



**HAL**  
open science

# A high-order nonconservative approach for hyperbolic equations in fluid dynamics

Remi Abgrall, P Bacigaluppi, S Tokareva

► **To cite this version:**

Remi Abgrall, P Bacigaluppi, S Tokareva. A high-order nonconservative approach for hyperbolic equations in fluid dynamics. 2017. hal-01476636v2

**HAL Id: hal-01476636**

**<https://hal.science/hal-01476636v2>**

Preprint submitted on 21 Mar 2018

**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.

# A high-order nonconservative approach for hyperbolic equations in fluid dynamics

R. Abgrall<sup>a</sup>, P. Bacigaluppi<sup>a</sup>, S. Tokareva<sup>a</sup>

<sup>a</sup>*Institute of Mathematics, University of Zurich, Switzerland*

---

## Abstract

It is well known, thanks to Lax-Wendroff theorem, that the local conservation of a numerical scheme for a conservative hyperbolic system is a simple and systematic way to guarantee that, if stable, a scheme will provide a sequence of solutions that will converge to a weak solution of the continuous problem. In [1], it is shown that a nonconservative scheme will not provide a good solution. The question of using, nevertheless, a nonconservative formulation of the system and getting the correct solution has been a long-standing debate. In this paper, we show how to get a relevant weak solution from a pressure-based formulation of the Euler equations of fluid mechanics. This is useful when dealing with nonlinear equations of state because it is easier to compute the internal energy from the pressure than the opposite. This makes it possible to get oscillation-free solutions, contrarily to classical conservative methods. An extension to multiphase flows is also discussed.

*Keywords:* Nonconservative formulation, residual distribution, conservation, fluid dynamics, Euler equations, multiphase flow systems

---

## 1. Introduction

According to the Lax-Wendroff theorem, it is well known that, when considering the numerical approximation of a system of hyperbolic PDEs written in conservative form, the numerical scheme must be written in conservation form, too. It is also known that, for a sequence of meshes with characteristic sizes tending to zero, if a sequence of solutions remains bounded and if its subsequence converges in some norm in  $L^p$ ,  $p \geq 1$ , then the limit solution is the weak solution of the original PDE. Moreover, if the scheme satisfies a discrete entropy inequality, then the limit solution will automatically satisfy an entropy inequality. If conservation is lost, then there is no hope to get any meaningful solution, see [1] for the analysis.

However, for engineering purposes, the conservative formulation of the behavior of a mechanical system is not necessarily the best one. Consider for example the Euler equations of fluid dynamics. The system of PDEs is

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ E \end{pmatrix} + \operatorname{div} \begin{pmatrix} \rho \mathbf{u} \\ \rho \mathbf{u} \otimes \mathbf{u} + p \operatorname{Id} \\ (E + p) \mathbf{u} \end{pmatrix} = 0, \quad (1)$$

supplemented by initial and boundary conditions. As usual,  $\rho$  stands for the density,  $\mathbf{u}$  for the velocity, and the total energy is

$$E = e + \frac{1}{2} \rho \mathbf{u}^2.$$

The pressure  $p$  is related to these variables via an equation of state (EOS):

$$p = p(\rho, e) = p\left(\rho, E - \frac{1}{2} \rho \mathbf{u}^2\right). \quad (2)$$

---

*Email addresses:* [remi.abgrall@math.uzh.ch](mailto:remi.abgrall@math.uzh.ch) (R. Abgrall), [paola.bacigaluppi@math.uzh.ch](mailto:paola.bacigaluppi@math.uzh.ch) (P. Bacigaluppi), [svetlana.tokareva@math.uzh.ch](mailto:svetlana.tokareva@math.uzh.ch) (S. Tokareva)

The system (1) is hyperbolic if  $\kappa := \frac{\partial p}{\partial e} > 0$  since the speed of sound  $c$  is defined by

$$c^2 = \kappa h, \quad h = \frac{e + p}{\rho}.$$

However, for engineering purposes, the relevant variables are not the conserved ones but rather the primitive ones, namely density, velocity and internal energy or pressure. When the solution is smooth, system (1) can be equivalently written as:

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ e \end{pmatrix} + \begin{pmatrix} \operatorname{div} \rho \mathbf{u} \\ \operatorname{div} (\rho \mathbf{u} \otimes \mathbf{u} + p \operatorname{Id}) \\ \mathbf{u} \cdot \nabla e + (e + p) \operatorname{div} \mathbf{u} \end{pmatrix} = 0 \quad (3)$$

or

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ p \end{pmatrix} + \begin{pmatrix} \operatorname{div} \rho \mathbf{u} \\ \operatorname{div} (\rho \mathbf{u} \otimes \mathbf{u} + p \operatorname{Id}) \\ \mathbf{u} \cdot \nabla p + \rho c^2 \operatorname{div} \mathbf{u} \end{pmatrix} = 0. \quad (4)$$

These equations are valid for smooth flows and cannot be considered for discontinuities. Nevertheless, there have been several attempts to solve the Euler equations with either formulation (3) or (4), including solutions with shocks. One example of such methods is Karni's hybrid scheme [2, 3], where formulation (4) is used along slip lines only thanks to a switch in the scheme. In any case, this method violates strict conservation.

This has been a long ongoing debate on how to, nevertheless, use formulations (3) or (4) for the numerical approximation of Euler equations valid for all kinds of flows, e. g. with complex equations of state. In the case of nonlinear equations of state, i. e. when the pressure explicitly depends on the density, the pressure obtained by the numerical scheme cannot be uniform across contact discontinuities. The reason for that behaviour is that, on one hand, the density is evaluated from the mass conservation, and, on the other hand, one evaluates the pressure via energy and density. If, in addition, we want the pressure to be constant across contact discontinuities, this puts a constraint that is in general not compatible with the updated densities, momentum and total energy, see [4] for a short discussion.

In our knowledge, in the Eulerian framework, there exists only one approach to this problem which is described in [5]. In this paper, we propose a simpler and more general framework for dealing with nonconservative formulations. We solve equations (3) or (4) in a way which is compatible with local conservation and the continuity of pressure and velocity across contact discontinuities. To achieve this, we rely on a finite volume formulation that uses residuals instead of fluxes. In a flux formulation, the unknowns are approximations of the average values of the conserved variables, and they are balanced by the sum of normal fluxes across the boundary of the control volume. This assumes that the control volume has a polygonal shape. In general, these control volumes are interpreted as cells of a dual mesh which is made of simplices. In the residual formulation, one starts by a mesh whose elements are simplices, and interprets the unknowns as approximations of the point values of the conserved variables. These unknowns, for any given degree of freedom, are then updated by a sum of the local residuals over all the elements that share this degree of freedom. Given any element  $K$ , the local conservation is recovered by requiring that the sum of the local residuals for that element is the normal flux over the boundary of  $K$  of some consistent approximation of the flux. It is easy to show that any flux formulation leads to a residual form, and the opposite is also true. However, the fluxes that are computed depend not only on the solution on both sides of the face of the control volume, see [6] for details.

The format of this paper is the following. In Section 2, we recall how one can get a residual distribution formulation for the system (3) that is equivalent to a flux formulation of (1). This enables us to get a relation on the increment of the energy that can be generalised for the residual formulation. In Section 3, we show how to use this principle, first on the energy-based formulation of the Euler equations (3) and then on the pressure-based formulation (4) for several kinds of equations of state. In Section 4, this is further generalised to multiphase flows where the phases may have very complex equations of state. Finally, we give some concluding remarks in Section 5.

## 2. From a conservative to a nonconservative formulation

### 2.1. A residual formulation of a finite volume scheme

The main advantage of the residual formulation can be understood from the one dimensional setting. Consider the problem:

$$\frac{\partial U}{\partial t} + \frac{\partial f(U)}{\partial x} = 0.$$

With standard notations, a generic finite volume scheme writes:

$$U_j^{n+1} = U_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2})$$

where  $\lambda = \Delta t/\Delta x$  and  $\hat{f}_{j+1/2}$  is the flux between the states  $U_j^n$  and  $U_{j+1}^n$ . High order accuracy in space amounts to tune the arguments of the flux, while high order in time can be reached via a strong stability preserving (SSP) scheme. To fix the conservation problem one must fix the scheme to recover a flux form, i.e to work directly with the fluxes. This is not easy from an algebraic point of view.

It is known, see for example [7] that any finite volume scheme can be rewritten in terms of a distribution of the residual. Consider for example, and for simplicity, the one dimensional case, its generalisation to any kind of control volume is straightforward, see again [7].

The residual formulation writes (in its simplest form) as

$$U_j^{n+1} = U_j^n - \lambda(\Phi_j^{j+1/2} + \Phi_j^{j-1/2}).$$

The conservation is recovered if for any element  $[x_j, x_{j+1}]$  one gets

$$\Phi_j^{j+1/2} + \Phi_{j+1}^{j+1/2} = f(U_{j+1}) - f(U_j) \quad (5)$$

for any order of accuracy. One can go from a flux formulation to a residual formulation by defining, for example:

$$\Phi_j^{j+1/2} = \hat{f}_{j+1/2} - f(U_j) \quad \text{and} \quad \Phi_{j+1}^{j+1/2} = f(U_{j+1}) - \hat{f}_{j+1/2}.$$

The local conservation is a consequence of relation (5). If we start from a nonconservative formulation in residual form, one can check the conservation if one can provide linear transformations of this residual to obtain a form satisfying (5).

### 2.2. Unsteady residual distribution formulation for the conservative case

Consider a multidimensional hyperbolic system in the form

$$\frac{\partial U}{\partial t} + \text{div } \mathbf{f}(U) = 0.$$

Recall the residual distribution approach from [8]. We start by a Runge-Kutta (RK) formulation: knowing  $U^n$ , we define:

$$\begin{aligned} U^{(0)} &= U^n, \\ \frac{U^{(1)} - U^{(0)}}{\Delta t} + \text{div } \mathbf{f}(U^{(0)}) &= 0, \\ \frac{U^{(2)} - U^{(0)}}{\Delta t} + \frac{1}{2} \left( \text{div } \mathbf{f}(U^{(0)}) + \text{div } \mathbf{f}(U^{(1)}) \right) &= 0, \\ U^{n+1} &= U^{(2)}. \end{aligned}$$

Next for  $l = 0, 1$  we rewrite each sub-step as:

$$\frac{U^{(l+1)} - U^{(0)}}{\Delta t} + \text{DIV } \mathcal{F}(U^{(l)}, U^{(0)}) = 0 \quad (6)$$

where

$$\text{DIV } \mathcal{F}(U^{(l)}, U^{(0)}) = \frac{1}{2} \left( \text{div } \mathbf{f}(U^{(0)}) + \text{div } \mathbf{f}(U^{(l)}) \right) \quad (7)$$

This will provide an approximation that is second order in time. Without loss of accuracy, we can also write it as

$$\text{DIV } \mathcal{F}(U^{(l)}, U^{(0)}) = \text{div } \mathbf{f} \left( \frac{U^{(0)} + U^{(l)}}{2} \right) \quad (8)$$

and the adapted modifications for the RK scheme.

We assume that the computational domain  $\Omega$  is covered by non-overlapping simplices  $\{K_j\}_{j \in \mathcal{J}}$ ,  $\Omega = \cup_{j \in \mathcal{J}} \{K_j\}$ . The elements  $K_j$  are segments in 1D, triangles/quadrilateral in 2D and tetrahedrons/hexahedrons in 3D. In order to simplify the notations, we denote by  $\mathbb{P}^1(K)$  the set of polynomials of degree 1 on triangles/tetrahedrons or by  $\mathbb{Q}^1(K)$  the set of polynomials of degree 1 on quadrilaterals/hexahedrons. Both guarantee a second order accurate approximation of any  $C^1$  function. We introduce the approximation space

$$V_h = \{u \in L^2(\Omega) : \text{for any } K_j, j \in \mathcal{J}, u|_{K_j} \in \mathbb{P}^1(K_j)\}.$$

We denote by  $\sigma$  the degrees of freedom, i.e. the vertices of the element  $K_j$ . Then the residual distribution approximation of (6) reads: for any degree of freedom  $\sigma$ , define first  $U^{(0)} = U^n$  and for  $l = 0, 1$ , do

$$|C_i| \frac{U_\sigma^{(l+1)} - U_\sigma^{(0)}}{\Delta t} + \sum_{K \ni \sigma} \Phi_\sigma^K(U_h^{(l)}, U_h^{(0)}) = 0 \quad (9)$$

with  $\Phi_\sigma^K(U_h^{(l)}, U_h^{(0)}) = \beta_\sigma^K(U_h^{(l)}, U_h^{(0)}) \left( \int_K \frac{U_h^{(l)} - U_h^{(0)}}{\Delta t} dx + \int_{\partial K} \mathcal{F}(U_h^{(l)}, U_h^{(0)}) \cdot \mathbf{n} d\Gamma \right)$

In (9), the parameter  $\beta_\sigma^K(U_h^{(l)}, U_h^{(0)})$  is a matrix (for Euler system of equations  $3 \times 3$  in 1D,  $4 \times 4$  in 2D, etc.), and it is constructed as a formal extension of the same construction for scalar problems that is guaranteed to have an  $L^\infty$  stability bound, see [8] for details.

To compute  $\beta_\sigma$ , we can proceed as follows.

- In (9), instead of the  $\beta$ -formulation, we first consider the Rusanov residual:

$$\Phi_\sigma^{K, Rus}(U_h^{(l)}, U_h^{(0)}) = \frac{1}{N_K} \left( \int_K \frac{U_h^{(l)} - U_h^{(0)}}{\Delta t} dx + \int_{\partial K} \mathcal{F}(U_h^{(l)}, U_h^{(0)}) \cdot \mathbf{n} d\Gamma \right) + \alpha_K \left( \frac{U_h^{(l)} + U_h^{(0)}}{2} - \widehat{U}_K \right) \quad (10)$$

with

$$\widehat{U}_K = \sum_{\sigma \in K} \frac{U_\sigma^{(l)} + U_\sigma^{(0)}}{2}.$$

In (10),  $N_K$  corresponds to the number of degrees of freedom in  $K$ .

- Then we consider the eigen-decomposition of  $A_{\mathbf{n}} = A(\widehat{U}_K)n_x + B(\widehat{U}_K)n_y$  (in 2D, while in 1D it would simply be  $A(\widehat{U}_K)$ ). The matrices  $A$  and  $B$  are the Jacobians of the  $x$ - and  $y$ - component of the flux  $\mathbf{f}$  with respect to the state  $\widehat{U}_K$ . Here, the vector  $\mathbf{n}$  is a unit vector in the direction of the velocity field when it is nonzero, or any arbitrary direction otherwise. Of course, this direction is not relevant in one dimension. The right eigenvectors of  $A_{\mathbf{n}}$  are denoted by  $\{\mathbf{r}_\xi\}$ . We denote by  $\{\ell_\xi\}$  the left eigenvectors of  $A_{\mathbf{n}}$ , so that any state  $X$  can be written as:

$$X = \sum_{\xi} \ell_\xi(X) \mathbf{r}_\xi.$$

- Then we decompose, for any degree of freedom  $\sigma$ :

$$\Phi_\sigma^{K,Rus}(U_h^{(l)}, U_h^{(0)}) = \sum_\xi \ell_\xi [\Phi_\sigma^{K,Rus}(U_h^{(l)}, U_h^{(0)})] \mathbf{r}_\xi$$

We note that for any  $\xi$ ,

$$\ell_\xi[\Phi^K] = \sum_{\sigma \in K} \ell_\xi[\Phi_\sigma^{K,Rus}(U_h^{(l)}, U_h^{(0)})],$$

where  $\Phi^K = \sum_{\sigma \in K} \Phi_\sigma^{K,Rus}$  is the total residual.

- The next step is to define  $\Phi_\sigma^{K,*}(U_h^{(l)}, U_h^{(0)})$  as:

$$\Phi_\sigma^{K,*}(U_h^{(l)}, U_h^{(0)}) = \sum_\xi \beta_\sigma^{K,*} \ell_\xi(\Phi^K) \mathbf{r}_\xi \quad (11a)$$

where:

$$\beta_\sigma^{K,*} = (1 - \Theta_\xi) \frac{\max\left(\frac{\ell_\xi(\Phi_\sigma^{K,Rus}(U_h^{(l)}, U_h^{(0)}))}{\ell_\xi[\Phi^K]}, 0\right)}{\sum_{\sigma' \in K} \max\left(\frac{\ell_\xi(\Phi_{\sigma'}^{K,Rus}(U_h^{(l)}, U_h^{(0)}))}{\ell_\xi[\Phi^K]}, 0\right)} + \Theta_\xi \ell_\xi(\Phi_\sigma^{K,Rus}(U_h^{(l)}, U_h^{(0)})) \quad (11b)$$

with

$$\Theta_\xi = \frac{|\ell_\xi(\Phi_\sigma^{K,Rus}(U_h^{(l)}, U_h^{(0)}))|}{\sum_{\sigma' \in K} |\ell_\xi(\Phi_{\sigma'}^{K,Rus}(U_h^{(l)}, U_h^{(0)}))|}. \quad (11c)$$

Note that  $\Theta_\xi \in [0, 1]$ . This guarantees that the scheme is second order in time and space and (formally) non-oscillatory, see [8, 9] for more details.

### 2.3. Discrete conservation for a nonconservative formulation

As a first exercise, let us show that any explicit scheme for (1) can be rewritten as an explicit scheme for (3) or (4). It is well known that

$$dE = de + \mathbf{u} \cdot d\mathbf{m} - \frac{1}{2} \|\mathbf{u}\|^2 d\rho,$$

where  $\mathbf{m} = \rho \mathbf{u}$  is the momentum, and thus the question is to see how one can use this relation for discrete problems.

To simplify the notations, we start from a simple Euler time stepping method, but it is important to note, that more general algorithms can be treated similarly. Setting  $U = (\rho, \mathbf{m}, E)^T$ , where  $\mathbf{m} = \rho \mathbf{u}$ , we start from the finite volume scheme

$$|C_i|(U_i^{n+1} - U_i^n) + \sum_{j \in \mathcal{V}_i} \hat{\mathbf{f}}_{ij} = 0 \quad (12)$$

where  $C_i$  is the area/volume of the control volume,  $\mathcal{V}_i$  is the set of neighbouring cells to cell  $i$ , and  $\hat{\mathbf{f}}_{ij}$  is the numerical flux between the cells  $C_i$  and  $C_j$ . In the following, since  $\hat{\mathbf{f}}_{ij}$  doesn't play any role, we write:

$$\delta \hat{\mathbf{f}} := \sum_{j \in \mathcal{V}_i} \hat{\mathbf{f}}_{ij}.$$

This finite volume scheme in the component-wise form reads:

$$|C_i|(\rho_i^{n+1} - \rho_i^n) + \delta \hat{f}_\rho = 0, \quad (13a)$$

$$|C_i|(\mathbf{m}_i^{n+1} - \mathbf{m}_i^n) + \delta \hat{\mathbf{f}}_{\mathbf{m}} = 0, \quad (13b)$$

$$|C_i|(E_i^{n+1} - E_i^n) + \delta \hat{f}_E = 0. \quad (13c)$$

Next, we introduce the  $\Delta$  operator

$$\Delta g = g^{n+1} - g^n \quad (14a)$$

and make the observation that

$$\Delta(gh) = g^{n+1}\Delta h + h^n\Delta g, \quad (14b)$$

which we can rewrite as

$$\Delta(gh) = \bar{g}\Delta h + \underline{h}\Delta g \quad (14c)$$

with

$$\bar{g} = g^{n+1}, \quad \underline{g} = g^n, \quad (14d)$$

so that

$$\overline{gh} = \bar{g} \bar{h} \quad \underline{gh} = \underline{g} \underline{h} \quad (14e)$$

From (14) we see that

$$\Delta E = \Delta e + \frac{1}{2}\Delta(\rho \mathbf{u}^2) = \Delta e + \frac{1}{2}\left(\overline{\rho \mathbf{u}} \cdot \Delta \mathbf{u} + \underline{\mathbf{u}} \cdot \Delta \mathbf{m}\right)$$

and since  $\overline{\rho \mathbf{u}} = \bar{\rho} \bar{\mathbf{u}}$  and  $\Delta \mathbf{m} = \bar{\rho} \Delta \mathbf{u} + \underline{\mathbf{u}} \Delta \rho$ , we can write

$$\Delta E = \Delta e + \frac{\bar{\mathbf{u}} + \underline{\mathbf{u}}}{2} \cdot \Delta \mathbf{m} - \frac{1}{2} \bar{\mathbf{u}} \cdot \underline{\mathbf{u}} \Delta \rho. \quad (15)$$

From this last equation we see that

$$|C_i|(e_j^{n+1} - e_j^n) + \delta \hat{f}_e = 0 \quad (16a)$$

with

$$\delta \hat{f}_e = \delta \hat{f}_E - \frac{\bar{\mathbf{u}} + \underline{\mathbf{u}}}{2} \cdot \delta \hat{\mathbf{f}}_{\mathbf{m}} + \frac{1}{2} \bar{\mathbf{u}} \cdot \underline{\mathbf{u}} \delta \hat{f}_\rho = 0, \quad (16b)$$

which coupled to (13a) and (13b) provides a consistent discretisation of (4) that is equivalent to the discretisation (13) of (1). The main problem is that (13a)-(13b)-(16) is identical to (13) with the same properties and same drawbacks, so no progress has yet been made.

Let us again consider the scheme (12), and in particular its flux term. It is known, see for example [7], that any finite volume scheme can be rewritten in terms of distribution of a residual. Consider for simplicity, the one dimensional case, its generalisation to any kind of control volume is straightforward, see again [7]. Relation (12), using standard notations, writes:

$$|C_j|(U_j^{n+1} - U_j^n) + \Delta t(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}) = 0,$$

i.e.

$$|C_j|(U_j^{n+1} - U_j^n) + \Delta t(\Phi_j^{j+1/2} + \Phi_j^{j-1/2}) = 0, \quad (17)$$

where we have set

$$\begin{aligned} \Phi_j^{j+1/2} &= \hat{f}_{j+1/2} - f(U_j) \\ \Phi_j^{j-1/2} &= f(U_j) - \hat{f}_{j-1/2}. \end{aligned}$$

Hence, on the element  $[x_j, x_{j+1}]$  we can define two residuals

$$\begin{aligned} \Phi_j^{j+1/2} &= \hat{f}_{j+1/2} - f(U_j) \\ \Phi_{j+1}^{j+1/2} &= f(U_{j+1}) - \hat{f}_{j+1/2} \end{aligned} \quad (18)$$

and we notice the conservation relation:

$$f(U_{j+1}) - f(U_j) = \Phi_j^{j+1/2} + \Phi_{j+1}^{j+1/2}. \quad (19)$$

Coming back to the system (13), written as a residual distribution scheme in the form (17), the residuals  $\Phi_\sigma^{j+1/2}$ ,  $\sigma = j, j+1$ , have 3 components, namely  $\Phi_{\sigma,\rho}^{j+1/2}$ ,  $\Phi_{\sigma,\mathbf{m}}^{j+1/2}$  and  $\Phi_{\sigma,E}^{j+1/2}$  for the mass, momentum and total energy, respectively. From (16a) and (16b) we see that we can define a residual for the internal energy as follows:

$$\Phi_{\sigma,e}^{j+1/2} = \Phi_{\sigma,E}^{j+1/2} - \frac{\bar{\mathbf{u}}_\sigma + \mathbf{u}_\sigma}{2} \cdot \Phi_{\sigma,\mathbf{m}}^{j+1/2} + \frac{1}{2} \bar{\mathbf{u}} \cdot \mathbf{u} \cdot \Phi_{\sigma,\rho}^{j+1/2}.$$

We obviously have

$$\sum_{\sigma=j}^{j+1} \Phi_{\sigma,e}^{j+1/2} = \sum_{\sigma=j}^{j+1} \Phi_{\sigma,E}^{j+1/2} - \sum_{\sigma=j}^{j+1} \frac{\bar{\mathbf{u}}_\sigma + \mathbf{u}_\sigma}{2} \cdot \Phi_{\sigma,\mathbf{m}}^{j+1/2} + \frac{1}{2} \sum_{\sigma=j}^{j+1} \bar{\mathbf{u}} \cdot \mathbf{u} \cdot \Phi_{\sigma,\rho}^{j+1/2}. \quad (20)$$

Defining the total residuals as

$$\Phi_\xi^{j+1/2} = \sum_{\sigma=j}^{j+1} \Phi_{\sigma,\xi}^{j+1/2}$$

for  $\xi = e$  and  $E$ , we see that (20) rewrites as

$$\Phi_e^{j+1/2} = \Phi_E^{j+1/2} - \sum_{\sigma=j}^{j+1} \frac{\bar{\mathbf{u}}_\sigma + \mathbf{u}_\sigma}{2} \cdot \Phi_{\sigma,\mathbf{m}}^{j+1/2} + \frac{1}{2} \sum_{\sigma=j}^{j+1} \bar{\mathbf{u}} \cdot \mathbf{u} \cdot \Phi_{\sigma,\rho}^{j+1/2}. \quad (21)$$

Relation (21) represents our target relation.

### 3. Conservative approximation of the Euler equations in primitive variables

If we are able to define the residuals  $\Phi_{j,e}^{j+1/2}$  and  $\Phi_{j+1,e}^{j+1/2}$  that satisfy (20)-(21), then using the same technique as in [10], we can show that if the conditions of the Lax-Wendroff theorem for the sequence of approximation hold, then the limit solution is a weak solution of (1). A sketch of the proof is recalled in AppendixA. In this section, we show how to achieve this requirement on system (3) using the residual formulation described in Section 2.2. To describe the principle of the method, we start with system (3) and we deal then later with system (4) in the case of non linear equations of state. The discussion will be general, however, the numerical experiments will be one dimensional in order to compare the results with exact solutions.

#### 3.1. Residual distribution formulation for unsteady problems in non-conservative form

To describe our method we first consider system (3) and start by considering the first order case, and then extend it to the second order in time and space.

##### 3.1.1. First order scheme

We start by a first order case in order to illustrate the method. The temporal scheme is a simple one step Euler scheme, and the space residual will also be first order accurate. We set  $U = (\rho, \mathbf{m}, e)^T$ , and knowing  $U^n$  at every degree of freedom,  $U^{n+1}$  is obtained by:

$$|C_\sigma|(U_\sigma^{n+1} - U_\sigma^n) + \Delta t \sum_{K \ni \sigma} \Phi_\sigma^K(U^n) = 0.$$

This scheme has the same format as the one of Section 2 that has been used to derive (21). To be specific, we calculate the Rusanov residual for the first two components of  $U$  by applying formula (10), which for the first order gives

$$\begin{aligned} \Phi_{\sigma,\rho}^{K,Rus} &= \frac{1}{N_K} \left( \int_{\partial K} \mathcal{F}_\rho(U^n) \cdot \mathbf{n} d\Gamma \right) + \alpha_K (\rho_\sigma^n - \hat{\rho}_K), \\ \Phi_{\sigma,\mathbf{m}}^{K,Rus} &= \frac{1}{N_K} \left( \int_{\partial K} \mathcal{F}_\mathbf{m}(U^n) \cdot \mathbf{n} d\Gamma \right) + \alpha_K (\mathbf{m}_\sigma^n - \hat{\mathbf{m}}_K). \end{aligned} \quad (22)$$



For the internal energy, we define the Rusanov residual as:

$$\Phi_{\sigma,e}^{K,Rus} = \frac{1}{N_K} \int_K (\mathbf{u} \cdot \nabla e + \rho h \operatorname{div} \mathbf{u})^{**} d\mathbf{x} + \alpha_K (e_\sigma^n - \hat{e}_K), \quad (23)$$

The choice of the approximation  $(\mathbf{u} \cdot \nabla e + \rho h \operatorname{div} \mathbf{u})^{**}$  in (23) is dictated only by the accuracy constraint, namely that for a smooth solution we must have

$$\int_K (\mathbf{u} \cdot \nabla e + \rho h \operatorname{div} \mathbf{u})^{**} d\mathbf{x} - \int_K (\mathbf{u} \cdot \nabla e + \rho h \operatorname{div} \mathbf{u}) d\mathbf{x} = O(h^{d+k}),$$

where  $k$  is the degree of the interpolation and  $d$  is the number of spacial dimensions, see [11] for more details.

Experimentally, one can see that under a CFL like constraint, the numerical solution converges, but the limit is not a weak solution of the problem, as expected. The reason is that the conservation relation (20) is not satisfied. Hence, from the density and the momentum equation, we can compute the velocity at time  $t_{n+1}$ , and then we modify the residual on the internal energy by setting:

$$\widetilde{\Phi}_{\sigma,e}^{K,Rus} := \Phi_{\sigma,e}^{K,Rus} + r_\sigma^K. \quad (24)$$

The perturbations  $r_\sigma^K$  are chosen such that the relation (20) is satisfied,

$$\sum_{\sigma \in K} \Phi_{\sigma,e}^{K,Rus} + \sum_{\sigma \in K} r_\sigma^K = \Phi_E^K - \sum_{\sigma \in K} \frac{\mathbf{u}_\sigma^{n+1} + \mathbf{u}_\sigma^n}{2} \cdot \Phi_{\sigma,\mathbf{m}}^{K,Rus} + \frac{1}{2} \sum_{\sigma \in K} \mathbf{u}_\sigma^n \cdot \mathbf{u}_\sigma^{n+1} \cdot \Phi_{\sigma,\rho}^{K,Rus}.$$

There is no reason to favour one degree of freedom over another, therefore we set:

$$\begin{aligned} r_\sigma^K = r^K = \frac{1}{N_K} & \left( \Phi_E^K - \sum_{\sigma \in K} \frac{\mathbf{u}_\sigma^{n+1} + \mathbf{u}_\sigma^n}{2} \cdot \Phi_{\sigma,\mathbf{m}}^{K,Rus} \right. \\ & \left. + \frac{1}{2} \sum_{\sigma \in K} \mathbf{u}_\sigma^n \cdot \mathbf{u}_\sigma^{n+1} \cdot \Phi_{\sigma,\rho}^{K,Rus} - \sum_{\sigma \in K} \Phi_{\sigma,e}^{K,Rus} \right). \end{aligned} \quad (25)$$

### 3.1.2. Second order scheme

We shall use scheme (6) to get second order of accuracy. One of its properties is that again each stage of the scheme looks like a first order Runge-Kutta scheme but with a modified residual since one needs to take into account both the time and space increments in relation (21). More specifically, we set  $U = (\rho, \mathbf{m}, e)^T$  and proceed as follows.

- For the first step: we first compute the velocity  $\mathbf{u}^{(1)}$  and relation (21) is written with

$$\bar{\mathbf{u}} = \mathbf{u}^{(1)}, \quad \underline{\mathbf{u}} = \mathbf{u}^{(0)} = \mathbf{u}^n,$$

- For the second step, knowing the states  $U^{(0)} = U^n$  and  $U^{(1)}$ , we first compute  $\mathbf{u}^{(2)} = \mathbf{u}^{n+1}$  and then

$$\bar{\mathbf{u}} = \mathbf{u}^{(2)}, \quad \underline{\mathbf{u}} = \mathbf{u}^{(1)}.$$

These modifications are mandatory to get proper conservation (see (14a)-(14b)).

This means that  $\Phi_E^K$  is now:

$$\Phi_E^K = \int_K \frac{E^{(l)} - E^{(0)}}{\Delta t} d\mathbf{x} + \int_{\partial K} \mathcal{F}_E(U^{(l)}, U^{(0)}) \cdot \mathbf{n} d\Gamma,$$

where the total energy flux  $\mathcal{F}_E$  is evaluated according to (7) or (8). The approximation  $(\mathbf{u} \cdot \nabla e + \rho h \operatorname{div} \mathbf{u})^{**}$  in (23) satisfies second order accuracy e.g. by means of arithmetic averages of the variables  $U^{(0)}$  and  $U^{(l)}$  with  $l = 0, 1$ . The correction is:

$$r_\sigma^K = r^K = \frac{1}{N_K} \left( \Phi_E^K - \sum_{\sigma \in K} \frac{\mathbf{u}_\sigma^{(l+1)} + \mathbf{u}_\sigma^{(l)}}{2} \cdot \Phi_{\sigma, \mathbf{m}}^{K, Rus} + \frac{1}{2} \sum_{\sigma \in K} \mathbf{u}_\sigma^{(l+1)} \cdot \mathbf{u}_\sigma^{(l)} \cdot \Phi_{\sigma, \rho}^{K, Rus} - \sum_{\sigma \in K} \Phi_{\sigma, e}^{K, Rus} \right). \quad (26)$$

Remark: Once again it is important to emphasize, that it is possible to use the velocity quantities at  $l+1$  to evaluate (26), since one first solves the mass and momentum equations for  $l+1$  and only then focuses on the nonconservative internal energy (or pressure) equation.

### 3.2. Ensuring conservation in the case of non-linear equations of state

In the introduction, we have mentioned the problem of evaluating the pressure across contact discontinuities for non linear equations of state. Consider an equation of state (EOS) to be

$$p = p(\rho, e)$$

and assume it to be such that

$$\kappa = \frac{\partial p}{\partial e} > 0.$$

Note that for a perfect gas  $\gamma = \kappa + 1$ , see [12] for details. Let us recall an example from [5]. We consider Cochran and Chan EOS given by

$$p = \Gamma \rho (\varepsilon - \varepsilon_0(\rho)) + p_0(\rho), \quad (27a)$$

with

$$\begin{aligned} \varepsilon_0(\rho) &= \frac{A_1}{\rho_0(E_1 - 1)} \left( \frac{\rho}{\rho_0} \right)^{E_1 - 1} - \frac{A_2}{\rho_0(E_2 - 1)} \left( \frac{\rho}{\rho_0} \right)^{E_2 - 1}, \\ p_0(\rho) &= A_1 \left( \frac{\rho}{\rho_0} \right)^{E_1} - A_2 \left( \frac{\rho}{\rho_0} \right)^{E_2}. \end{aligned} \quad (27b)$$

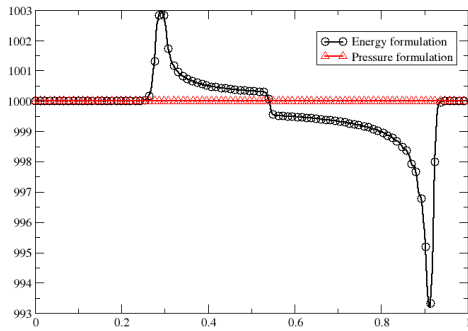
We set up a shock tube problem with the EOS parameters listed in Table 1 and consider piecewise-constant initial data given by

- $u = 1000$  [m/s],  $p = 20 \cdot 10^9$  [Pa]
- $\rho_L = 1134$ ,  $\rho_R = 500$  [kg/m<sup>3</sup>]

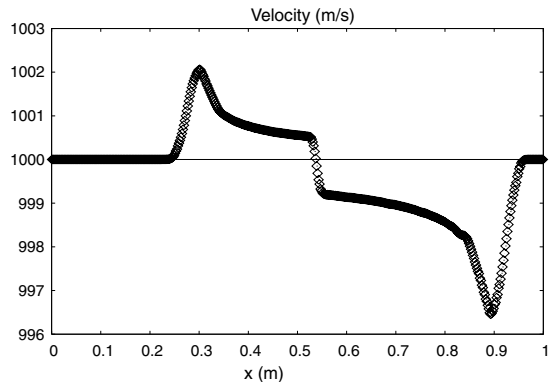
$\rho_0$ (kg/m <sup>3</sup> )	$A_1$ (GPa)	$E_1$	$A_2$ (GPa)	$E_2$	$\Gamma$
1134	0.819181	4.52969	1.50835	1.42144	1.19

Table 1: Parameters for the Cochran & Chan EOS (27).

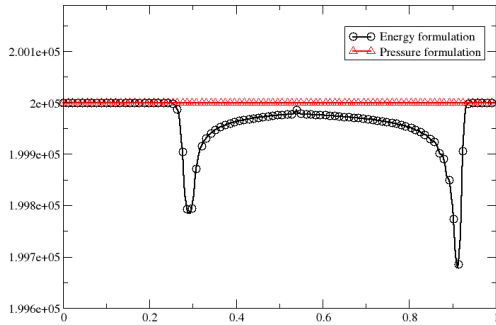
Using the conservative scheme described in Section 3 we get the results displayed in Figure 1. This results are compared to those of [5], which have been obtained with a HLLC scheme with a MUSCL extrapolation on the physical variables.



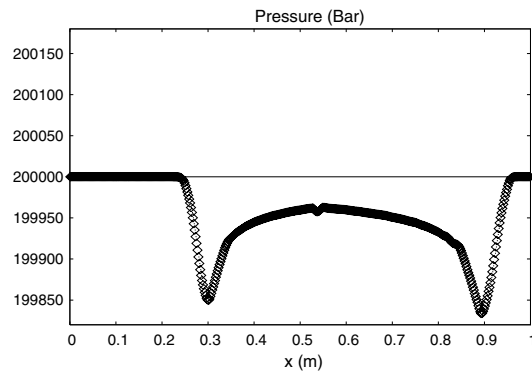
(a) velocity



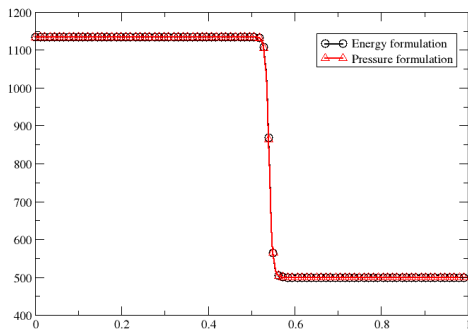
(b) velocity, taken from [5]



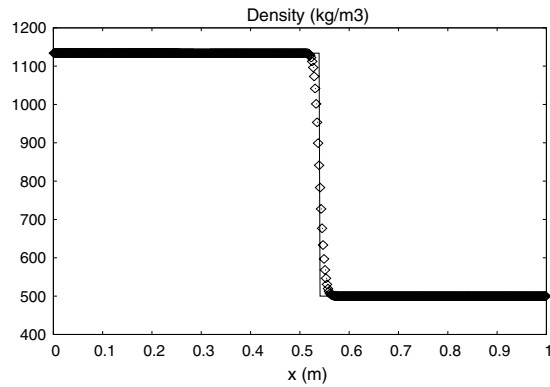
(c) pressure



(d) pressure, taken from [5]



(e) density



(f) density, taken from [5]

Figure 1: Results for a contact discontinuity for Cochran and Chan EOS. The energy and pressure formulations are considered on the left. The results are compared with those obtained for an HLLC scheme of second order (taken from [5]) on the right column.

In both cases, we see that the pressure and the velocity do not stay uniform as they should be. It is well

known that even with the Godunov scheme, the evaluation of the pressure across the contact discontinuity can be problematic, even for single fluids, see [4, 13] for the analysis. The reason is that the equation of state that relates the internal energy to the density and the pressure can be highly nonlinear. The internal energy is obtained from the total energy and the kinetic energy, and thus the pressure via:

$$e = \langle E \rangle - \frac{1}{2} \frac{\langle \rho \mathbf{u} \rangle^2}{\langle \rho \rangle}$$

where we have put *conserved variables* emphasised by using the notation  $\langle \cdot \rangle$ . The problem is that across a contact, the pressure is uniform, and in the case of a highly nonlinear EOS, there is no reason that the relation

$$\langle E \rangle - \frac{1}{2} \frac{\langle \rho \mathbf{u} \rangle^2}{\langle \rho \rangle} = e(\langle \rho \rangle, p)$$

will guarantee a uniform pressure when  $\langle \rho \rangle$  changes across a contact discontinuity. Up to our knowledge, the only Eulerian method that provides correct values is described in [5]. It is however quite complicated and tuned for Cartesian meshes.

A way to solve this issue is to start from (4) and use the same ideas as before. At the continuous level, we have

$$de = \frac{\partial e}{\partial \rho} d\rho + \frac{\partial e}{\partial p} dp, \quad (28)$$

therefore

$$dE = \frac{\partial e}{\partial \rho} d\rho + \frac{\partial e}{\partial p} dp + \mathbf{u} \cdot d(\mathbf{m}) - \frac{\mathbf{u}^2}{2} d\rho,$$

and this is the relation to mimic in the numerical scheme. Hence, we wish to satisfy

$$\begin{aligned} \sum_{\sigma \in K} \left( \frac{\widetilde{\partial e}}{\partial \rho_\sigma} \Phi_{\sigma, \rho}^K + \frac{\widetilde{\partial e}}{\partial p_\sigma} (\Phi_{\sigma, p}^K + r_\sigma^p) \right) + \sum_{\sigma \in K} \bar{\mathbf{u}}_\sigma \cdot \Phi_{\sigma, \mathbf{m}}^K - \sum_{\sigma \in K} \frac{\bar{\mathbf{u}}_\sigma^2}{2} \cdot \Phi_{\sigma, \rho}^K &= \Phi_E^K \\ &= \int_K (E^{(l)} - E^{(0)}) d\mathbf{x} + \Delta t \int_{\partial K} \mathcal{F}_E(U^{(l)}, U^{(0)}) \cdot \mathbf{n} d\Gamma, \end{aligned} \quad (29)$$

where  $\frac{\widetilde{\partial e}}{\partial \rho_\sigma}$ ,  $\frac{\widetilde{\partial e}}{\partial p_\sigma}$  are approximations of  $\frac{\partial e}{\partial \rho}$  and  $\frac{\partial e}{\partial p}$  at the degree of freedom  $\sigma$  which need to be determined and  $r_\sigma^p$  corresponds to the corrections on the pressure residuals. As before, we will assume that  $r_\sigma^p = r^p$  is independent of  $\sigma$ . Note that we only perturb the pressure residual and not the density residual. The reason is that we wish to keep an explicit scheme: first we update the density, then the velocity and finally the pressure. If we can omit this constraint, more freedom can be obtained.

Relation (29) can be rewritten as:

$$\begin{aligned} \sum_{\sigma \in K} \frac{\widetilde{\partial e}}{\partial p_\sigma} r_\sigma^p &= \Phi_E^K - \left[ \sum_{\sigma \in K} \bar{\mathbf{u}}_\sigma \cdot \Phi_{\sigma, \mathbf{m}}^K - \sum_{\sigma \in K} \frac{\bar{\mathbf{u}}_\sigma^2}{2} \cdot \Phi_{\sigma, \rho}^K + \sum_{\sigma \in K} \left( \frac{\widetilde{\partial e}}{\partial \rho_\sigma} \Phi_{\sigma, \rho}^K + \frac{\widetilde{\partial e}}{\partial p_\sigma} \Phi_{\sigma, p}^K \right) \right] \\ &:= \Delta_e. \end{aligned} \quad (30)$$

The question is how to define the terms  $\frac{\widetilde{\partial e}}{\partial \rho_\sigma}$  and  $\frac{\widetilde{\partial e}}{\partial p_\sigma}$ . Once this is done we can get  $r^p$ . One constraint is that, if initially the pressure and the velocity are uniform, we keep this property at the next time step. We will start to see how relation (29) behaves in case of a 1D flow with uniform velocities and pressures. We start by the first order scheme, and then go to the second order.

The goal is to find 'good' approximations of partial derivatives of the internal energy (28) so that the correction for the pressure vanishes for uniform pressures, without violating conservation. For this, we consider the behaviour of the internal energy increment. For any  $\sigma$ , we can write, for any  $\lambda \in \mathbb{R}$

$$e(p_\sigma^{(l+1)}, \rho_\sigma^{(l+1)}) - e(p_\sigma^{(l)}, \rho_\sigma^{(l)}) = \frac{\widetilde{\partial e}}{\partial p_{\sigma, \lambda}} (p_\sigma^{(l+1)} - p_\sigma^{(l)}) + \frac{\widetilde{\partial e}}{\partial \rho_{\sigma, \lambda}} (\rho_\sigma^{(l+1)} - \rho_\sigma^{(l)}),$$

with

$$\begin{aligned}\frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} &= \lambda \frac{e(p_\sigma^{(l+1)}, \rho_\sigma^{(l+1)}) - e(p_\sigma^{(l)}, \rho_\sigma^{(l+1)})}{p_\sigma^{(l+1)} - p_\sigma^{(l)}} + (1 - \lambda) \frac{e(p_\sigma^{(l+1)}, \rho_\sigma^{(l)}) - e(p_\sigma^{(l)}, \rho_\sigma^{(l)})}{p_\sigma^{(l+1)} - p_\sigma^{(l)}}, \\ \frac{\widetilde{\partial e}}{\partial \rho_{\sigma,\lambda}} &= \lambda \frac{e(p_\sigma^{(l)}, \rho_\sigma^{(l+1)}) - e(p_\sigma^{(l)}, \rho_\sigma^{(l)})}{\rho_\sigma^{(l+1)} - \rho_\sigma^{(l)}} + (1 - \lambda) \frac{e(p_\sigma^{(l+1)}, \rho_\sigma^{(l+1)}) - e(p_\sigma^{(l+1)}, \rho_\sigma^{(l)})}{\rho_\sigma^{(l+1)} - \rho_\sigma^{(l)}}.\end{aligned}\quad (31)$$

When  $p$  and  $u$  are uniform, then the right hand side of (29) reduces to

$$\Phi_E^K = \int_K (e^{(l)} - e^{(0)}) d\mathbf{x} + \Delta t \int_{\partial K} \frac{e^{(l)} \mathbf{u}^{(l)} + e^{(0)} \mathbf{u}^{(0)}}{2} \cdot \mathbf{n} d\Gamma.$$

If for any  $\lambda$  we have  $\frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} = 0$ , then  $\Delta_e = 0$ . Otherwise, one can find  $\lambda \in \mathbb{R}$  such that if

$$\frac{e(p_\sigma^{(l+1)}, \rho_\sigma^{(l+1)}) - e(p_\sigma^{(l)}, \rho_\sigma^{(l+1)})}{p_\sigma^{(l+1)} - p_\sigma^{(l)}} \neq \frac{e(p_\sigma^{(l+1)}, \rho_\sigma^{(l)}) - e(p_\sigma^{(l)}, \rho_\sigma^{(l)})}{p_\sigma^{(l+1)} - p_\sigma^{(l)}},$$

then

$$\sum_{\sigma \in K} \frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} r_\sigma^p = \Delta_e \quad (32a)$$

and since  $\frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} \neq 0$ , we can find  $r_\sigma^p = r^p \in \mathbb{R}$  such that

$$r^p \left( \sum_\sigma \frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} \right) = \Phi^{E,K} - \left[ \sum_{\sigma \in K} \bar{\mathbf{u}}_\sigma \cdot \Phi_{\sigma,\mathbf{m}}^K - \sum_{\sigma \in K} \frac{\bar{\mathbf{u}}_\sigma^2}{2} \Phi_{\sigma,\rho}^K + \sum_{\sigma \in K} \left( \frac{\widetilde{\partial e}}{\partial \rho_{\sigma,\lambda}} \Phi_{\sigma,\rho}^K + \frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} \Phi_{\sigma,p}^K \right) \right]. \quad (32b)$$

Clearly, we have to set  $r^p = 0$ , if  $p$  is uniform so that uniformity is preserved at the next time step.

It is important to note, that (31) is explicit in pressure in the case that the chosen equation of state guarantees that

$$\frac{\partial p}{\partial \rho} = \psi(\rho), \quad \frac{\partial p}{\partial e} = \eta(\rho). \quad (33)$$

This is for example always achieved when using the Mie-Grüneisen EOS

$$p = \Gamma \rho (\varepsilon - \varepsilon_0(\rho)) + p_0(\rho),$$

as done in the current work with the Cochran Chan EOS. In case the reader might be interested in an EOS not fulfilling (33), it is possible to simplify (31) and consider

$$\frac{\widetilde{\partial e}}{\partial p_{\sigma,\lambda}} = \frac{\partial e}{\partial p}(p_\sigma^{(l)}, \rho_\sigma^{(l)}), \quad \frac{\widetilde{\partial e}}{\partial \rho_{\sigma,\lambda}} = \frac{\partial e}{\partial \rho}(p_\sigma^{(l)}, \rho_\sigma^{(l)}). \quad (34)$$

To detect contact discontinuities and preserve the uniformity of pressure and velocity, in the one-dimensional case we proceed as follows:

- if

$$\max \left( \frac{|\max_{\sigma \in K} \bar{u}_\sigma - \min_{\sigma \in K} \bar{u}_\sigma|}{|\max_{\sigma \in K} \bar{u}_\sigma| + |\min_{\sigma \in K} \bar{u}_\sigma| + \varepsilon_1}, \frac{|\max_{\sigma \in K} p_\sigma^{(l)} - \min_{\sigma \in K} p_\sigma^{(l)}|}{|\max_{\sigma \in K} p_\sigma^{(l)}| + |\min_{\sigma \in K} p_\sigma^{(l)}| + \varepsilon_1} \right) \leq \varepsilon, \quad (35)$$

then we set  $r^p = 0$ , as this would mean that we are going across a contact wave.

- In any other case,  $r^p$  is evaluated from (32b).

### 3.3. Numerical results

System (1) has been tested on a set of very demanding benchmark problems with two different typologies of equations of state.

#### 3.3.1. Perfect gas EOS

The first test consists in a shock tube on a domain  $x = [0, 1]m$  with a diaphragm located at  $x_d = 0.5m$  at the initial conditions left and right of  $x_d$

- $\rho_l = 100 [\frac{kg}{m^3}]$ ,  $u_l = 0 [\frac{m}{s}]$ ,  $p_l = 10^9 [Pa]$ ,
- $\rho_r = 1 [\frac{kg}{m^3}]$ ,  $u_r = 0 [\frac{m}{s}]$ ,  $p_r = 10^5 [Pa]$ .

The closing equations of state for system (1) are for a perfect gas and read  $p(\rho, e) = (\gamma - 1)\rho e$ , with  $\gamma = 1.4$ .

The results are shown at  $t = 45 \mu s$  and have been obtained with a grid of  $N = 5000$  nodes and a CFL = 0.5. The choice of a high number of cells for this test case is intended to show how the proposed numerical approximation converges to the exact solution for a very strong "Sod"-like problem.

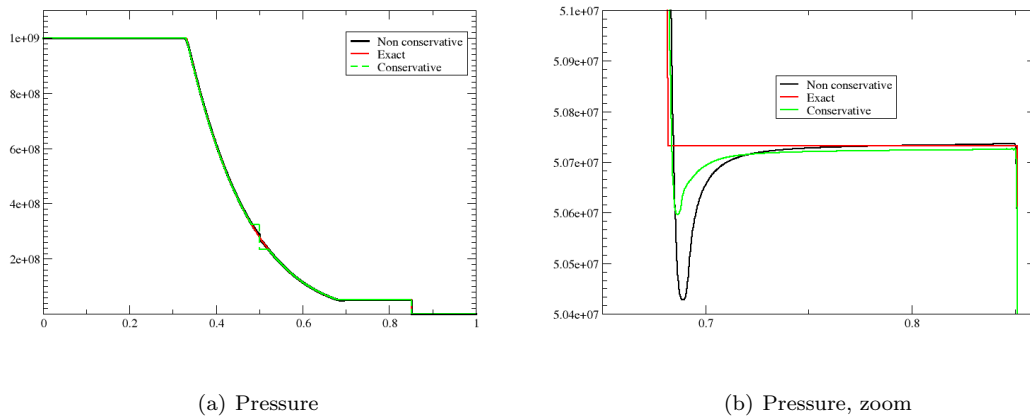
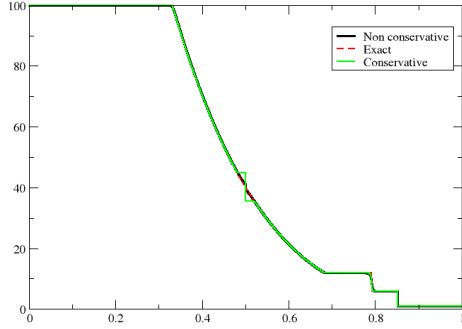
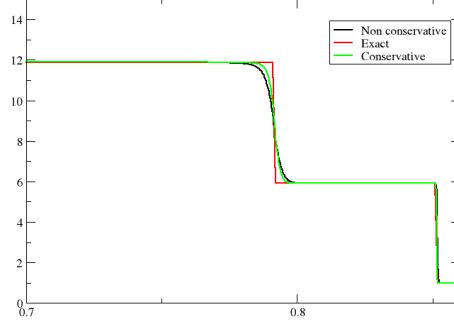


Figure 2: Strong shock tube problem with perfect gas EOS for the Euler equations. Comparison with the exact solution and both the conservative and nonconservative versions of the scheme. Display of the (a) pressure distribution along the shock tube and (b) a zoom on the pressure between  $x = [0.65, 0.86]m$  of (a).

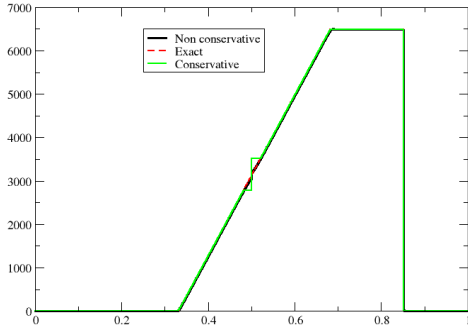
Figures 2 and 3 show the comparison of the results given by the exact solution and the conservative and nonconservative approximations presented in this work. The behaviour of these solutions show a good overlap. Both the conservative and nonconservative results are characterized by a glitch at the sonic point, which is less pronounced in the nonconservative case. The glitch itself can be easily corrected by adding some entropy correction, see e. g. [14] for a possible fix, but what matters in these results is that, this difference is also due to the fact, that the nonconservative approximation results in a slightly more diffused solution. This can be particularly seen in the zooms of the pressure (Fig. 2(c)), the density and of the velocity (Fig. 3(b) and (d)).



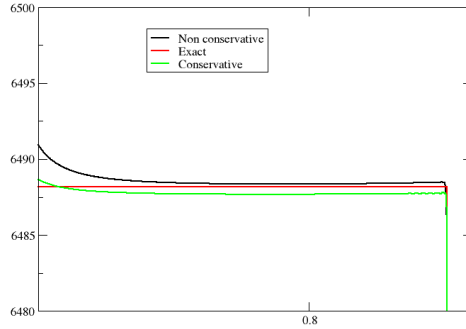
(a) Density



(b) Density, zoom



(c) Velocity



(d) Velocity, zoom

Figure 3: Strong shock tube problem with perfect gas EOS for the Euler equations. Comparison with the exact solution and both the conservative and nonconservative versions of the scheme. Display of (a) the density distribution along the shock tube, (b) a zoom on the density between  $x = [0.65, 0.86]m$ , (c) the velocity distribution and (d) a zoom on the velocity between  $x = [0.7, 0.86]m$ .

### 3.3.2. Nonlinear EOS

In order to check the quality of the approximation of the nonconservative scheme both for the pressure and energy formulations, a second testcase on a Riemann problem with a strong discontinuity has been evaluated with the choice of the Cochran-Chan EOS, as described in Section 3.2. The values for the EOS are those of Table 1 and we set<sup>1</sup>  $\varepsilon_1 = \varepsilon = 10^{-6}$  in (35). The considered domain  $x = [0, 1]m$  is split at  $x_d = 0.5m$  in a left and right state, where initially the values are set to

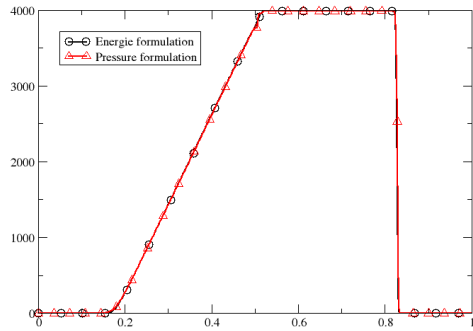
- $\rho_l = 1134 \left[ \frac{kg}{m^3} \right], u_l = 0 \left[ \frac{m}{s} \right], p_l = 2 \cdot 10^{10} [Pa]$
- $\rho_r = 120 \left[ \frac{kg}{m^3} \right], u_r = 0 \left[ \frac{m}{s} \right], p_r = 2 \cdot 10^5 [Pa]$ .

The solution displayed in Figure 4 at a final time  $50 \cdot 10^{-6}s$  show an excellent approximation of the contact discontinuity wave for both nonconservative approximations. It is interesting to notice how the

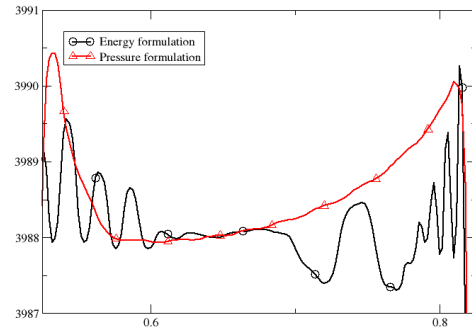
<sup>1</sup>Note that the results are not sensitive to the choice of  $\varepsilon$ .

pressure formulation is less oscillatory than the energy one. In general, for both approximations, the shock propagates at the same speed.

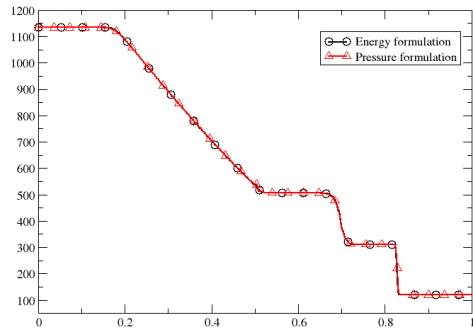




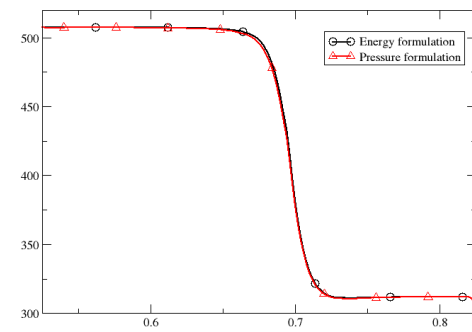
(a) Velocity



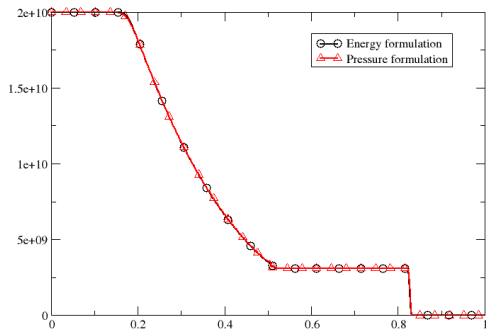
(b) Velocity, zoom



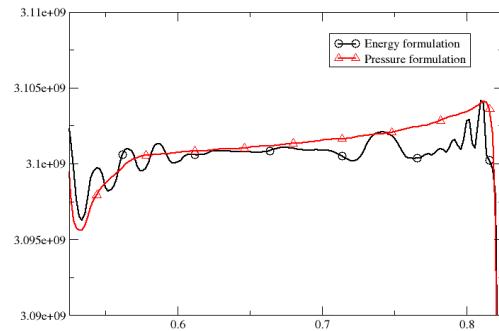
(c) Density



(d) Density, zoom



(e) Pressure



(f) Pressure, zoom

Figure 4: Riemann problem for the Euler system with Cochran-Chan EOS. Comparison of the solutions obtained from the pressure and energy formulations. On the right, zooms of the solutions around the contact are provided.

## 4. Extension to multiphase flows

### 4.1. Kapila's five equation model

Let us consider a different set of equations in the framework of compressible multiphase flows given by the five equation model of Kapila et al. [15] and shown in [16] to be the formal limit of the Baer and Nunziato model [17] when the relaxation parameters simultaneously tend to infinity though being proportional.

$$\frac{\partial \alpha_1}{\partial t} + u \cdot \nabla \alpha_1 = K \operatorname{div} u, \quad K := \frac{\rho_2 c_2^2 - \rho_1 c_1^2}{\frac{\rho_1 c_1^2}{\alpha_1} + \frac{\rho_2 c_2^2}{\alpha_2}} \quad (36a)$$

$$\frac{\partial(\alpha_1 \rho_1)}{\partial t} + \operatorname{div}(\alpha_1 \rho_1 u) = 0 \quad (36b)$$

$$\frac{\partial(\alpha_2 \rho_2)}{\partial t} + \operatorname{div}(\alpha_2 \rho_2 u) = 0 \quad (36c)$$

$$\frac{\partial(\rho u)}{\partial t} + \operatorname{div}(\rho u^2 + p) = 0 \quad (36d)$$

$$\frac{\partial E}{\partial t} + \operatorname{div}\left((E + p)u\right) = 0 \quad (36e)$$

In this two-phase system,  $\alpha_1$  is the volume fraction of phase  $\Sigma_1$ , while the volume of the second phase  $\Sigma_2$  is given by  $\alpha_2 = 1 - \alpha_1$ . The density of the phase  $\Sigma_1$  (respectively  $\Sigma_2$ ) is  $\rho_1$  (respectively  $\rho_2$ ). The mixture density is given by  $\rho = \alpha_1 \rho_1 + \alpha_2 \rho_2$ . We assume a single velocity  $u$  and a single pressure  $p$ . This allows to consider mixture quantities for the energy and momentum conservation laws, while the mass conservation law is described for each phase separately. The internal energy of each phase  $\Sigma_j$  is given by  $e_j = e_j(p, \rho_j)$  with  $j = 1, 2$ , and the mixture internal energy reads  $e = \alpha_1 e_1 + \alpha_2 e_2$ . The total energy  $E$  is the sum of the internal energy and the kinetic energy. In (36a),  $c_j$  represents the speed of sound of phase  $\Sigma_j$  and, in general, this transport equation is nonconservative. System (36) is a hyperbolic model and the mixture speed of sound  $c$  is defined via

$$\frac{1}{\rho c^2} = \frac{\alpha_1}{\rho_1 c_1^2} + \frac{\alpha_2}{\rho_2 c_2^2}. \quad (37)$$

with  $c_j$  being the speed of sound of a phase  $j$ .

The Baer and Nunziato seven equation model [17], from which (36) has been derived, considers two phases which are described by a set of a mass, momentum and energy conservation laws for each phase and an additional transport equation which links the two sets of equations in terms of the volume fractions. In case of mechanical relaxation, which means that we assume a very large interface between the two phases, we can consider the pressures of each phase to be identical. The same also holds for the velocities. Following [16], since the pressures of the two phases are equal, their Lagrangian derivatives are equal, too. Therefore, it is possible to write that the entropies  $s_1, s_2$  are constant and that  $p_1(\rho_1, s_1) = p_2(\rho_2, s_2)$ , leading to

$$c_2^2 \frac{d\rho_2}{dt} = c_1^2 \frac{d\rho_1}{dt} = \frac{\rho_1 c_1^2}{\alpha_1} \frac{d\alpha_1}{dt} + \rho_1 c_1^2 \frac{\partial u}{\partial x} = \frac{\rho_2 c_2^2}{\alpha_2} \frac{d\alpha_2}{dt} + \rho_2 c_2^2 \frac{\partial u}{\partial x}$$

This allows to reformulate the transport equation as

$$\frac{d\alpha_1}{dt} = \frac{\rho_2 c_2^2 - \rho_1 c_1^2}{\frac{\rho_1 c_1^2}{\alpha_1} + \frac{\rho_2 c_2^2}{\alpha_2}}.$$

and to reduce the original seven equations model of [17] to the one of [15].

In the following, in order to be able to rewrite the system (36) in terms of primitive variables, we need the differential relations linking the pressure and the internal energy to the densities,  $\alpha_j \rho_j$  and the volume fraction  $\alpha_1$ , since these are independent parameters. To achieve this, we start from:

$$dp_j = \kappa_j de_j + \chi_j d\rho_j, \quad \kappa_j = \left. \frac{\partial p_j}{\partial e_j} \right|_{\rho_j}, \quad \chi_j = \left. \frac{\partial p_j}{\partial \rho_j} \right|_{e_j}.$$

Since  $e = \alpha_1 e_1 + \alpha_2 e_2$ ,  $p_1 = p_2 = p$  and  $d(\alpha_j \rho_j) = \alpha_j d\rho_j + \rho_j d\alpha_j$ , we get:

$$\begin{aligned}
de &= \sum_j \alpha_j de_j + \sum_j e_j d\alpha_j \\
&= \left( \sum_j \frac{\alpha_j}{\kappa_j} \right) dp - \sum_j \frac{\chi_j}{\kappa_j} \alpha_j d\rho_j + \sum_j e_j d\alpha_j \\
&= \left( \sum_j \frac{\alpha_j}{\kappa_j} \right) dp - \sum_j \frac{\chi_j}{\kappa_j} (d(\alpha_j \rho_j) - \rho_j d\alpha_j) + \sum_j e_j d\alpha_j \\
&= \left( \sum_j \frac{\alpha_j}{\kappa_j} \right) dp - \sum_j \frac{\chi_j}{\kappa_j} d(\alpha_j \rho_j) + \sum_j \left( e_j + \rho_j \frac{\chi_j}{\kappa_j} \right) d\alpha_j.
\end{aligned}$$

We rewrite

$$\begin{aligned}
e_j + \rho_j \frac{\chi_j}{\kappa_j} &= \frac{\rho_j}{\kappa_j} \left( \kappa_j \frac{e_j + p}{\rho_j} + \chi_j \right) - p = \frac{\rho_j c_j^2}{\kappa_j} - p, \\
dp &= \kappa de + \sum_j \frac{\partial p}{\partial(\alpha_j \rho_j)} d(\alpha_j \rho_j) + \frac{\partial p}{\partial \alpha_1} d\alpha_1
\end{aligned} \tag{38}$$

with

$$\begin{aligned}
\frac{1}{\kappa} &= \sum_j \frac{\alpha_j}{\kappa_j} \\
\frac{\partial p}{\partial(\alpha_j \rho_j)} &= \chi_j \frac{\kappa}{\kappa_j} \\
\frac{\partial p}{\partial \alpha_1} &= -\kappa \left( \frac{\rho_1 c_1^2}{\kappa_1} - \frac{\rho_2 c_2^2}{\kappa_2} + e_1 - e_2 \right)
\end{aligned} \tag{39}$$

## 4.2. Numerical results

Similarly as in the previous section, we test the system (36) on two different test cases with different equations of state and physical properties.

### 4.2.1. Stiffened Gas EOS

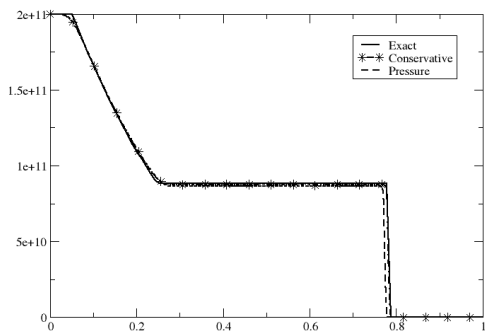
To test the capabilities of the nonconservative approximation in the context of compressible multiphase flows, a very severe benchmark problem with high differences in the pressure along a shock tube for epoxy and spinel has been taken from literature [18–21].

The considered Riemann problem has a domain of  $1m$  length and a discontinuity at  $x_d = 0.6m$ . The parameters for each phase are shown in Table 2 and we use the stiffened gas EOS for both phases, which is part of the Mie Grüneisen EOS family and reads  $p = (\gamma - 1)\rho_j e_j - \gamma p_\infty$ . A mixture of epoxy and spinel is set up, with  $u = u_1 = u_2 = 0 \left[ \frac{m}{s} \right]$  both on the left and on the right of the discontinuity, while on the left the pressure is set to  $p_L = p_{1L} = p_{2L} = 2 \cdot 10^{11} [Pa]$  and on the right to  $p_R = p_{1R} = p_{2R} = 1 \cdot 10^5 [Pa]$ .

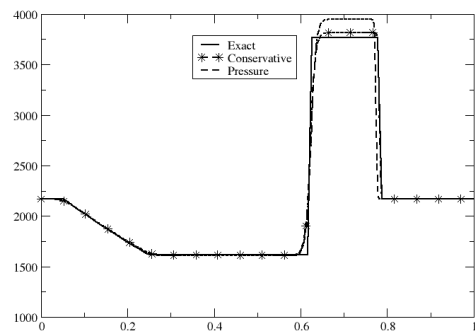
Phase	Fluid	$\alpha_1$	$\rho \left[ \frac{kg}{m^3} \right]$	$\gamma$	$P_\infty [Pa]$
1	Epoxy	0.5954	1185	2.43	$5.3 \cdot 10^9$
2	Spinel	0.4046	3622	1.62	$141 \cdot 10^9$

Table 2: Initial fluid properties and parameters for the test case.

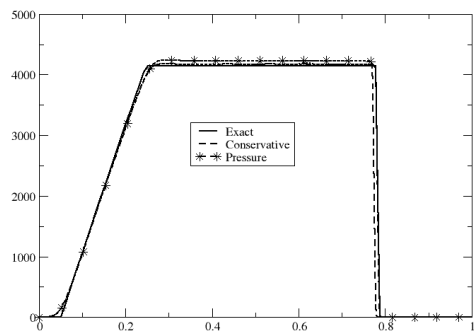
The final time is  $t = 29 \mu s$  and the CFL number is set to 0.5. The results are displayed in Figures 5 and 6.



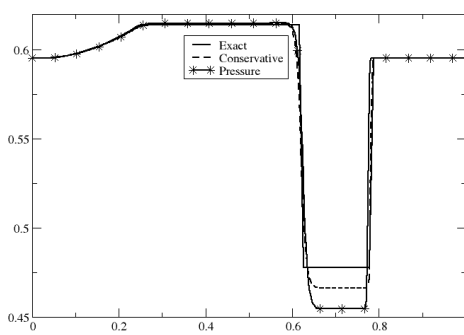
(a) Pressure



(b) Density

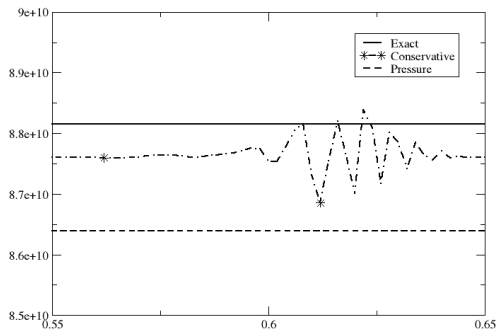


(c) Velocity

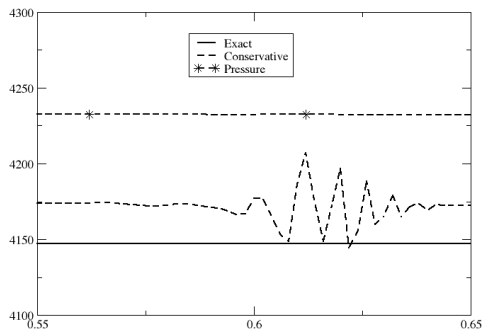


(d)  $\alpha_1$

Figure 5: Comparison between the exact solution and numerical solution given by the classical conservative total energy and nonconservative pressure formulation for the stiffened gas EOS.



(a) Pressure



(b) Velocity

Figure 6: Comparison between the exact solution and the numerical solution given by the classical conservative total energy and nonconservative pressure formulation for the stiffened gas EOS. Zoom of the region of the contact discontinuity.

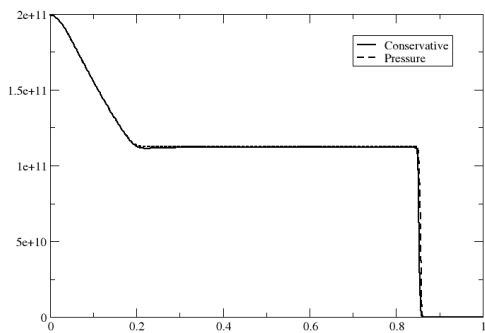
The obtained solutions show an excellent contact discontinuity approximation. The difference in the plateaus is caused by the nonconservative form of the systems, which affects the numerical dissipation, while the rarefaction waves are identical. Note however that the complexity of the pressure formulation is lower: the pressure is a primary parameter, and it doesn't need to be computed from the internal energy. In other words, there is no need to get  $p$  knowing the mass fraction and the densities via the formula

$$\varepsilon = Y_1 \varepsilon_1(\rho_1, p) + Y_2 \varepsilon_2(\rho_2, p).$$

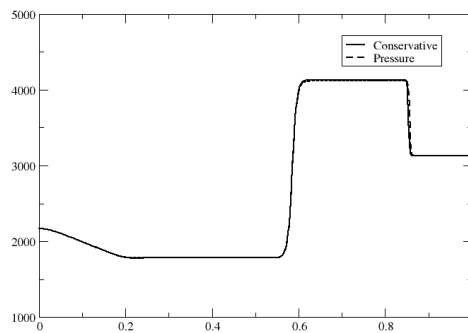
The inversion of the latter relation can be costly for nonlinear EOS. Our goal is certainly not to get the best possible solution in this context, and there exist methods that provide better results on this kind of problems, see e. g. [22, 23]. Our goal is only to show the versatility of our approach.

#### 4.2.2. Mixed stiffened gas and Cochran-Chan EOS

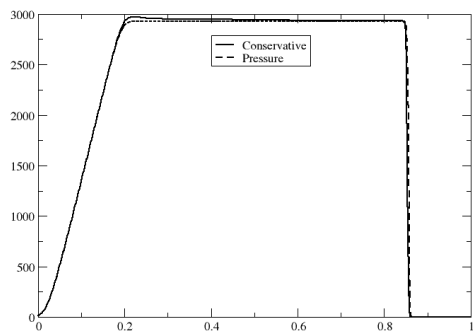
The following test case is a more challenging variation of the previous test case of section 4.2.1. Let us assume the first phase to be described by a stiffened gas EOS with parameters  $\gamma_1 = 2.43$ ,  $p_\infty = 5.3 \cdot 10^9$ , and the other one by the Cochran-Chan EOS with the parameters given in Table 1. We initially set the same conditions as in section 4.2.1 and change only the volume fractions of the involved phases:  $\alpha_l = 0.5954$  and  $\alpha_r = 0.2$ . The results are displayed in Figures 7 and 8.



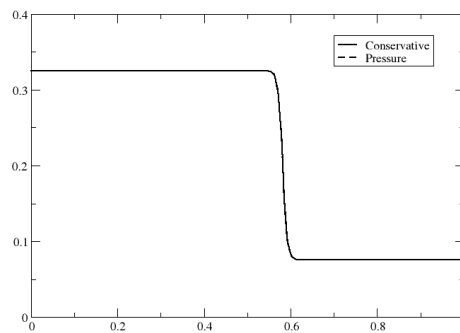
(a) Pressure



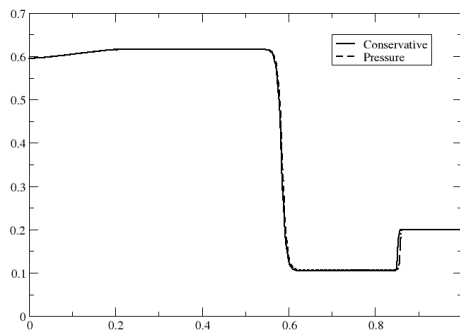
(b) Density



(c) Velocity



(d)  $Y_1$



(e)  $\alpha_1$

Figure 7: Comparison between the exact solution and the numerical solution given by the classical conservative total energy and nonconservative pressure formulation for a mixture of stiffened gas and Cochran-Chan EOS.

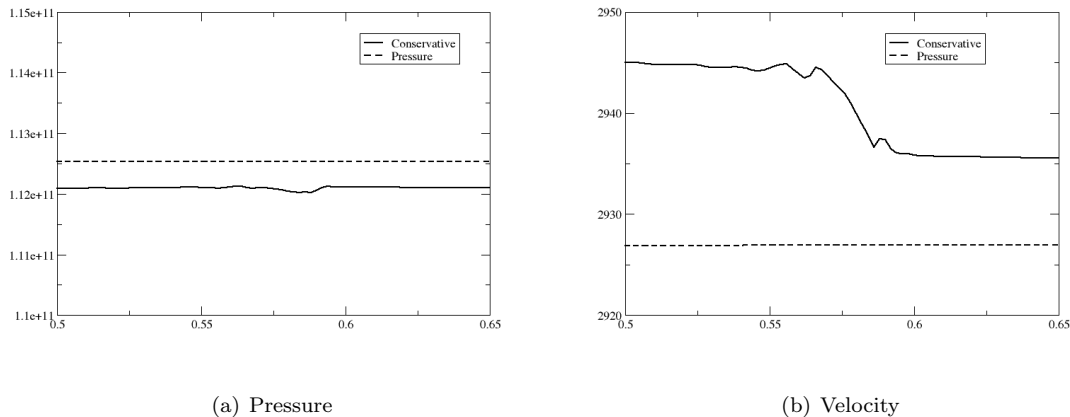


Figure 8: Comparison between the exact solution and the numerical solution given by the classical conservative total energy and nonconservative pressure formulation for a mixture of stiffened gas and Cochran-Chan EOS. Zoom of the contact discontinuity region.

The obtained solutions show again an excellent contact discontinuity approximation without any oscillations. The fact that the plateaus are on different levels is not unexpected, as pointed out before in 4.2.1.

## 5. Conclusions

In this paper, we have described a technique that enables the usage of a nonconservative formulation which nevertheless guarantees the conservation of the involved quantities. This method uses a residual distribution discretization with simple additional conditions which are able to guarantee the convergence to the correct weak solution. The emphasis is put on nonlinear equations of states where the pressure depends nonlinearly on the density. The presented approximation is then generalized to a multiphase system. The numerical tests have been done in one dimension in order to compare the results to the exact solutions. Extension to multidimensional problems can be done following the lines of [8] for the second order. Extension to higher order of accuracy could easily be obtained using the ideas of [24, 25], where the Runge-Kutta type timestepping scheme that we use here can be reinterpreted as a particular version of a deferred correction method. In this case, the relation (21) stays the same provided the residuals are defined according to [24] and the residual on the internal energy behaves like  $O(h^{k+d})$ , where  $k$  is the expected order and  $d$  the spatial dimension. We emphasize that the algebra remains identical. This will be demonstrated in a forthcoming paper.

## Acknowledgements.

P. B. and S. T. have been funded by SNSF project # 200021\_153604. R. A. has been funded in part by the same grant.

## Appendix A. A remark on the convergence to a weak solution

There are two ways of showing this. We provide the main idea for the system (3). Assuming that we have a scheme for this system that satisfies (21), and to simplify the derivations, we shall use the first order time scheme. Then one can define a residual for the total energy by simply setting

$$\Phi_{\sigma,E}^K := \Phi_{\sigma,e}^K + \frac{\bar{\mathbf{u}}_\sigma + \mathbf{u}_\sigma}{2} \cdot \Phi_{\sigma,m}^K - \frac{1}{2} \bar{\mathbf{u}}_\sigma \cdot \underline{\mathbf{u}}_\sigma \cdot \Phi_{\sigma,\rho}^K.$$

By construction we have

$$\sum_{\sigma \in K} \Phi_{\sigma,E}^K = \Phi_E^K.$$

This shows that the sequence of solutions will converge to a weak solution. Using the results of [6], one can compute numerical fluxes for the density, momentum and total energy, so that local conservation is guaranteed. It is also simple to extend the proof of the Lax-Wendroff type theorem of [10] using the same conditions.

## References

- [1] T. Y. Hou, P. G. Le Floch, Why nonconservative schemes converge to wrong solutions: Error analysis., *Math. Comput.* 62 (206) (1994) 497–530. doi:10.2307/2153520.
- [2] S. Karni, Multicomponent flow calculations by a consistent primitive algorithm., *J. Comput. Phys.* 112 (1) (1994) 31–43. doi:10.1006/jcph.1994.1080.
- [3] S. Karni, Hybrid multifluid algorithms., *SIAM J. Sci. Comput.* 17 (5) (1996) 1019–1039. doi:10.1137/S106482759528003X.
- [4] R. Abgrall, Handbook on Numerical Methods for Hyperbolic Problems, Applied and Modern Issues, Vol. XVIII of Handbook on Numerical Analysis, Springer Verlag, 2017, Ch. Some failures of Riemann solvers, pp. 351–360.
- [5] R. Saurel, E. Franquet, E. Daniel, O. Le Metayer, A relaxation-projection method for compressible flows. I: The numerical equation of state for the Euler equations., *J. Comput. Phys.* 223 (2) (2007) 822–845. doi:10.1016/j.jcp.2006.10.004.
- [6] R. Abgrall, On a class of high order schemes for hyperbolic problems, in: S. Y. Jang, Y. R. Kim, D.-W. Lee, I. Yie (Eds.), Proceedings of the international Conference of Mathematicians, Seoul 2014, Vol. IV, invited lectures, Seoul, 2015, pp. 699–726.
- [7] R. Abgrall, Toward the ultimate conservative scheme: Following the quest., *J. Comput. Phys.* 167 (2) (2001) 277–315. doi:10.1006/jcph.2000.6672.
- [8] M. Ricchiuto, R. Abgrall, Explicit Runge-Kutta residual distribution schemes for time dependent problems: second order case., *J. Comput. Phys.* 229 (16) (2010) 5653–5691. doi:10.1016/j.jcp.2010.04.002.
- [9] R. Abgrall, Essentially non-oscillatory residual distribution schemes for hyperbolic problems., *J. Comput. Phys.* 214 (2) (2006) 773–808. doi:10.1016/j.jcp.2005.10.034.
- [10] R. Abgrall, P. L. Roe, High-order fluctuation schemes on triangular meshes., *J. Sci. Comput.* 19 (1-3) (2003) 3–36. doi:10.1023/A:1025335421202.
- [11] R. Abgrall, A. Larat, M. Ricchiuto, Construction of very high order residual distribution schemes for steady inviscid flow problems on hybrid unstructured meshes., *J. Comput. Phys.* 230 (11) (2011) 4103–4136. doi:10.1016/j.jcp.2010.07.035.
- [12] R. Menikoff, B. Plohr, The Riemann problem for fluid flow of real materials, *Rev. Mod. Phys.* 61 (1).
- [13] E. F. Toro, Anomalies of conservative methods: analysis, numerical evidence and possible cures, *Comput. Fluid Dyn. J.* 11 (2) (2002) 128–143.
- [14] K. Sermeus, H. Deconinck, An entropy fix for multi-dimensional upwind residual distribution schemes, *Computers and Fluids* 34 (4-5) (2005) 617–640.
- [15] A. Kapila, R. Menikoff, J. Bdzil, S. Son, D. Stewart, Two-phase modeling of deflagration-to-detonation transition in granular materials: reduced equations., *Phys. Fluids* 13 (10) (2001) 23. doi:10.1063/1.1398042.
- [16] A. Murrone, H. Guillard, A five equation reduced model for compressible two-phase flow problems., *J. Comput. Phys.* 202 (2) (2005) 664–698. doi:10.1016/j.jcp.2004.07.019.
- [17] M. Baer, J. Nunziato, A two-phase mixture theory for the deflagration-to-detonation transition (DDT) in reactive granular materials., *Int. J. Multiphase Flow* 12 (1986) 861–889. doi:10.1016/0301-9322(86)90033-9.
- [18] S. Marsh, LASL Shock Hugoniot Data, Los Alamos Scientific Laboratory Series on Dynamic Material Properties, Vol 5, University of California Press, 1980.
- [19] F. Petitpas, E. Franquet, R. Saurel, O. L. Metayer, A relaxation-projection method for compressible flows. part ii: Artificial heat exchanges for multiphase shocks, *Journal of Computational Physics* 225 (2) (2007) 2214–2248.
- [20] R. Saurel, O. L. Métayer, J. Massoni, S. Gavrilyuk, Shock jump relations for multiphase mixtures with stiff mechanical relaxation, *Shock Waves* 16 (3) (2007) 209–232. URL <http://dx.doi.org/10.1007/s00193-006-0065-7>
- [21] F. Petitpas, R. Saurel, E. Franquet, A. Chinnayya, Modelling detonation waves in condensed energetic materials: multiphase cj conditions and multidimensional computations, *Shock Waves* 19 (5) (2009) 377–401. URL <http://dx.doi.org/10.1007/s00193-009-0217-7>
- [22] S. Schoch, N. Nikiforakis, B. Lee, R. Saurel, Multi-phase simulation of ammonium nitrate emulsion detonations., *Combustion and Flame* 160 (9) (2013) 1883–1899.
- [23] R. Abgrall, H. Kumar, Numerical approximation of a compressible multiphase system., *Commun. Comput. Phys.* 15 (5) (2014) 1237–1265.
- [24] R. Abgrall, High order schemes for hyperbolic problems using continuous finite elements representation and avoiding mass matrices., *Journal of Scientific Computing* in press, <https://hal.archives-ouvertes.fr/hal-01445543v2>.
- [25] R. Abgrall, P. Bacigaluppi, S. Tokareva, High-order residual distribution schemes for the multidimensional unsteady euler equations, in preparation.