



**HAL**  
open science

# Sensitivity analysis and numerical diffusion effects for hyperbolic PDE systems with discontinuous solutions. The case of barotropic Euler equations in Lagrangian coordinates

Christophe Chalons, Régis Duvigneau, Camilla Fiorini

► **To cite this version:**

Christophe Chalons, Régis Duvigneau, Camilla Fiorini. Sensitivity analysis and numerical diffusion effects for hyperbolic PDE systems with discontinuous solutions. The case of barotropic Euler equations in Lagrangian coordinates. 2017. hal-01589337v1

**HAL Id: hal-01589337**

**<https://inria.hal.science/hal-01589337v1>**

Preprint submitted on 18 Sep 2017 (v1), last revised 14 Dec 2018 (v2)

**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.

# Sensitivity analysis and numerical diffusion effects for hyperbolic PDE systems with discontinuous solutions. The case of barotropic Euler equations in Lagrangian coordinates

C. Chalons, R. Duvigneau, C. Fiorini

July 20, 2017

## Abstract

Sensitivity analysis (SA) is the study of how the output of a mathematical model is affected by changes in the inputs. SA is widely studied, due to its many applications: uncertainty quantification, quick evaluation of close solutions, and optimization, to name but a few. In this work we show that the classic SA techniques, in particular the continuous sensitivity equation (CSE) method, cannot be used if the mathematical model is a system of hyperbolic partial differential equations (PDEs) with discontinuous solutions. The problem arises from the fact that the CSE method requires the differentiation of the state variable: if the latter is discontinuous, this in turns generates Dirac delta functions in the sensitivity. The focus of the first part of this work is to define a system of sensitivity equations valid also in case of discontinuous state: in order to do that, we add a correction term based on the Rankine-Hugoniot conditions. In the second part of this work we illustrate with some numerical tests how some classical finite volume schemes (an exact Godunov method and a Roe-type method) do not converge to the analytical solution for the sensitivity, due to the numerical diffusion: for this reason, we present an anti-diffusive numerical scheme, which provides the correct results for the sensitivity. In this work we carry out the computation in detail for the barotropic Euler equations in Lagrangian coordinates, but the approach is general and can be applied to any hyperbolic system with discontinuous solutions.

## 1 Introduction

Sensitivity analysis (SA) concerns the quantification of changes in Partial Differential Equations (PDEs) solution due to perturbations in the model input. It is obviously a valuable tool for engineering applications, which allows to quantify the physical response of a system to any change of parameter values (geometry, boundary conditions, etc). The underlying concepts have been used for a long time in optimal design methods, which replace now the traditional “trials and errors” approach, to determine rigorously the optimal parameters of a system. Beside optimization, SA methods can also be carried out to measure the performance loss due to an unexpected perturbation of the operational conditions, in the framework of uncertainty quantification approaches. Finally, SA methods can likewise be employed to monitor and explore interactively neighboring solutions for a negligible computational expense.

Two strategies can be implemented to compute such sensitivities. The first and most popular one is the *adjoint equation method* [16, 19, 20], which introduces additional *adjoint* variables to compute the derivative of any functional output with respect to all input parameters. The

---

This work is an extension of the proceedings [5].

adjoint equation is independent from the input parameters, yielding this approach very efficient for optimization problems involving a large number of design parameters. However, the adjoint equation should be solved backwards in time, which could lead to practical difficulties for unsteady problems. The second strategy is the *sensitivity equation method* [2, 9, 8, 15], which allows to compute the derivative of the PDE solution itself, at any location and time, with respect to a single input parameter. This approach is therefore better suited to the exploration of neighboring solutions. Note that it relies on a forward time integration and, thus, seems to be well adapted to time-dependent hyperbolic PDE systems, which constitute the context of the present work.

However, a major difficulty arises when one considers the application of the sensitivity equation method to hyperbolic PDE systems: in such a context, discontinuities in the solution can appear, even if the initial solution is regular. This corresponds to the generation of shock waves in compressible flows for instance. Such discontinuities in the solution lead to specific issues regarding sensitivity analysis, because they correspond to the presence of Dirac functions in the sensitivity fields. This question has been explored in [1, 19] with a theoretical viewpoint, and more recently in [11, 12] with a numerical viewpoint. Indeed, the capture of the Dirac peak in the sensitivity solution by numerical schemes is intractable in practice. Therefore, a modification of the sensitivity system was proposed in the later references, to remove the Dirac functions from the numerical sensitivity solution. More specifically, a modification of HLL Riemann solver used to evaluate fluxes in a finite-volume method was proposed, in the context of Saint-Venant equations.

In this work, we adopt a similar point of view: we propose to define and approximate numerically a modified system of sensitivity equations which is valid also when the state is discontinuous. To correct some shortcomings appearing in the sensitivity solution, reported in [11, 12], some alternative formulations to define the sensitivity solution update are studied, ranging from Godunov method, first- and second-order Roe-type solvers, to an “anti-diffusive” scheme. These methods are tested numerically in the context of the barotropic Euler equations in Lagrangian coordinates, but the proposed formulations are general and could be used for other hyperbolic systems of conservation laws.

This article is organized as follows: in the first sections, we present the state equations, derive the sensitivity equations and detail the exact resolution of the Riemann problem for the state and sensitivities. Then, the modification of the sensitivity equations to account for the Dirac functions is presented and the implementation for two classical methods (Godunov and Roe-type solvers) is examined. Some numerical tests are conducted, which exhibit grid-convergence issues. Finally, an “anti-diffusive” scheme is introduced to alleviate these difficulties and the accuracy of the proposed approach is demonstrated for some cases.

## 2 Problem description

As already mentioned in the previous section, standard SA methods can be used only if the solution  $\mathbf{U}$  is regular enough [1], which is usually not the case for hyperbolic systems of the general form

$$\begin{cases} \partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, & x \in \mathbb{R}, \quad t > 0, \\ \mathbf{U}(x, 0) = \mathbf{U}_0(x). \end{cases}$$

In fact, it is well known this kind of systems can have discontinuous solutions, regardless of the regularity of the initial condition  $\mathbf{U}_0$ . If the state  $\mathbf{U}$  is discontinuous, the sensitivity  $\mathbf{U}_a = \partial_a \mathbf{U}$  will exhibit Dirac delta functions. Here and throughout this work,  $a$  denotes the parameter of interest which may vary and induce a non trivial sensitivity. We remark that we consider a scalar parameter  $a$  only to simplify the notation: everything we do in this work can be done for a vector of parameters  $\mathbf{a}$  as well, with no added complexity from a theoretical point of view.

In this work, we consider the barotropic Euler equations in Lagrangian coordinates, i.e. the  $p$ -system; however, everything can be extended to any hyperbolic system. The choice of the  $p$ -system is motivated by the fact that, although quite simple, it presents all the main features of hyperbolic systems: this allows us to solve the state problem easily and to focus on the sensitivity problem.

The system writes:

$$\begin{cases} \partial_t \tau - \partial_x u = 0, \\ \partial_t u + \partial_x p(\tau) = 0, \end{cases} \quad (1)$$

where  $\tau > 0$  is the co-volume (i.e.  $\tau = \frac{1}{\rho}$ , a  $\rho$  is the density of the fluid),  $u$  is the Lagrangian velocity and the pressure  $p(\tau)$  is a function only of  $\tau$ . We assume  $p'(\tau) < 0$  and  $p''(\tau) > 0$ . The Jacobian matrix of the system is the following:

$$\mathbf{M}(\tau, u) = \begin{bmatrix} 0 & -1 \\ p'(\tau) & 0 \end{bmatrix} \quad (2)$$

and its eigenvalues are real and distinct  $\lambda_{\pm} = \pm c$ , where  $c = \sqrt{-p'(\tau)}$  is the Lagrangian sound speed. Therefore  $\mathbf{M}$  is  $\mathbb{R}$ -diagonalisable, and (1) is strictly hyperbolic. In this work we will consider  $p(\tau) = \tau^{-\gamma}$ , where  $\gamma = 1.4$  is the heat capacity ratio.

If we consider *smooth* solutions of (1), we can apply the CSE method, differentiate (1) with respect to  $a$  and obtain the following sensitivity equations:

$$\begin{cases} \partial_t \tau_a - \partial_x u_a = 0, \\ \partial_t u_a + \partial_x (p'(\tau) \tau_a) = 0. \end{cases} \quad (3)$$

In order to introduce a more compact notation, we define the state and sensitivity vectors and their fluxes:

$$\mathbf{U} = \begin{bmatrix} \tau \\ u \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} -u \\ p(\tau) \end{bmatrix}, \quad \mathbf{U}_a = \begin{bmatrix} \tau_a \\ u_a \end{bmatrix}, \quad \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = \begin{bmatrix} -u_a \\ p'(\tau) \tau_a \end{bmatrix},$$

and rewrite the systems (1) and (3) in a vectorial form:

$$\begin{cases} \partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \\ \partial_t \mathbf{U}_a + \partial_x \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = 0. \end{cases} \quad (4)$$

The Jacobian matrix of the global system (4) writes:

$$\mathbf{A}(\mathbf{V}) = \frac{\partial \mathbf{G}}{\partial \mathbf{V}} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ p'(\tau) & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ p''(\tau) \tau_a & 0 & p'(\tau) & 0 \end{bmatrix}, \quad \text{with } \mathbf{V} = \begin{bmatrix} \mathbf{U} \\ \mathbf{U}_a \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{F} \\ \mathbf{F}_a \end{bmatrix}.$$

One can remark that  $\mathbf{A}(\mathbf{V})$  is a lower triangular block matrix whose diagonal blocks are identical to each other and to the state system's Jacobian matrix. We observe that the global system (4) has two repeated eigenvalues, the same  $\lambda_{\pm}$  as the original system (1), and that the matrix  $\mathbf{A}(\mathbf{V})$  is not  $\mathbb{R}$ -diagonalisable as soon as  $\tau_a \neq 0$ . This means that the global system (4) is only weakly hyperbolic. Therefore, the system (4) as it is will provide us, in case of discontinuous state  $\mathbf{U}$ , with a sensitivity  $\mathbf{U}_a$  presenting Dirac delta functions, in addition to the usual discontinuity, so that these solutions have to be interpreted in the sense of measures. We refer for instance the reader to the following papers, and the references therein: [4, 7, 10, 17, 18, 22]. However, sensitivities with Dirac delta functions are unusable for many applications. For this reason, we add a correction term to the sensitivity equations, as done in [11]. The definition of a proper correction term is the subject of the next section.

### 3 Source term

In this section, we aim at proposing a new version of (4) which is also valid for discontinuous solutions of the state variable  $\mathbf{U}$ . Recall indeed that (4) has been derived assuming formally that

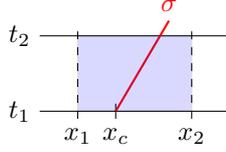


Figure 1: Control volume in light blue.

the solution is smooth whilst hyperbolic equations are well known to develop discontinuities in finite time even for smooth initial data  $\mathbf{U}(x, t = 0) = \mathbf{U}_0(x)$ . In order to compensate the Dirac delta functions that appear in the solutions  $\mathbf{U}_a$  of (4) when  $\mathbf{U}$  is discontinuous, we add to (3) a source term  $\mathbf{S}$  of the following form:

$$\mathbf{S} = \sum_{k=1}^{N_s} \delta_k \boldsymbol{\rho}_k, \quad (5)$$

where  $N_s$  is the number of discontinuities in the state solution  $\mathbf{U}$ ,  $\boldsymbol{\rho}_k$  is the amplitude of the  $k$ -th correction (to be computed), and  $\delta_k$  is the Dirac delta function  $\delta_k = \delta(x - x_{s,k})$ , where  $x_{s,k}$  is the position of the  $k$ -th discontinuity. The new version of (4) we are going to consider thus writes:

$$\begin{cases} \partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \\ \partial_t \mathbf{U}_a + \partial_x \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = \mathbf{S}. \end{cases} \quad (6)$$

Let us motivate our choice and define  $\boldsymbol{\rho}_k$  by considering a control volume  $(x_1, x_2) \times (t_1, t_2)$  as in Figure 1, which contains only the  $k$ -th discontinuity, propagating at speed  $\sigma_k$ . We integrate the second equation of (6) over the control volume:

$$\int_{x_1}^{x_2} \int_{t_1}^{t_2} \partial_t \mathbf{U}_a + \partial_x \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) dx dt = \int_{x_1}^{x_2} \int_{t_1}^{t_2} \delta(x - x_{s,k}) \boldsymbol{\rho}_k(t) dx dt,$$

and we obtain

$$\int_{x_1}^{x_2} \mathbf{U}_a(x, t_2) - \mathbf{U}_a(x, t_1) dx + \int_{t_1}^{t_2} \mathbf{F}_a(x_2, t) - \mathbf{F}_a(x_1, t) dt = \int_{t_1}^{t_2} \boldsymbol{\rho}_k(t) dt, \quad (7)$$

where we used the simplified notation  $\mathbf{F}_a(x, t) = \mathbf{F}_a(\mathbf{U}(x, t), \mathbf{U}_a(x, t))$ . We divide (7) by  $(t_2 - t_1)$  and as the size of the control volume tends to zero we have:

$$\boldsymbol{\rho}_k(t) = (\mathbf{U}_a^- - \mathbf{U}_a^+) \sigma_k + \mathbf{F}_a^+ - \mathbf{F}_a^-, \quad (8)$$

where the plus (respectively minus) stands for the value of the variables to the right (respectively left) of the discontinuity. The relation (8) gives a natural meaning of  $\boldsymbol{\rho}_k$  in terms of a defect measure of the Rankine-Hugoniot conditions for (3). Now, we want to define  $\boldsymbol{\rho}_k$  so that the new model with the source term is valid also in the case of discontinuous state. We start from the Rankine-Hugoniot conditions for the state variable  $\mathbf{U}$  across a discontinuity:

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

and we suggest to differentiate them with respect to the parameter  $a$ . As we do that, we should consider the fact that the Rankine-Hugoniot conditions are valid only at the discontinuity location  $x_{k,s}(t)$ , which depends on the parameter  $a$ . Therefore, we obtain

$$\begin{aligned} & (\mathbf{U}_a^- - \mathbf{U}_a^+) \sigma_k + (\mathbf{U}^- - \mathbf{U}^+) \sigma_{k,a} + \sigma_k (\nabla \mathbf{U}^+ - \nabla \mathbf{U}^-) \partial_a x_{k,s}(t) = \\ & = \mathbf{F}_a^- - \mathbf{F}_a^+ + \left( \frac{\partial \mathbf{F}(\mathbf{U}^+)}{\partial \mathbf{U}} \nabla \mathbf{U}^+ - \frac{\partial \mathbf{F}(\mathbf{U}^-)}{\partial \mathbf{U}} \nabla \mathbf{U}^- \right) \partial_a x_{k,s}(t). \end{aligned}$$

The terms depending on  $\partial_a x_{k,s}(t)$  are very difficult to estimate. However we remark that, thanks to the presence of the gradients, they are zero if we consider that the solution  $\mathbf{U}$  is constant in the left and right neighbourhoods of the shock. This is verified in a standard first order finite volume approach. We obtain therefore a simpler formula:

$$(\mathbf{U}_a^- - \mathbf{U}_a^+) \sigma_k + (\mathbf{U}^- - \mathbf{U}^+) \sigma_{k,a} = \mathbf{F}_a^- - \mathbf{F}_a^+, \quad (9)$$

with  $\sigma_{k,a} = \partial_a \sigma_k$ . Comparing the latter with (8), one is thus led to set

$$\boldsymbol{\rho}_k(t) = \sigma_{k,a} (\mathbf{U}^+ - \mathbf{U}^-). \quad (10)$$

Our choice is of course valid for each  $k$ -discontinuity of the state solution, leading us to definition (5) where the sum is taken over the number of discontinuities.

Note also that by construction, if a triple  $(\mathbf{U}^-, \mathbf{U}^+, \sigma)$  is associated with an admissible discontinuity with a left (respectively right) state  $\mathbf{U}^-$  (respectively  $\mathbf{U}^+$ ) and  $\sigma$  is the speed of propagation, then the triple  $(\mathbf{U}_a^-, \mathbf{U}_a^+, \sigma)$  with  $\mathbf{U}_a^- = \partial_a \mathbf{U}^-$  and  $\mathbf{U}_a^+ = \partial_a \mathbf{U}^+$  is also admissible in the sense that it satisfies the generalised Rankine-Hugoniot relations imposed by  $\mathbf{S}$ . In other words, the sensitivity solution of (6) is obtained by differentiating the state solution with respect to  $a$  when the solution is smooth or discontinuous with constant left and right states. As far as the initial condition is concerned, we have:

$$\begin{pmatrix} \mathbf{U}(x, t=0) \\ \mathbf{U}_a(x, t=0) \end{pmatrix} = \begin{pmatrix} \mathbf{U}_0(x) \\ \partial_a \mathbf{U}_0(x) \end{pmatrix}.$$

## 4 Exact solution of the Riemann problem

In this section, we present the exact resolution of the Riemann problem for the state and the sensitivity, associated with the initial data:

$$\mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L & x < x_c, \\ \mathbf{U}_R & x > x_c, \end{cases} \quad \mathbf{U}_a(x, 0) = \begin{cases} \mathbf{U}_{a,L} & x < x_c, \\ \mathbf{U}_{a,R} & x > x_c, \end{cases}$$

for a given  $x_c$ . First, we compute the solution of the state system (1), which is well-known but necessary to solve (4). Then we derive it with respect to the parameter of interest  $a$  to obtain the sensitivity. As we will see, the sensitivity exhibits interesting and non trivial behaviours.

### 4.1 The state variable

We recall that the eigenvalues of the Jacobian matrix of the state system are:

$$\lambda_1(\mathbf{U}) = -\sqrt{-p'(\tau)} \quad \text{and} \quad \lambda_2(\mathbf{U}) = \sqrt{-p'(\tau)},$$

and the eigenvectors:

$$\mathbf{r}_1(\mathbf{U}) = \begin{bmatrix} \frac{2\sqrt{-p'(\tau)}}{p''(\tau)} \\ \frac{2p'(\tau)}{p''(\tau)} \end{bmatrix} \quad \text{and} \quad \mathbf{r}_2(\mathbf{U}) = \begin{bmatrix} -\frac{2\sqrt{-p'(\tau)}}{p''(\tau)} \\ \frac{2p'(\tau)}{p''(\tau)} \end{bmatrix},$$

which are chosen in such a way that  $\nabla \lambda_i \cdot \mathbf{r}_i = 1$ . Since the couples  $(\lambda_i, \mathbf{r}_i)$  are both genuinely non linear, the waves associated can be either shocks or rarefaction waves. The structure of the analytical solution of the state is resumed in Figure 2 and it consists of two waves, whose speeds can be computed exactly.

In order to give more details on this structure, which will be necessary to explain the structure of the sensitivity, let us consider the plane  $(\tau, u)$  and the points  $\mathbf{U}_L$  and  $\mathbf{U}_R$ : starting from  $\mathbf{U}_L$

we need to reach  $\mathbf{U}_R$  passing from an intermediate state  $\mathbf{U}^*$  using shocks and rarefaction waves, see Figure 2 for the notations. First, we compute which points  $\mathbf{U}$  are reachable through a shock of speed  $\sigma$  from  $\mathbf{U}_L$ . Across a shock, the Rankine-Hugoniot conditions are valid, therefore:

$$\begin{cases} -u + u_L = \sigma(\tau - \tau_L) \\ p(\tau) - p(\tau_L) = \sigma(u - u_L). \end{cases} \quad (11)$$

Finding  $\sigma$  from the first equation and replacing it in the second one, one has:

$$(u - u_L)^2 = -(p(\tau) - p(\tau_L))(\tau - \tau_L),$$

and we observe that the right-hand side is always positive because  $p'(\tau) < 0$ , therefore:

$$u = u_L \pm \sqrt{-(p(\tau) - p(\tau_L))(\tau - \tau_L)}. \quad (12)$$

In order to chose the sign in (12), we use the Lax conditions:

$$\lambda_1(\mathbf{U}) < \sigma < \lambda_1(\mathbf{U}_L),$$

which implies that  $\sigma$  is negative (because  $\lambda_1$  is negative), and:

$$-\sqrt{-p'(\tau)} < -\sqrt{-p'(\tau_L)} \Rightarrow \tau < \tau_L,$$

where we used the hypothesis  $p''(\tau) > 0$ . Both  $\sigma$  and  $(\tau - \tau_L)$  are negative, therefore their product is positive and from the first equation of (11) we can say that  $u_L > u$ . We can conclude that the sign in (12) is a minus and therefore the points reachable from  $\mathbf{U}_L$  through a shock are on the curve of equation:

$$u = u_L - \sqrt{-(p(\tau) - p(\tau_L))(\tau - \tau_L)}. \quad (13)$$

Repeating everything for the 2-wave one finds that the states  $\mathbf{U}$  reachable from  $\mathbf{U}_R$  through a shock are those on the curve of equation:

$$u = u_R + \sqrt{-(p(\tau) - p(\tau_R))(\tau - \tau_R)}. \quad (14)$$

We now repeat the same procedure to compute the set of points reachable through a rarefaction wave from  $\mathbf{U}_L$  and from  $\mathbf{U}_R$  as we did for the shock.

The self-similar solution  $\xi \mapsto \tilde{\mathbf{U}}_1(\xi)$  in a 1-rarefaction wave starting from  $\mathbf{U}_L$  respects the following equation:

$$\begin{cases} \tilde{\mathbf{U}}_1'(\xi) = \mathbf{r}_1(\tilde{\mathbf{U}}_1(\xi)), \\ \tilde{\mathbf{U}}_1(\xi_0) = \mathbf{U}_L, \end{cases} \quad (15)$$

with  $\xi = \lambda_1(\mathbf{U}(\xi))$  for all  $\xi$ , which gives the following system:

$$\begin{cases} \tilde{\tau}_1'(\xi) = \frac{2\sqrt{-p'(\tilde{\tau}_1)}}{p''(\tilde{\tau}_1)}, & \tilde{\tau}_1(\xi_0) = \tau_L, \\ \tilde{u}_1'(\xi) = -\frac{2p'(\tilde{\tau}_1)}{p''(\tilde{\tau}_1)}, & \tilde{u}_1(\xi_0) = u_L. \end{cases} \quad (16)$$

The first equation can be rewritten as

$$(p'(\tilde{\tau}_1))' = 2\sqrt{-p'(\tilde{\tau}_1)},$$

and its solution writes  $p'(\tilde{\tau}_1) = -\xi^2$ , therefore:

$$\tilde{\tau}_1 = (p')^{-1}(-\xi^2).$$

After a change of variable, the second equation writes:

$$\frac{d\tilde{u}_1}{d\tilde{\tau}_1} = \sqrt{-p'(\tilde{\tau}_1)},$$

hence, one has the curve of points reachable through a 1-rarefaction starting from  $\mathbf{U}_L$ :

$$\tilde{u}_1 = u_L + \int_{\tau_L}^{\tilde{\tau}_1} \sqrt{-p'(\tau)} d\tau.$$

If  $p(\tau) = \tau^{-\gamma}$ , one would find:

$$\tilde{\tau}_1 = \left( \frac{\gamma t^2}{(x - x_c)^2} \right)^{\frac{1}{\gamma+1}}, \quad \tilde{u}_1 = u_L + \frac{2\sqrt{\gamma}}{1-\gamma} \left( \tilde{\tau}_1^{\frac{1-\gamma}{2}} - \tau_L^{\frac{1-\gamma}{2}} \right). \quad (17)$$

Repeating exactly the same, one can find the points reachable through a 2-rarefaction starting from  $\mathbf{U}_R$ :

$$\tilde{u}_2 = u_R + \int_{\tilde{\tau}_2}^{\tau_R} \sqrt{-p'(\tau)} d\tau,$$

hence for our choice of  $p$ :

$$\tilde{u}_2 = u_R + \frac{2\sqrt{\gamma}}{1-\gamma} \left( \tau_R^{\frac{1-\gamma}{2}} - \tilde{\tau}_2^{\frac{1-\gamma}{2}} \right), \quad \tilde{\tau}_2 = \left( \frac{\gamma t^2}{(x - x_c)^2} \right)^{\frac{1}{\gamma+1}}. \quad (18)$$

Finally, in order to compute  $\mathbf{U}^*$  we define the two following 1- and 2-wave curves:

$$g_1(\tau; \mathbf{U}_L) = \begin{cases} u_L - \sqrt{-(\tau^{-\gamma} - \tau_L^{-\gamma})(\tau - \tau_L)} & \text{if } \tau \leq \tau_L, \\ u_L + \frac{2\sqrt{\gamma}}{1-\gamma} (\tau^{\frac{1-\gamma}{2}} - \tau_L^{\frac{1-\gamma}{2}}) & \text{if } \tau > \tau_L, \end{cases} \quad (19)$$

$$g_2(\tau; \mathbf{U}_R) = \begin{cases} u_R + \sqrt{-(\tau^{-\gamma} - \tau_R^{-\gamma})(\tau - \tau_R)} & \text{if } \tau \leq \tau_R, \\ u_R + \frac{2\sqrt{\gamma}}{1-\gamma} (\tau_R^{\frac{1-\gamma}{2}} - \tau^{\frac{1-\gamma}{2}}) & \text{if } \tau > \tau_R, \end{cases} \quad (20)$$

which are smooth functions whose derivatives with respect to  $\tau$  will be denoted  $g'_i$ . The intermediate state  $\tau^*$  is defined as the intersection between  $g_1$  and  $g_2$ , and one has:  $u^* = g_1(\tau^*; \mathbf{U}_L) = g_2(\tau^*; \mathbf{U}_R)$ . Newton's method can be used to compute  $\tau^*$ . We remark that there is no intersection between  $g_1$  and  $g_2$  under the following condition:

$$u_L - u_R < \frac{2\sqrt{\gamma}}{1-\gamma} \left( \tau_L^{\frac{1-\gamma}{2}} + \tau_R^{\frac{1-\gamma}{2}} \right).$$

## 4.2 The sensitivity variable

As already explained, to compute the sensitivity we differentiate with respect to  $a$  the state solution, this means that the initial data for the state and the sensitivity are linked by the following relation:

$$\mathbf{U}_{a,L} = \frac{\partial \mathbf{U}_L}{\partial a} \quad \mathbf{U}_{a,R} = \frac{\partial \mathbf{U}_R}{\partial a}.$$

Furthermore, the sensitivity has the same two-wave structure as the state and the waves travel at the same speed as for the state. Therefore, we need to compute the derivative of  $\mathbf{U}^*$  and  $\tilde{\mathbf{U}}$  and this concludes the computation of the analytical sensitivity. To compute  $\tau_a^*$  and  $u_a^*$ , we differentiate, with respect to  $a$  the following equality:

$$g_1(\tau^*; \mathbf{U}_L) = g_2(\tau^*; \mathbf{U}_R)$$

and we obtain (recall that  $\mathbf{U}_L = (\tau_L, u_L)^T$  and  $\mathbf{U}_R = (\tau_R, u_R)^T$ ):

$$\begin{aligned} u_a^* &= g'_1(\tau^*; \mathbf{U}_L)\tau_a^* + \frac{\partial g_1}{\partial \tau_L}(\tau^*; \mathbf{U}_L)\tau_{a,L} + \frac{\partial g_1}{\partial u_L}(\tau^*; \mathbf{U}_L)u_{a,L} = \\ &= g'_2(\tau^*; \mathbf{U}_R)\tau_a^* + \frac{\partial g_2}{\partial \tau_R}(\tau^*; \mathbf{U}_R)\tau_{a,R} + \frac{\partial g_2}{\partial u_R}(\tau^*; \mathbf{U}_R)u_{a,R}, \end{aligned}$$

which gives an explicit solution (although dependent on  $\mathbf{U}^*$ ) for  $\tau_a^*$ :

$$\tau_a^* = \frac{\frac{\partial g_2}{\partial \tau_R}(\tau^*; \mathbf{U}_R)\tau_{a,R} + \frac{\partial g_2}{\partial u_R}(\tau^*; \mathbf{U}_R)u_{a,R} - \frac{\partial g_1}{\partial \tau_L}(\tau^*; \mathbf{U}_L)\tau_{a,L} - \frac{\partial g_1}{\partial u_L}(\tau^*; \mathbf{U}_L)u_{a,L}}{g'_1(\tau^*; \mathbf{U}_L) - g'_2(\tau^*; \mathbf{U}_R)}.$$

Finally, we differentiate the state solution in the rarefaction  $\tilde{\mathbf{U}}$  given by (17)-(18).

### 4.3 Examples

In the numerical section of this work, we will consider as parameter of interest the initial data, which means that  $a$  can either be  $\tau_L$ ,  $u_L$ ,  $\tau_R$ ,  $u_R$  or a combination of them, and from (17) and (18) one can observe that  $\tilde{\tau}_i$  does not depend on those parameters, therefore:

$$\tilde{\tau}_{a,i} = \frac{\partial \tilde{\tau}_i}{\partial a} = 0 \quad i = 1, 2.$$

Concerning the sensitivity of the velocity, one obtains:

$$\tilde{u}_{a,1} = \frac{\partial \tilde{u}_1}{\partial a} = u_{a,L} - \sqrt{\gamma} \tau_L^{-\frac{1+\gamma}{2}} \tau_{a,L}, \quad \tilde{u}_{a,2} = \frac{\partial \tilde{u}_2}{\partial a} = u_{a,R} + \sqrt{\gamma} \tau_R^{-\frac{1+\gamma}{2}} \tau_{a,R}.$$

Interestingly, we remark that the sensitivity is constant in the rarefaction zone of the state variable, which means that for the sensitivity this zone corresponds to at most two discontinuities propagating with velocities given by the extreme left and right velocities of the rarefaction in the state variable, see Figure 3. This simplification is due to the fact that we are considering a reduced Euler system, under barotropic conditions (cf. [13]). In particular, there are two cases:

- (i) if the state presents a 1-rarefaction (respectively a 2-rarefaction) and the parameter of interest  $a$  is  $\tau_L$  (respectively,  $\tau_R$ ), the wave associated with the rarefaction in the sensitivity splits in two discontinuities, as explained above (cf. Figure 4).
- (ii) if the parameter of interest is  $u_L$  (or  $u_R$ ) we have  $\tilde{u}_{a,1} = u_{a,L}$  and  $\tilde{u}_{a,2} = u_{a,R}$ , therefore the wave associated with the rarefaction becomes a single discontinuity for the sensitivity, travelling at the more internal velocity of the state rarefaction wave (cf. Figure 5).

## 5 Classical numerical schemes

The aim of this section is to design relevant numerical schemes for (6). As we will see, this task is not easy and requires a nice discretisation of  $\mathbf{S}$  in order to avoid Dirac delta functions and it is necessary to control numerical diffusion across the shocks where this term is active. Only under these conditions we will get a perfect agreement between exact and numerical solutions. Let us first introduce our notation, although quit classical: we use a constant space step  $\Delta x$  and a varying time step  $\Delta t^n$ . The mesh interfaces are denoted  $x_{j+1/2} = j\Delta x$ , the cells  $C_j = [x_{j-1/2}, x_{j+1/2}]$ , the cell centres  $x_j$  and the intermediate times  $t^{n+1} = t^n + \Delta t^n$ , where  $\Delta t^n$  is chosen according to the usual CFL condition. In the following subsections, we will briefly introduce two classical schemes for the state, Godunov and Roe method, and we will adapt them to the sensitivity.

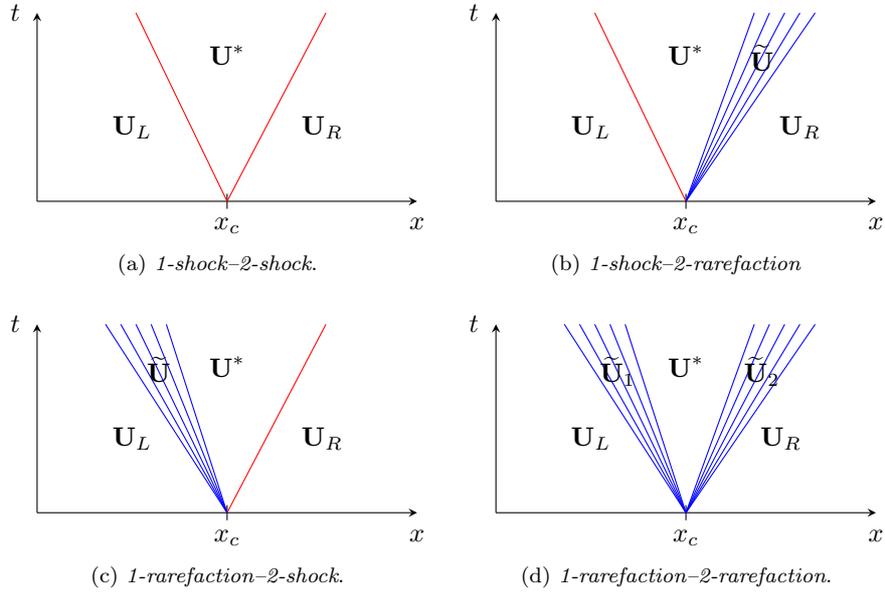


Figure 2: Configurations for the state variable  $\mathbf{U}$ .

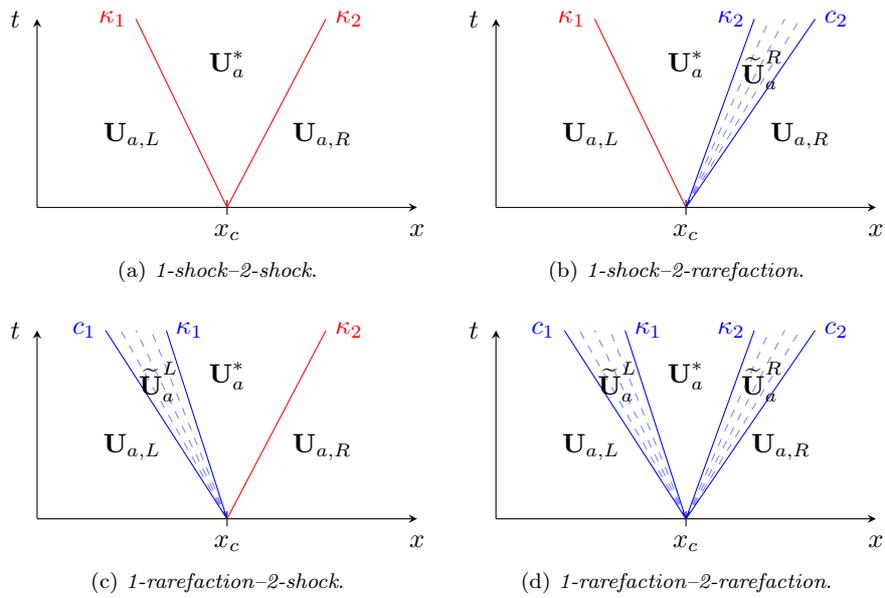


Figure 3: Corresponding configurations for the sensitivity  $\mathbf{U}_a$ .

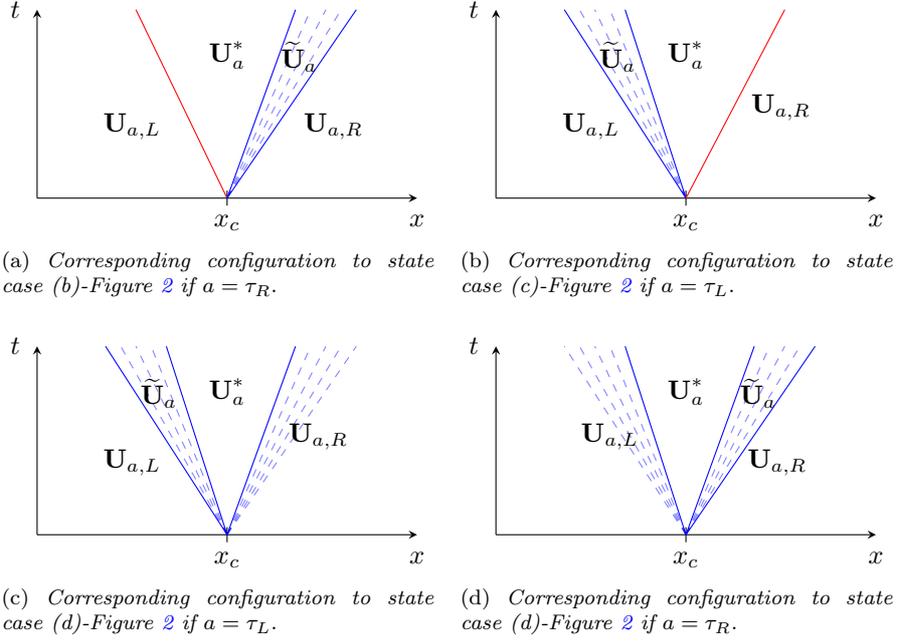


Figure 4: Corresponding configurations for the sensitivity  $\mathbf{U}_a$ , example (i).

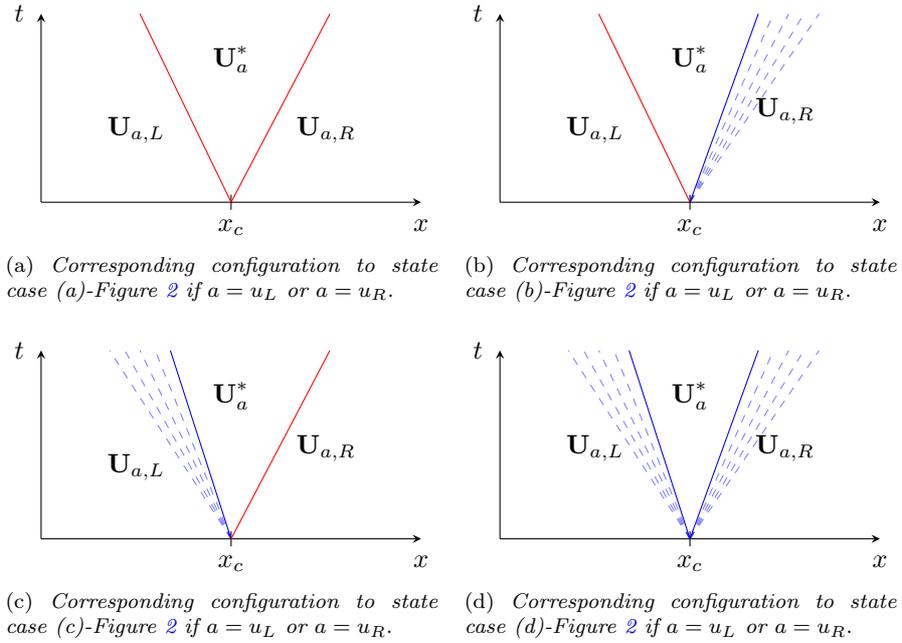


Figure 5: Corresponding configurations for the sensitivity  $\mathbf{U}_a$ , example (ii).

## 5.1 The Godunov method

In this section, we present an exact Godunov-type method. Being the state equations (1) conservative, the classic update formula can be used:

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \frac{\Delta t}{\Delta x} (\mathbf{F}(\mathbf{U}_{j+1/2}^*) - \mathbf{F}(\mathbf{U}_{j-1/2}^*)), \quad (21)$$

where  $\mathbf{U}_{j-1/2}^*$  is the exact intermediate state known implicitly from (19)-(20), with  $\mathbf{U}_L = \mathbf{U}_{j-1}$  and  $\mathbf{U}_R = \mathbf{U}_j$  [21].

The update formula (21) cannot be applied to the sensitivity variable, because of the source term. However, as explained in section 4.2, the structure of the sensitivity is made of discontinuities only. Therefore, we can directly compute the average on each cell, if the slopes of the red lines and the solid blue lines in Figure 3 are known at each interface  $j - 1/2$ . The slopes of the red lines are computed from the Rankine-Hugoniot conditions, while the ones of the blue lines are the eigenvalues evaluated in the correct state. We obtain the following formulas (cf. Figure 3 for the notations):

$$\begin{aligned} \kappa_{1,j-1/2} &= \begin{cases} \frac{u_{j-1/2}^* - u_{j-1}}{\tau_{j-1} - \tau_{j-1/2}^*} & \text{if 1-shock at interface } j - 1/2, \\ \lambda_1(\mathbf{U}_{j-1/2}^*) & \text{if 1-rarefaction at interface } j - 1/2, \end{cases} \\ \kappa_{2,j-1/2} &= \begin{cases} \frac{u_{j-1/2}^* - u_j}{\tau_j - \tau_{j-1/2}^*} & \text{if 2-shock at interface } j - 1/2, \\ \lambda_2(\mathbf{U}_{j-1/2}^*) & \text{if 2-rarefaction at interface } j - 1/2, \end{cases} \\ c_{1,j-1/2} &= \begin{cases} \kappa_{1,j-1/2} & \text{if 1-shock at interface } j - 1/2, \\ \lambda_1(\mathbf{U}_{j-1}) & \text{if 2-rarefaction at interface } j - 1/2, \end{cases} \\ c_{2,j-1/2} &= \begin{cases} \kappa_{2,j-1/2} & \text{if 2-shock at interface } j - 1/2, \\ \lambda_2(\mathbf{U}_j) & \text{if 2-rarefaction at interface } j - 1/2. \end{cases} \end{aligned}$$

Then, the update formula for the sensitivity writes:

$$\begin{aligned} \mathbf{U}_{a,j}^{n+1} &= \mathbf{U}_{a,j}^n + \frac{\Delta t}{\Delta x} \left( \kappa_{2,j-1/2} (\mathbf{U}_{a,j-1/2}^* - \tilde{\mathbf{U}}_{a,j-1/2}^R) + c_{2,j-1/2} (\tilde{\mathbf{U}}_{a,j-1/2}^R - \mathbf{U}_{a,j}^n) \right. \\ &\quad \left. - \kappa_{1,j+1/2} (\mathbf{U}_{a,j+1/2}^* - \tilde{\mathbf{U}}_{a,j+1/2}^L) - c_{1,j-1/2} (\tilde{\mathbf{U}}_{a,j+1/2}^L - \mathbf{U}_{a,j}^n) \right), \end{aligned} \quad (22)$$

where the intermediate states  $\mathbf{U}_{a,j-1/2}^*$ ,  $\mathbf{U}_{a,j+1/2}^*$ ,  $\tilde{\mathbf{U}}_{a,j-1/2}^R$  and  $\tilde{\mathbf{U}}_{a,j+1/2}^L$  are known analytically, from section 4.2. We remark that the source term is encompassed in (22), since (22) comes from the exact Riemann solver of (6).

## 5.2 A Roe-type method

### First order

In this section we illustrate a Roe-type Riemann solver, consisting of three constant states (which we denote  $\mathbf{U}_L$ ,  $\mathbf{U}^*$  and  $\mathbf{U}_R$  for the state, and  $\mathbf{U}_{a,L}$ ,  $\mathbf{U}_a^*$  and  $\mathbf{U}_{a,R}$  for the sensitivity), connected by two discontinuities travelling at velocities

$$\lambda_{L,j-1/2}^{ROE} = -\sqrt{\frac{p(\tau_{j-1}^n) - p(\tau_j^n)}{\tau_{j-1}^n - \tau_j^n}}, \quad \lambda_{R,j-1/2}^{ROE} = \sqrt{\frac{p(\tau_{j-1}^n) - p(\tau_j^n)}{\tau_{j-1}^n - \tau_j^n}}$$

if  $\tau_{j-1}^n \neq \tau_j^n$  and  $\mp \sqrt{-p'(\tau_j^n)}$  otherwise. In the following, we will use the notation  $\lambda_{j-1/2}^{ROE} = \lambda_{R,j-1/2}^{ROE} = -\lambda_{L,j-1/2}^{ROE}$ . The Harten, Lax and van Leer consistency relations [14] for the state at the interface  $j - 1/2$  write:

$$\mathbf{U}_{j-1/2}^* = \frac{1}{2}(\mathbf{U}_{j-1}^n + \mathbf{U}_j^n) - \frac{\mathbf{F}(\mathbf{U}_j^n) - \mathbf{F}(\mathbf{U}_{j-1}^n)}{2\lambda_{j-1/2}^{ROE}}. \quad (23)$$

Since  $\mathbf{U}_{j-1/2}^*$  and  $\lambda_{j-1/2}^{ROE}$  are known at each interface, we can write the following update formula for the state:

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n + \frac{\Delta t}{\Delta x} (\lambda_{j-1/2}^{ROE} (\mathbf{U}_{j-1/2}^* - \mathbf{U}_j^n) + \lambda_{j+1/2}^{ROE} (\mathbf{U}_{j+1/2}^* - \mathbf{U}_j^n)). \quad (24)$$

Writing the integral conditions for the sensitivity with the source term, one obtains:

$$\mathbf{F}_a(\mathbf{U}_j^n, \mathbf{U}_{a,j}^n) - \mathbf{F}_a(\mathbf{U}_{j-1}^n, \mathbf{U}_{a,j-1}^n) - \Delta x \mathbf{S}_{j-1/2}^n = \lambda_{j-1/2}^{ROE} (\mathbf{U}_{a,j}^n + \mathbf{U}_{a,j-1}^n) - 2\lambda_{j-1/2}^{ROE} \mathbf{U}_{a,j-1/2}^*,$$

from which we have the following form for  $\mathbf{U}_{a,j-1/2}^*$ :

$$\mathbf{U}_{a,j-1/2}^* = \frac{\mathbf{U}_{a,j-1}^n + \mathbf{U}_{a,j}^n}{2} - \frac{\mathbf{F}_a(\mathbf{U}_j^n, \mathbf{U}_{a,j}^n) - \mathbf{F}_a(\mathbf{U}_{j-1}^n, \mathbf{U}_{a,j}^n)}{2\lambda_{j-1/2}^{ROE}} + \frac{\Delta x \mathbf{S}_{j-1/2}^n}{2\lambda_{j-1/2}^{ROE}}.$$

The source term is discretised as follows:

$$\mathbf{S}_{j-1/2}^n = \lambda_{a,j-1/2}^{ROE} \left( (\mathbf{U}_{j-1}^n - \mathbf{U}_{j-1/2}^*) \frac{d_{1,j-1}}{\Delta x} + (\mathbf{U}_j^n - \mathbf{U}_{j-1/2}^*) \frac{d_{2,j}}{\Delta x} \right), \quad (25)$$

where  $d_{\ell,j}$  is a shock detector which is equal to 1 if there is an  $\ell$ -shock in the  $j$ -th cell, it is zero elsewhere and  $d_{\ell,j}/\Delta x$  approximates numerically the Dirac  $\delta_k$  in the definition of the source term (5). In this work, we use a very simple shock detector: in section 4.1 we showed that the velocity  $u$  is decreasing across a shock, whilst the co-volume  $\tau$  decreases across a 1-shock, and it increases across a 2-shock. Based on this, we set:

$$d_{1,j} = \begin{cases} 1 & \text{if } u_j > u_{j+1/2}^* \text{ and } \tau_j > \tau_{j+1/2}^*, \\ 0 & \text{otherwise,} \end{cases}$$

$$d_{2,j} = \begin{cases} 1 & \text{if } u_j < u_{j-1/2}^* \text{ and } \tau_j > \tau_{j-1/2}^*, \\ 0 & \text{otherwise.} \end{cases}$$

Finally,  $\mathbf{U}_{a,j-1/2}^*$  is computed as follows:

$$\begin{aligned} \mathbf{U}_{a,j-1/2}^* &= \frac{1}{2} (\mathbf{U}_{a,j-1}^n + \mathbf{U}_{a,j}^n) - \frac{\mathbf{F}_a(\mathbf{U}_j^n, \mathbf{U}_{a,j}^n) - \mathbf{F}_a(\mathbf{U}_{j-1}^n, \mathbf{U}_{a,j}^n)}{2\lambda_{j-1/2}^{ROE}} \\ &\quad + \frac{\lambda_{a,j-1/2}^{ROE}}{2\lambda_{j-1/2}^{ROE}} \left( (\mathbf{U}_{j-1}^n - \mathbf{U}_{j-1/2}^*) d_{1,j-1} + (\mathbf{U}_j^n - \mathbf{U}_{j-1/2}^*) d_{2,j} \right). \end{aligned} \quad (26)$$

We remark that the discretisation of the source term (25) is such that  $\mathbf{U}_{a,j-1/2}^* = \partial_a \mathbf{U}_{j-1/2}^*$ , in fact differentiating (23) with respect to  $a$ , one finds:

$$\partial_a \mathbf{U}_{j-1/2}^* = \frac{\mathbf{U}_{a,j-1}^n + \mathbf{U}_{a,j}^n}{2} - \frac{\mathbf{F}_a(\mathbf{U}_j^n, \mathbf{U}_{a,j}^n) - \mathbf{F}_a(\mathbf{U}_{j-1}^n, \mathbf{U}_{a,j}^n)}{2\lambda_{j-1/2}^{ROE}} + \frac{\mathbf{F}(\mathbf{U}_j^n) - \mathbf{F}(\mathbf{U}_{j-1}^n)}{2\lambda_{j-1/2}^{ROE}} \frac{\lambda_{a,j-1/2}^{ROE}}{\lambda_{j-1/2}^{ROE}}.$$

Using again (23) one has:

$$\mathbf{U}_a^* = \frac{\mathbf{U}_{a,j-1}^n + \mathbf{U}_{a,j}^n}{2} - \frac{\mathbf{F}_a(\mathbf{U}_j^n, \mathbf{U}_{a,j}^n) - \mathbf{F}_a(\mathbf{U}_{j-1}^n, \mathbf{U}_{a,j}^n)}{2\lambda_{j-1/2}^{ROE}} + \frac{\lambda_{a,j-1/2}^{ROE}}{2\lambda_{j-1/2}^{ROE}} (\mathbf{U}_{j-1}^n + \mathbf{U}_j^n - 2\mathbf{U}_{j-1/2}^*),$$

which is equal to (26), once the shock detectors are added. Furthermore, the definition (26) encompasses the source term, which means that we can use the update formula (24) for the sensitivity, too.

Finally, in order to prepare the second order extension in space, we define a residual as follows:

$$\mathbf{R}_j^I(\mathbf{U}^n) = \lambda_{j-1/2}^{ROE} (\mathbf{U}_{j-1/2}^* - \mathbf{U}_j^n) + \lambda_{j+1/2}^{ROE} (\mathbf{U}_{j+1/2}^* - \mathbf{U}_j^n), \quad (27)$$

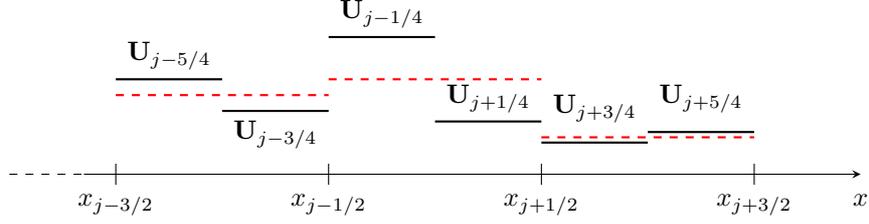


Figure 6: Second order discretisation. In red, the corresponding first order discretisation.

where  $\mathbf{R}_j^I(\mathbf{U}^n)$  is a more compact notation for  $\mathbf{R}^I(\mathbf{U}_{j-1}, \mathbf{U}_j, \mathbf{U}_{j+1})$ . This allows us to write the update formulas in the following way:

$$\begin{cases} \mathbf{U}_j^{n+1} = \mathbf{U}_j^n + \frac{\Delta t}{\Delta x} \mathbf{R}_j^I(\mathbf{U}^n), \\ \mathbf{U}_{a,j}^{n+1} = \mathbf{U}_{a,j}^n + \frac{\Delta t}{\Delta x} \mathbf{R}_j^I(\mathbf{U}_a^n). \end{cases}$$

Furthermore, it will be useful for the numerical schemes introduced hereafter.

## Second order

We extend this scheme to the second order: for the time discretisation we use a two-step Runge-Kutta method, whilst in space we propose a MUSCL-type scheme with some minor modifications in order to have a second order discretisation of the source term. In particular, we remark that (10) is valid only if the solution  $\mathbf{U}$  is locally constant to the left and to the right of the shock, which is true for a first order approximation but not for a second order, in which, classically, the numerical solution is a piecewise affine function. To overcome this problem, we suggest to consider the numerical solution to be a piecewise constant function on half of every cell (cf. [3], section 2.8): the value in the left half (respectively right half) of the  $j$ -th cell is denoted  $U_{j-1/4}^n$  (respectively  $U_{j+1/4}^n$ ), as shown in Figure 6 and they are computed as in a classical MUSCL approach:

$$\mathbf{U}_{j\pm 1/4}^n = \mathbf{U}_j \pm \Delta \mathbf{U}_j^n,$$

and a usual choice for  $\Delta \mathbf{U}_j^n$  is given by a slope-limiter procedure. In this work we use the so-called minmod limiter:

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

where the function minmod is defined as follows:

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

This interpretation of the second order allows us to define the source term as we did for the first order, however we need to consider an additional Riemann problem for each cell. This leads to the following definition of the residual:

$$\mathbf{R}_j^{II}(\mathbf{U}^n) = (\lambda_{j-1/2}^{ROE}(\mathbf{U}_{j-1/2}^* - \mathbf{U}_{j-1/4}^n) + \lambda_{j+1/2}^{ROE}(\mathbf{U}_{j+1/2}^* - \mathbf{U}_{j+1/4}^n)) + \lambda_j^{ROE}(2\mathbf{U}_j^* - \mathbf{U}_{j-1/4} + \mathbf{U}_{j+1/4}),$$

where all the  $\lambda^{ROE}$  and the  $\mathbf{U}^*$  are computed from the extrapolated values  $\mathbf{U}_{j\pm 1/4}^n$ . Finally, the second order scheme writes:

$$\begin{cases} \mathbf{U}_j^{n+1/2} = \mathbf{U}_j^n + \frac{\Delta t}{2\Delta x} \mathbf{R}^{II}(\mathbf{U}^n), \\ \mathbf{U}_j^{n+1} = \mathbf{U}_j^n + \frac{\Delta t}{\Delta x} \mathbf{R}^{II}(\mathbf{U}^{n+1/2}), \end{cases} \quad \begin{cases} \mathbf{U}_{a,j}^{n+1/2} = \mathbf{U}_{a,j}^n + \frac{\Delta t}{2\Delta x} \mathbf{R}^{II}(\mathbf{U}_a^n), \\ \mathbf{U}_{a,j}^{n+1} = \mathbf{U}_{a,j}^n + \frac{\Delta t}{\Delta x} \mathbf{R}^{II}(\mathbf{U}_a^{n+1/2}). \end{cases}$$

## 6 Numerical results

We present some numerical results obtained with the schemes described in the previous section, on a spatial domain is  $(0, 1)$ , with final time  $T = 0.03$ . We consider Riemann problems with  $x_c = 0.5$ .

First, we consider a 1-shock–2-rarefaction case, with the following initial conditions for the state:

$$\mathbf{U}_L = \begin{pmatrix} 0.7 \\ 0 \end{pmatrix}, \quad \mathbf{U}_R = \begin{pmatrix} 0.2 \\ 0 \end{pmatrix}.$$

The parameter of interest is  $a = u_L$ , so that the initial conditions for the sensitivity are:

$$\mathbf{U}_{a,L} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{U}_{a,R} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Figure 7 shows the state variables  $u$  and  $\tau$  and their sensitivities  $u_a$  and  $\tau_a$  at the final time  $T$ . Since the state is a quite classical problem, it is not surprising that all the methods provide very similar solutions one to another. As for the sensitivity, we remark that the modified formulation is able to remove the peak which approximate the Dirac delta function, located at  $x \approx 0.4$  and evident in the scheme without correction term, whose label is “ $\mathbf{S} = \mathbf{0}$ ” in Figures 7-8. However, even with the addition of the source term, the sensitivity solution have two issues: first, the discontinuity associated with the state rarefaction is not well captured; secondly, the value of the plateau in the star zone is not the analytical one. Out of these two problems, the first is the less important one, for two reasons: the fact that the state rarefaction splits into two discontinuity for the sensitivity is typical to the PDEs system considered, it does not happen, for instance, in the case of the complete Euler system; furthermore, the numerical solution converges to the analytical one as  $\Delta x$  goes to 0, meaning that this issue can be solved by using a finer mesh or a higher-order scheme. The second problem is more critical and we believe that numerical diffusion is the cause of it. In Figure 9 we plot the convergence curves of the all the schemes for each variable: all the methods converge as expected for the state variable; however, for the sensitivity the error seems to be convergent for coarser meshes, but it reaches a plateau for finer ones. This can be explained if we split the error into two parts: the part concentrated in the rarefaction zone, which is the bigger one in the coarse meshes, converges; however when this part reaches the same order of magnitude as the error in the star zone, which is constant, the plateau is reached.

The second test case here presented is an isolated 2-shock for the state as well as for the sensitivity. In order to have an isolated shock we choose the following initial data:

$$\mathbf{U}_L = \begin{pmatrix} 0.2 \\ g_2(\tau_L; \mathbf{U}_R) \end{pmatrix} \simeq \begin{pmatrix} 0.2 \\ -1.56 \end{pmatrix}, \quad \mathbf{U}_R = \begin{pmatrix} 0.5 \\ -3 \end{pmatrix},$$

where  $g_2$  is the 2–wave curve defined in (20). As parameter of interest  $a$  we choose the arc length of the curve  $g_2$ , which yields the following initial data for the sensitivity:

$$\mathbf{U}_{a,L} = \begin{pmatrix} 1 \\ g_2'(\tau_L; \mathbf{U}_R) \end{pmatrix} \simeq \begin{pmatrix} 1 \\ -9.35 \end{pmatrix}, \quad \mathbf{U}_{a,R} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Figure 10 shows the results for the state and the sensitivity obtained with a mesh  $\Delta x = 10^{-3}$ : one can notice a spurious wave in the state which does not affect the value in the star zone. However, in the sensitivity this spurious wave is amplified; moreover, the value in the star zone is not correct. Considering the fact that the approximate Riemann solver of Roe is exact in the case of an isolated shock (as well as the exact Godunov solver), the error is necessarily introduced in the average step of the numerical methods and therefore it is due to the numerical diffusion which comes along with the average operation. For this reason, in the next section we introduce a scheme without numerical diffusion in the shock.

## 7 An anti-diffusive Roe-type numerical scheme

Since we believe that the failure of the previous schemes is caused by the numerical diffusion in the shock, we present a scheme which does not have any numerical diffusion in the shock. The

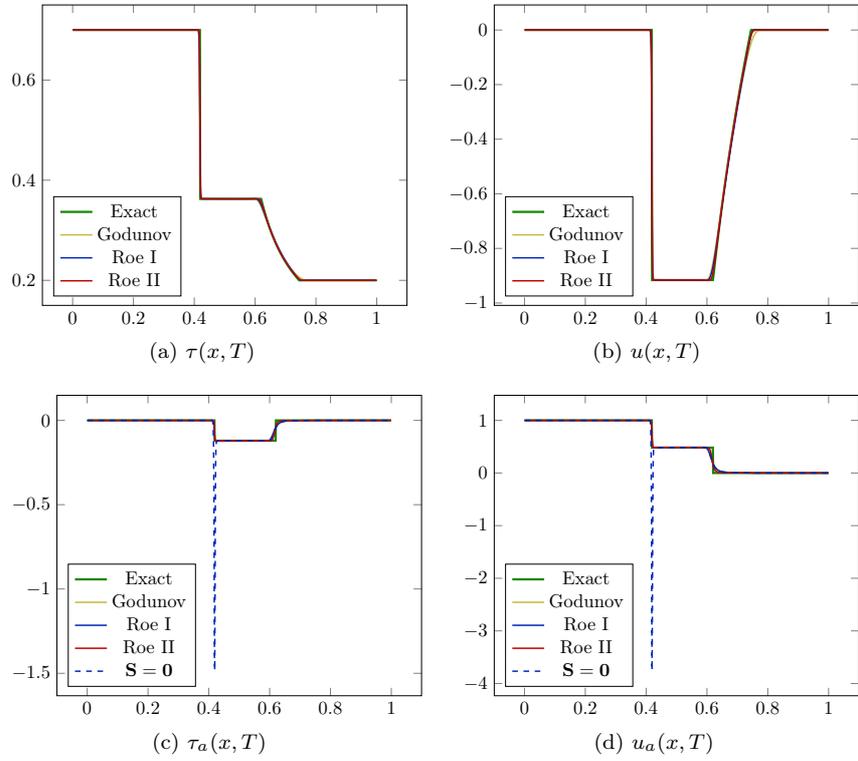


Figure 7: Classical finite volume schemes.

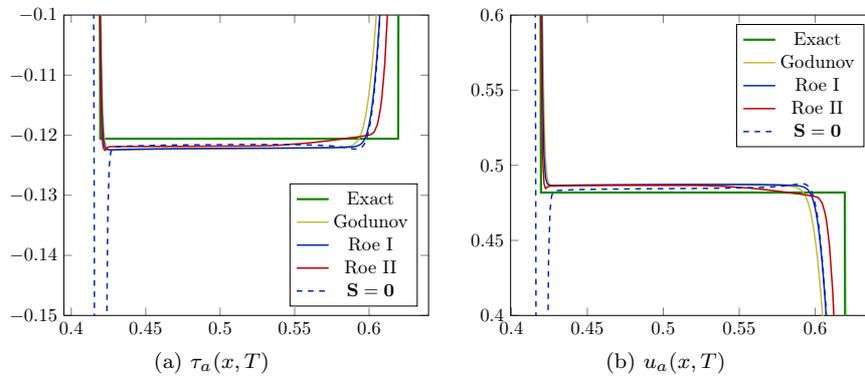


Figure 8: Classical finite volume schemes for sensitivities - zoom.

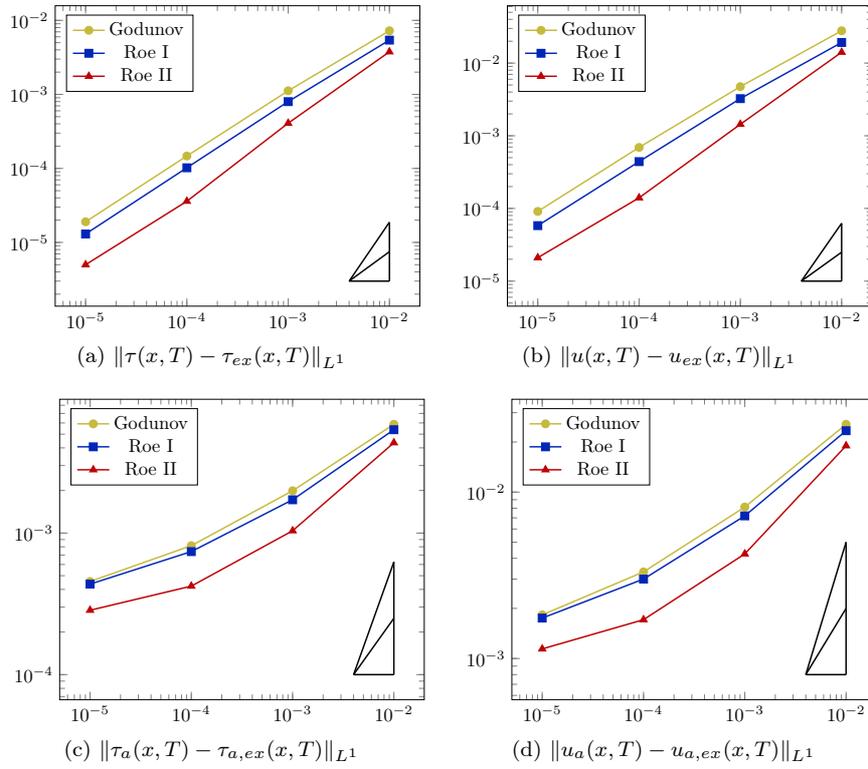


Figure 9: Convergence of the classical finite volume schemes.

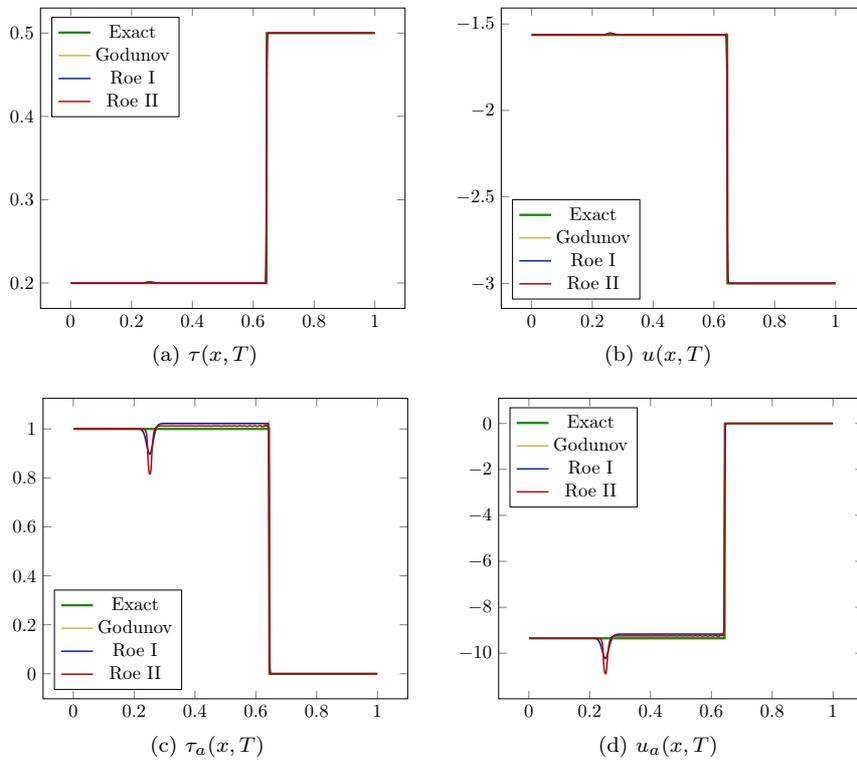


Figure 10: Test case: isolated shock.

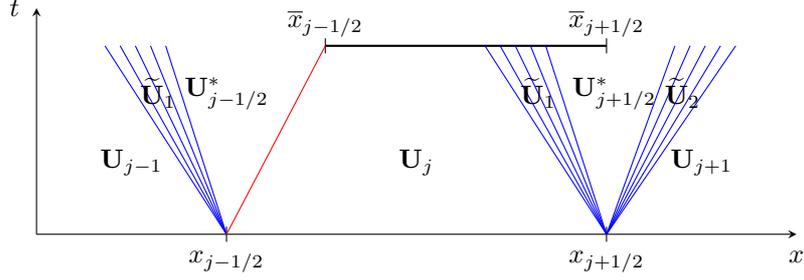


Figure 11: Definition of the temporary staggered mesh.

scheme was first introduced in [6] and here we adapt it to the sensitivity problem. It is a modified Godunov method and it can be coupled with any Riemann solver, in this work we couple it with the Roe-type method proposed in the previous section. In fact, the first step is to solve a Riemann problem at each interface, as for a standard Godunov method. The difference between the two methods is in the average step: instead of averaging on the cells  $[x_{j-1/2}, x_{j+1/2}]$ , a new temporary mesh is defined, whose  $j$ -th cell is denoted  $[\bar{x}_{j-1/2}^n, \bar{x}_{j+1/2}^n]$ , and the average is performed on this mesh. The new mesh is non uniform and it is defined as follows:

$$\bar{x}_{j-1/2}^n = x_{j-1/2} + \sigma_{j-1/2}^n \Delta t^n,$$

where  $\sigma_{j-1/2}^n$  is a proper speed and it depends on the problem. The average operation on the modified mesh provides us with a piecewise constant solution on the new mesh, which we denote  $\bar{U}_j^{n+1}$ . The final step of this method is to go back to the initial mesh, i.e. compute  $U_j^{n+1}$  starting from  $\bar{U}_j^{n+1}$ , and this is done using a sampling technique: the value of the solution on the  $j$ -th cell at time  $t^{n+1}$ ,  $U_j^{n+1}$ , is chosen randomly among  $\bar{U}_{j-1}^{n+1}$ ,  $\bar{U}_j^{n+1}$ , and  $\bar{U}_{j+1}^{n+1}$ , in agreement with their rate of presence in the cell. More precisely, given a random sequence  $(\alpha_n)$  varying in  $(0, 1)$ , the choice is the following:

$$U_j^{n+1} = \begin{cases} \bar{U}_{j-1}^{n+1} & \text{if } \alpha_{n+1} \in \left(0, \frac{\Delta t}{\Delta x} \max(\sigma_{j-1/2}^n, 0)\right), \\ \bar{U}_j^{n+1} & \text{if } \alpha_{n+1} \in \left[\frac{\Delta t}{\Delta x} \max(\sigma_{j-1/2}^n, 0), 1 + \frac{\Delta t}{\Delta x} \min(\sigma_{j+1/2}^n, 0)\right), \\ \bar{U}_{j+1}^{n+1} & \text{if } \alpha_{n+1} \in \left[1 + \frac{\Delta t}{\Delta x} \min(\sigma_{j+1/2}^n, 0), 1\right). \end{cases} \quad (28)$$

The sampling technique mimics the classical averaging if  $(\alpha_n)$  is a well distributed random sequence, for instance  $\alpha_n \sim \mathcal{U}(0, 1)$ , or if it is a deterministic low discrepancy sequence, such as the van der Corput sequence (cf. [6]):

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

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

Our choice for  $\sigma_{j+1/2}^n$  is the following:

$$\sigma_{j+1/2}^n = \begin{cases} \lambda_{j+1/2}^n & u_j > u_{j+1} \text{ and } \tau_j < \tau_{j+1}, \\ -\lambda_{j+1/2}^n & u_j > u_{j+1} \text{ and } \tau_j > \tau_{j+1}, \\ 0 & \text{otherwise.} \end{cases} \quad (29)$$

If  $u$  is increasing, which means that a rarefaction is expected, the mesh is not modified, whilst in case of expected shock the mesh follows it: in this way one never performs the average across a shock and therefore there is no numerical diffusion (cf. Figure 11).

**Remark.** Considering only the initial (non moving) mesh, we remark that this method can also be understood as solving the following two-step problem:

$$\begin{cases} \partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) - \sigma \partial_x \mathbf{U} = 0, \\ \partial_t \mathbf{U} + \sigma \partial_x \mathbf{U} = 0. \end{cases} \quad (30)$$

The first step is equivalent to solving the Riemann problems at each interface and performing the average on the initial uniform mesh, whilst the second step is equivalent to the sampling (28).

### First order formulas

Here, as already said, we couple this anti-diffusive approach with the same Roe-type approximate Riemann solver we presented in the previous section, so we define the following residual:

$$\tilde{\mathbf{R}}_j^I(\mathbf{U}^n) = \mathbf{R}_j^I(\mathbf{U}^n) - \sigma_{j-1/2}^n \mathbf{U}_{j-1/2}^* + \sigma_{j+1/2}^n \mathbf{U}_{j+1/2}^*,$$

where in the definition of  $\sigma_{j-1/2}$  we use  $\lambda_{j-1/2} = \lambda_{j-1/2}^{ROE}$ . Then, the scheme writes:

$$\begin{cases} \bar{\mathbf{U}}_j^{n+1} = \frac{\Delta x}{\Delta x_j} \mathbf{U}_j^n + \frac{\Delta t}{\Delta x_j} \tilde{\mathbf{R}}_j^I(\mathbf{U}^n), \\ +(28). \end{cases} \quad (31)$$

### Second order formulas

One can also couple this approach with the second order Roe-type scheme from the previous section. As for the first order, we define the new residual as follows:

$$\tilde{\mathbf{R}}_j^{II}(\mathbf{U}^n) = \mathbf{R}_j^{II}(\mathbf{U}^n) - \sigma_{j-1/2}^n \mathbf{U}_{j-1/2}^* + \sigma_{j+1/2}^n \mathbf{U}_{j+1/2}^*.$$

Then the scheme writes:

$$\begin{cases} \mathbf{U}_j^{n+1/2} = \mathbf{U}_j^n + \frac{\Delta x - \Delta x_j}{\Delta x_j} \mathbf{U}_j^n + \frac{\Delta t}{2\Delta x_j} \tilde{\mathbf{R}}_j^{II}(\mathbf{U}^n), \\ \bar{\mathbf{U}}_j^{n+1} = \mathbf{U}_j^n + \frac{\Delta x - \Delta x_j}{\Delta x_j} \mathbf{U}_j^{n+1/2} + \frac{\Delta t}{\Delta x_j} \tilde{\mathbf{R}}_j^{II}(\mathbf{U}^{n+1/2}), \\ +(28). \end{cases} \quad (32)$$

From the two-step problem (30) point of view, the discretisation (32) is a second order discretisation of the first step followed by the second step, i.e. the sampling technique, which remains unvaried.

## 8 Numerical results of the anti-diffusive method

The results of the anti-diffusive method are shown in Figures 12-13. As one can see, removing the numerical diffusion in the shock for the state variables allows us to be more precise in the definition of the source term which, in turns, provides us with better solution for the sensitivity: the plateau in the star zone is correct. Furthermore, we show in Figure 14 the convergence results of the classical Roe-type schemes with diffusion compared to the same schemes without diffusion: the latter show a good convergence rate even for the sensitivity variables.

We now present another test case with initial data:

$$\mathbf{U}_L = \begin{pmatrix} 0.7 \\ 0 \end{pmatrix}, \quad \mathbf{U}_R = \begin{pmatrix} 0.2 \\ 0 \end{pmatrix}, \quad \mathbf{U}_{a,L} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{U}_{a,R} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

therefore the parameter of interest  $a$  is in this case  $\tau_R$ . The initial data for the state is the same as in the previous test case, meaning that we are in configuration (b) of Figure 2 and, since  $a = \tau_R$ , the rarefaction wave splits into two discontinuities for the sensitivity as shown in Figure 5-(a). For this test case we chose a bigger final time ( $T = 0.07$ ) so that the two extremes of the rarefaction

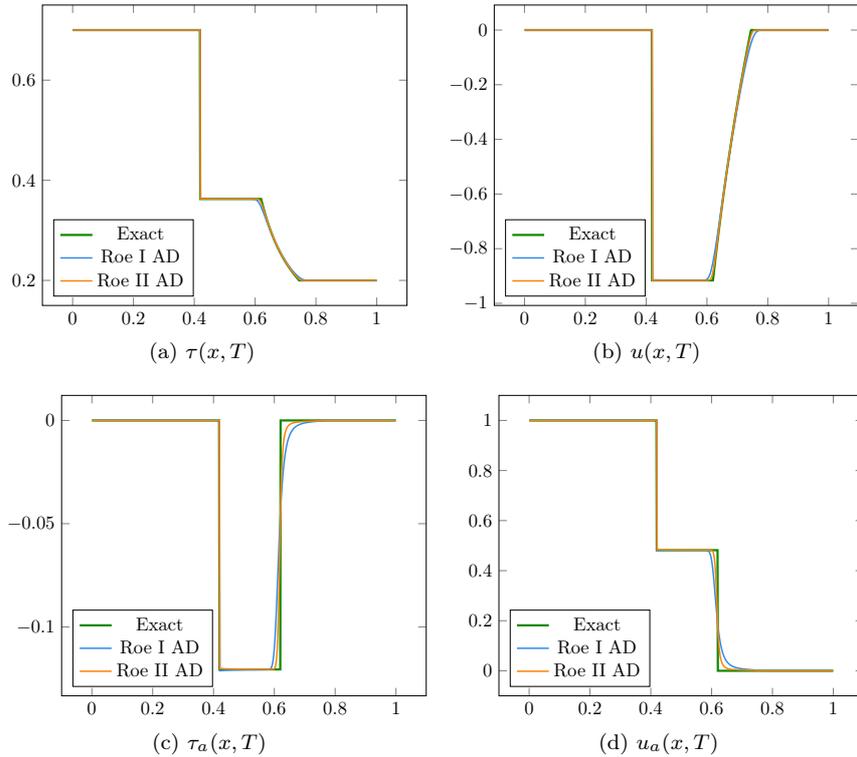


Figure 12: Roe-type schemes without numerical diffusion.

wave could be well separated, in order to attenuate the effect of the numerical diffusion in the middle. We also changed the starting point of the discontinuity ( $x_c = 0.3$ ) in order to have the second discontinuity associated with the rarefaction still in the domain at the final time. The results shown in Figure 15 are obtained with a mesh  $\Delta x = 10^{-4}$ : even in this particular case, with three discontinuities, we are able to approximate well the sensitivity provided that the mesh is fine enough.

## 9 Conclusion and discussion

The first goal of this work was to define a sensitivity system providing a solution which is suitable for applications, i.e. without peaks approximating the Dirac delta function: this was achieved by adding a properly defined source term. However, the numerical results presented in section 6 show that the numerical diffusion plays a very important role in the discretisation of the sensitivity system, to such an extent that classical finite volume schemes do not converge to the analytical solution: in particular, the value of the plateau in the star zone is not correct. To overcome this problem, we propose a numerical scheme based on sampling techniques, which does not have any numerical diffusion: with this scheme we are able to discretise more precisely the source term and to obtain a correct solution for the sensitivity, too.

Currently, we are extending this to the complete Euler system. This will allow us to tackle more realistic problems and to use the sensitivity analysis for the applications described in the introduction, such as optimization and uncertainty quantification.

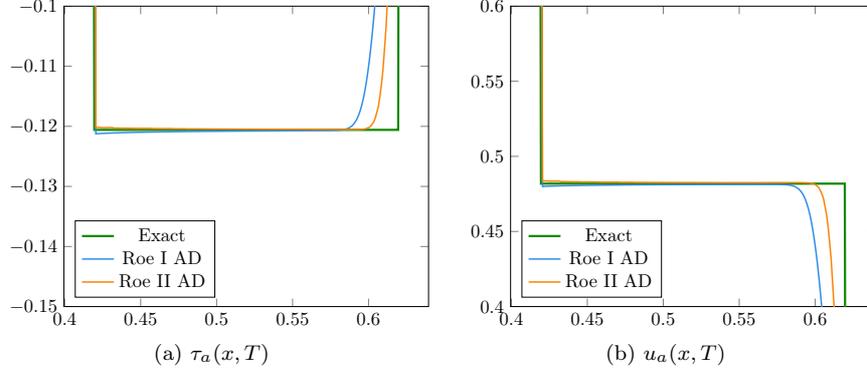


Figure 13: Roe-type schemes without numerical diffusion for the sensitivity - zoom.

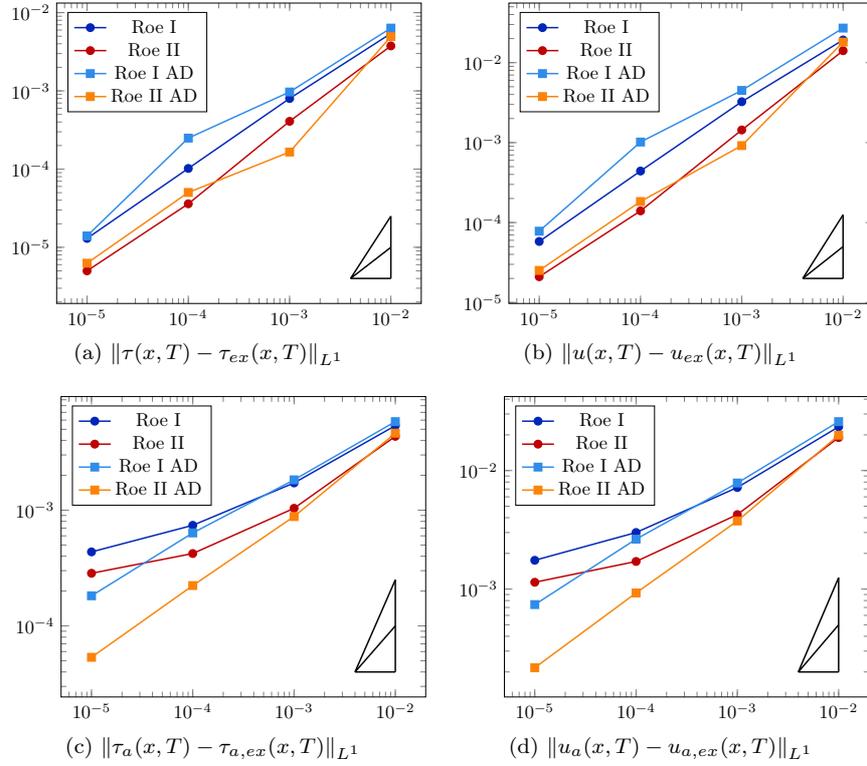


Figure 14: Convergence of Roe-type schemes, with and without numerical diffusion.

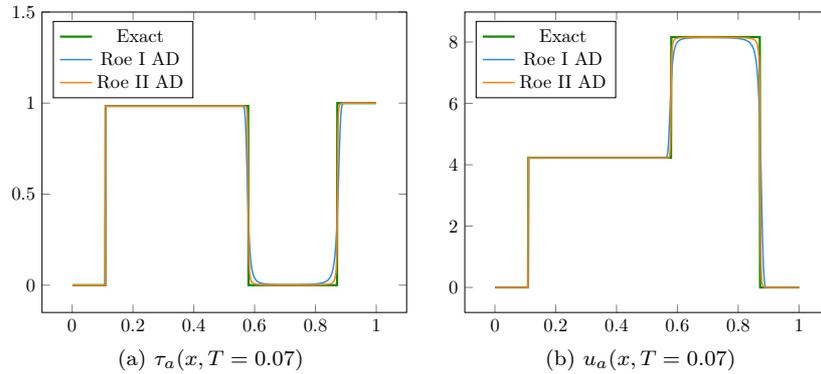


Figure 15: Test case shock-rarefaction,  $a = \tau_R$ : sensitivity.  $\Delta x = 10^{-4}$ ,  $T = 0.07$ .

## References

- [1] 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.
- [2] J. Borggaard and J. Burns. A PDE sensitivity equation method for optimal aerodynamic design. *Journal of Computational Physics*, 136(2):366 – 384, 1997.
- [3] F. Bouchut. *Nonlinear stability of finite Volume Methods for hyperbolic conservation laws: And Well-Balanced schemes for sources*. Springer Science & Business Media, 2004.
- [4] F. Bouchut, S. Jin, and X. Li. Numerical approximations of pressureless and isothermal gas dynamics. *SIAM Journal on Numerical Analysis*, 41(1):135–158, 2003.
- [5] C. Chalons, R. Duvigneau, and C. Fiorini. Sensitivity analysis for the euler equations in lagrangian coordinates. In *International Conference on Finite Volumes for Complex Applications*, pages 71–79. Springer, 2017.
- [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] C. Chalons, D. Kah, and M. Massot. Beyond pressureless gas dynamics: quadrature-based moment models. *Communications in Mathematical Sciences*, 10(4):1241–1272, 2012.
- [8] 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.
- [9] 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.
- [10] A. Forestier and P. Le Floch. Multivalued solutions to some non-linear and non-strictly hyperbolic systems. *Japan journal of industrial and applied mathematics*, 9(1):1, 1992.
- [11] 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.
- [12] 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.
- [13] 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.

- [14] A. Harten, P. D. Lax, and B. Van Leer. On upstream differencing and godunov-type schemes for hyperbolic conservation laws. In *Upwind and High-Resolution Schemes*, pages 53–79. Springer, 1997.
- [15] 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.
- [16] A. Jameson. Aerodynamic design via control theory. *Journal of Scientific Computing*, 3(97-0101):233–260, 1988.
- [17] K. T. Joseph. A riemann problem whose viscosity solutions contain  $\delta$ -measures. *Asymptotic Analysis*, 7(2):105–120, 1993.
- [18] P. G. LeFloch. An existence and uniqueness result for two nonstrictly hyperbolic systems, ima volumes in math. and its appl. 27. *Nonlinear evolution equations that change type*, pages 126–138, 1990.
- [19] B. Mohammadi and O. Pironneau. *Applied Optimal Shape Design for Fluids*. Oxford University Press, 2001.
- [20] O. Pironneau. On optimum design in fluid mechanics. *J. Fluid Mechanics*, (64), 1974.
- [21] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer Science & Business Media, 2013.
- [22] H. Yang and Y. Zhang. New developments of delta shock waves and its applications in systems of conservation laws. *Journal of Differential Equations*, 252(11):5951–5993, 2012.