} Q_{j'} \tag{67}$$

  where $V(j)$ is the set of all neighbooring cells around j (those having at least a common node with j), see Figure 1.

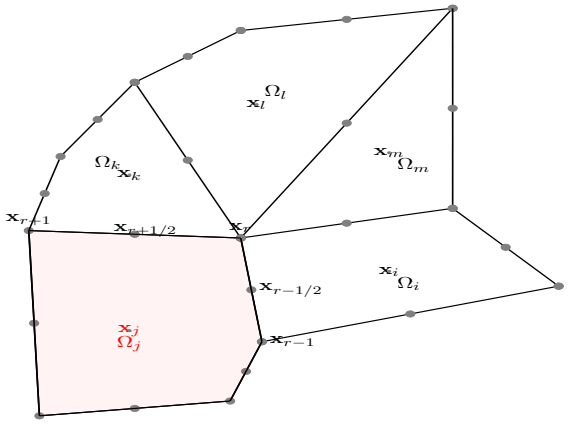

Figure 1. Neighborood of generic cell j: V_j a cell sharing a node or an edge with cell j, quadrature boundary points $\{x_r\}_{r\in\partial\Omega_j}$,$\{x_{r+1/2}\}_{r+1/2\in\partial\Omega_j}$.

- If $\lambda_{\epsilon} = 0$ then $\lambda_{\mathbf{U}} = 0$ (first order taking also $\lambda_{\rho} = 0$). If $\lambda_{\epsilon} = 1$ then $\lambda_{\mathbf{U}} = 1$ (unlimited scheme taking also $\lambda_{\rho} = 1$). Now, in case of $\lambda_{\epsilon} \in]0,1[$ then $\lambda_{\mathbf{U}} > \lambda_{\epsilon}$ (see Figure 2).
- we emphasize that in our limitation process, we need to define **only one limiter** for both the internal energy ϵ in a direct way and all the component of the velocity $\mathbf{U}$ through (64).

Note that the reconstruction of density appearing at the denominator of the velocity (66) and of the specific internal energy (65) is itself limited by any of the two different cases:

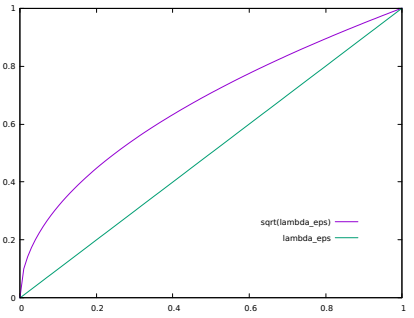

Figure 2. Behavior of the induced limiter of velocity $\lambda_{\mathbf{U}}$ with respect to limiter of internal specific energy λ_ϵ on [0,1]

- a hierarchical strategy λ_ρ^l for l^{th} order term $\nabla^l\rho$:

$$P^\lambda(\mathbf{x},\rho) = \overline{\rho}_j^\lambda + \sum_{l=1}^{n} \frac{1}{l!}\lambda_\rho^l(\nabla^l\rho)_j(\mathbf{x}-\mathbf{x}_j,...,\mathbf{x}-\mathbf{x}_j), \tag{68}$$

  with

$$\overline{\rho}_j^\lambda = \rho_j - \sum_{k=1}^{n} c_j^{k,\lambda}(\rho) \tag{69}$$

  and :

$$c_j^{k,\lambda}(Q) := \frac{1}{k!}\frac{1}{|\Omega_j|}\int_{\Omega_j} \lambda_Q^k(\nabla^k Q)_j(\mathbf{x}-\mathbf{x}_j,...,\mathbf{x}-\mathbf{x}_j)d\mathbf{x}. \tag{70}$$

- or the same function λ_ρ is used for all high-order terms (in the same way as (65) (66)):

$$P^\lambda(\mathbf{x},\rho) = \overline{\rho}_j^\lambda + \lambda_\rho \sum_{l=1}^{n} \frac{1}{l!}(\nabla^l\rho)_j(\mathbf{x}-\mathbf{x}_j,...,\mathbf{x}-\mathbf{x}_j), \tag{71}$$

  with (see (13)):

$$\overline{\rho}_j^\lambda = \rho_j - \lambda_\rho \sum_{k=1}^{n} c_j^k(\rho) \tag{72}$$

Here, λ_ρ in (71) or each λ_ρ^i in (68) are designed to obtain at least the positivity criterion $P^\lambda(\mathbf{x}_l,\rho) > 0$ for all $\mathbf{x}_l$ quadrature points of $\partial\Omega_j$ but also (67) (with $Q = \rho$).

Remark 11. We recover exactly the second and third-order cases of [15]. In (65) (66), we have the following expression for $HP\big(\mathbf{x},[\rho]_0^{n-1},[\frac{1}{2}|\mathbf{U}|^2]_1^{n-1}\big)$

- For n=1: $HP\big(\mathbf{x}\big) \equiv 0$
- For n=2:

$$HP\big(\mathbf{x},\rho_j,(\nabla\mathbf{U})_j\big) = \frac{1}{2}\,{}^t(\mathbf{x}-\mathbf{x}_j)(\rho_j({}^t\nabla\mathbf{U})_j(\nabla\mathbf{U})_j)(\mathbf{x}-\mathbf{x}_j) - \frac{1}{2}\frac{1}{|\Omega_j|}\int_{\Omega_j} {}^t(\mathbf{x}-\mathbf{x}_j)(\rho_j({}^t\nabla\mathbf{U})_j(\nabla\mathbf{U})_j)(\mathbf{x}-\mathbf{x}_j)d\mathbf{x}$$

which are indeed 2-homogeneous polynomials w.r.t high-order velocity terms.

3.1. *Choices of cell by cell limiter $\lambda_\rho, \lambda_\epsilon$*

In view of (10), we need to evaluate high-order fluxes wherever we consider Riemann problems. Many limiter for the density (71) and for the internal energy (63) (that induce velocity limiters (64)) can be investigated. Here we use a minmod/Barth-Jespersen like limiter constructed in the spirit of (71). For each cell we define a scalar cell limiter λ_j^ρ :

- $P^{\lambda}(\mathbf{x},\rho)$, the high-order limited reconstruction of the density:

$$P^{\lambda}(\mathbf{x},\rho) = \rho_j + \lambda_j^{\rho}(P(\mathbf{x},\rho) - \rho_j). \tag{73}$$

Now, we explain the definition of λ_j^{ρ}:

(1) Evaluation of (unlimited) $P(\mathbf{x},\rho)$ at $\mathbf{x}_l$ coordinates of quadrature points (on boundary cell Ω_j nodes and edges) where theRiemann problem is posed, see (6), (10): the associated values $P(\mathbf{x}_l,\rho)$ are denoted by $P_{j,l}$. We also define:

$$\begin{cases} M_j^{\rho,\partial\Omega_j} := \max\limits_{\mathbf{x}_l\in\partial\Omega_j} P_{j,l} \\ m_j^{\rho,\partial\Omega_j} := \min\limits_{\mathbf{x}_l\in\partial\Omega_j} P_{j,l} \end{cases} \tag{74}$$

(2) Evaluation of (unlimited) $P(\mathbf{x},\rho)$ at $\mathbf{x}_{j'}$ centroïd of neighboring cell $\Omega_{j'}$, j' belongs to the neighborhood V_j of j ($j \notin V_j$) see Figure 1, the associated values $P(\mathbf{x}_{j'},\rho)$ are denoted $P_{j,j'}$, we also define:

$$\begin{cases} M_j^{\rho,V_j} := \max\limits_{j'\in V_j} P_{j,j'}, \\ m_j^{\rho,V_j} := \min\limits_{j'\in V_j} P_{j,j'}. \end{cases} \tag{75}$$

(3) Computation of extremal values of previous reconstructed values:

$$\begin{cases} M\rho_j := \max(M_j^{\rho,\partial\Omega_j}, M_j^{\rho,V_j}), \\ m\rho_j := \min(m_j^{\rho,\partial\Omega_j}, m_j^{\rho,V_j}). \end{cases} \tag{76}$$

(4) Computation of extremal bounds of solution in neighborhood V_j:

$$\begin{cases} \rho_j^{max} := \max(\rho_j, \max\limits_{j'\in V_j}\rho_{j'}), \\ \rho_j^{min} := \min(\rho_j, \min\limits_{j'\in V_j}\rho_{j'}). \end{cases} \tag{77}$$

(5) Evaluation of min and max gap to obtain spatial local bound preservation:

$$\begin{cases} s_j^{\rho,max} := \begin{cases} 1, & if\ |M\rho_j - \rho_j| \le \delta, \qquad (\delta \ll 1) \\ \frac{\rho_j^{max}-\rho_j}{M\rho_j-\rho_j}, & else. \end{cases} \\ s_j^{\rho,min} := \begin{cases} 1, & if\ |m\rho_j - \rho_j| \le \delta, \\ \frac{\rho_j^{min}-\rho_j}{m\rho_j-\rho_j}, & else. \end{cases} \end{cases} \tag{78}$$

Finally, we take :

$$\lambda_j^{\rho} = \max(0, \min(1, \min(s_j^{\rho,min}, s_j^{\rho,max}))). \tag{79}$$

- $R^{\lambda}(\mathbf{x},\epsilon)$, the high-order limited reconstruction for internal energy (see (65)), we adopt the same approach replacing $P^{\lambda}(\mathbf{x},\rho)$ (resp. ρ_j) by $R^{\lambda}(\mathbf{x},\epsilon)$ (resp. ϵ_j). We obtain

$$\lambda_j^{\epsilon} = \max(0, \min(1, \min(s_j^{\epsilon,min}, s_j^{\epsilon,max}))), \tag{80}$$

and deduce from (65)

$$R^{\lambda}(\mathbf{x},\epsilon) = \epsilon_j + \lambda_j^{\epsilon}\hat{R}(\mathbf{x},\epsilon)$$

- $\boldsymbol{R}^{\lambda}(\mathbf{x},\mathbf{U})$, the high-order induced limited reconstruction for velocity (66):

$$\boldsymbol{R}^{\lambda}(\mathbf{x},\mathbf{U}) = \mathbf{U}_j + \lambda_j^{\mathbf{U}}\hat{\boldsymbol{R}}(\mathbf{x},\mathbf{U}) \overset{(64)}{=} \mathbf{U}_j + \sqrt{\lambda_j^{\epsilon}}\hat{\boldsymbol{R}}(\mathbf{x},\mathbf{U})$$

Remark 12.

- We emphasize here that other *a priori* limitation strategies are possible, and will be investigated in future works.
- In case of the weaker positivity condition on a target quantity $f = \rho$ (density) or $f = \epsilon$ (specific internal energy), we just set $s_j^{f,max} = 1$ and $f_j^{min} = 0$ (or $f_j^{min} = \eta$ meaning that $\rho \geq \eta$ (or $\epsilon \geq \eta$) for example $\eta = 10^{-15}$) in (78). Nevertheless, we have noticed that in some strong shock test cases, imposing only positivity may lead to oscillations.
- Step (2) has been added to (76) and gives better results in some test cases.
- In order to obtain both $R^\lambda(\mathbf{x}, \rho) > \rho_{min}$ and $R^\lambda(\mathbf{x}, \epsilon) > \epsilon_{min}$ at each quadrature point in (79) (65), we need of course that the constant term (giving by finite volume unknowns at each time step) satisfies:

$$\rho_j > \rho_{min} \quad \text{and} \quad \epsilon_j > \epsilon_{min} \tag{81}$$

  This is essentially linked to qualitative property of the underlying numerical flux of having enough numerical viscosity and also that the underlying first order scheme (with first order reconstruction) fullfills constraints (81).

4. Numerical results

This section is devoted to numerical test cases involving high-order (second and third-order here) reconstruction with the limitation (71) (65) (66). The high-order time integration is obtained with RK3-TVD [25]. For the tests we present here, we take $\delta = 10^{-15}$ in (78). We use a composite extension (node and edges) of Roe (with entropy correction) and Rusanov numerical fluxes [8, 28] (see [15] and (10) with $\theta = \frac{2}{3}$ or $\theta = \frac{\pi}{4}$). For sanity check, we test the overall reconstruction limiter-free $\lambda^\rho \equiv 1$ in (68) (or (71)), and $\lambda^\epsilon \equiv 1$ in (65) (and then (66)) on a regular solution then on a series of classical non regular problems. As highlighted by the last point of remark (12), in order to obtain physical valid constraints (81) (if not fullfill by an high-order scheme), we apply an APITALI process (see section 4.4 of [15] and [14, 16]). For the approximation of discontinuous solutions, the a-posteriori control permits to obtain such constraints. Especially, in the next sub-section 4.2.4, in order to make comparisons, some computations are done without *a priori* limiters ($\lambda^\rho \equiv 1$ (71) and $\lambda^\epsilon \equiv 1$ (65) and then (66)). In this case, we observe the following qualitative behavior:

(1) If we use *a posteriori* strategy by just imposing positivity (81), the numerical solution at final time exhibits some oscillations (see Figure 8 and Figure 9).
(2) If we do not use *a posteriori* strategy, the code may crash in very few time steps by non physical states.

4.1. *Third-order checking on regular Taylor-Green Vortex solution*

This standard test case is used to assess the expected (high-) order of a numerical method. The domain is $[-5, 5] \times [-5, 5]$, and the analytic regular solution is given by:

$$\rho(t, \mathbf{x}) = (T_\infty + \delta T)^{1/(\gamma-1)}, \ \mathbf{U}(t, \mathbf{x}) = \mathbf{U}_\infty + \delta \mathbf{U}, \ p(t, \mathbf{x}) = \rho^\gamma, \tag{82}$$

$$\begin{aligned} \text{with} \qquad \delta \mathbf{U}(t, \mathbf{x}) &= \frac{\beta}{2\pi} \exp(1 - r^2)(-x_2, x_1) \\ \delta T(t, \mathbf{x}) &= -\frac{(\gamma-1)\beta^2}{8\gamma\pi^2} \exp(1 - r^2) \end{aligned} \tag{83}$$

and $T_\infty = 1$, $\mathbf{U}_\infty = \mathbf{0}$ (stationnary), $\beta = 5$, $\gamma = 1.4$, $\mathbf{x} = (x_1, x_2)$ and $r^2 = |\mathbf{x}|^2$, $t_f = 1$.

We note (see Figure 3) that the strategy based on gathering the high-order terms in reconstructed fields: velocity $\mathbf{U}$ and internal energy ϵ (by the way of Leibniz formula) gives indeed a third order convergence (without limitation). We also compare results with the value $\theta = 2/3$ (Simpson formula)

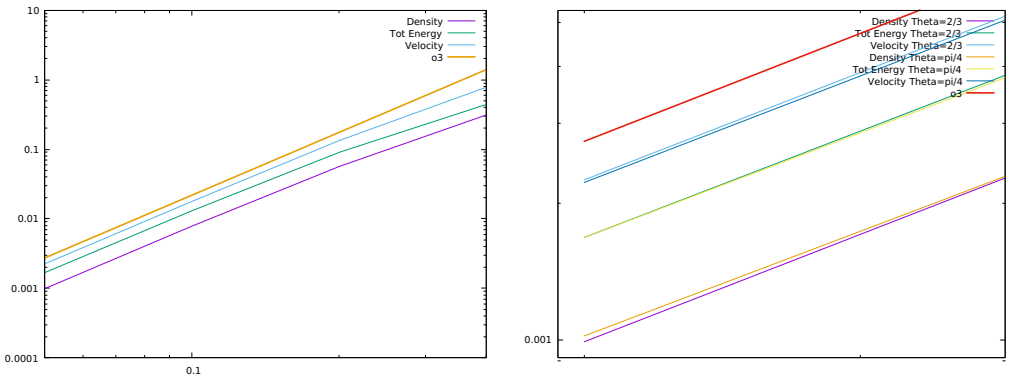

Figure 3. Convergence of third order composite Rusanov limiter-free with $\theta = 2/3$ and $\theta = \pi/4$.

to the value $\theta = \pi/4$ (the conical degenerate case, see [15]). Although only the former gives a third/fourth order quadrature formula for the fluxes, we observe for both cases approximately the same error and order convergence rate.

4.2. *1D test cases with non-smooth solution*

4.2.1. *Sod Shock Tube*

The standard Riemann problem of Sod shock tube is solved to asses the good behavior of limitation process, conservation property (discontinuity location and level of intermediate states for weak solutions) of the schemes. It initialy consists of a perfect gas ($\gamma = 1.4$) at rest with left state $(\rho, u, P)_{Left} = (1, 0, 1)$ and a right state $(\rho, u, P)_{Right} = (.125, 0, 0.1)$.

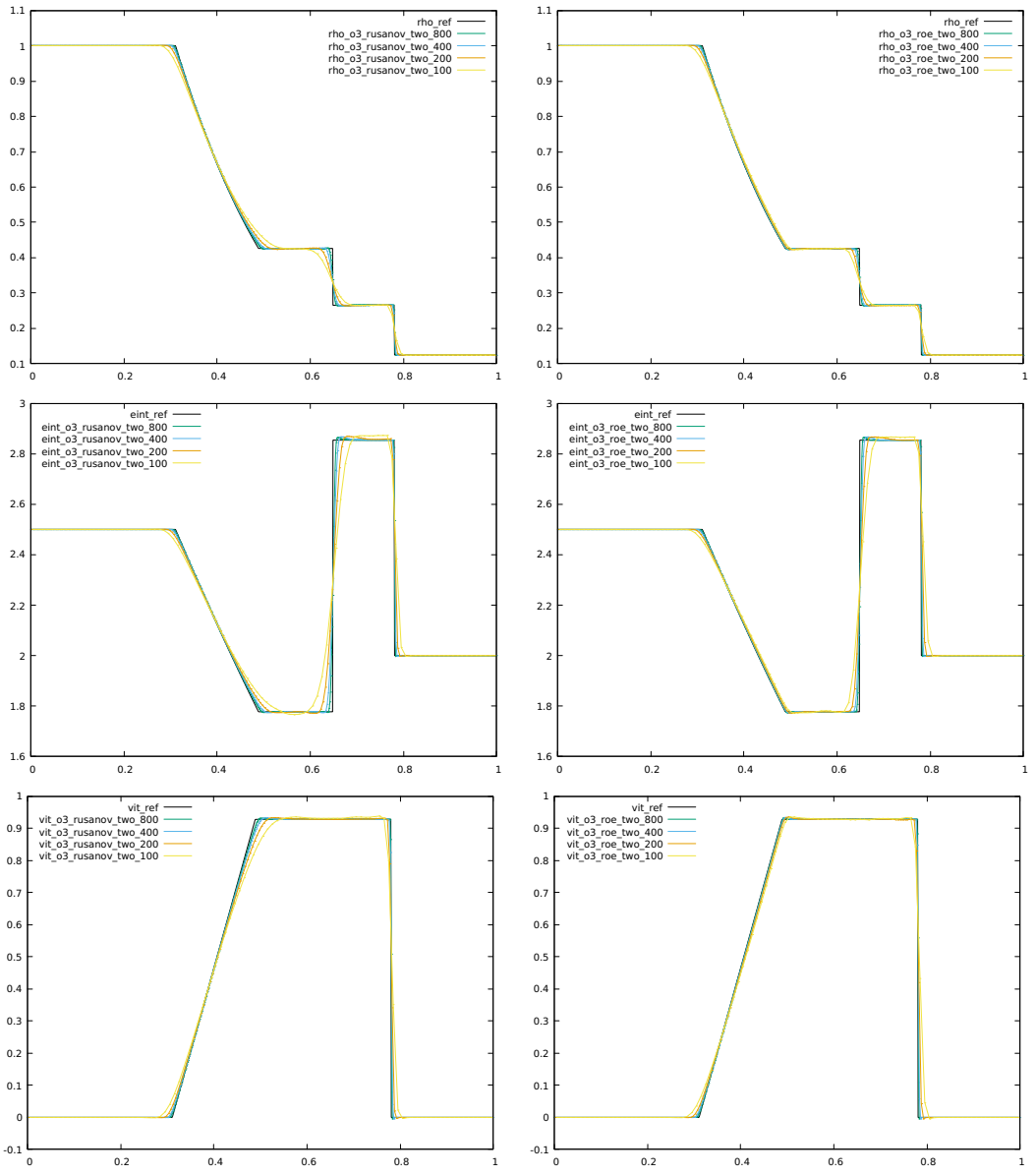

Figure 4. Mesh convergence of third-order reconstruction and induced limitation on Sod shock tube (Rusanov (Left) / Roe (Right)): density (Top), internal energy (Center), velocity (Below).

For the Sod shock tube (cf Figure 4), we observe a good behavior of the induced limitation for the third order reconstruction, especially for the velocity field for which the limitation is linked with those of internal energy (cf (66)). In Figure 5, for two different numerical composite fluxes, we observe the benefits made by increasing the order (one, two and three) of the reconstruction along with the induced limitation.

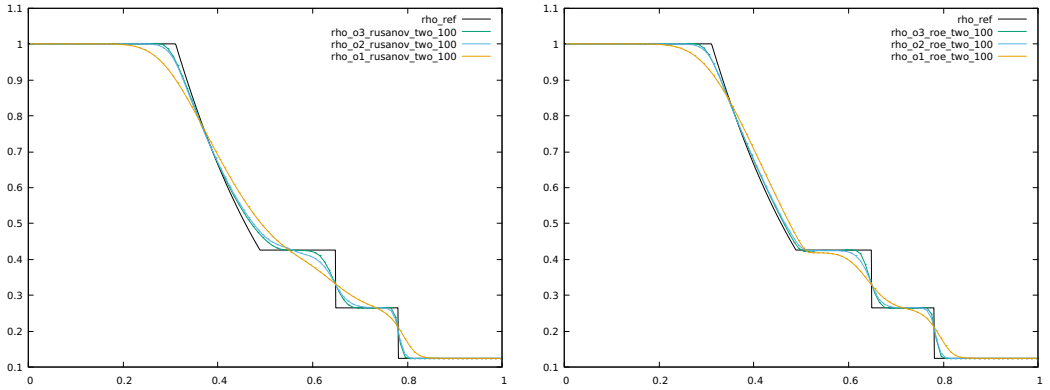

Figure 5. Comparison on 100 cells (for density) of order's 1, 2 and 3 with reconstruction and induced limitation (71)(79) (65)(80) (66) on Sod shock tube (Rusanov (Left) / Roe (Right)).

4.2.2. *Osher-Shu test case*

This test case [26] requires the scheme to be both high-order to reproduce the smooth physical high frequencies but also to have good limitation process to resolve a discontinuity. The initial conditions contains a moving Mach 3 shock wave that interacts with periodic perturbations in density:

$$(\rho, u, P) = \begin{array}{ll} (3.857143, 2.629369, 10.333333), & -5 \leq x_1 \leq -4, \\ (1 + 0.2\sin(5x_1), 0, 1), & -4 < x_1 \leq 5. \end{array} \tag{84}$$

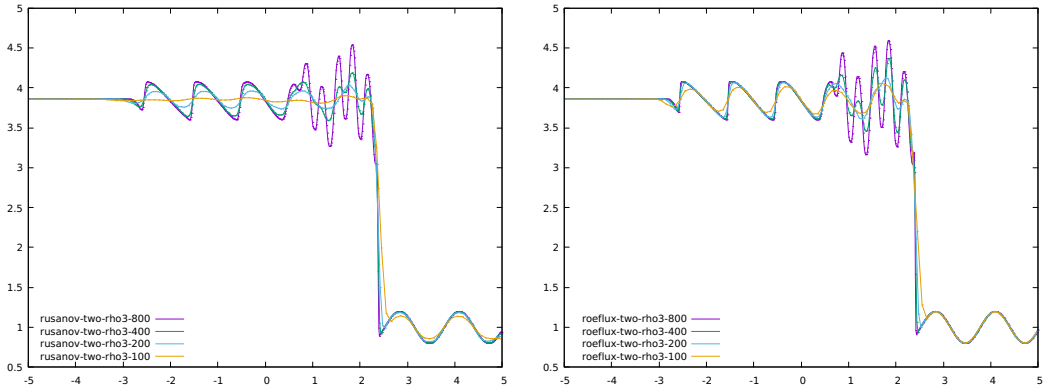

Figure 6. Mesh convergence of third-order schemes with induced limitation (71) (65) (66) on Osher-Shu test case (Rusanov (Left) / Roe (Right))

For the Osher-Shu test case (Figure 6), we note that the convergence rate of the Rusanov scheme is slower than the convergence rate of Roe scheme. Indeed for a coarse mesh (100 cells), even for third order reconstruction, the composite Rusanov scheme damps all high frequencies of the solution ($x_1 \in [0,2]$). They are almost flattened (compared to composite Roe flux). Obviously, for finer mesh (800 cells), this difference between both schemes is less significant.

4.2.3. *Blast waves test case*

The blast waves test case proposed by Colella/Woodward [5] assesses the robustness of the overall numerical machinery. The solution results from the interaction of two reflecting shocks. A robust limitation strategy is mandatory (if not then negative internal energy occurs quasi immediately).

$$(\rho, u, P) = \begin{cases} (1, 0, 1000), & 0 \leq x_1 \leq 0.1, \\ (1, 0, 10^{-2}), & 0.1 < x_1 < 0.9, \\ (1, 0, 100), & 0.9 \leq x_1 \leq 1. \end{cases} \tag{85}$$

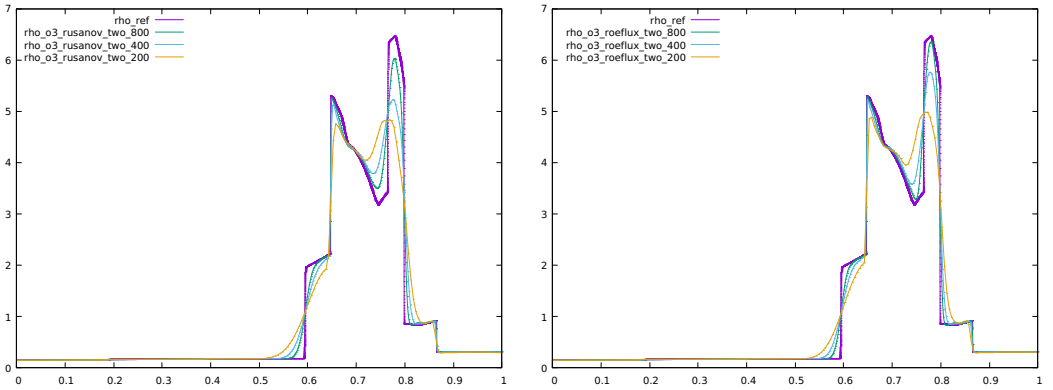

Figure 7. Mesh convergence of third-order schemes with induced limitation (71)(79) (65)(80) (66) for the Colella/Woodward's blast waves problem (Rusanov (Left) / Roe (Right)).

For the blast waves test case (Figure 7), we observe less smeared convergence of Roe scheme with respect to the Rusanov but the numerical results do not exhibit Gibbs phenomenum. However, the resolution of contact discontinuity at $x_1 \simeq 0.6$ could be enhanced by using an HLLC composite flux or by other ad-hoc technics.

4.2.4. *Effect of* a priori *induced limitation of section 3*

In this paragraph, we test the behavior with and without *a priori* induced limitation strategy in (71)(79) (65) (80) (66) (see Figure 8).

Sod.

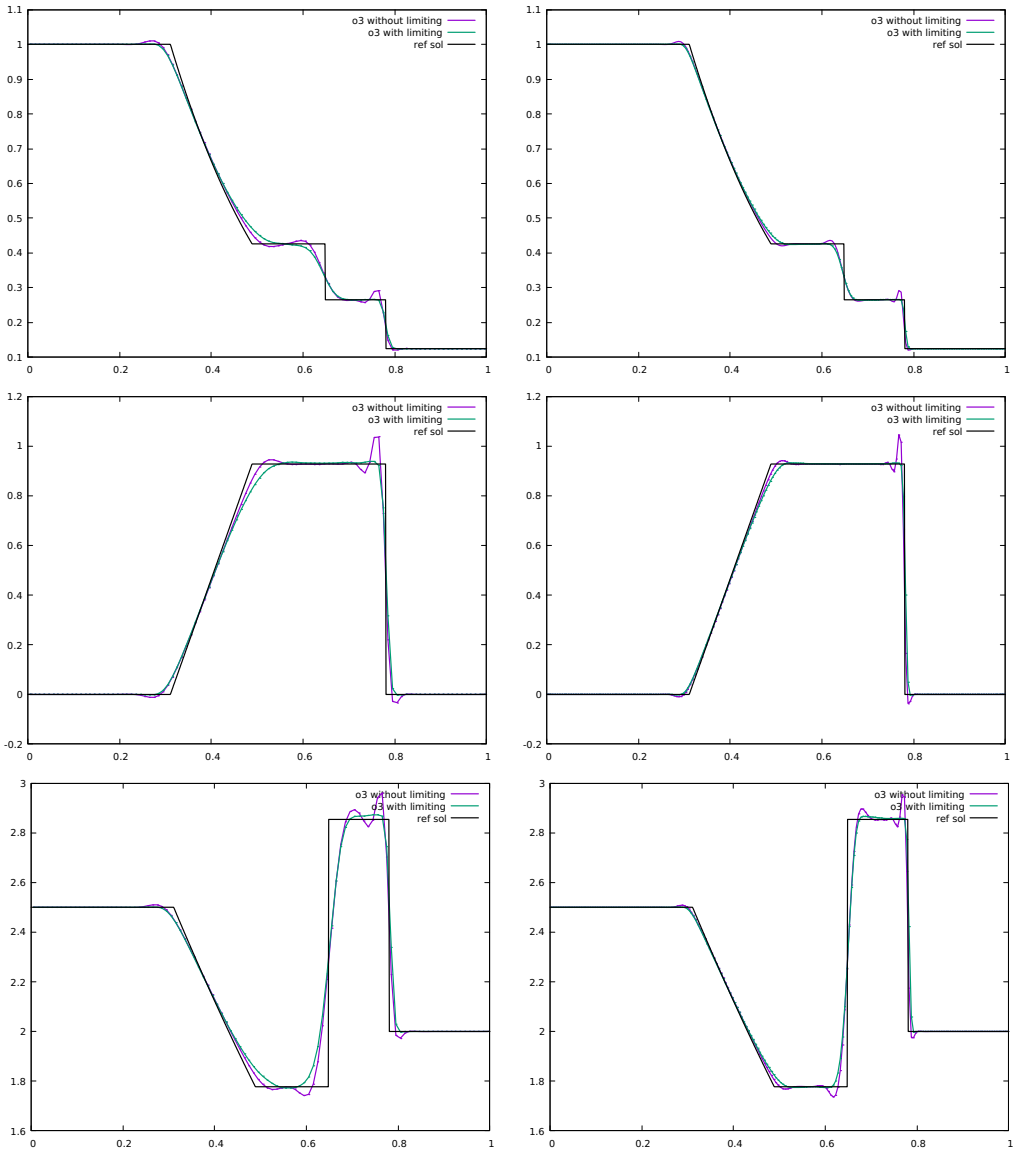

Figure 8. Sod shock tube: From Top to Bottom: Density, Velocity, Internal Energy. Comparison (third order Rusanov composite scheme) with and without a-priori induced limitation (71)(79) (65) (80) (66) on 100 cells (left) and 200 cells (right).

Noh. The Noh test case need of using special care for high-order methods. This is mainly because of the internal energy is nearly zero. More precisely, for this shock tube problem, the data are $\gamma = \frac{5}{3}$, $t_f = 0.6$ and the initial condition is uniform:

$$(\rho, u, P) = (1, -1, 10^{-13}) \tag{86}$$

On the left side, a symmetry is imposed, and an inflow on the right side (see Figure 9).

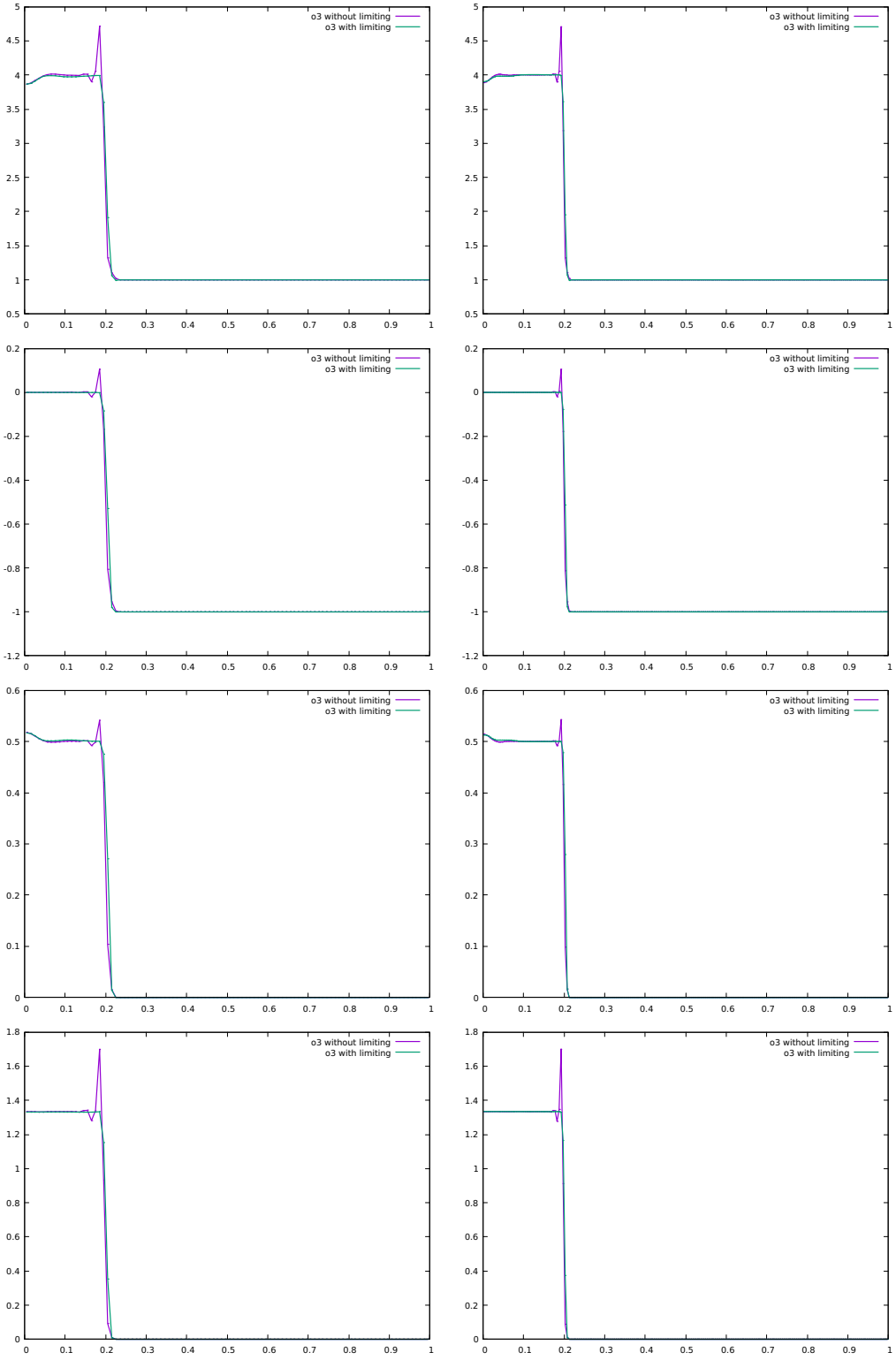

Figure 9. Noh test: From Top to Bottom: Density, Velocity, Internal Energy, Pressure. Comparison (third order Rusanov composite scheme) with and without a-priori induced limitation (71)(79) (65) (80) (66) on 100 cells (left) and 200 cells (right).

Results show the good capability of the induced limitation described in previous section 3 do deal with discontinuities (see e.g. velocity or pressure in Noh test case Figure 9). We recall that no direct limitation of the velocity reconstruction is used, it is only deduce from the limitation of an ad-hoc internal energy reconstruction and (64). An *a posteriori* limitation is mandatory (resp. useless) in order to reach final time in case of the induce *a priori* limitation strategy is not used (resp. used).

4.3. *2D test cases with non-smooth solution*

4.3.1. *Lax-Liu (case 12)*

In two dimensional case, Lax-Liu proposed some 2D-Riemann problems with initial data resulting in different Shock/Contact/Rarefaction configurations (see also [19] for some schemes comparison). For the case 12 of [18], the data on each quadrants are (see Figure 10 Top Left):

$$\begin{aligned} (\rho,\mathbf{U},P)_2 &= (1,(0.7276,0),1), & (\rho,\mathbf{U},P)_1 &= (.5313,(0,0),0.4),\\ (\rho,\mathbf{U},P)_3 &= (.8,(0,0),1), & (\rho,\mathbf{U},P)_4 &= (1,(0,0.7276),1). \end{aligned} \tag{87}$$

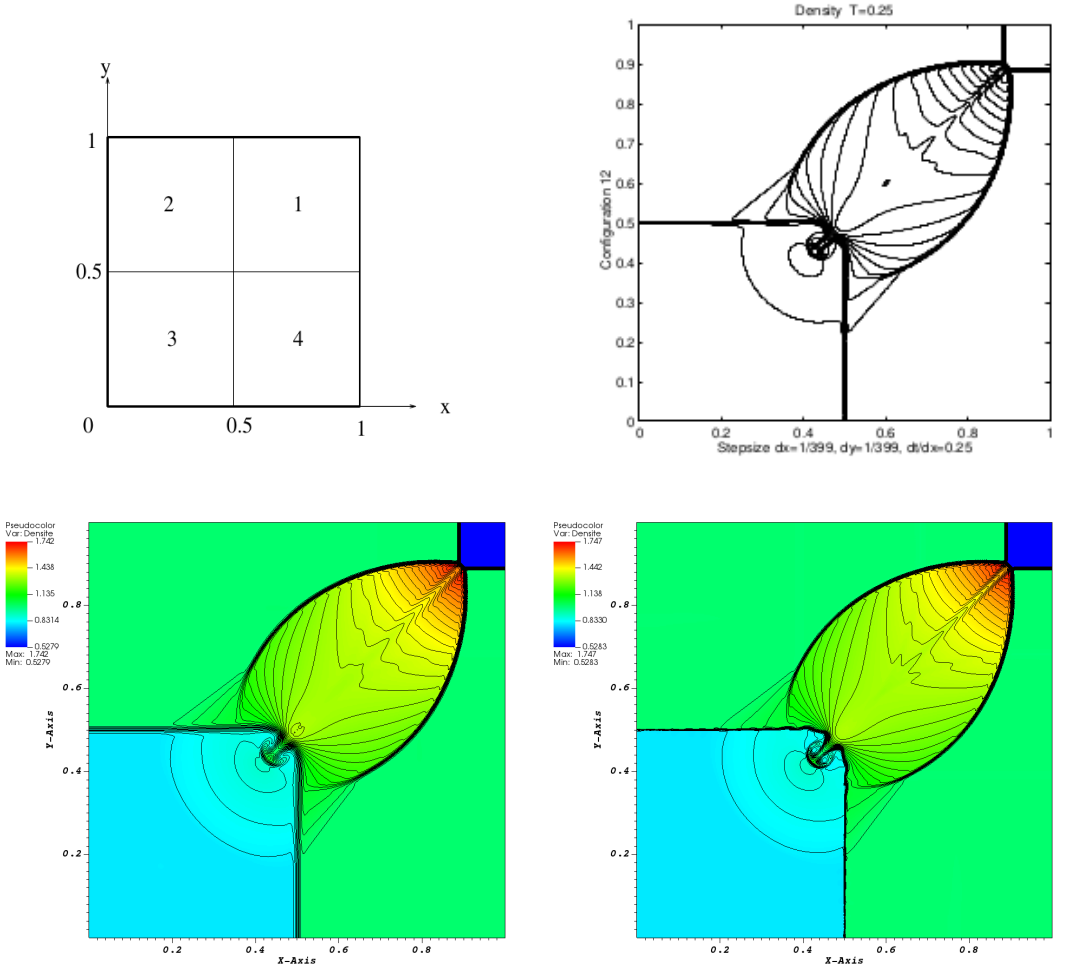

Figure 10. Density for third-order schemes with induced limitation (71)(79) (65) (80) (66) for two dimensional Riemann problem (configuration 12 of [18] Top Left and (87) with a reference solution on Top Right see in [19]) (Rusanov (Left) / Roe (Right)) with 400x400 cells.

We observe in Figure 10 a good agreement with respect to manufactured reference solution. Here, we emphasize that for the velocity treatment, we do not need to use an ad-hoc procedure for vectorial limitation such as VIP [16, 21].

4.3.2. *Double Mach Reflection*

For the test case proposed in [5], we compare results (see Figure 11) of our proposed limitation strategy with two numerical composite fluxes Rusanov and Roe. At final time $t_f = 2$, we note that

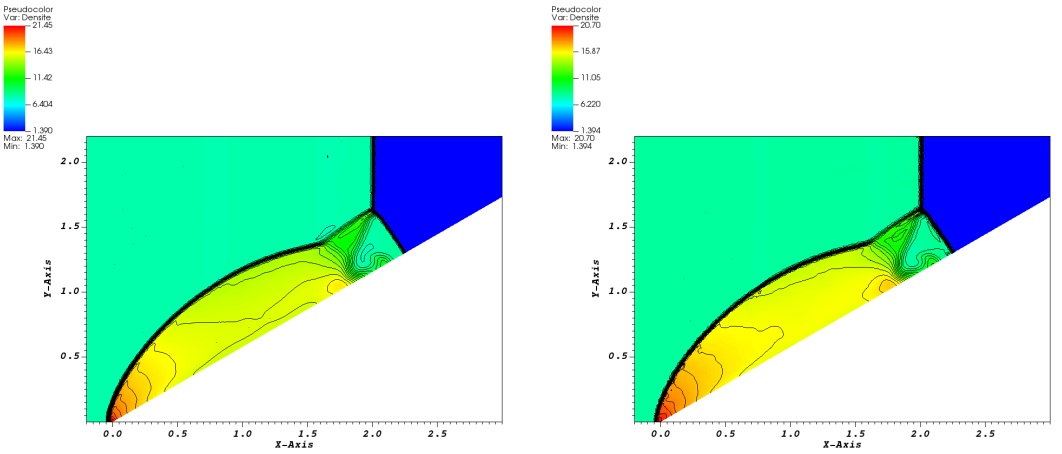

Figure 11. Density for third-order schemes with induced limitation (71)(79) (65)(80) (66) for the double Mach reflection problem with composite schemes (Rusanov (Left) / Roe (Right)) on 40201 unstructured triangles.

upper shock at $x_1 = 2$ is well resolved (without oscillation and at the right position). Here, we do not observe the "jet" near the oblic reflection axis that may occur with standard edge flux schemes (carbuncle phenomenum).

5. Conclusion

In this paper, we propose an *a priori* limitation process in the reconstruction step for the d-dimensional compressible Euler equations of gas dynamics. Taking an arbitrary order polynomial reconstruction of conservative quantities of this system, we have obtained sufficient stability condition on density $\rho > 0$ and on the internal specific energy $\epsilon > 0$. The main idea is to use the multi-variate (arbitrary order) Leibniz formula for primal specific quantities: the velocity $\mathbf{U}$, the specific total energy E. The high-order kinetic energy part is deduced (again by Leibniz formula) from high-order terms coming from those of velocity. Finally, using an explicit form of the non linear rational reconstruction $R(\mathbf{x}, \epsilon)$, it is possible to obtain a direct limitation (for ϵ) that induces a limitation for the velocity reconstruction $\boldsymbol{R}(\mathbf{x}, \mathbf{U})$.

Thanks

The author want to thanks the referee for valuable remarks and comments, helping the better understanding of the document. And also thank to S. Del Pino and X. Blanc for their carefull reading of the paper.

References

[1] Abgrall, R. , *On Essentially Non-oscillatory Schemes on Unstructured Meshes: Analysis and Implementation* , J. Comput. Phys.,114:45-58, 1994.

[2] Barth, T. and Jespersen, D., *The design and application of upwind schemes on unstructured meshes*, American Institute of Aeronautics and Astronautics, pp 366, 1989.

[3] Bernard-Champmartin A., Hoch P., Seguin N. ,*Stabilité locale et montée en ordre pour la reconstruction de quantités volumes finis sur maillages coniques non-structurés en dimension 2*, Research report, CEA, CEA/DAM/DIF, Bruyères-le-Châtel, France; Univ-Rennes1; Université Paris 6, March; 2020. https://hal.archives-ouvertes.fr/hal-02497832.

[4] Clain, S. and Diot, S. and Loubère, R., *A high-order finite volume method for systems of conservation laws Multi-dimensional Optimal Order Detection (MOOD)*, J. Comput. Phys., 230:4028-4050, 2011.

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

[6] Després, B., *Quadratic stability of flux limiters*, ESAIM: M2AN, 57:395-422, 2023.

[7] Dumbser, M., Munz, CD.,*Building Blocks for Arbitrary High Order Discontinuous Galerkin Schemes*, J Sci Comput., 27, 215–230, 2006.

[8] Godlewski, E. and Raviart, P.A. , *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, Applied Mathematical Sciences 118, Springer, 1996.

[9] Goodman, J. and LeVeque, R.J., *On the accuracy of stable schemes for 2D scalar conservation laws.*, Math. Comp. 45, 171, 15-21, 1985.

[10] Guermond, J.-L. and Maier, M. and Popov, B. and Tomas, I., *Second-order invariant domain preserving approximation of the compressible Navier-Stokes equations*, Comput. Methods Applied Mech. Engin., 375, 2021.

[11] Harten, A., *High resolution schemes for hyperbolic conservation laws*, J. Comput. Phys., 49 (2): 357-393, 1983.

[12] Harten A., and Osher S., *Uniformly high-order accurate non-oscillatory schemes I*, SIAM J. Numer. Anal., 24 (1987), no. 2, 279-309.

[13] Harten, A. and Engquist, B. and Osher, S. and Chakravarthy, S.R. , *Uniformly High Order Accurate Essentially Non-oscillatory Schemes, III*, J. Comput. Phys., vol 131, 3-47, 1997.

[14] Hoch, P., *An Arbitrary Lagrangian-Eulerian strategy to solve compressible fluid flows*, HAL, URL:https://hal.science/hal-00366858, 2009.

[15] Hoch, P., *Nodal extension of Approximate Riemann Solvers and nonlinear high order reconstruction for finite volume method on unstructured polygonal and conical meshes: the homogeneous case*, https://hal.archives-ouvertes.fr/hal-03585115/, HAL, 2022.

[16] Hoch, P. and Labourasse, E., *A frame invariant and maximum principle enforcing second-order extension for cell-centered ALE schemes based on local convex hull preservation*, Int. J. Numer. Meth. Fluids 2014; 76:1043–1063.

[17] Kuzmin, D. and Löhner, R. and Turek, S., *Flux-corrected transport. Principles, algorithms and applications*, Springer, 2005.

[18] Lax, P.D. and Liu, X-D, *Solution of Two-Dimensional Riemann Problems of Gas Dynamics by Positive Schemes*, SIAM Journal on Scientific Computing, (19), 319-340, 1998.

[19] Liska, R. and Wendroff, B., *Comparison of several difference schemes on 1D and 2D test problems for the Euler equations*,URL : http://www-troja.fjfi.cvut.cz/˜liska/CompareEuler/compare.pdf, nov 2001.

[20] Liu, X.-D. and Osher, S. and Chan, T., *Weighted essentially nonoscillatory schemes*, J. Comput. Phys., (115), pp.200–212, 1994.

[21] Luttwak, G. and Falcovitz, J., *Slope limiting for vectors: a novel vector limiting algorithm*, International Journal for Numerical Methods in Fluids 2011; 65:1365–1375.

[22] Maire, P.H. and Shu, C.W. and Vilar, F., *Positivity-preserving cell-centered Lagrangian schemes for multi-material compressible flow problems: Part II the two-dimensional case*, J. Comput. Phys., 312:416-442, 2016.

[23] Park, J.S. and Yoon, S.H. and Kim, C., *Multi-dimensional limiting process for hyperbolic conservation laws on unstructured grids*, J. Comput. Phys., (229),pp 788-812, 2010.

[24] Shu, C.-W., *Bound-preserving high order finite volume schemes for conservation laws and convection-diffusion equations*, in Finite volumes for complex applications VIII, vol. 199 of Springer, 2017, pp. 3-14.

[25] Shu, C.-W. and Osher, S., *Efficient implementation of essentially nonoscillatory shock-capturing schemes*, J. Comput. Phys., (77), pp 439-471, 1988.

[26] Shu, C.-W and Osher, S., *Efficient implementation of essentially non-oscillatory shock-capturing schemes, II.*, Journal of Computational Physics, 83:32–78, 1989.
[27] Sweby, P.K., *High resolution schemes using flux-limiters for hyperbolic conservation laws*, SIAM J. Numer. Anal., 21 (5): 995-1011, 1984.
[28] Toro, E., *Riemann Solvers and Numerical Methods for Fluid Dynamics*, A Practical Introduction, Springer, 2009.
[29] Vanderheiden, W. and Kashiwa, B., *Compatible fluxes from van Leer advection*, J. Comput. Phys., (146), pp 1-28, 1998.
[30] Zhang, X. and Shu, C.W., *Positivity-preserving high order finite difference WENO schemes for compressible Euler equations*, J. Comput. Phys., (231), pp 2245-2258, 2012.