A modified sensitivity equation method for the Euler equations in presence of shocks

Camilla Fiorini, Christophe Chalons, Régis Duvigneau

A modified sensitivity equation method for the Euler equations in presence of shocks

Camilla Fiorini[1], Christophe Chalons[2], and Régis Duvigneau[3]

[1]LJLL, Sorbonne Université, 4 place Jussieu, 75005 Paris, France.
[2]LMV, Université de Versailles St-Quentin-en-Yvelines, 45 avenue des États-Unis, 78000 Versailles, France.
[3]Université Côte d'Azur, Inria, CNRS, LJAD, 2004 route des lucioles, 06902 Sophia-Antipolis, France.

Abstract

The Continuous Sensitivity Equation (CSE) method allows to quantify how changes in the input of a Partial Differential Equation (PDE) model affect the outputs, by solving additional PDEs obtained by differentiating the model. However, this method cannot be used directly in the framework of hyperbolic PDE systems with discontinuous solution, because it yields Dirac delta functions in the sensitivity solution at the location of state discontinuities. This difficulty is well known from theoretical viewpoint, but only a few works can be found in the literature regarding the possible numerical treatment. Therefore, we investigate in this study how classical numerical schemes for compressible Euler equations can be modified to account for shocks when computing the sensitivity solution. In particular, we propose the introduction of a source term, that allows to remove the spikes associated to the Dirac delta functions in the numerical solution. Numerical studies exhibit a strong impact of the numerical diffusion on the accuracy of this strategy. Therefore, we propose an anti-diffusive numerical scheme coupled with the approximate Riemann solver of Roe for the state problem. For the sensitivity problem, two different numerical schemes are implemented and compared: one which takes into account the contact wave and another that neglects it. The effects of the numerical diffusion on the convergence of the schemes with respect to the grid are discussed. Finally, an application to uncertainty propagation is investigated and the different numerical schemes are compared.

1 Introduction

The study of how changes in the inputs of a model affect the outputs is critical for several engineering processes, such as design optimization or uncertainty quantification. This task is usually referred as *sensitivity analysis* (SA) and can be done in many ways, depending on the nature of the model, the amplitude of the perturbations considered, their deterministic or stochastic nature, etc [23]. In the present work, we consider only systems governed be Partial Differential Equations (PDEs) and we focus on the estimation of the derivative of the PDE solution with respect to an input parameter. This approach is intrinsically local and only make sense for perturbations of small amplitude, especially for highly non-linear models. The estimation of the derivative is achieved by solving a set of additional PDEs obtained by differentiating the original PDE model with respect to a single input parameter of interest. This approach is referred as the sensitivity equation method for PDE models [4], and is closely related to the linear perturbation method [23].

In the specific case of PDE models, there are two main classes of methods to compute the sensitivities: the *discretise-then-differentiate* approach and the *differentiate-then-discretise* one. Both strategies have advantages and disadvantages and both are valid and are suitable for different applications. The *differentiate-then-discretise* approach is usually considered as more flexible, because it does not require the knowledge of how the original PDE model is solved, and is qualified as *non-intrusive*. On the contrary, the *discretise-then-differentiate* approach necessitates the knowledge of the discretized

equations, but yields a set of consistent derivatives. A detailed comparison between the two for optimization problems is done in [19]. In this work, we focus on the *differentiate-then-discretise* approach, referred in this context as the Continuous Sensitivity Equation (CSE) method [4, 13, 12, 22]. Therefore, the sensitivity equations are obtained by *formally* differentiating the PDE model with respect to the parameter of interest, and then by exchanging the derivatives with respect to the parameter with the ones in space and time, yielding a new system of PDEs that should be discretized and solved numerically.

However, this method works only under certain assumptions of regularity of the state solution, which may not be verified in the hyperbolic framework. In fact, if this technique is directly applied to hyperbolic equations in case of discontinuous solutions, Dirac delta functions will appear in the sensitivity. This question has been explored in [3, 26] with a theoretical viewpoint, and more recently in [16, 17, 5] with a numerical viewpoint. While some authors have adapted their numerical strategies to handle the Dirac delta functions in the solution [9, 11, 15], others have proposed a modification of the sensitivity system to "remove" the spikes from the numerical sensitivity solution[16, 17], while maintaining the original solution in the regular regions. This is mainly motivated by the observation that the spikes can difficultly be seized numerically, even if they are physical, and do not interact well with classical numerical schemes [2, 18, 26]. We already contributed to these investigations, in particular in the context of the barotropic Euler system in Lagrangian coordinates, i.e. the $p-$system [5]. In this paper, we extend the proposed methodology to the complete compressible Euler system. Firstly, to remove the barotropic condition the additional energy equation must be considered, and this leads to the presence of a third wave, which is a contact discontinuity. Secondly, in Lagrangian coordinates the sign of the speed of the waves is known, which is not the case in Eulerian coordinates. These facts lead to a slightly more complicated design of the numerical schemes. Numerical results show that the numerical diffusion plays an important role in this framework, so particular attention is given to the design of anti-diffusive numerical schemes. Finally, another objective of this paper is to investigate, for a simple problem of uncertainty propagation, the impact of removing the spikes in the sensitivity solution and compare the different schemes designed in this context.

The paper is organised as follows: in the first sections, we introduce the state equations and derive the sensitivity equations. Then, the modification of the sensitivity equations to account for the Dirac delta functions is presented and the new sensitivity system is introduced. Next, we detail the exact resolution of the Riemann problem for the state and sensitivity in a specific case, known as the Sod shock tube problem. Some diffusive and anti-diffusive numerical schemes are illustrated: in particular, for the state a Roe Riemann solver is proposed, and two different schemes are designed for the sensitivity. Some numerical convergence tests are conducted, which exhibit grid-convergence issues of the diffusive schemes and a faster convergence for the anti-diffusive ones. Finally, an uncertainty quantification problem is defined and the results of the diffusive and anti-diffusive, with and without correction term, are compared to the results of the Monte Carlo method.

1.1 The state system

The Euler system writes:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0, \\ \partial_t(\rho E) + \partial_x(u(\rho E + p)) = 0, \end{cases} \tag{1}$$

where ρ is the density, u is the velocity, ρE the total energy per volume unit, and p the pressure. The system is closed by the following algebraic equation:

$$p = (\gamma - 1)\left(\rho E - \frac{1}{2}\rho u^2\right), \tag{2}$$

where $\gamma = 1.4$ is the heat capacity ratio. We introduce two other quantities which will be useful in the following: the total enthalpy $H = E + \frac{p}{\rho}$ and the speed of sound $c = \sqrt{(\gamma-1)(H - \frac{1}{2}u^2)}$. We can rewrite the system (1) in the vectorial form:

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \tag{3}$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho E \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(\rho E + p) \end{bmatrix} = \begin{bmatrix} w_2 \\ \frac{w_2^2}{w_1} + (\gamma - 1)\left(w_3 - \frac{1}{2}\frac{w_2^2}{w_1}\right) \\ \gamma \frac{w_2 w_3}{w_1} - \frac{(\gamma-1)}{2}\frac{w_2^3}{w_1^2} \end{bmatrix}.$$

One can also write (1) in the nonconservative form:

$$\partial_t \mathbf{U} + \mathbf{A}(\mathbf{U})\partial_x \mathbf{U} = 0, \tag{4}$$

where the Jacobian matrix $\mathbf{A}$ writes:

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{2}u^2 & (3-\gamma)u & \gamma-1 \\ \frac{\gamma-2}{2}u^3 - \frac{c^2 u}{\gamma-1} & \frac{3-2\gamma}{2}u^2 + \frac{c^2}{\gamma-1} & \gamma u \end{bmatrix},$$

its eigenvalues are $\lambda_1 = u - c$, $\lambda_2 = u$, and $\lambda_3 = u + c$ and its eigenvectors are:

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ u-c \\ H - uc \end{bmatrix}, \qquad \mathbf{r}_2 = \begin{bmatrix} 1 \\ u \\ \frac{u^2}{2} \end{bmatrix}, \qquad \mathbf{r}_3 = \begin{bmatrix} 1 \\ u+c \\ H + uc \end{bmatrix}.$$

Therefore $\mathbf{A}$ is $\mathbb{R}$-diagonalisable and the system (1) is strictly hyperbolic. At last, (3) will be supplemented with a given initial data $\mathbf{U}(x, t = 0) = \mathbf{U}_0(x)$, $\forall x \in \mathbb{R}$.

1.2 The sensitivity system

Considering only smooth solutions of (1), one can apply the Continuous Sensitivity Equation (CSE) [22, 4, 13] method which consists in differentiating (1) with respect to the parameter of interest a. One can then formally exchange the derivatives in time and space with the ones with respect to a (see [3] for the theoretical aspects) and obtain the following sensitivity system:

$$\begin{cases} \partial_t \rho_a + \partial_x(\rho u)_a = 0, \\ \partial_t(\rho u)_a + \partial_x(\rho_a u^2 + 2\rho u u_a + p_a) = 0, \\ \partial_t(\rho E)_a + \partial_x(u_a(\rho E + p) + u((\rho E)_a + p_a)) = 0, \end{cases} \tag{5}$$

which can be written in vectorial form as

$$\partial_t \mathbf{U}_a + \partial_x \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = 0, \tag{6}$$

where we used the following notation:

$$\mathbf{U}_a = \partial_a \mathbf{U} = \begin{bmatrix} \rho_a \\ (\rho u)_a \\ (\rho E)_a \end{bmatrix}, \qquad \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = \partial_a \mathbf{F}(\mathbf{U}) = \begin{bmatrix} (\rho u)_a \\ \rho_a u^2 + 2\rho u u_a + p_a \\ u_a(\rho E + p) + u((\rho E)_a + p_a) \end{bmatrix}.$$

Note that differentiating (2) one has:

$$p_a = (\gamma - 1)((\rho E)_a - \frac{1}{2}\rho_a u^2 - \rho_a u u_a)$$

which acts as a closure relation for (5). The initial data for the sensitivity is $\mathbf{U}_a(x, t = 0) = \partial_a \mathbf{U}_0(x)$.

1.3 The global system

In order to write the global system, i.e. the state and sensitivity system, in a more compact way, we introduce the following vectors:

$$\mathbf{V} = \begin{bmatrix} \mathbf{U} \\ \mathbf{U}_a \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \end{bmatrix},$$

$$\mathbf{G}(\mathbf{V}) = \begin{bmatrix} \mathbf{F}(\mathbf{U}) \\ \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) \end{bmatrix} = \begin{bmatrix} w_2 \\ \frac{w_2^2}{w_1} + (\gamma - 1)\left(w_3 - \frac{1}{2}\frac{w_2^2}{w_1}\right) \\ \gamma \frac{w_2 w_3}{w_1} - \frac{(\gamma-1)}{2}\frac{w_2^3}{w_1^2}) \\ w_5 \\ \frac{\gamma-3}{2}\frac{w_2^2 w_4}{w_1^2} - (\gamma - 3)\frac{w_2 w_5}{w_1} + (\gamma - 1)w_6 \\ \gamma \frac{w_3 w_5}{w_1} - \gamma \frac{w_2 w_3 w_4}{w_1^2} - \frac{3}{2}(\gamma - 1)\frac{w_2^2 w_5}{w_1} + (\gamma - 1)\frac{w_2^3 w_4}{w_1^3} + \gamma \frac{w_2 w_6}{w_1} \end{bmatrix}.$$

Therefore, the complete system writes:

$$\begin{cases} \partial_t \mathbf{V} + \partial_x \mathbf{G}(\mathbf{V}) = 0, \\ \mathbf{V}(x,0) = \mathbf{V}_0(x), \end{cases} \tag{7}$$

with $\mathbf{V}_0(x) = (\mathbf{U}_0(x), \partial_a \mathbf{U}_0(x))^t$. The Jacobian matrix of the complete system has the following form:

$$\frac{\partial \mathbf{G}(\mathbf{V})}{\partial \mathbf{V}} = \mathbf{M}(\mathbf{V}) = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B} & \mathbf{A} \end{bmatrix}$$

where $\mathbf{A}$ is the Jacobian matrix of the state system and $\mathbf{B}$ writes:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ (\gamma-3)uu_a & (3-\gamma)u_a & 0 \\ (\star) & (\bullet) & \gamma u_a \end{bmatrix}, \tag{8}$$

with

$$(\star) = -\frac{c^2}{\gamma-1}\frac{p_a}{p}u + \frac{3}{2}(\gamma-2)u^2u_a + \frac{c^2}{\gamma-1}\frac{u\rho_a}{\rho} - \frac{c^2}{\gamma-1}u_a + \gamma\frac{u^3\rho_a}{\rho},$$

and

$$(\bullet) = \frac{\gamma}{2}u^2\rho_a - \frac{c^2}{\gamma-1}\rho_a + \frac{6-5\gamma}{2}\frac{u^2\rho_a}{\rho} + (3-2\gamma)uu_a + 3(\gamma-1)\frac{u\rho_a}{\rho^2} + \frac{c^2}{\gamma-1}\frac{p_a}{p}.$$

The matrix $\mathbf{M}$ has three repeated eigenvalues, which are the eigenvalues of the matrix $\mathbf{A}$. More precisely, one can prove the following result.

Proposition. 1. *The global system (7) is weakly hyperbolic.*

Proof. A system of the form (7) is weakly hyperbolic if its Jacobian matrix has real eigenvalues and it is not $\mathbb{R}$-diagonalisable. We want to investigate whether or not the matrix $\mathbf{M}$ is $\mathbb{R}$-diagonalisable. A matrix is diagonalisable if and only if its minimal polynomial splits in distinct roots. Since the characteristic polynomial of the matrix $\mathbf{M}$ is the following:

$$p_M(x) = (x-\lambda_1)^2(x-\lambda_2)^2(x-\lambda_3)^2, \tag{9}$$

the minimal polynomial, in order to have distinct roots, can be at most of degree 3. Therefore, if $\mathbf{M}$ is diagonalisable, it must be:

$$(\mathbf{M}-\lambda_1 I_6)(\mathbf{M}-\lambda_2 I_6)(\mathbf{M}-\lambda_3 I_6) = 0. \tag{10}$$

Let us write (10) by blocks:

$$\begin{bmatrix} A-\lambda_1 I_3 & 0 \\ B & A-\lambda_1 I_3 \end{bmatrix}\begin{bmatrix} A-\lambda_2 I_3 & 0 \\ B & A-\lambda_2 I_3 \end{bmatrix}\begin{bmatrix} A-\lambda_3 I_3 & 0 \\ B & A-\lambda_3 I_3 \end{bmatrix} = 0 \tag{11}$$

Developing the left-hand side products one obtains the following matrix:

$$\begin{bmatrix} (A-\lambda_1 I_3)(A-\lambda_2 I_3)(A-\lambda_3 I_3) & 0 \\ (\blacksquare) & (A-\lambda_1 I_3)(A-\lambda_2 I_3)(A-\lambda_3 I_3) \end{bmatrix}, \tag{12}$$

where

$$(\blacksquare) = B(A-\lambda_2 I_3)(A-\lambda_3 I_3) + (A-\lambda_1 I_3)B(A-\lambda_3 I_3) + (A-\lambda_1 I_3)(A-\lambda_2 I_3)B.$$

The top-left and bottom-right coefficients are equal to the characteristic polynomial of $\mathbf{A}$ evaluated in $\mathbf{A}$, thus they are zero. Therefore, the matrix $\mathbf{M}$ is diagonalisable if and only if $(\blacksquare) = 0$. Let us compute the coefficient $(1,1)$ of $(\blacksquare)$:

$$(\blacksquare)_{(1,1)} = 0 + [c-u,\ 1,\ 0]\begin{bmatrix} 0 \\ (3-\gamma)u^2u_a - \frac{(\gamma-3)^2}{2}u^2u_a \\ \diamondsuit \end{bmatrix} + \\ + [c-u,\ 1,\ 0]\begin{bmatrix} (\gamma-3)uu_a \\ (\gamma-2)(3-\gamma)u^2u_a + (\gamma-1)(\star) \\ \triangle \end{bmatrix} = \\ = -\frac{3}{2}(\gamma-1)u^2u_a + (\gamma-3)cuu_a - c^2u\frac{p_a}{p} + c^2u\frac{\rho_a}{\rho} - c^2u_a + \gamma(\gamma-1)u^3\frac{\rho_a}{\rho},$$

where there is no need to specify $\diamondsuit$ and $\triangle$. There is no reason why the quantity should be always be zero. Therefore, the matrix is not diagonalisable and the complete system is not hyperbolic in general. However, as the eigenvalues are real, the system is weakly hyperbolic. □

2 Source term

The sensitivity system (5) was derived assuming that the state solution $\mathbf{U}$ is regular. However, this is not generally true for hyperbolic systems such as the one considered [3]: if the state is discontinuous, the sensitivity exhibits Dirac delta functions. As said earlier, different choices are possible: some authors have tried to adapt their numerical schemes in order to deal with the Dirac functions [9, 11, 15], some others added a correction term to the sensitivity equations [16, 17]. We decide to adopt the second strategy, as done in [5], that leads to an accurate sensitivity almost everywhere in the domain, except for the discontinuity points. The correction term that we add to the sensitivity equations has the following form:

$$\mathbf{S} = \sum_{k=1}^{N_s} \delta_k \boldsymbol{\rho}_k, \tag{13}$$

where N_s is the number of discontinuities, which can be either shocks or contact discontinuities, $\delta_k = \delta(x - x_{k,s}(t))$ is the Dirac delta function with $x_{k,s}(t)$ position of the $k-$th shock and $\boldsymbol{\rho}_k$ is the amplitude of the $k-$th correction. To compute the amplitude $\boldsymbol{\rho}_k(t)$, we start by integrating the sensitivity equations with the source term on a control volume which contains a single discontinuity travelling at speed σ_k. As the control volume goes to zero, one has:

$$\boldsymbol{\rho}_k = (\mathbf{U}^-_{a,k} - \mathbf{U}^+_{a,k})\sigma_k + \mathbf{F}_a(\mathbf{U}^+_k, \mathbf{U}^+_{a,k}) - \mathbf{F}_a(\mathbf{U}^-_k, \mathbf{U}^-_{a,k}), \tag{14}$$

where $\mathbf{U}^+_{k,a}$ (respectively $\mathbf{U}^-_{k,a}$) is the value of the sensitivity to the right (respectively left) of the $k-$th discontinuity. Then, one writes the Rankine-Hugoniot relations associated with (3)

$$-\sigma_k(\mathbf{U}^+_k - \mathbf{U}^-_k) + \mathbf{F}(\mathbf{U}^+_k) - \mathbf{F}(\mathbf{U}^-_k) = 0,$$

where $\mathbf{U}^+_k$ (respectively $\mathbf{U}^-_k$) is the value of the state to the right (respectively left) of the $k-$th discontinuity. If we differentiate these conditions with respect to a, we obtain:

$$\begin{aligned}(\mathbf{U}^-_{k,a} - \mathbf{U}^+_{k,a})\sigma_k + (\mathbf{U}^-_k - \mathbf{U}^+_k)\sigma_{k,a} + \sigma_k(\partial_x \mathbf{U}^+ -_k \partial_x \mathbf{U}^-_k)\partial_a x_{k,s}(t) = \\ = \mathbf{F}_a(\mathbf{U}^-_k, \mathbf{U}^-_{a,k}) - \mathbf{F}_a(\mathbf{U}^+_k, \mathbf{U}^+_{a,k}) + \left(\frac{\partial \mathbf{F}(\mathbf{U}^+_k)}{\partial \mathbf{U}}\partial_x \mathbf{U}^+_k - \frac{\partial \mathbf{F}(\mathbf{U}^-_k)}{\partial \mathbf{U}}\partial_x \mathbf{U}^-_k\right)\partial_a x_{k,s}(t),\end{aligned} \tag{15}$$

where $\sigma_{k,a} := \partial_a \sigma_k$. Replacing (15) into (14), one obtains the following definition of $\boldsymbol{\rho}_k$, which does not depend on the sensitivity itself:

$$\boldsymbol{\rho}_k = (\mathbf{U}^+ - \mathbf{U}^-)\sigma_{k,a} + \sigma_k(\partial_x \mathbf{U}^+ - \partial_x \mathbf{U}^-)\partial_a x_{k,s}(t) - \left(\frac{\partial \mathbf{F}(\mathbf{U}^+)}{\partial \mathbf{U}}\partial_x \mathbf{U}^+ - \frac{\partial \mathbf{F}(\mathbf{U}^-)}{\partial \mathbf{U}}\partial_x \mathbf{U}^-\right)\partial_a x_{k,s}(t). \tag{16}$$

The terms depending on $\partial_a x_{k,s}(t)$ are very difficult to estimate. However, one can remark that all these terms contain $\partial_x \mathbf{U}^\pm_k$: therefore, in the following sections, when we design first order finite volume schemes for the sensitivity we will consider the simpler expression

$$\boldsymbol{\rho}_k(t) = \sigma_{k,a}(\mathbf{U}^+_k - \mathbf{U}^-_k), \tag{17}$$

since the solution, in a first order finite volume framework, is piecewise constant on the cells and therefore $\partial_x \mathbf{U}^\pm_k = 0$. Replacing (17) into (13) gives the following definition of the source term (valid only for piecewise constant functions):

$$\mathbf{S} = \sum_{k=1}^{N_s} \sigma_{k,a}(\mathbf{U}^+_k - \mathbf{U}^-_k)\delta_k. \tag{18}$$

A special treatment, which will be detailed later, is necessary for a second or higher order discretisation, where the discrete solution is not constant within each cell and therefore $\partial_x \mathbf{U}^\pm_k \neq 0$.

The new system thus writes:

$$\begin{cases}\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0 \\ \partial_t \mathbf{U}_a + \partial_x \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = \mathbf{S}.\end{cases} \tag{19}$$

In the next section, we design a numerical scheme to approximate the solution of (19). The analytical solution for a given initial data of Riemann type is detailed in appendix A.

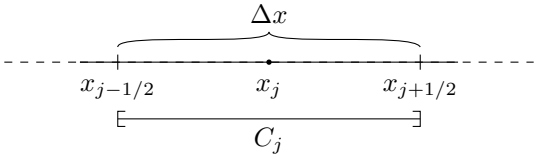


Figure 1: Spatial discretisation.

3 Numerical methods

In this section we consider the numerical approximation of (19). We derive first and second order Roe-type numerical schemes and we pay particular attention to the numerical diffusion effects induced by these approaches. Indeed and as we will see it may prevent the numerical solution from converging to the correct solution. We consider a uniform grid in space with a constant step Δx, x_j is the center of the $j-$th cell C_j, whose extrema are $x_{j-1/2}$ and $x_{j+1/2}$ (cf. Figure 1). We use an adaptive time step Δt^n, chosen according to a CFL condition, and the intermediate times are $t^{n+1} = t^n + \Delta t^n$. We indicate with $\mathbf{V}_j^n = (\mathbf{U}_j^n, \mathbf{U}_{a,j}^n)^t$ the average value of the state and the sensitivity in the cell C_j at time t^n.

We use Godunov-type schemes, which consist of two main steps: first, one solves the Riemann problem at each interface $x_{j-1/2}$ at time t^n, obtaining in this way a solution at time t^{n+1}, $\mathbf{v}(x, t^{n+1}) = (\mathbf{u}(x, t^{n+1}), \mathbf{u}_a(x, t^{n+1}))^t$; the second step is to project $\mathbf{v}(x, t^{n+1})$ in order to obtain a piecewise constant solution on the mesh. How to compute $\mathbf{v}(x, t)$ is the topic of the next subsections: first we describe an approximate solver for the state and then two for the sensitivity. Different choices for the solution of the Riemann problem lead to different numerical schemes. Then, we explain two different projection techniques: the classical one and an anti-diffusive one. Finally, we explain how to extend the schemes to higher order in space, focusing in particular on the second order.

3.1 Riemann solver for the state

First, we consider the state system, for which the classical numerical schemes can be used: in this work we used the approximate Riemann solver of Roe, because it has the property of being exact for an isolated shock and we want to be as precise as possible in the shocks. In addition, we remark that it would not be possible to use a solver with only one intermediate star state, such as HLL (Harten, Lax and van Leer [21]), because of the definition of the source term (17): two intermediate states are necessary in order to be able to compute the correction term across the contact discontinuity (cf. Figures 2-3 for the structure of different solvers).

The main idea of the Roe scheme is to replace the Jacobian matrix $\mathbf{A}(\mathbf{U})$ in (4) with a constant matrix $\mathbf{A}(\mathbf{U}_L, \mathbf{U}_R)$, obtaining in this way a linearised system, whose solution to the Riemann problem can be computed exactly. For the Euler system, a proper linearisation is provided by Roe in the original paper [28]. Furthermore, there is no need to assemble the matrix, it is sufficient to know its eigenvalues and eigenvectors, which are the following:

$$\lambda_1^{ROE} = \tilde{u} - \tilde{c}, \quad \lambda_2^{ROE} = \tilde{u}, \quad \lambda_3^{ROE} = \tilde{u} + \tilde{c},$$

$$\tilde{\mathbf{r}}_1 = \begin{pmatrix} 1 \\ \tilde{u} - \tilde{c} \\ \tilde{H} - \tilde{u}\tilde{c} \end{pmatrix}, \quad \tilde{\mathbf{r}}_2 = \begin{pmatrix} 1 \\ \tilde{u} \\ \frac{\tilde{u}^2}{2} \end{pmatrix}, \quad \tilde{\mathbf{r}}_3 = \begin{pmatrix} 1 \\ \tilde{u} + \tilde{c} \\ \tilde{H} + \tilde{u}\tilde{c} \end{pmatrix}.$$

The quantities denoted with a tilde are *Roe averaged* quantities defined as follows:

$$\tilde{u} = \frac{\sqrt{\rho_L}u_L + \sqrt{\rho_R}u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \quad \tilde{H} = \frac{\sqrt{\rho_L}H_L + \sqrt{\rho_R}H_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \quad \tilde{c} = \sqrt{(\gamma - 1)\left(\tilde{H} - \frac{1}{2}\tilde{u}^2\right)}.$$

Therefore, the Roe solver consists of four constant states ($\mathbf{U}_L$, $\mathbf{U}_L^*$, $\mathbf{U}_R^*$, and $\mathbf{U}_R$, cf. Figure 2) connected by three discontinuities travelling at speeds λ_i^{ROE}. To compute the star states $\mathbf{U}_L^*$ and $\mathbf{U}_R^*$, first we decompose the jump $\mathbf{U}_R - \mathbf{U}_L$ along the eigenvectors of the Jacobian matrix $\mathbf{A}$:

$$\Delta\mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{i=1}^{3} \alpha_i \tilde{\mathbf{r}}_i. \tag{20}$$

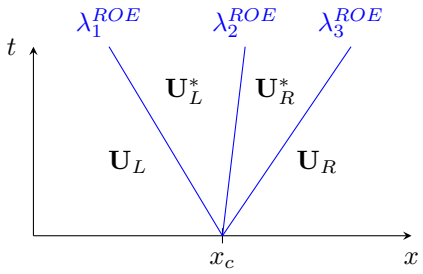


Figure 2: Structure of the Roe solver for the state.

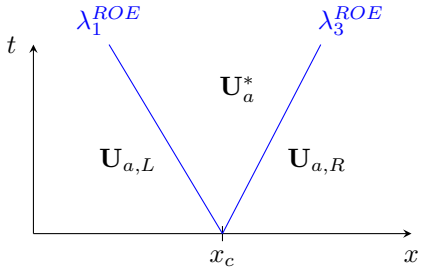


Figure 3: Structure of the HLL-type solver for the sensitivity.

The relation (20) is used to compute the coefficients α_i, then one has:

$$\mathbf{U}_L^* = \mathbf{U}_L + \alpha_1 \tilde{\mathbf{r}}_1 = \mathbf{U}_R - \alpha_2 \tilde{\mathbf{r}}_2 - \alpha_3 \tilde{\mathbf{r}}_3, \quad \mathbf{U}_R^* = \mathbf{U}_R - \alpha_3 \tilde{\mathbf{r}}_3 = \mathbf{U}_L + \alpha_1 \tilde{\mathbf{r}}_1 + \alpha_2 \tilde{\mathbf{r}}_2. \tag{21}$$

Once all the quantities $\mathbf{U}_L^*$, $\mathbf{U}_R^*$, and λ_ℓ^{ROE} are known at each interface $x_{j-1/2}$, $\mathbf{u}(x, t^{n+1})$ can be built by juxtaposition of the solutions of each Riemann problem.

It is well known that, in case of transonic rarefaction, the Roe solver provides a non-entropic solution. To overcome this problem, we implemented the entropic fix proposed in [20].

3.2 Riemann solvers for the sensitivity

For the sensitivity we propose two different strategies. Indeed and as explained in the previous section, for the state it is necessary to use a Riemann solver with two different star states, in order to be able to compute the source term across the contact discontinuity. However, for the sensitivity an HLL-type approach can be used, which gives a first strategy. Another possible strategy is to keep for the sensitivity the same structure as for the state, and therefore to have an HLLC-type scheme (Harten, Lax and van Leer Contact [29]). A third possibility which we will not analyse here, explored in detail in [2], is to rewrite the sensitivity flux in such a way that the same Roe Riemann solver used for the state can be applied for the sensitivity. Let us now describe the two possibilities considered in detail.

HLL-type scheme

The first Riemann solver proposed for the sensitivity has a simpler structure than the state solver: we neglect the contact discontinuity, therefore the solver consists only of three constant states ($\mathbf{U}_{a,L}$, $\mathbf{U}_a^*$, and $\mathbf{U}_{a,R}$) connected by two discontinuities travelling at speeds λ_1^{ROE} and λ_3^{ROE} (cf. Figure 3). The star value of the sensitivity $\mathbf{U}_a^*$ at the interface $j-1/2$ can be computed directly from the Harten, Lax and van Leer conditions [21] applied to system of conservation laws with source terms. We get:

$$\begin{aligned} \mathbf{U}_{a,j-1/2}^* = \frac{1}{\lambda_3^{ROE} - \lambda_1^{ROE}} \Big(& \lambda_3^{ROE} \mathbf{U}_{a,j}^n - \lambda_1^{ROE} \mathbf{U}_{a,j-1}^n \\ & - \mathbf{F}_a(\mathbf{U}_j, \mathbf{U}_{a,j}) + \mathbf{F}_a(\mathbf{U}_{j-1}, \mathbf{U}_{a,j-1}) + \mathbf{S}_{j-1/2} \Big), \end{aligned} \tag{22}$$

where the source term is discretised as follows:

$$\mathbf{S}_{j-1/2} = \partial_a \lambda^{ROE}_{1,j-1/2}(\mathbf{U}^*_{L,j-1/2} - \mathbf{U}_{j-1}) d_{1,j-1/2} + \partial_a \lambda^{ROE}_{2,j-1/2}(\mathbf{U}^*_{R,j-1/2} - \mathbf{U}^*_{L,j-1/2}) + \partial_a \lambda^{ROE}_{3,j-1/2}(\mathbf{U}_j - \mathbf{U}^*_{R,j-1/2}) d_{3,j-1/2},$$

where $d_{\ell,j-1/2}$ are shock detectors, $d_{\ell,j-1/2} = 1$ if there is an $\ell-$shock at the interface $j-1/2$, it is zero otherwise. They are based on the fact that the velocity u is always decreasing across a shock, whilst the density ρ is increasing across a $1-$shock and it is decreasing across a $3-$shock:

$$d_{1,j-1/2} = \begin{cases} 1 & \text{if } \rho_j > \rho_{j-1} \text{ and } u_j < u_{j-1}, \\ 0 & \text{otherwise,} \end{cases} \qquad d_{3,j-1/2} = \begin{cases} 1 & \text{if } \rho_j < \rho_{j-1} \text{ and } u_j < u_{j-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, we remark that there is no need for a contact detector because it is known that the middle wave is always a contact discontinuity.

Such a discretisation of the source term comes directly from (18), if one considers the fact that a Riemann problem can have at most three discontinuities.

HLLC-type scheme

Another possible approach for the sensitivity is to keep the same structure as for the state (cf. Figure 2), with the same speeds of propagation for the three discontinuities. We need to compute the two intermediate constant states $\mathbf{U}^*_{a,L}$ and $\mathbf{U}^*_{a,R}$. Again, a possible strategy to compute $\mathbf{U}^*_{a,L}$ and $\mathbf{U}^*_{a,R}$ is to follow the Harten, Lax and van Leer formalism with source term and to impose the following linear system, made of Rankine-Hugoniot jump relations:

$$\begin{cases} -\lambda_1(\rho^*_{a,L} - \rho_{a,L}) + (\rho u)^*_{a,L} - (\rho u)_{a,L} = \partial_a \lambda_1(\rho^*_L - \rho_L), \\ -\lambda_2(\rho^*_{a,R} - \rho^*_{a,L}) + (\rho u)^*_{a,R} - (\rho u)^*_{a,L} = \partial_a \lambda_2(\rho^*_R - \rho^*_L), \\ -\lambda_3(\rho_{a,R} - \rho^*_{a,R}) + (\rho u)_{a,R} - (\rho u)^*_{a,R} = \partial_a \lambda_3(\rho_R - \rho^*_R), \\ \frac{(\gamma-3)}{2}\tilde{u}^2(\rho^*_{a,R} - \rho^*_{a,L}) + (2-\gamma)\tilde{u}((\rho u)^*_{a,R} - (\rho u)^*_{a,L}) \\ \qquad +(\gamma-1)((\rho E)^*_{a,R} - (\rho E)^*_{a,L}) = \partial_a \lambda_2((\rho u)^*_R - (\rho u)^*_L), \\ (\lambda_2 - \lambda_1)(\rho u)^*_{a,L} + (\lambda_3 - \lambda_2)(\rho u)^*_{a,R} + \lambda_1(\rho u)_{a,L} - \lambda_3(\rho u)_{a,R} \\ \qquad +\mathbf{F}_{a,R}|_2 - \mathbf{F}_{a,L}|_2 = \Delta x \mathbf{S}|_2, \\ (\lambda_2 - \lambda_1)(\rho E)^*_{a,L} + (\lambda_3 - \lambda_2)(\rho E)^*_{a,R} + \lambda_1(\rho E)_{a,L} - \lambda_3(\rho E)_{a,R} \\ \qquad +\mathbf{F}_{a,R}|_3 - \mathbf{F}_{a,L}|_3 = \Delta x \mathbf{S}|_3, \end{cases} \tag{23}$$

where $\lambda_1 = \tilde{u} - \tilde{c}$, $\lambda_2 = \tilde{u}$, and $\lambda_3 = \tilde{u} + \tilde{c}$. The first three equations are the Rankine-Hugoniot condition on ρ across the three waves, differentiated with respect to a. Note that summing up these equations gives the integral condition of the Harten, Lax and van Leer formalism of the density variable. The fourth equation is the Rankine-Hugoniot condition on ρu for the linearised system differentiated with respect to a; the last two equations are the integral conditions on the sensitivities $(\rho u)_a$ and $(\rho E)_a$. If we define the following vectors

$$\mathbf{x} = (\rho^*_{a,L}, \rho^*_{a,R}, (\rho u)^*_{a,L}, (\rho u)^*_{a,R}, (\rho E)^*_{a,L}, (\rho E)^*_{a,R})^t$$

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{pmatrix} = \begin{pmatrix} \partial_a \lambda_1(\rho^*_L - \rho_L) + (\rho u)_{a,L} - \lambda_1 \rho_{a,L} \\ \partial_a \lambda_2(\rho^*_R - \rho^*_L) \\ \partial_a \lambda_3(\rho_R - \rho^*_R) - (\rho u)_{a,R} + \lambda_3 \rho_{a,R} \\ \partial_a \lambda_2((\rho u)^*_R - (\rho u)^*_L) \\ \Delta x \mathbf{S}|_2 - \lambda_1(\rho u)_{a,L} + \lambda_3(\rho u)_{a,R} - \mathbf{F}_{a,R}|_2 + \mathbf{F}_{a,L}|_2 \\ \Delta x \mathbf{S}|_3 - \lambda_1(\rho E)_{a,L} + \lambda_3(\rho E)_{a,R} - \mathbf{F}_{a,R}|_3 + \mathbf{F}_{a,L}|_3 \end{pmatrix}$$

the system can be rewritten as:

$$\mathcal{A}\mathbf{x} = \mathbf{b},$$

where $\mathcal{A}$ is the following matrix:

$$\mathcal{A} = \begin{pmatrix} -\lambda_1 & 0 & 1 & 0 & 0 & 0 \\ \lambda_2 & -\lambda_2 & -1 & 1 & 0 & 0 \\ 0 & \lambda_3 & 0 & -1 & 0 & 0 \\ -\frac{(\gamma-3)}{2}\tilde{u}^2 & \frac{(\gamma-3)}{2}\tilde{u}^2 & -(2-\gamma)\tilde{u} & (2-\gamma)\tilde{u} & -(\gamma-1) & (\gamma-1) \\ 0 & 0 & \tilde{c} & \tilde{c} & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{c} & \tilde{c} \end{pmatrix}$$

and we have $\det(\mathcal{A}) = 4\tilde{c}^4(\gamma-1) \neq 0$. The solution of the system has the following form:

$$\mathbf{x} = \begin{pmatrix} \frac{(2\tilde{c}+\tilde{u})b_1+(\tilde{c}+\tilde{u})b_2+\tilde{u}b_3-b_5}{2\tilde{c}^2} \\ \frac{-\tilde{u}b_1+(\tilde{c}-\tilde{u})b_2+(2\tilde{c}-\tilde{u})b_3+b_5}{2\tilde{c}^2} \\ \frac{(\tilde{u}^2+\tilde{c}\tilde{u})b_1+(\tilde{u}^2-\tilde{c}^2)b_2+(\tilde{u}^2-\tilde{c})\tilde{u}b_3+(\tilde{c}-\tilde{u})b_5}{2\tilde{c}^2} \\ \frac{-(\tilde{u}^2+\tilde{c}\tilde{u})b_1+(\tilde{c}^2-\tilde{u}^2)b_2+(\tilde{c}\tilde{u}-\tilde{u}^2)b_3+(\tilde{c}+\tilde{u})b_5}{2\tilde{c}^2} \\ \frac{(\gamma-1)(\tilde{u}^3+\tilde{c}\tilde{u}^2)b_1+((\gamma-1)\tilde{u}^3+2(2-\gamma)\tilde{c}^2\tilde{u})b_2+(\gamma-1)(\tilde{u}^3-\tilde{c}\tilde{u}^2)b_3-2\tilde{c}^2b_4-(\gamma-1)\tilde{u}^2b_5+2(\gamma-1)\tilde{c}b_6}{4(\gamma-1)\tilde{c}^2} \\ \frac{-(\gamma-1)(\tilde{u}^3+\tilde{c}\tilde{u}^2)b_1-((\gamma-1)\tilde{u}^3+2(2-\gamma)\tilde{c}^2\tilde{u})b_2+(\gamma-1)(\tilde{c}\tilde{u}^2-\tilde{u}^3)b_3+2\tilde{c}^2b_4+(\gamma-1)\tilde{u}^2b_5+2(\gamma-1)\tilde{c}b_6}{4(\gamma-1)\tilde{c}^2} \end{pmatrix}$$

An alternative strategy to compute $\mathbf{U}^*_{a,L}$ and $\mathbf{U}^*_{a,R}$ is to differentiate with respect to a the following relations:

$$\mathbf{U}^*_L = \mathbf{U}_L + \alpha_1\mathbf{r}_1, \qquad \mathbf{U}^*_R = \mathbf{U}_R - \alpha_3\mathbf{r}_3, \tag{24}$$

obtaining

$$\mathbf{U}^*_{a,L} = \mathbf{U}_{a,L} + \alpha_{a,1}\mathbf{r}_1 + \alpha_1\mathbf{r}_{a,1}, \qquad \mathbf{U}^*_{a,R} = \mathbf{U}_{a,R} - \alpha_{a,3}\mathbf{r}_3 - \alpha_3\mathbf{r}_{a,3}, \tag{25}$$

with

$$\mathbf{r}_1 = \begin{pmatrix} 1 \\ \tilde{u}-\tilde{c} \\ \tilde{H}-\tilde{u}\tilde{c} \end{pmatrix}, \quad \mathbf{r}_{a,1} = \begin{pmatrix} 0 \\ \tilde{u}_a-\tilde{c}_a \\ \tilde{H}_a-\tilde{u}_a\tilde{c}-\tilde{u}\tilde{c}_a \end{pmatrix},$$

$$\mathbf{r}_2 = \begin{pmatrix} 1 \\ \tilde{u} \\ \frac{\tilde{u}^2}{2} \end{pmatrix}, \quad \mathbf{r}_{a,2} = \begin{pmatrix} 0 \\ \tilde{u}_a \\ \tilde{u}\tilde{u}_a \end{pmatrix},$$

$$\mathbf{r}_3 = \begin{pmatrix} 1 \\ \tilde{u}+\tilde{c} \\ \tilde{H}+\tilde{u}\tilde{c} \end{pmatrix}, \quad \mathbf{r}_{a,3} = \begin{pmatrix} 0 \\ \tilde{u}_a+\tilde{c}_a \\ \tilde{H}_a+\tilde{u}_a\tilde{c}+\tilde{u}\tilde{c}_a \end{pmatrix},$$

$$\begin{cases} \alpha_2 = \dfrac{\gamma-1}{\tilde{c}^2}\Big[(\rho_R-\rho_L)(\tilde{H}-\tilde{u}^2)+\tilde{u}\Big((\rho u)_R-(\rho u)_L\Big)-\Big((\rho E)_R-(\rho E)_L\Big)\Big], \\ \alpha_1 = \dfrac{1}{\tilde{c}}\Big[(\rho_R-\rho_L)(\tilde{u}+\tilde{c})-\Big((\rho u)_R-(\rho u)_L\Big)-\tilde{c}\alpha_2\Big], \\ \alpha_3 = (\rho_R-\rho_L)-(\alpha_1+\alpha_2), \end{cases}$$

$$\begin{cases} \alpha_{a,2} &= -\dfrac{2\tilde{c}_a(\gamma-1)}{\tilde{c}^3}\Big[(\rho_R-\rho_L)(\tilde{H}-\tilde{u}^2)+\tilde{u}\Big((\rho u)_R-(\rho u)_L\Big)-\Big((\rho E)_R-(\rho E)_L\Big)\Big] \\ &\quad +\dfrac{\gamma-1}{\tilde{c}^2}\Big[(\rho_{a,R}-\rho_{a,L})(\tilde{H}-\tilde{u}^2)+(\rho_R-\rho_L)(\tilde{H}_a-2\tilde{u}\tilde{u}_a) \\ &\quad +\tilde{u}_a\Big((\rho u)_R-(\rho u)_L\Big)-\Big((\rho E)_R-(\rho E)_L\Big)+\tilde{u}\Big((\rho u)_{a,R}-(\rho u)_{a,L}\Big)-\Big((\rho E)_{a,R}-(\rho E)_{a,L}\Big)\Big], \\ \alpha_{a,1} &= -\dfrac{\tilde{c}_a}{\tilde{c}^2}\Big[(\rho_R-\rho_L)(\tilde{u}+\tilde{c})-((\rho u)_R-(\rho u)_L)-\tilde{c}\alpha_2\Big] \\ &\quad +\dfrac{1}{2\tilde{c}}\Big[(\rho_{a,R}-\rho_{a,L})(\tilde{u}+\tilde{c})+(\rho_R-\rho_L)(\tilde{u}_a+\tilde{c}_a)-((\rho u)_{a,R}-(\rho u)_{a,L})-\tilde{c}_a\alpha_2-\tilde{c}\alpha_{a,2}\Big], \\ \alpha_{a,3} &= (\rho_{a,R}-\rho_{a,L})-(\alpha_{a,1}+\alpha_{a,2}). \end{cases}$$

The next proposition states that the two strategies to define $\mathbf{U}^*_{a,L}$ and $\mathbf{U}^*_{a,R}$ are equivalent.

Proposition. 2. *The star sensitivities (25) solve the system (23).*

Proof. We will prove that the star sensitivities defined in (25) satisfy the system (23).

1. First equation. Writing the first coefficient of (24) one easily finds $\rho^*_L - \rho_L = \alpha_1$, and writing the first two coefficient of (25) one finds:

$$\rho^*_{a,L} - \rho_{a,L} = \alpha_{a,1}, \qquad (\rho u)^*_{a,L} - (\rho u)_{a,L} = \alpha_{a,1}(\tilde{u}-\tilde{c}) + \alpha_1(\tilde{u}_a - \tilde{c}_a).$$

We now replace these three expressions in the first equation of (23) and we obtain:

$$-\lambda_1\alpha_{a,1} + \alpha_{a,1}(\tilde{u} - \tilde{c}) + \alpha_1(\tilde{u}_a - \tilde{c}_a) = \partial_a\lambda_1\alpha_1,$$

which is always verified, since $\lambda_1 = \tilde{u} - \tilde{c}$.

2. Second equation. We recall that

$$\mathbf{U}_R - \mathbf{U}_L = \sum_{i=1}^{3} \alpha_i \mathbf{r}_i, \qquad \mathbf{U}_{a,R} - \mathbf{U}_{a,L} = \sum_{i=1}^{3} \alpha_{a,i}\mathbf{r}_i + \alpha_i \mathbf{r}_{a,i}.$$

Therefore, one has:

$$\mathbf{U}^*_R - \mathbf{U}^*_L = \alpha_2\mathbf{r}_2, \qquad \mathbf{U}^*_{a,R} - \mathbf{U}^*_{a,L} = \alpha_{a,2}\mathbf{r}_2 + \alpha_2\mathbf{r}_{a,2},$$

which gives us the following relations:

$$\rho^*_R - \rho^*_L = \alpha_2, \quad \rho^*_{a,R} - \rho^*_{a,L} = \alpha_{a,2}, \quad (\rho u)^*_{a,R} - (\rho u)^*_{a,L} = \alpha_{a,2}\tilde{u} + \alpha_2\tilde{u}_a.$$

We now replace them in the second equation of (23) and we obtain:

$$-\lambda_2\alpha_{a,2} + \alpha_{a,2}\tilde{u} + \alpha_2\tilde{u}_a = \partial_a\lambda_2\alpha_2,$$

which is always verified, since $\lambda_2 = \tilde{u}$.

3. Third equation. As we did for the first two equations, one can find the three following expressions:

$$\rho_R - \rho^*_R = \alpha_3, \quad \rho_{a,R} - \rho^*_{a,R} = \alpha_{a,3}, \quad (\rho u)_{a,R} - (\rho u)^*_{a,R} = \alpha_{a,3}(\tilde{u} + \tilde{c}) + \alpha_2(\tilde{u}_a + \tilde{c}_a).$$

By replacing them in the third equation of (23) one can easily check that the equation is always verified, since $\lambda_3 = \tilde{u} + \tilde{c}$.

4. Fourth equation. As we did for the previous equations, one can find the three following expressions:

$$(\rho u)^*_R - (\rho u)^*_L = \alpha_2\tilde{u}, \quad \rho^*_{a,R} - \rho^*_{a,L} = \alpha_{a,2}, \quad (\rho u)^*_{a,R} - (\rho u)^*_{a,L} = \alpha_{a,2}\tilde{u} + \alpha_2\tilde{u}_a,$$
$$(\rho E)^*_{a,R} - (\rho E)^*_{a,L} = \alpha_{a,2}\frac{\tilde{u}^2}{2} + \alpha_2\tilde{u}\tilde{u}_a.$$

By replacing them in the fourth equation of (23) one can easily check that the equation is always verified, since $\lambda_2 = \tilde{u}$.

5. Fifth and sixth equations. The last two equations are the last two components of the following vectorial equation:

$$(\lambda_2 - \lambda_1)\mathbf{U}^*_{a,L} + (\lambda_3 - \lambda_2)\mathbf{U}^*_{a,R} + \lambda_1\mathbf{U}_{a,L} - \lambda_3\mathbf{U}_{a,R} + \mathbf{F}_{a,R} - \mathbf{F}_{a,L} = \Delta x\mathbf{S},$$

which can be rewritten as:

$$\lambda_1(\mathbf{U}_{a,L} - \mathbf{U}^*_{a,L}) + \lambda_2(\mathbf{U}^*_{a,L} - \mathbf{U}^*_{a,R}) + \lambda_3(\mathbf{U}^*_{a,R} - \mathbf{U}_{a,R}) + \mathbf{F}_{a,R} - \mathbf{F}_{a,L} = \Delta x\mathbf{S}.$$

Replacing the definitions (25) one finds:

$$-\lambda_1(\alpha_{a,1}\mathbf{r}_1 + \alpha_1\mathbf{r}_{a,1}) - \lambda_2(\alpha_{a,2}\mathbf{r}_2 + \alpha_2\mathbf{r}_{a,2}) - \lambda_3(\alpha_{a,3}\mathbf{r}_3 + \alpha_3\mathbf{r}_{a,3}) + \mathbf{F}_{a,R} - \mathbf{F}_{a,L} = \Delta x\mathbf{S}.$$

We recall that by definition of Roe fluxes, one has:

$$\mathbf{F}_R - \mathbf{F}_L = \sum_{i=1}^{3} \alpha_i\lambda_i\mathbf{r}_i \Rightarrow \mathbf{F}_{a,R} - \mathbf{F}_{a,L} = \sum_{i=1}^{3} \alpha_{a,i}\lambda_i\mathbf{r}_i + \alpha_i\lambda_{a,i}\mathbf{r}_i + \alpha_i\lambda_i\mathbf{r}_{a,i}.$$

Therefore, we obtain:

$$\Delta x\mathbf{S} = \sum_{i=1}^{3} \alpha_i\lambda_{a,i}\mathbf{r}_i = \lambda_{a,1}(\mathbf{U}^*_L - \mathbf{U}_L) + \lambda_{a,2}(\mathbf{U}^*_R - \mathbf{U}^*_L) + \lambda_{a,3}(\mathbf{U}_R - \mathbf{U}^*_R),$$

which is consistent with our discretisation of the source term.

□

Finally, one can obtain $\mathbf{u}_a(x, t^{n+1})$ by juxtaposition, as we did for the state.

3.3 Projection step

The projection step is usually performed by averaging on the cell the solution $\mathbf{v}(x, t^{n+1})$, whose components $\mathbf{u}(x, t^{n+1})$ and $\mathbf{u}_a(x, t^{n+1})$ can be computed as described in the previous sections.

$$\mathbf{V}_j^{n+1} = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{v}(x, t^{n+1}) dx. \tag{26}$$

We remark that the integral (26) is easy to compute, $\mathbf{v}$ being piecewise constant. However, this projection method introduces numerical diffusion. As shown in [5], numerical diffusion plays a fundamental role in the discretisation of the sensitivity, especially across shocks. For this reason, we propose another projection method, introduced in [6] and inspired by Glimm's method [14, 7]. First, we define a staggered mesh, whose cells will be denoted $\overline{C}_j^n$, as follows:

$$\overline{C}_j^n = (\bar{x}_{j-1/2}^n, \bar{x}_{j+1/2}^n), \quad \bar{x}_{j-1/2}^n = x_{j-1/2} + \sigma_{j-1/2}^n \Delta t^n,$$

where $\sigma_{j-1/2}^n$ is a proper speed, defined in order to avoid averaging across a shock. Numerical results show that there is no need to move the mesh for the contact discontinuity (cf. section 4). The definition of $\sigma_{j-1/2}^n$ is the following:

$$\sigma_{j-1/2}^n = \begin{cases} \lambda_{1,j-1/2}^{ROE} & \text{if } d_{1,j-1/2} = 1, \\ \lambda_{3,j-1/2}^{ROE} & \text{if } d_{3,j-1/2} = 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $d_{\ell,j-1/2}$ are the shock detectors defined earlier.

The second step is to perform the average on the staggered mesh, obtaining in this way an intermediate solution $\overline{\mathbf{V}}_j^{n+1}$:

$$\overline{\mathbf{V}}_j^{n+1} = \frac{1}{\Delta x_j^n} \int_{\bar{x}_{j-1/2}}^{\bar{x}_{j+1/2}} \mathbf{v}(x, t^{n+1}) dx, \tag{27}$$

where $\Delta x_j^n = \bar{x}_{j+1/2} - \bar{x}_{j-1/2}$. Finally, the last step is a sampling step, in order to go back to the initial uniform grid. Let (β_n) be a random sequence varying in $(0, 1)$, for instance $\beta_n \sim \mathcal{U}([0, 1])$; then:

$$\mathbf{V}_j^{n+1} = \begin{cases} \overline{\mathbf{V}}_{j-1}^{n+1} & \text{if } \beta_{n+1} \in \left(0, \frac{\Delta t}{\Delta x} \max(\sigma_{j-1/2}^n, 0)\right), \\ \overline{\mathbf{V}}_j^{n+1} & \text{if } \beta_{n+1} \in \left[\frac{\Delta t}{\Delta x} \max(\sigma_{j-1/2}^n, 0), 1 + \frac{\Delta t}{\Delta x} \min(\sigma_{j+1/2}^n, 0)\right), \\ \overline{\mathbf{V}}_{j+1}^{n+1} & \text{if } \beta_{n+1} \in \left[1 + \frac{\Delta t}{\Delta x} \min(\sigma_{j+1/2}^n, 0), 1\right). \end{cases} \tag{28}$$

We remark that one β_n is drawn at each time step and it is the same for all the cells. The method is proven to be convergent even if a low discrepancy deterministic sequence is used. In this work, we use the van der Corput sequence (cf. [6]):

$$\beta_n = \sum_{k=0}^{m} i_k 2^{-(k+1)}, \quad n = \sum_{k=0}^{m} i_k 2^k,$$

where $i_k = 0, 1$ is the binary expansion of the integers.

3.4 Second order extension

In this section, we extend to the second order the schemes presented above. In time, we use a standard two steps Runge-Kutta method, whilst in space we use a MUSCL-type (Monotonic Upstream-centered Scheme for Conservation Laws, [31]) approach, inspired from [5]. In a few words (we refer to [5] for more details) the main idea of a MUSCL-type scheme is to consider in replacement of a constant value $\mathbf{V}_j^n$ in each cell, a higher order polynomial $\mathbf{V}_j^n(x)$, $x \in [x_{j-1/2}, x_{j+1/2}]$. The edge values $\mathbf{V}_j^n(x_{j+1/2})$, $\mathbf{V}_{j+1}^n(x_{x+1/2})$ are used as left and right values for the Riemann problem at the interface $j + 1/2$; the Riemann problem is then solved as explained in the previous section. However, the definition of the source term (13)-(17) is valid only if the state is piecewise constant (cf. [5]). Therefore, we suggest to consider a piecewise constant state on half of each cell: these two constant values will be denoted $\mathbf{V}_{j\pm1/4}^n$ and correspond to the edge values $\mathbf{V}_j^n(x_{j\pm1/2})$ (see Figure 4). In this work, we compute the edge values with a standard approach:

$$\mathbf{V}_{j\pm1/4}^n = \mathbf{V}_j^n \pm \Delta \mathbf{V}_j^n,$$

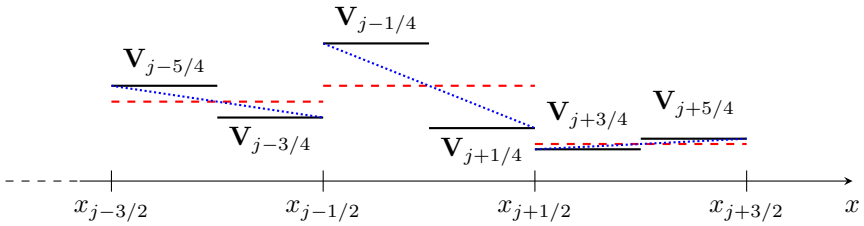


Figure 4: MUSCL discretisation. Dashed red line: first order discretisation. Dotted blue line: classical second order discretisation. Solid black line: second order discretisation used in this work.

	Diffusive	Anti-diffusive
1st order	ROE I	ROE I AD
2nd order	ROE II	ROE II AD

(a) State schemes.

	Diffusive	Anti-diffusive
1st order	HLL I	HLL I AD
2nd order	HLL II	HLL II AD

(b) Sensitivity HLL schemes.

	Diffusive	Anti-diffusive
1st order	HLLC I	HLLC I AD
2nd order	HLLC II	HLLC II AD

(c) Sensitivity HLLC schemes.

Table 1: Numerical schemes summary.

and usual choice for $\Delta\mathbf{V}_j^n$ is to use a slope-limiter procedure, for instance:

$$\Delta\mathbf{V}_j^n = \frac{1}{2}\text{minmod}(\mathbf{V}_{j+1}^n - \mathbf{V}_j^n, \mathbf{V}_j^n - \mathbf{V}_{j-1}^n),$$

where

$$\text{minmod}(a,b) = \begin{cases} \text{sgn}(a)\min(|a|,|b|) & \text{if } ab > 0, \\ 0 & \text{otherwise.} \end{cases}$$

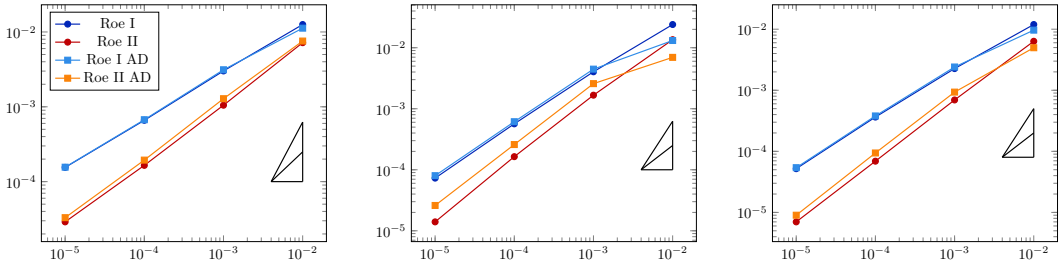
We remark that this approach leads to an additional Riemann problem in the middle of the cell: in this way we are able to extend the scheme accuracy to second-order, while keeping piecewise constant representations in the cells, which is necessary to make the terms containing $\partial_x \mathbf{U}^{\pm}$ vanishing in (15).

3.5 Summary

Here, we briefly sum up all the ingredients introduced in this section and we clarify how they can be combined to obtain different numerical schemes:

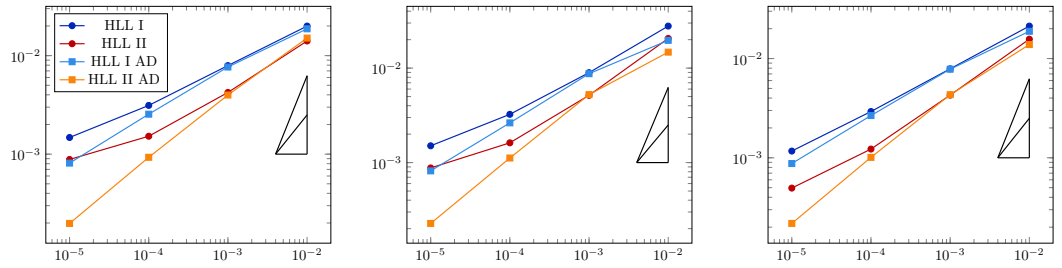
- Order of the scheme. In this paper we focused on first and second order schemes, but higher order can be used bearing in mind that the state needs to be piecewise constant for the definition of the source term to be valid.
- Riemann solver for the state. In this paper we proposed the Roe Riemann solver for the state, but the only constraint is to use a solver with two intermediate states $\mathbf{U}_L^*$ and $\mathbf{U}_R^*$ (for instance, HLL could not be used for the state).
- Riemann solver for the sensitivity. We designed two different numerical schemes for the sensitivity: an HLL and HLLC-type scheme.
- Type of projection. Either the classical projection or the anti-diffusive projection can be used. Numerical results in the next section will show that, for this problem, diffusive scheme does not converge to the analytical solution.

Finally, we remark that these four choices are independent of each other. In table (1) we summarize all the combinations used in this work, with the labels used in the next sections.



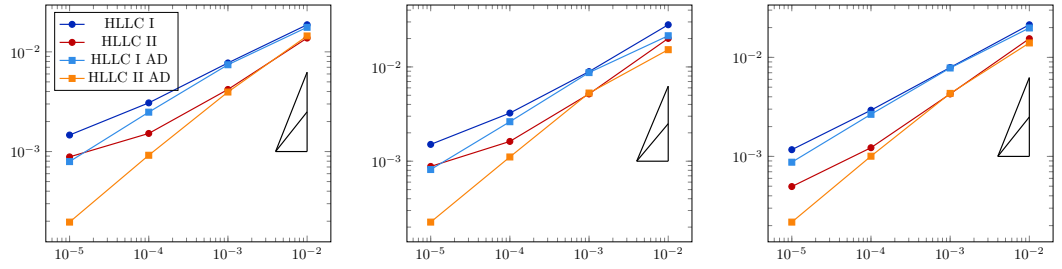
(a) $\|\rho(x,T) - \rho_{ex}(x,T)\|_{L^1}$ (b) $\|u(x,T) - u_{ex}(x,T)\|_{L^1}$ (c) $\|p(x,T) - p_{ex}(x,T)\|_{L^1}$

Figure 5: Convergence test for the state. The Roman numerals I and II stand for the order of the scheme. AD stands for anti-diffusive.



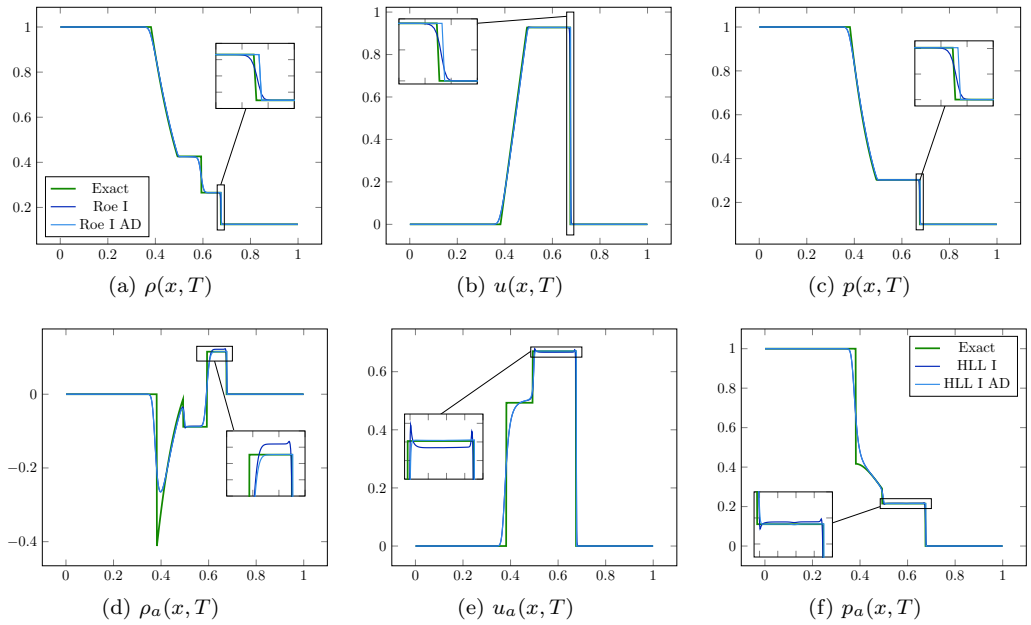
(a) $\|\rho_a(x,T) - \rho_{a,ex}(x,T)\|_{L^1}$ (b) $\|u_a(x,T) - u_{a,ex}(x,T)\|_{L^1}$ (c) $\|p_a(x,T) - p_{a,ex}(x,T)\|_{L^1}$

Figure 6: Convergence test for the sensitivity - HLL-type scheme.



(a) $\|\rho_a(x,T) - \rho_{a,ex}(x,T)\|_{L^1}$ (b) $\|u_a(x,T) - u_{a,ex}(x,T)\|_{L^1}$ (c) $\|p_a(x,T) - p_{a,ex}(x,T)\|_{L^1}$

Figure 7: Convergence test for the sensitivity - HLLC-type scheme.



(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

(d) $\rho_a(x,T)$ (e) $u_a(x,T)$ (f) $p_a(x,T)$

Figure 8: First order schemes, with and without numerical diffusion. HLL-type scheme for the sensitivity.

4 Convergence tests for the numerical schemes

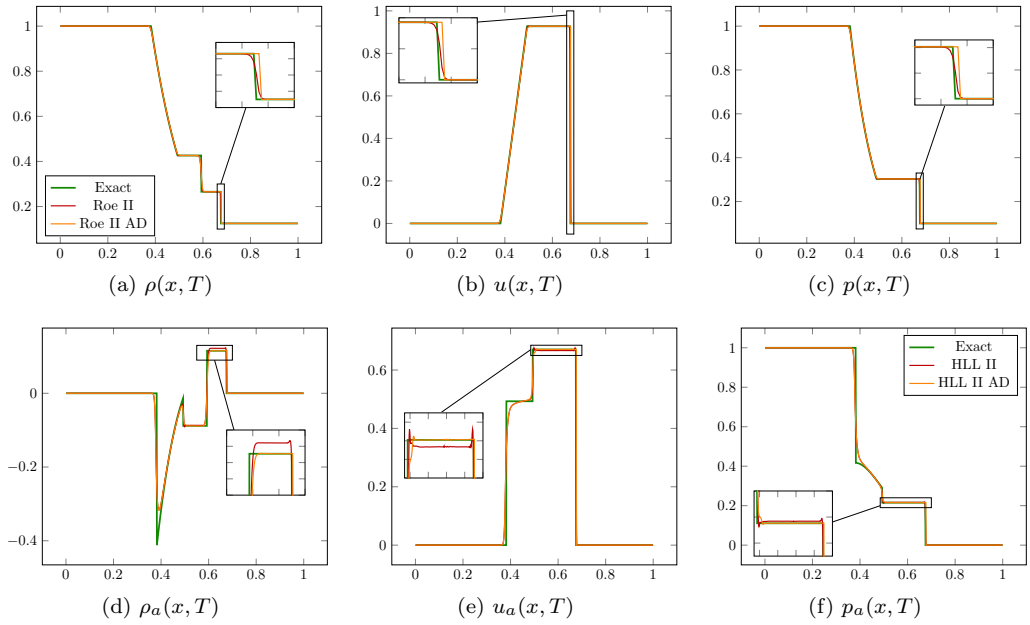
We consider the Riemann problem described in appendix A. The initial data for the state on the physical variables is the following:

$$\rho_L = 1,\ u_L = 0,\ p_L = 1, \qquad \rho_R = 0.125,\ u_R = 0,\ p_R = 0.1.$$

We consider as parameter of interest $a = p_L$, therefore the initial data for the sensitivity is:

$$\rho_{a,L} = \rho_{a,R} = u_{a,L} = u_{a,R} = p_{a,R} = 0, \quad p_{a,L} = 1.$$

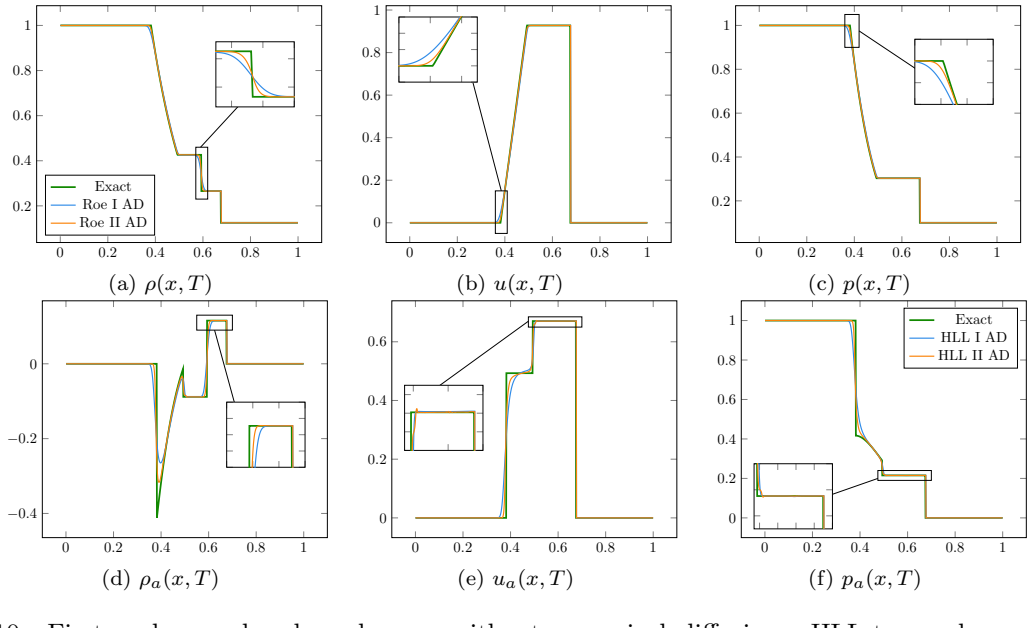
In Figures 5-6-7 we show the convergence of the different numerical schemes presented in Section 3. Figure 5 shows the convergence for the state: the rate of convergence is the expected one; one can remark that the antidiffusive schemes are slightly less precise than the diffusive ones. In Figures 6-7 we plot the error for the sensitivity, first with the HLL-type scheme (Figure 6) and then with the HLLC-type scheme (Figure 7): considering two different star regions for the sensitivity does not seem to make much difference; however one can remark the same effect shown in [5] for a simpler system: the diffusive schemes do not converge for the sensitivity, this is especially evident for the variable ρ_a. In Figure 8 we plot the solution at the final time $T = 0.1$, obtained with a mesh $\Delta x = 10^{-3}$ with the first order schemes, both diffusive and antidiffusive (for the sensitivity, the HLL-type scheme has been used): one can notice that the plateau in the right-star zone is not properly captured by the diffusive scheme. This does not change as one refines the mesh, nor with a higher order scheme, as one can see from Figure 9. In Figure 10 we compare the antidiffusive schemes, first and second order: for the state, the difference is noticeable mainly in the contact discontinuity (therefore only for ρ), whilst for the sensitivity the difference is significant in the neighbourhood of the discontinuities before and after the rarefaction. Finally, in Figure 11 we compare the HLL and the HLLC-type schemes for the sensitivity: as anticipated by the error plots, the two schemes are almost equivalent in terms of results; the difference between the solutions provided by the two second order schemes in $L^\infty-$norm is 0.0096, 0.0139, and 0.0062 respectively for ρ_a, u_a, and p_a and this is why they are almost indistinguishable in Figure 11. For this reason, the use of HLL-type scheme is preferable, being less expensive from a computational point of view and less complicated to implement.



(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

(d) $\rho_a(x,T)$ (e) $u_a(x,T)$ (f) $p_a(x,T)$

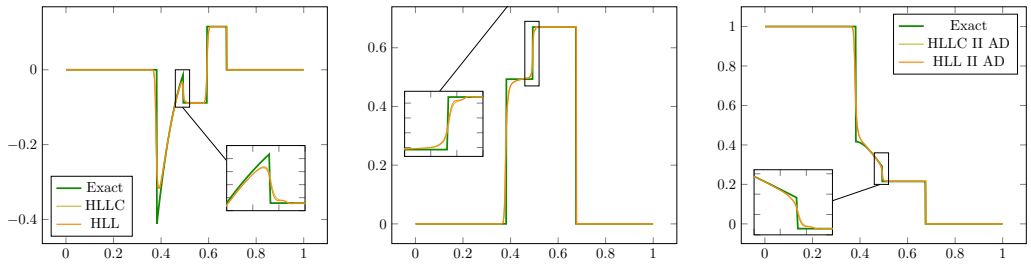
Figure 9: Second order schemes, with and without numerical diffusion. HLL-type scheme for the sensitivity.



(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

(d) $\rho_a(x,T)$ (e) $u_a(x,T)$ (f) $p_a(x,T)$

Figure 10: First and second order schemes, without numerical diffusion. HLL-type scheme for the sensitivity.



(a) $\rho_a(x,T)$ (b) $u_a(x,T)$ (c) $p_a(x,T)$

Figure 11: Second order antidiffusive schemes: HLL and HLLC comparison.

5 Uncertainty Quantification

5.1 Problem description

In this section, we show how the estimated sensitivities can be used for uncertainty propagation, i.e. estimate statistical moments of the solution accounting for some uncertain parameters [27, 30, 10]. We want to provide a demonstration of one of the many possible applications of the CSE method, in order to underline the potentials of the proposed approach and its limitations, too. We also intend to quantify the impact of removing the spikes from the sensitivity solution. Many techniques have been developed during the last decades to propagate uncertainty through PDE models: these methods can be either probabilistic or deterministic. The proposed method based on derivatives estimation falls into the second category, while the most well-known of these techniques, the Monte Carlo method, is in the first. Other techniques are for instance polynomial chaos [32, 34, 24, 11], or the random space partition [1]. A very good review and comparison of many techniques with applications to fluid dynamics can be found in [33]. A typical objective of uncertainty propagation is to determine a confidence interval for the output of a model, in our case $\mathbf{U}$, given the uncertainty on the input parameters. This estimation is part of the broader domain of Uncertainty Quantification (UQ), which also includes the identification of the most critical uncertain parameters, their ranking, and the analysis of the variability of the output.

In this work, we compare the Monte Carlo approach and a sensitivity-based estimation. In the following, X will stand for one of the variables, i.e. X can either be ρ, u or p, and X_a the corresponding sensitivity. We use the notation μ_X to indicate the average of the variable X and σ_X^2 for its variance. Once this two quantities are known, one can build a confidence interval for the variable X as:

$$CI_X = [\mu_X - \kappa\sigma_X, \mu_X + \kappa\sigma_X]. \tag{29}$$

We remark that (29) is valid only for gaussian data. The coefficient κ regulates the amplitude of the interval and it is related to the probability for the variable X to actually fall in the interval. For instance, the choice $\kappa = 1.96$ provides a 95% confidence interval, while $\kappa = 2.58$ a 99% one.

Monte Carlo method. Here we briefly introduce the Monte Carlo method, for more details see for instance [8]. The Monte Carlo method is a probabilistic technique: to obtain an estimate of the average and of the standard deviation one needs to perform multiple random simulations. Let $\mathbf{a}$ be the vector of uncertain parameters, with a known distribution. Then, N random samples $\mathbf{a}_i$ are drawn from this distribution, and for each $\mathbf{a}_i$ the corresponding solution X_i is computed. Then, the unbiased average and variance estimators are used:

$$\mu_X = \frac{1}{N}\sum_{i=1}^{N} X_i, \qquad \sigma_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\mu_X - X_i)^2.$$

These estimates are good if N is sufficiently large: the slow convergence, and therefore the high computational cost, is probably the main limitation of the Monte Carlo method.

Sensitivity-based method. Once the sensitivities of the solution with respect to the input parameters are known, a deterministic estimation of the average μ_X and of the variance σ_X^2 of the

output X can be easily obtained. Let $\mu_{\mathbf{a}}$ be the average of the uncertain vector $\mathbf{a}$ and $\sigma_{\mathbf{a}}$ the covariance matrix:

$$\mu_{\mathbf{a}} = \begin{bmatrix} \mu_{a_1} \\ \vdots \\ \mu_{a_M} \end{bmatrix}, \qquad \sigma_{\mathbf{a}} = \begin{bmatrix} \sigma^2_{a_1} & \mathrm{cov}(a_1, a_2) & \dots & \mathrm{cov}(a_1, a_M) \\ \mathrm{cov}(a_1, a_2) & \sigma^2_{a_2} & \dots & \mathrm{cov}(a_2, a_M) \\ \vdots & & \ddots & \vdots \\ \mathrm{cov}(a_1, a_M) & \dots & & \sigma^2_{a_M} \end{bmatrix},$$

where M is the number of uncertain parameters, μ_{a_i} the average of the $i-$th parameter, $\sigma^2_{a_i}$ its variance and $\mathrm{cov}(\cdot, \cdot)$ the covariance. Let us consider the first order Taylor expansion for the variable X with respect to the vector of parameters $\mathbf{a}$:

$$X(\mathbf{a}) = X(\mu_{\mathbf{a}}) + \sum_{i=1}^{M} (a_i - \mu_{a_i}) X_{a_i}(\mu_{\mathbf{a}}) + o(\|\mathbf{a}\|^2).$$

Then computing the average, since $X(\mu_{\mathbf{a}})$ and $X_{a_i}(\mu_{\mathbf{a}})$ are not random variables, at first order one gets:

$$\mu_X = E[X(\mathbf{a})] = X(\mu_{\mathbf{a}}) + \sum_{i=1}^{M} X_{a_i}(\mu_{\mathbf{a}}) E[a_i - \mu_{a_i}] = X(\mu_{\mathbf{a}}),$$

because $E[(a_i - \mu_{a_i})] = 0$. In the same way, one can compute the variance:

$$\sigma^2_X = E[(X(\mathbf{a}) - \mu_X)^2] = E\left[\left(\sum_{i=1}^{M} X_{a_i}(\mu_{\mathbf{a}})(a_i - \mu_{a_i}) \right)^2 \right] =$$

$$= \sum_{i=1}^{M} X^2_{a_i}(\mu_{\mathbf{a}}) E[(a_i - \mu_{a_i})^2] + \sum_{\substack{i,j=1 \\ i \neq j}}^{M} X_{a_i}(\mu_{\mathbf{a}}) X_{a_j}(\mu_{\mathbf{a}}) E[(a_i - \mu_{a_i})(a_j - \mu_{a_j})].$$

Therefore, we obtain the following first order estimates of the average and the variance of the variable X:

$$\mu_X = X(\mu_{\mathbf{a}}), \quad \sigma^2_X = \sum_{i=1}^{M} X^2_{a_i} \sigma^2_{a_i} + \sum_{\substack{i,j=1 \\ i \neq j}}^{M} X_{a_i} X_{a_j} \mathrm{cov}(a_i, a_j).$$

Higher order estimates require higher order sensitivities [25].

5.2 Numerical results

We applied the uncertainty propagation techniques described in the previous subsection to the test case described in appendix A. The uncertain parameters are the left and right values of the physical variables for the state, i.e.:

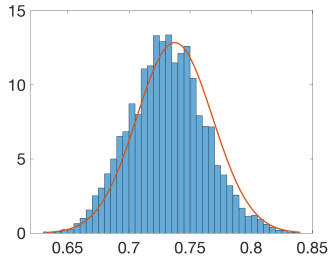
$$\mathbf{a} = (\rho_L, \rho_R, u_L, u_R, p_L, p_R)^t,$$

and have a Gaussian distribution with the following average and covariance matrix:

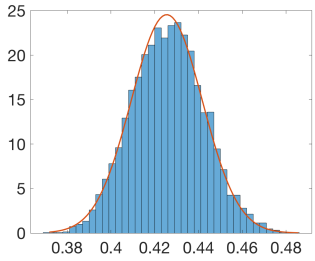
$$\mu_{\mathbf{a}} = (1, 0.125, 0, 0, 1, 0.1)^t, \quad \sigma_{\mathbf{a}} = \mathrm{diag}(0.001, 0.000125, 0.0001, 0.0001, 0.001, 0.0001).$$

This choice means that all the parameters are uncorrelated and we chose as their variance the 0.1% of their average, except for the velocity, whose average is 0.

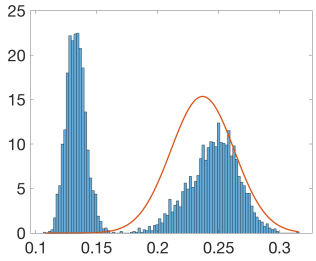
We remark that the fact that the parameters follow a Gaussian distribution does not say anything about the distribution of the output of the model. Therefore, before using the confidence interval (29), one should check that the output is Gaussian, too. In Figures-12-13-14, three histograms are shown for each physical variable: for $x = 0.35$ (i.e. in the middle of the rarefaction wave), for $x = 0.6$ (i.e. in the middle plateau) and for $x = 0.85$ (i.e. close to the shock position). The histograms are obtained by computing the analytical solution for 5000 different values of the parameter vector $\mathbf{a}$, which are sampled from its distribution. The histograms are then normalised with respect to the probability density function, using the MATLAB option 'normalization', 'pdf'. As one can see, the output is Gaussian when far from the shock; close to the shock, two distinct groups of points can be identified. The distribution predicted with the SA is drawn in red: the prediction is wrong in the neighbourhood of the shock, as as; it fits perfectly in the plateau and it is slightly shifted for the rarefaction. This shift is due to the prediction for the average, which is based entirely on the state



(a) In the rarefaction ($x = 0.35$)

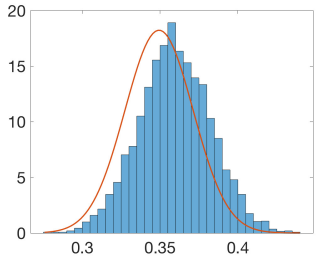


(b) In the plateau ($x = 0.6$)

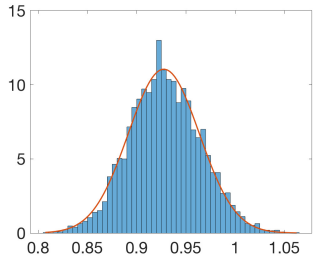


(c) In the shock neighbourhood ($x = 0.85$)

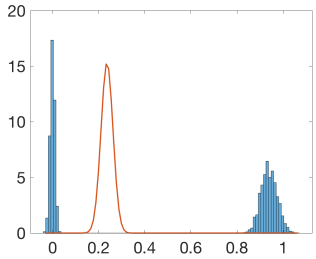
Figure 12: Distribution of $\rho(x, T; \mathbf{a})$ for three different values of x at final time $T = 0.2$



(a) In the rarefaction ($x = 0.35$)



(b) In the plateau ($x = 0.6$)

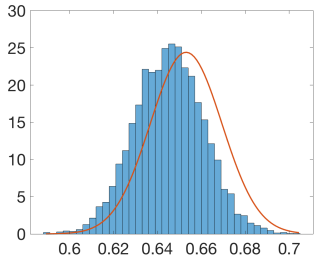


(c) In the shock neighbourhood ($x = 0.85$)

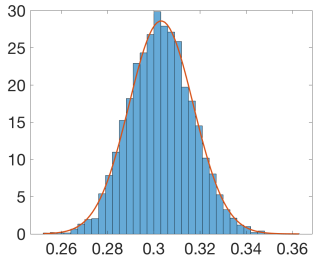
Figure 13: Distribution of $u(x, T; \mathbf{a})$ for three different values of x at final time $T = 0.2$

and caused by the fact that it is only a first order approximation. However, one can remark that the variance is correctly estimated using the sensitivities. In the following, we use the expression (29), expecting however a loss of precision in the neighbourhood of the shock.

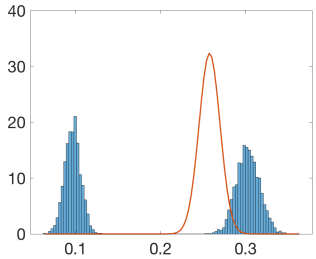
In Figure 15 we show the results of the Monte Carlo approach: the average and the average plus and minus twice the standard deviation (i.e. $\kappa = 2$) are plotted in red, five samples are plotted in black. These results are obtained with $N = 1000$ samples, on a mesh with $\Delta x = 10^{-3}$ using a Roe first order diffusive scheme. As one can see, the average process smudges the shock and the standard deviation is bigger in that zone. In fact, this area of large variance around the shock location is related to the delta Dirac function in sensitivity solution, but this peak is "smooth" due to the fact that Monte-Carlo approach accounts for flow non-linearities and is not limited to the first-order estimate of the perturbation. In Figures 16-17 we show the results of the sensitivity-based approach, with $\Delta x = 10^{-3}$ and the diffusive first order scheme, when the sensitivity is computed without the correction term (13): the spikes in the neighbourhood of the shock are very different



(a) In the rarefaction ($x = 0.35$)

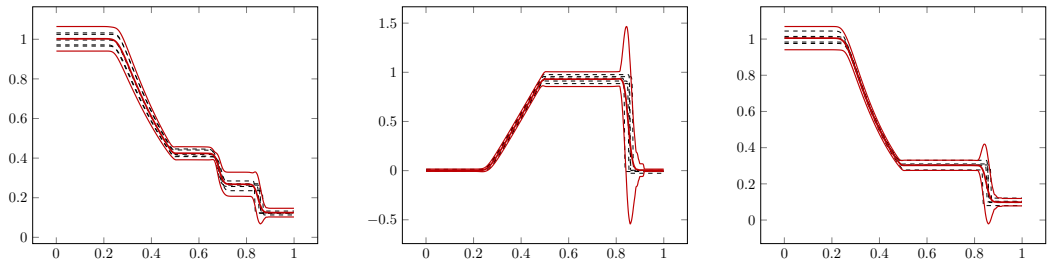


(b) In the plateau ($x = 0.6$)



(c) In the shock neighbourhood ($x = 0.85$)

Figure 14: Distribution of $p(x, T; \mathbf{a})$ for three different values of x at final time $T = 0.2$



(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

Figure 15: Monte Carlo approach. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines.

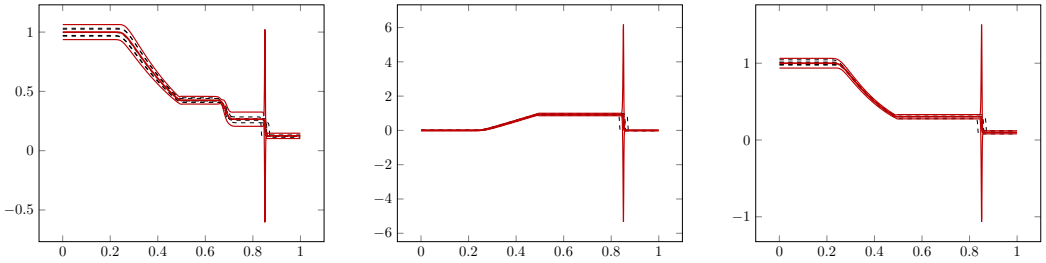
with respect to the ones we get with the Monte Carlo approach. As explained above, these spikes are due to the default of high-order terms in the Taylor expansion. On one hand, these peaks lead to non-physical values for the solution (in particular, the confidence intervals contains negative values for the pressure and for the density); on the other hand, they do not enlarge sufficiently the zone to contain the majority of the samples: one can observe that four out of five samples fall out of the predicted interval in the neighbourhood of the shock. Nevertheless, we observe that the non-linear effects are located in a small region around the shock and elsewhere confidence interval is well estimated by the sensitivity-based method. The results obtained with the corrected sensitivities are shown in Figure 18: the confidence interval obtained correspond to the ones obtained with the Monte Carlo approach, apart for the shock zone. Of course, the sensitivity-based approach does not capture the uncertainty in the neighborhood of the shock, because it neglects the dependence of the speed of the shock on the parameters. This is why most of the samples fall out of the zone predicted with the sensitivity-based approach, and it is the case with and without correction. However, the correction avoids non-physical values in the confidence interval. Although the sensitivity-based approach is not able to account for non-linear effects, contrary to the Monte-Carlo method, it is far less expensive: the Monte Carlo approach requires 1000 solutions of the state, whilst the sensitivity-based approach only one solution of the state and as many solution of the sensitivity as the number of uncertain parameters, in this case 6. Therefore, this approach can still be interesting for computationally-demanding problems, for which the use of the Monte-Carlo method cannot be envisaged.

Finally, in Figure 19 we show the results obtained with the anti-diffusive scheme: the difference with respect to the diffusive scheme is not significant. This is a good news for possible future developments in $2D$: the anti-diffusive scheme is very difficult to adapt in higher dimensional spaces; in fact the Glimm method has been proven not to work in a two-dimensional space. With these results, we underline how the numerical diffusion plays an important role in the convergence of the scheme, but it is not so significant for the final application.

The numerical results of the SA method applied to an uncertainty propagation problem show the potential and the limits of the method: it is really affordable from a computational point of view, with the trade off of being less precise than, for instance, a Monte Carlo method in the discontinuous zones. If one accepts this loss of precision, the proposed approach remains accurate in the regular zone and is highly competitive thanks to its very low computational cost.

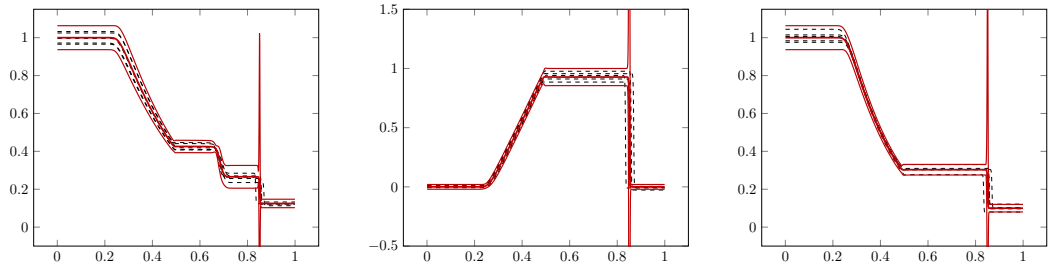
6 Conclusion

In this work, we extended to the complete Euler system the method proposed in [5] for the $p-$system. The definition of the source term does not differ significantly and the same shock detectors can be used in this case. We remark that a contact discontinuity detector is not necessary, since the middle wave is always a contact discontinuity. However, in this more complex case, the form of the proposed source term precludes the application of some well-known and widely used numerical schemes such as the HLL scheme, and all the approximate Riemann solvers with only one middle state. The numerical results show that the numerical diffusion in the shocks plays an important role and corrupts the convergence to the correct solution even for the complete system. However, we remark that the expected convergence rate can be achieved without removing the numerical diffusion in the contact discontinuity, which simplifies the definition of the staggered mesh. Currently, we are extending this



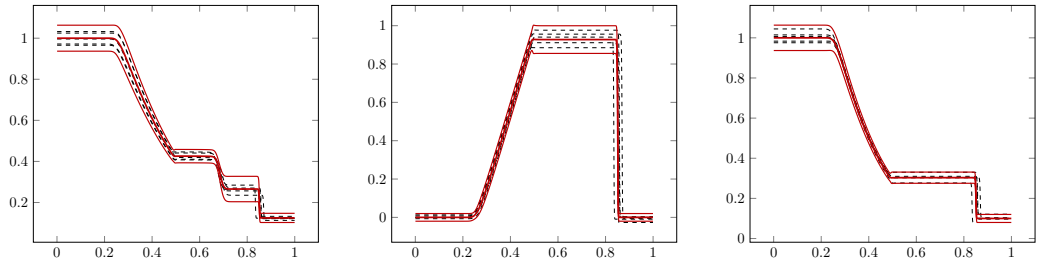
(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

Figure 16: SA approach without correction. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines



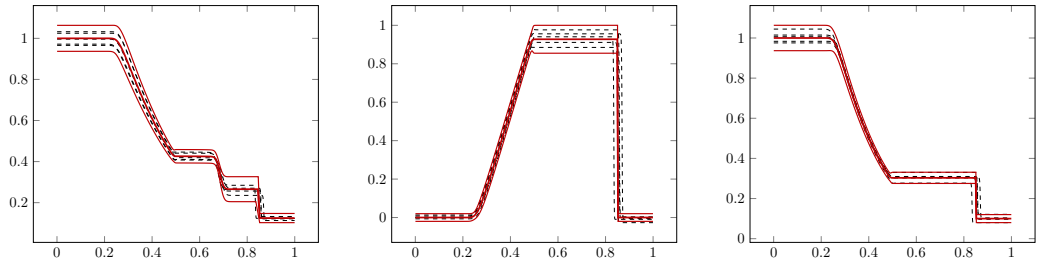
(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

Figure 17: SA approach without correction. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines - zoom.



(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

Figure 18: SA approach with correction. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines



(a) $\rho(x,T)$ (b) $u(x,T)$ (c) $p(x,T)$

Figure 19: SA approach with correction, anti-diffusive scheme. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines

to the quasi 1D Euler system and we are dealing with some applications, such as optimization and uncertainty quantification: the results obtained in those frameworks show the importance of the correction term.

References

[1] R. Abgrall and P. M. Congedo. A semi-intrusive deterministic approach to uncertainty quantification in non-linear fluid flow problems. *J. Comput. Physics*, 2012.

[2] J. R. Appel. *Sensitivity calculations for conservation laws with application to discontinuous fluid flows*. PhD thesis, PhD thesis, Virginia Polytechnic Institute and State University, 1997.

[3] C. Bardos and O. Pironneau. A formalism for the differentiation of conservation laws. *Compte rendu de l'Académie des Sciences*, 335(10):839–845, 2002.

[4] J. Borggaard and J. Burns. A PDE sensitivity equation method for optimal aerodynamic design. *Journal of Computational Physics*, 136(2):366 – 384, 1997.

[5] C. Chalons, R. Duvigneau, and C. Fiorini. Sensitivity analysis and numerical diffusion effects for hyperbolic PDE systems with discontinuous solutions. The case of barotropic Euler equations in Lagrangian coordinates. *SIAM Journal on Scientific Computing*, 2018. To appear.

[6] C. Chalons and P. Goatin. Godunov scheme and sampling technique for computing phase transitions in traffic flow modeling. *Interfaces and Free Boundaries*, 10(2):197–221, 2008.

[7] A. J. Chorin. Random choice solution of hyperbolic systems. *Journal of Computational Physics*, 22(4):517–533, 1976.

[8] P. R. Christian and G. Casella. Monte Carlo statistical methods, 1999.

[9] J.-M. Clarisse, S. Jaouen, and P.-A. Raviart. A Godunov-type method in Lagrangian coordinates for computing linearly-perturbed planar-symmetric flows of gas dynamics. *Journal of Computational Physics*, 198(1):80 – 105, 2004.

[10] C. Delenne. Propagation de la sensibilité dans les modèles hydrodynamiques., 2014.

[11] B. Després, G. Poëtte, and D. Lucor. *Robust Uncertainty Propagation in Systems of Conservation Laws with the Entropy Closure Method*, pages 105–149. Springer International Publishing, Cham, 2013.

[12] R. Duvigneau and D. Pelletier. A sensitivity equation method for fast evaluation of nearby flows and uncertainty analysis for shape parameters. *Int. J. of CFD*, 20(7):497–512, August 2006.

[13] R. Duvigneau, D. Pelletier, and J. Borggaard. An improved continuous sensitivity equation method for optimal shape design in mixed convection. *Numerical Heat Transfer part B : Fundamentals*, 50(1):1–24, July 2006.

[14] J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Communications on pure and applied mathematics*, 18(4):697–715, 1965.

[15] E. Godlewski and P.-A. Raviart. The linearized stability of solutions of nonlinear hyperbolic systems of conservation laws: A general numerical approach. *Mathematics and Computers in Simulation*, 50(1):77 – 95, 1999.

[16] V. Guinot. Upwind finite volume solution of sensitivity equations for hyperbolic systems of conservation laws with discontinuous solutions. *Computers & Fluids*, 38(9):1697–1709, 2009.

[17] V. Guinot, C. Delenne, and B. Cappelaere. An approximate riemann solver for sensitivity equations with discontinuous solutions. *Advances in Water Resources*, 32(1):61–77, 2009.

[18] V. Guinot, M. Leménager, and B. Cappelaere. Sensitivity equations for hyperbolic conservation law-based flow models. *Advances in water resources*, 30(9):1943–1961, 2007.

[19] M. D. Gunzburger. *Perspectives in flow control and optimization*, volume 5. Siam, 2003.

[20] A. Harten and J. M. Hyman. Self adjusting grid methods for one-dimensional hyperbolic conservation laws. *Journal of computational Physics*, 50(2):235–269, 1983.

[21] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM review*, 25(1):35–61, 1983.

[22] H. Hristova, S. Etienne, D. Pelletier, and J. Borggaard. A continuous sensitivity equation method for time-dependent incompressible laminar flows. *Int. J. for Numerical Methods in Fluids*, 50:817–844, 2004.

[23] B. Iooss and P. Lemaître. *A Review on Global Sensitivity Analysis Methods*, pages 101–122. Springer US, 2015.

[24] O.M. Knio and O.P. Le Maitre. Uncertainty propagation in CFD using polynomial chaos decomposition. *Fluid Dynamics Research*, 38(9):616–640, September 2006.

[25] M. Martinelli and R. Duvigneau. On the use of second-order derivative and metamodel-based monte-carlo for uncertainty estimation in aerodynamics. *Computers and Fluids*, 37(6), 2010.

[26] B. Mohammadi and O. Pironneau. *Applied Optimal Shape Design for Fluids*. Oxford University Press, 2001.

[27] M.M. Putko, P.A. Newman, A.C. Taylor, and L.L. Green. Approach for uncertainty propagation and robust design in cfd using sensitivity derivatives. In *15th AIAA Computational Fluid Dynamics Conference*, Anaheim, CA, June 2001.

[28] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of computational physics*, 43(2):357–372, 1981.

[29] E. F. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the HLL-Riemann solver. *Shock waves*, 4(1):25–34, 1994.

[30] É. Turgeon, D. Pelletier, and J. Borggaard. Sensitivity and uncertainty analysis for variable property flows. In *39th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, Jan. 2001.

[31] B. Van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *Journal of computational Physics*, 32(1):101–136, 1979.

[32] R. Walters. Towards stochastic fluid mechanics via polynomial chaos. In *41st AIAA Aerospace Sciences Meeting and Exhibit, Reno, USA*, 2003.

[33] R. W. Walters and L. Huyse. Uncertainty analysis for fluid mechanics with applications. Technical report, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION HAMPTON VA LANGLEY RESEARCH CENTER, 2002.

[34] D.B. Xiu and G.E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of Computational Physics*, (187):137–167, 2003.

A Solution of the Riemann problem

In this appendix, we write the exact solution for the system (19) in a specific case (cf. [2]), which was used as a test case to check the convergence of the numerical schemes proposed. We consider a Riemann problem, i.e.:

$$\mathbf{V}_0(x) = \begin{cases} \mathbf{V}_L & x < x_c, \\ \mathbf{V}_R & x > x_c. \end{cases}$$

The general solution for this kind of problem is quite complicated, especially for the sensitivity (the last three components of $\mathbf{V}$). First, we study the state (the first three components of $\mathbf{V}$): the pair $(\lambda_2, \mathbf{r}_2)$ is linearly degenerated, i.e. $\nabla\lambda_2 \cdot \mathbf{r}_2 = 0$, therefore the middle wave is always a contact discontinuity; concerning the $1-$wave and the $3-$wave, they are genuinely nonlinear therefore they can either be shocks or rarefaction waves. In Figure 20 we show the structure of the state in the case rarefaction-contact-shock. Concerning the sensitivity, it has the same structure as the state (cf. Figure 21 in the case rarefaction-contact-shock): the middle wave is always a contact wave, and the $1-$ and $2-$wave are of the same type as for the state. The only difference is that the sensitivity presents discontinuities in the two extrema of the rarefaction fan (and this is why in Figure 21 the external lines of the rarefaction fan are thicker).

In the following, we illustrate this analysis of the wave structure by giving the detailed solution for the state and for the sensitivity in a specific case. The initial data for the state on the physical variables is the following:

$$\rho_L = 1,\ u_L = 0,\ p_L = 1, \qquad \rho_R = 0.125,\ u_R = 0,\ p_R = 0.1.$$

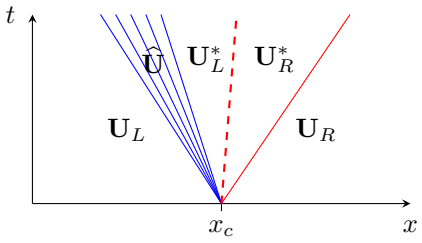


Figure 20: Structure of the solution for the Riemann problem for the state.

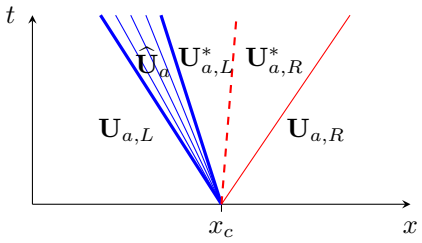


Figure 21: Structure of the solution for the Riemann problem for the sensitivity.

We consider as parameter of interest $a = p_L$, therefore the initial data for the sensitivity is:

$$\rho_{a,L} = \rho_{a,R} = u_{a,L} = u_{a,R} = p_{a,R} = 0, \quad p_{a,L} = 1.$$

This choice of initial data leads to the structure in Figures 20-21, for the state as well as for the sensitivity: the $1-$wave is a rarefaction and the $3-$wave is a shock. For the notation, please refer to Figure 20 for the state and Figure 21 for the sensitivity. Let us now give the exact formulas for the state and for the sensitivity.

State solution: the exact solution for the physical variables is given in [2]. Every variable is given as a function of the pressure in the right-star zone p_R^*, which is computed numerically from the following implicit relation:

$$p_L = p_R^* \left(1 - \frac{(\gamma-1)\frac{c_R}{c_L}\left(\frac{p_R^*}{p_R} - 1\right)}{\sqrt{2\gamma\left(2\gamma + (\gamma+1)(\frac{p_R^*}{p_R} - 1)\right)}} \right)^{-\frac{2\gamma}{\gamma-1}}, \tag{30}$$

where $c_\ell = \sqrt{\dfrac{\gamma p_\ell}{\rho_\ell}}$, with $\ell = L, R$. In the star regions, we have:

$$p_L^* = p_R^* = p^*,$$

$$u_L^* = u_R^* = u^* = c_R \left(\frac{p^*}{p_R} - 1\right) \sqrt{\frac{2}{\gamma(\gamma+1)\frac{p^*}{p_R} + \gamma(\gamma-1)}},$$

because the velocity u and the pressure p are Riemann invariants across the $2-$wave; as for the density ρ, we have:

$$\rho_R^* = \rho_R \frac{p^*}{p_R} \left(\frac{1 + \frac{\gamma-1}{\gamma+1}\frac{p_R}{p^*}}{1 + \frac{\gamma-1}{\gamma+1}\frac{p^*}{p_R}} \right),$$

$$\rho_L^* = \rho_L \left(\frac{p^*}{p_L}\right)^{\frac{1}{\gamma}}.$$

In the rarefaction wave, we have:

$$\hat{u}(x,t) = \frac{2(u^* - u_L)}{(\gamma+1)u^*} \left(\frac{x - x_c}{t}\right) + 2\frac{c_L u^* - u_L\left(c_L - \frac{\gamma+1}{2}u^*\right)}{(\gamma+1)u^*},$$

$$\hat{\rho}(x,t) = \rho_L \left(1 - (\gamma-1)\frac{\hat{u}(x,t)}{2c_L}\right)^{\frac{2}{\gamma-1}},$$

$$\hat{p}(x,t) = p_L \left(1 - (\gamma-1)\frac{\hat{u}(x,t)}{2c_L}\right)^{\frac{2\gamma}{\gamma-1}}.$$

Finally, the solution writes:

$$\mathbf{U}(x,t) = \begin{cases} \mathbf{U}_L & x - x_c < -c_L t, \\ \widehat{\mathbf{U}} & -c_L t < x - x_c < \left(\frac{\gamma+1}{2}u^* - c_L\right) t, \\ \mathbf{U}_L^* & \left(\frac{\gamma+1}{2}u^* - c_L\right) t < x - x_c < u^* t, \\ \mathbf{U}_R^* & u^* t < x - x_c < c_R \sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma}\frac{p^*}{p_R}}\, t, \\ \mathbf{U}_R & x - x_c > c_R \sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma}\frac{p^*}{p_R}}\, t. \end{cases} \tag{31}$$

Sensitivity solution: here we are solving the second part of system 19, that is the one with the source term. The source term was designed in such a way that the solution for the sensitivity is the derivative of the state solution *in the regular zones* and there are no Dirac delta function where the state is discontinuous. Therefore, by differentiating (30) with respect to a, one obtains the following explicit formula for $p_{a,R}^*$:

$$p_{a,R}^* = p_{a,L}^* = p_a^* = \frac{1 + \Theta^{\frac{1-3\gamma}{\gamma-1}}\Xi p^*}{\Theta^{-\frac{2\gamma}{\gamma-1}} + \Theta^{\frac{1-3\gamma}{\gamma-1}}(\Lambda - \Psi)p^*},$$

where:

$$\Theta = 1 - \frac{(\gamma-1)c_R\left(\frac{p^*}{p_R} - 1\right)}{c_L\sqrt{4\gamma^2 + 2\gamma(\gamma-1)\left(\frac{p^*}{p_R} - 1\right)}},$$

$$\Xi = \frac{c_R\left(\frac{p^*}{p_R} - 1\right)c_{a,R}\sqrt{2\gamma}}{c_L^2\sqrt{2\gamma + (\gamma+1)\left(\frac{p^*}{p_R} - 1\right)}},$$

$$\Lambda = \frac{\sqrt{2\gamma}\,c_R}{c_L p_R\sqrt{2\gamma + (\gamma+1)\left(\frac{p^*}{p_R} - 1\right)}},$$

$$\Psi = \frac{\gamma(\gamma+1)c_R\left(\frac{p^*}{p_R} - 1\right)}{c_L p_R\sqrt{2\gamma}\left(2\gamma + (\gamma+1)\left(\frac{p^*}{p_R} - 1\right)\right)^{\frac{3}{2}}}.$$

In the star regions, by differentiating the corresponding state, one finds:

$$u_a^* = \frac{2c_{a,L}}{\gamma-1}\left(1 - \left(\frac{p^*}{p_L}\right)^{\frac{\gamma-1}{2\gamma}}\right) - \frac{c_L}{\gamma}\left(\frac{p^*}{p_L}\right)^{\frac{-\gamma-1}{2\gamma}}\left(\frac{p_L p_a^* - p^*}{p_L^2}\right),$$

$$\rho_{a,R}^* = \frac{\rho_R p_a^*}{p_R}\frac{\left(1 + \frac{\gamma-1}{\gamma+1}\frac{p_R}{p^*}\right)}{\left(1 + \frac{\gamma-1}{\gamma+1}\frac{p^*}{p_R}\right)} + \rho_R\frac{p^*}{p_R}\frac{\gamma-1}{\gamma+1}\left(\frac{-\frac{p_R p_a^*}{p^{*2}}\left(1 + \frac{\gamma-1}{\gamma+1}\frac{p^*}{p_R}\right) - \frac{p_a^*}{p_R}\left(1 + \frac{\gamma-1}{\gamma+1}\frac{p_R}{p^*}\right)}{(1 + \frac{\gamma-1}{\gamma+1}\frac{p^*}{p_R})^2}\right),$$

$$\rho_{a,L}^* = \frac{\rho_L}{\gamma}\frac{p_L p_a^* - p^*}{p_L^2}\left(\frac{p^*}{p_L}\right)^{\frac{1-\gamma}{\gamma}}.$$

Finally, in the rarefaction:

$$\hat{u}_a(x,t) = \frac{2u_L u^*}{(\gamma+1)u^{*2}}\frac{x - x_c}{t} + 2\frac{c_{a,L}u^{*2} - c_{a,L}u_L u^* + c_L u_L u_a^*}{(\gamma+1)u^{*2}},$$

$$\hat{\rho}_a(x,t) = -\rho_L\left(\frac{\hat{u}_a(x,t)c_L - \hat{u}(x,t)c_{a,L}}{c_L^2}\right)\left(1 - \frac{(\gamma-1)\hat{u}(x,t)}{2c_L}\right)^{\frac{3-\gamma}{\gamma-1}},$$

$$\hat{p}_a(x,t) = \left(1 - \frac{(\gamma-1)\hat{u}(x,t)}{2c_L}\right)^{\frac{2\gamma}{\gamma-1}} - p_L\gamma\left(\frac{\hat{u}_a(x,t)c_L - \hat{u}(x,t)c_{a,L}}{c_L^2}\right)\left(1 - \frac{(\gamma-1)\hat{u}(x,t)}{2c_L}\right)^{\frac{\gamma+1}{\gamma-1}}.$$

The sensitivity has the same structure as the state, therefore:

$$\mathbf{U}_a(x,t) = \begin{cases} \mathbf{U}_{a,L} & x - x_c < -c_L t, \\ \widehat{\mathbf{U}}_a\left(\frac{x-x_c}{t}\right) & -c_L t < x - x_c < \left(\frac{\gamma+1}{2}u^* - c_L\right)t, \\ \mathbf{U}^*_{a,L} & \left(\frac{\gamma+1}{2}u^* - c_L\right)t < x - x_c < u^* t, \\ \mathbf{U}^*_{a,R} & u^* t < x - x_c < c_R\sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma}\frac{p^*}{p_R}}\,t, \\ \mathbf{U}_{a,R} & x - x_c > c_R\sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma}\frac{p^*}{p_R}}\,t. \end{cases} \tag{32}$$

We remark that if one writes the Rankine-Hugoniot conditions across the shock one finds:

$$-c_R\sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma}\frac{p^*}{p_R}}\,(\mathbf{U}_{a,R} - \mathbf{U}^*_{a,R}) + \mathbf{F}_a(\mathbf{U}_R, \mathbf{U}_{a,R}) - \mathbf{F}_a(\mathbf{U}^*_R, \mathbf{U}^*_{a,R}) = \mathbf{S}.$$