



**HAL**  
open science

# Sensitivity equation method for Euler equations in presence of shocks applied to uncertainty quantification

Camilla Fiorini, Christophe Chalons, Régis Duvigneau

► **To cite this version:**

Camilla Fiorini, Christophe Chalons, Régis Duvigneau. Sensitivity equation method for Euler equations in presence of shocks applied to uncertainty quantification. 2018. hal-01817815v1

**HAL Id: hal-01817815**

**<https://inria.hal.science/hal-01817815v1>**

Preprint submitted on 18 Jun 2018 (v1), last revised 11 Dec 2019 (v3)

**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 equation method for Euler equations in presence of shocks applied to uncertainty quantification

Camilla Fiorini, Christophe Chalons, Régis Duvigneau

June 8, 2018

## Abstract

Sensitivity analysis (SA) is the study of how changes in the input of a model affect the outputs. Standard sensitivity analysis techniques, such as the continuous sensitivity equation (CSE) 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: an HLL-type scheme and an HLLC-type scheme. The effects of the numerical diffusion on the convergence of the schemes with respect to the grid are discussed. Finally, an application of SA to uncertainty quantification is investigated and the different numerical schemes are compared.

## 1 Introduction

Sensitivity analysis (SA) is the study of how changes in the inputs of a model affect the outputs. The sensitivity itself is defined as the derivative of the state, i.e. the solution to the PDE model considered, with respect to one or several parameters of interest. SA is obviously a valuable tool for engineering applications, since it allows to quantify the physical response of a system to any change of parameter values: one of its straightforward application in this direction is uncertainty quantification (UQ). Other applications of SA are optimization [4, 11] and quick evaluation of close solutions [10].

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. A detailed comparison between the two for optimization problems is done in [15]. In this work, we mainly focus on the differentiate-then-discretise approach, and in particular on the continuous sensitivity equation (CSE) method [4, 11, 10, 18], which allows to compute the derivative of the PDE solution itself, at any location and time, with respect to a single input parameter (or possibly a vector of parameters): this is done by formally differentiating the state system 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, obtaining in such a way a new system of PDEs, the sensitivity system.

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, 21] with a theoretical viewpoint, and more recently in [13, 14, 5] with a numerical viewpoint, where a modification of the sensitivity system was proposed, to “remove” the spikes from the numerical sensitivity solution. We 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 a simple problem of uncertainty quantification 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:

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

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

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

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:

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

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:

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

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:

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

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) [18, 4, 11] 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:

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

which can be written in vectorial form as

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

where we used the following notation:

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

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 nothing but  $\mathbf{U}_a(x, t = 0) = \partial_a \mathbf{U}_0(x)$ .

## 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:

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

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

Therefore, the complete system writes:

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

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:

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

with

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

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:

$$p_M(x) = (x - \lambda_1)^2(x - \lambda_2)^2(x - \lambda_3)^2, \quad (9)$$

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:

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

Let us write (10) by blocks:

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

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

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

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  $\blacksquare$ :

$$\begin{aligned} \blacksquare_{(1,1)} &= 0 + [c - u, 1, 0] \begin{bmatrix} 0 \\ (3 - \gamma)u^2 u_a - \frac{(\gamma-3)^2}{2} u^2 u_a \\ \diamond \end{bmatrix} + \\ &+ [c - u, 1, 0] \begin{bmatrix} (\gamma - 3)uu_a \\ (\gamma - 2)(3 - \gamma)u^2 u_a + (\gamma - 1)(\star) \\ \Delta \end{bmatrix} = \\ &= -\frac{3}{2}(\gamma - 1)u^2 u_a + (\gamma - 3)cuu_a - c^2 u \frac{p_a}{p} + c^2 u \frac{\rho_a}{\rho} - c^2 u_a + \gamma(\gamma - 1)u^3 \frac{\rho_a}{\rho}, \end{aligned}$$

where there is no need to specify  $\diamond$  and  $\Delta$ . 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.  $\square$

## 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]. Therefore, we add a correction term in the Rankine-Hugoniot conditions of (5), as done in [5] in order to make this system valid also in the present framework of hyperbolic systems with possibly discontinuous solutions:

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

where  $N_s$  is the number of discontinuities, which can be either shocks or contact discontinuities,  $\delta_k = \delta(x - x_{k,s})$  is the Dirac delta function with  $x_{k,s}$  position of the  $k$ -th shock and  $\rho_k$  is the amplitude of the  $k$ -th correction which is

$$\rho_k(t) = \sigma_{k,a}(\mathbf{U}^+ - \mathbf{U}^-). \quad (14)$$

The definition of the source term comes from the derivation with respect to  $a$  of the Rankine-Hugoniot relations associated with (3):

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

which gives:

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

where the terms with  $\partial_a x_s(t)$  are very difficult to estimate, however they are all zero in a first order finite volume framework, therefore they can be neglected. A special treatment, which will be detailed later, is necessary for a second order discretisation. The new system thus writes:

$$\begin{cases} \partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0 \\ \partial_t \mathbf{U}_a + \partial_x \mathbf{F}_a(\mathbf{U}, \mathbf{U}_a) = \mathbf{S}. \end{cases} \quad (16)$$

Before going to the design of a numerical scheme to approximate the solution of (16), we briefly discuss in the next section the typical structure of these solutions and compare it with the ones of (3). We specify the solution for a given initial data of Riemann type.

### 3 Riemann problem

In this section, we write the exact solution for the system (16) in a specific case (cf. [2]), which will be 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 couple  $(\lambda_2, \mathbf{r}_2)$  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 1 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 2 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 2 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:

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

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

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

This choice of initial data leads to the structure in Figures 1-2, 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 1 for the state and Figure 2 for the sensitivity. Let us now give the exact formulas for the state and for the sensitivity.

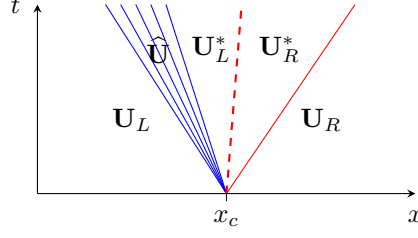


Figure 1 – Structure of the solution for the Riemann problem for the state.

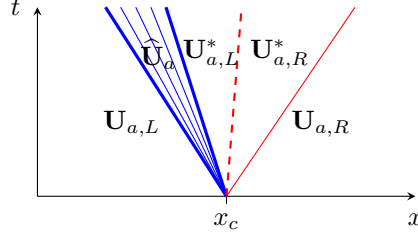


Figure 2 – Structure of the solution for the Riemann problem for the sensitivity.

*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:

$$p_L = p_R^* \left( 1 - \frac{(\gamma - 1) \frac{c_R}{c_L} \left( \frac{p_R^*}{p_R} - 1 \right)}{\sqrt{2\gamma \left( 2\gamma + (\gamma + 1) \left( \frac{p_R^*}{p_R} - 1 \right) \right)}} \right)^{-\frac{2\gamma}{\gamma-1}}, \quad (17)$$

where  $c_\ell = \sqrt{\frac{\gamma p_\ell}{\rho_\ell}}$ , 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  $\rho$ , 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:

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

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

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

Finally, the solution writes:

$$\mathbf{U}(x, t) = \begin{cases} \mathbf{U}_L & x - x_c < -c_L t, \\ \widehat{\mathbf{U}} & -c_L t < x - x_c < \left(\frac{\gamma+1}{2}u^* - c_L\right)t, \\ \mathbf{U}_L^* & \left(\frac{\gamma+1}{2}u^* - c_L\right)t < x - x_c < u^*t, \\ \mathbf{U}_R^* & u^*t < x - x_c < c_R \sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma} \frac{p^*}{p_R}} t, \\ \mathbf{U}_R & x - x_c > c_R \sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma} \frac{p^*}{p_R}} t. \end{cases} \quad (18)$$

*Sensitivity solution:* by differentiating (17) 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:

$$\begin{aligned} \Theta &= 1 - \frac{(\gamma-1)c_R \left(\frac{p^*}{p_R} - 1\right)}{c_L \sqrt{4\gamma^2 + 2\gamma(\gamma-1) \left(\frac{p^*}{p_R} - 1\right)}}, \\ \Xi &= \frac{c_R \left(\frac{p^*}{p_R} - 1\right) c_{a,R} \sqrt{2\gamma}}{c_L^2 \sqrt{2\gamma + (\gamma+1) \left(\frac{p^*}{p_R} - 1\right)}}, \\ \Lambda &= \frac{\sqrt{2\gamma} c_R}{c_L p_R \sqrt{2\gamma + (\gamma+1) \left(\frac{p^*}{p_R} - 1\right)}}, \\ \Psi &= \frac{\gamma(\gamma+1)c_R \left(\frac{p^*}{p_R} - 1\right)}{c_L p_R \sqrt{2\gamma} \left(2\gamma + (\gamma+1) \left(\frac{p^*}{p_R} - 1\right)\right)^{\frac{3}{2}}}. \end{aligned}$$

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

$$\begin{aligned} u_a^* &= \frac{2c_{a,L}}{\gamma-1} \left(1 - \left(\frac{p^*}{p_L}\right)^{\frac{\gamma-1}{2\gamma}}\right) - \frac{c_L}{\gamma} \left(\frac{p^*}{p_L}\right)^{-\frac{\gamma-1}{2\gamma}} \left(\frac{p_L p_a^* - p^*}{p_L^2}\right), \\ \rho_{a,R}^* &= \frac{\rho_R p_a^*}{p_R} \frac{\left(1 + \frac{\gamma-1}{\gamma+1} \frac{p_R}{p^*}\right)}{\left(1 + \frac{\gamma-1}{\gamma+1} \frac{p^*}{p_R}\right)} + \rho_R \frac{p^*}{p_R} \frac{\gamma-1}{\gamma+1} \left(\frac{-\frac{p_R p_a^*}{p^{*2}} \left(1 + \frac{\gamma-1}{\gamma+1} \frac{p^*}{p_R}\right) - \frac{p_a^*}{p_R} \left(1 + \frac{\gamma-1}{\gamma+1} \frac{p_R}{p^*}\right)}{\left(1 + \frac{\gamma-1}{\gamma+1} \frac{p^*}{p_R}\right)^2}\right), \\ \rho_{a,L}^* &= \frac{\rho_L}{\gamma} \frac{p_L p_a^* - p^*}{p_L^2} \left(\frac{p^*}{p_L}\right)^{\frac{1-\gamma}{\gamma}}. \end{aligned}$$

Finally, in the rarefaction:

$$\begin{aligned} \hat{u}_a(x, t) &= \frac{2u_L u^*}{(\gamma+1)u^{*2}} \frac{x - x_c}{t} + 2 \frac{c_{a,L} u^{*2} - c_{a,L} u_L u^* + c_L u_L u_a^*}{(\gamma+1)u^{*2}}, \\ \hat{\rho}_a(x, t) &= -\rho_L \left(\frac{\hat{u}_a(x, t) c_L - \hat{u}(x, t) c_{a,L}}{c_L^2}\right) \left(1 - \frac{(\gamma-1)\hat{u}(x, t)}{2c_L}\right)^{\frac{3-\gamma}{\gamma-1}}, \\ \hat{p}_a(x, t) &= \left(1 - \frac{(\gamma-1)\hat{u}(x, t)}{2c_L}\right)^{\frac{2\gamma}{\gamma-1}} - p_L \gamma \left(\frac{\hat{u}_a(x, t) c_L - \hat{u}(x, t) c_{a,L}}{c_L^2}\right) \left(1 - \frac{(\gamma-1)\hat{u}(x, t)}{2c_L}\right)^{\frac{\gamma+1}{\gamma-1}}. \end{aligned}$$

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

$$\mathbf{U}_a(x, t) = \begin{cases} \mathbf{U}_{a,L} & x - x_c < -c_L t, \\ \widehat{\mathbf{U}}_a \left(\frac{x-x_c}{t}\right) & -c_L t < x - x_c < \left(\frac{\gamma+1}{2}u^* - c_L\right)t, \\ \mathbf{U}_{a,L}^* & \left(\frac{\gamma+1}{2}u^* - c_L\right)t < x - x_c < u^*t, \\ \mathbf{U}_{a,R}^* & u^*t < x - x_c < c_R \sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma} \frac{p^*}{p_R}} t, \\ \mathbf{U}_{a,R} & x - x_c > c_R \sqrt{\frac{\gamma-1}{2\gamma} + \frac{\gamma+1}{2\gamma} \frac{p^*}{p_R}} t. \end{cases} \quad (19)$$



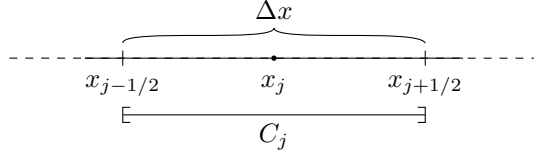


Figure 3 – Spatial discretisation.

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

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

## 4 Numerical methods

In this section we consider the numerical approximation of (16). 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  $\Delta x$ ,  $x_j$  is the center of the  $j$ -th cell  $C_j$ , whose extrema are  $x_{j-1/2}$  and  $x_{j+1/2}$  (cf. Figure 3). We use an adaptive time step  $\Delta t^n$ , 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: different choices for the solution of the Riemann problem lead to different numerical schemes.

### 4.1 Projection step

The projection step is usually performed by averaging the solution  $\mathbf{v}(x, t^{n+1})$  on the cell:

$$\mathbf{V}_j^{n+1} = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{v}(x, t^{n+1}) dx. \quad (20)$$

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 [12, 7]. First, we define a staggered mesh, whose cells will be denoted  $\bar{C}_j^n$ , as follows:

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

where  $\sigma_{j-1/2}^n$  is a proper speed. In this case, we choose it equal to zero if no shock is expected at the interface  $j-1/2$ , whilst it is equal to the speed of the shock, if there is one. More details on the choice of  $\sigma_{j-1/2}^n$  will be provided in the next subsections. The second step is to perform the average on the staggered mesh, obtaining in this way an intermediate solution  $\bar{\mathbf{V}}_j^{n+1}$ :

$$\bar{\mathbf{V}}_j^{n+1} = \frac{1}{\Delta x_j^n} \int_{\bar{x}_{j-1/2}^n}^{\bar{x}_{j+1/2}^n} \mathbf{v}(x, t^{n+1}) dx, \quad (21)$$

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

$$\mathbf{V}_j^{n+1} = \begin{cases} \bar{\mathbf{V}}_{j-1}^{n+1} & \text{if } \alpha_{n+1} \in (0, \frac{\Delta t}{\Delta x} \max(\sigma_{j-1/2}^n, 0)), \\ \bar{\mathbf{V}}_j^{n+1} & \text{if } \alpha_{n+1} \in [\frac{\Delta t}{\Delta x} \max(\sigma_{j-1/2}^n, 0), 1 + \frac{\Delta t}{\Delta x} \min(\sigma_{j+1/2}^n, 0)], \\ \bar{\mathbf{V}}_{j+1}^{n+1} & \text{if } \alpha_{n+1} \in [1 + \frac{\Delta t}{\Delta x} \min(\sigma_{j+1/2}^n, 0), 1]. \end{cases} \quad (22)$$

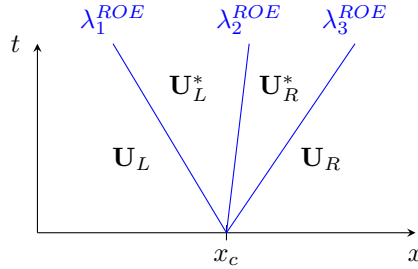


Figure 4 – Structure of the Roe solver for the state.

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]):

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

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

In the next subsections we present some Riemann solvers for the state and for the sensitivity, and their two variations: the diffusive standard version with the projection step (20), and their anti-diffusive version (21)-(22) and we compare their results. The strategy is to solve the state system separately from the sensitivity system, since the global system (7) is only weakly hyperbolic and some of the standard approaches require strict hyperbolicity. Moreover, notice that the state system (3) evolves in an independent way with respect to the sensitivity system (6).

## 4.2 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. It turns out that this property plays an important role in the anti-diffusive approach (21)-(22). Indeed, it makes the method equivalent to the random choice Glimm scheme in the specific case of an isolated shock wave, and therefore is convergent to the correct solution (still in this very specific case). In addition, we remark that it would not be possible to use a solver with only one intermediate star state, such as HLL, because of the definition of the source term (14): two intermediate states are necessary in order to be able to compute the correction term across the contact discontinuity.

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 [23]. 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 4) connected by three discontinuities travelling at speeds  $\lambda_i^{ROE}$ . 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  $\mathbf{A}$ :

$$\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{i=1}^3 \alpha_i \tilde{\mathbf{r}}_i. \quad (23)$$

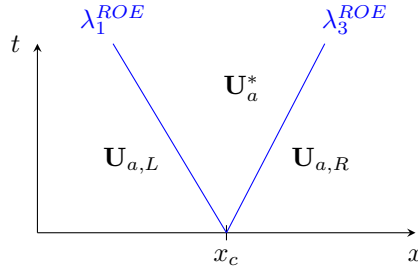


Figure 5 – Structure of the HLL-type solver for the sensitivity.

The relation (23) is used to compute the coefficients  $\alpha_i$ , then one has:

$$\mathbf{U}_L^* = \mathbf{U}_L + \alpha_1 \tilde{\mathbf{r}}_1 = \mathbf{U}_R - \alpha_2 \tilde{\mathbf{r}}_2 - \alpha_3 \tilde{\mathbf{r}}_3, \quad \mathbf{U}_R^* = \mathbf{U}_R - \alpha_3 \tilde{\mathbf{r}}_3 = \mathbf{U}_L + \alpha_1 \tilde{\mathbf{r}}_1 + \alpha_2 \tilde{\mathbf{r}}_2. \quad (24)$$

Once all the quantities  $\mathbf{U}_L^*$ ,  $\mathbf{U}_R^*$ , and  $\lambda_\ell^{ROE}$  are known at each interface  $x_{j-1/2}$ ,  $\mathbf{w}(x, t^{n+1})$  is defined and the integrals (20) or (21) can easily be computed, since  $\mathbf{w}(x, t^{n+1})$  is a piecewise constant.

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 [16].

#### 4.2.1 Definition of $\sigma_{j-1/2}^n$

As already said,  $\sigma_{j-1/2}^n$  is 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 5). 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 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  $\rho$  is increasing across a 1-shock and it is decreasing across a 3-shock:

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

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.

### 4.3 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. 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  $\lambda_1^{ROE}$  and  $\lambda_3^{ROE}$  (cf. Figure 5). 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 [17] applied to system of conservation laws with source terms. We get:

$$\begin{aligned} \mathbf{U}_{a,j-1/2}^* &= \frac{1}{\lambda_3^{ROE} - \lambda_1^{ROE}} \left( \lambda_3^{ROE} \mathbf{U}_{a,j}^n - \lambda_1^{ROE} \mathbf{U}_{a,j-1}^n \right. \\ &\quad \left. - \mathbf{F}_a(\mathbf{U}_j, \mathbf{U}_{a,j}) + \mathbf{F}_a(\mathbf{U}_{j-1}, \mathbf{U}_{a,j-1}) + \mathbf{S}_{j-1/2} \right), \end{aligned} \quad (25)$$

where the source term is naturally discretised as follows:

$$\begin{aligned} \mathbf{S}_{j-1/2} &= \partial_a \lambda_{1,j-1/2}^{ROE} (\mathbf{U}_{L,j-1/2}^* - \mathbf{U}_{j-1}) d_{1,j-1/2} + \partial_a \lambda_{2,j-1/2}^{ROE} (\mathbf{U}_{R,j-1/2}^* - \mathbf{U}_{L,j-1/2}^*) \\ &\quad + \partial_a \lambda_{3,j-1/2}^{ROE} (\mathbf{U}_j - \mathbf{U}_{R,j-1/2}^*) d_{3,j-1/2}. \end{aligned}$$

### HLLC-type scheme

Another possible approach for the sensitivity is to keep the same structure as for the state (cf. Figure 4), 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:

$$\begin{cases} -\lambda_1(\rho_{a,L}^* - \rho_{a,L}) + (\rho u)_{a,L}^* - (\rho u)_{a,L} = \partial_a \lambda_1(\rho_L^* - \rho_L), \\ -\lambda_2(\rho_{a,R}^* - \rho_{a,L}^*) + (\rho u)_{a,R}^* - (\rho u)_{a,L}^* = \partial_a \lambda_2(\rho_R^* - \rho_L^*), \\ -\lambda_3(\rho_{a,R} - \rho_{a,R}^*) + (\rho u)_{a,R} - (\rho u)_{a,R}^* = \partial_a \lambda_3(\rho_R - \rho_R^*), \\ \frac{(\gamma-3)}{2} \tilde{u}^2(\rho_{a,R}^* - \rho_{a,L}^*) + (2-\gamma)\tilde{u}((\rho u)_{a,R}^* - (\rho u)_{a,L}^*) \\ \quad + (\gamma-1)((\rho E)_{a,R}^* - (\rho E)_{a,L}^*) = \partial_a \lambda_2((\rho u)_R^* - (\rho u)_L^*), \\ (\lambda_2 - \lambda_1)(\rho u)_{a,L}^* + (\lambda_3 - \lambda_2)(\rho u)_{a,R}^* + \lambda_1(\rho u)_{a,L} - \lambda_3(\rho u)_{a,R} \\ \quad + \mathbf{F}_{a,R}|_2 - \mathbf{F}_{a,L}|_2 = \Delta x \mathbf{S}|_2, \\ (\lambda_2 - \lambda_1)(\rho E)_{a,L}^* + (\lambda_3 - \lambda_2)(\rho E)_{a,R}^* + \lambda_1(\rho E)_{a,L} - \lambda_3(\rho E)_{a,R} \\ \quad + \mathbf{F}_{a,R}|_3 - \mathbf{F}_{a,L}|_3 = \Delta x \mathbf{S}|_3, \end{cases} \quad (26)$$

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  $\rho$  across the three waves, derived 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  $\rho u$  for the linearised system derived 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

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

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:

$$\mathbf{U}_L^* = \mathbf{U}_L + \alpha_1 \mathbf{r}_1, \quad \mathbf{U}_R^* = \mathbf{U}_R - \alpha_3 \mathbf{r}_3, \quad (27)$$

obtaining

$$\mathbf{U}_{a,L}^* = \mathbf{U}_{a,L} + \alpha_{a,1} \mathbf{r}_1 + \alpha_1 \mathbf{r}_{a,1}, \quad \mathbf{U}_{a,R}^* = \mathbf{U}_{a,R} - \alpha_{a,3} \mathbf{r}_3 - \alpha_3 \mathbf{r}_{a,3}, \quad (28)$$

with

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

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

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

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

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

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 (28) solve the system (26).*

*Proof.* We will prove that the star sensitivities defined in (28) satisfy the system (26).

1. First equation. Writing the first coefficient of (27) one easily finds  $\rho_L^* - \rho_L = \alpha_1$ , and writing the first two coefficient of (28) one finds:

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

We now replace these three expressions in the first equation of (26) 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

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

Therefore, one has:

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

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 (26) 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 (26) 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:

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

By replacing them in the fourth equation of (26) 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 (28) 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 coherent with our discretisation of the source term.  $\square$

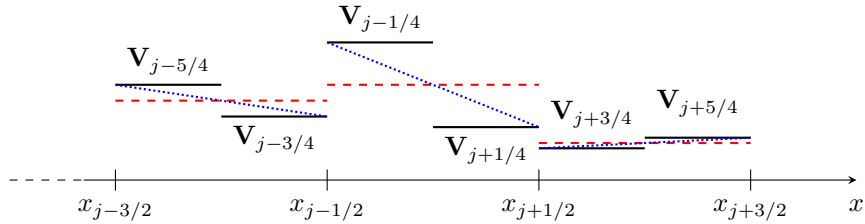


Figure 6 – MUSCL discretisation. Dashed red line: first order discretisation. Dotted blue line: classical second order discretisation. Solid black line: second order discretisation used in this work.

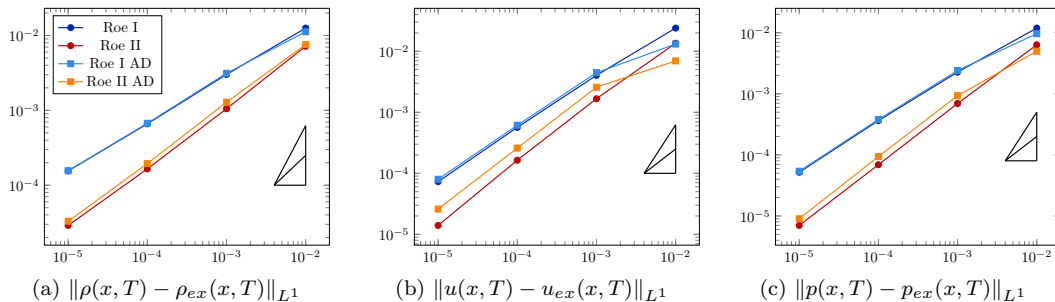


Figure 7 – Convergence test for the state. The Roman numerals I and II stand for the order of the scheme. AD stands for anti-diffusive.

#### 4.4 Second order MUSCL-type 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 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  $\mathbf{V}_j^n(x_{j+1/2})$ ,  $\mathbf{V}_{j+1}^n(x_{j+1/2})$  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)-(14) 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  $\mathbf{V}_{j\pm 1/4}^n$  and correspond to the edge values  $\mathbf{V}_j^n(x_{j\pm 1/2})$  (see Figure 6). In this work, we compute the edge values with a standard approach:

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

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

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

where

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

We remark that this approach leads to an additional Riemann problem in the middle of the cell: in this way we are able to account for the neglected terms in (15).

## 5 Convergence tests for the numerical schemes

We consider the Riemann problem of section 3. In Figures 7-8-9 we show the convergence of the different numerical schemes presented in Section 4. Figure 7 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 8-9 we plot the error for the sensitivity, first with

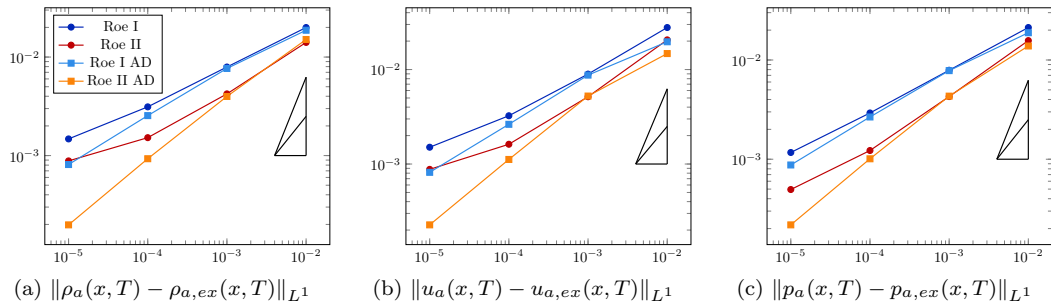


Figure 8 – Convergence test for the sensitivity - HLL-type scheme.

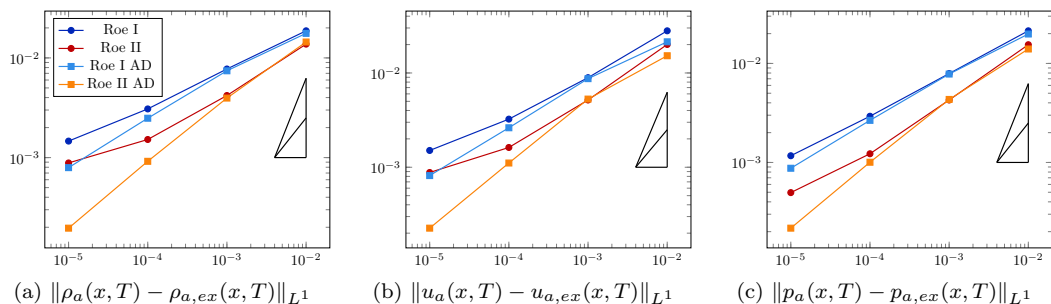


Figure 9 – Convergence test for the sensitivity - HLLC-type scheme.

the HLL-type scheme (Figure 8) and then with the HLLC-type scheme (Figure 9): 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  $\rho_a$ . In Figure 10 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 11. In Figure 12 we compare the antidiffusive schemes, first and second order: for the state, the difference is noticeable mainly in the contact discontinuity (therefore only for  $\rho$ ), whilst for the sensitivity the difference is significant in the neighbourhood of the discontinuities before and after the rarefaction. Finally, in Figure 13 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. 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.

## 6 Uncertainty Quantification

### 6.1 Problem description

In this section, we show how SA can be used for uncertainty quantification (UQ), cf. [22, 24, 9]. Many UQ techniques have been developed during the last decades: these methods can be either probabilistic or deterministic. SA falls into the second category while the most well-known of these techniques, the Monte Carlo method, is in the first. Other UQ techniques are polynomial chaos [25, 27, 19] and the random space partition [1]. A very good review and comparison of many techniques with applications to fluid dynamics can be found in [26]. The main aim of UQ is to determine a confidence interval for the output of a model, in our case  $\mathbf{U}$ , given the uncertainty and the error on the input parameters. According to the AIAA definition, error has a deterministic nature, while uncertainty a stochastic one. Furthermore, uncertainty can be categorised in aleatoric uncertainty and epistemic uncertainty.

In this work, we compare two different UQ methods: Monte Carlo and sensitivity analysis. Both methods aim at providing statistical quantities like moments (mean, variance, ...) of the output of



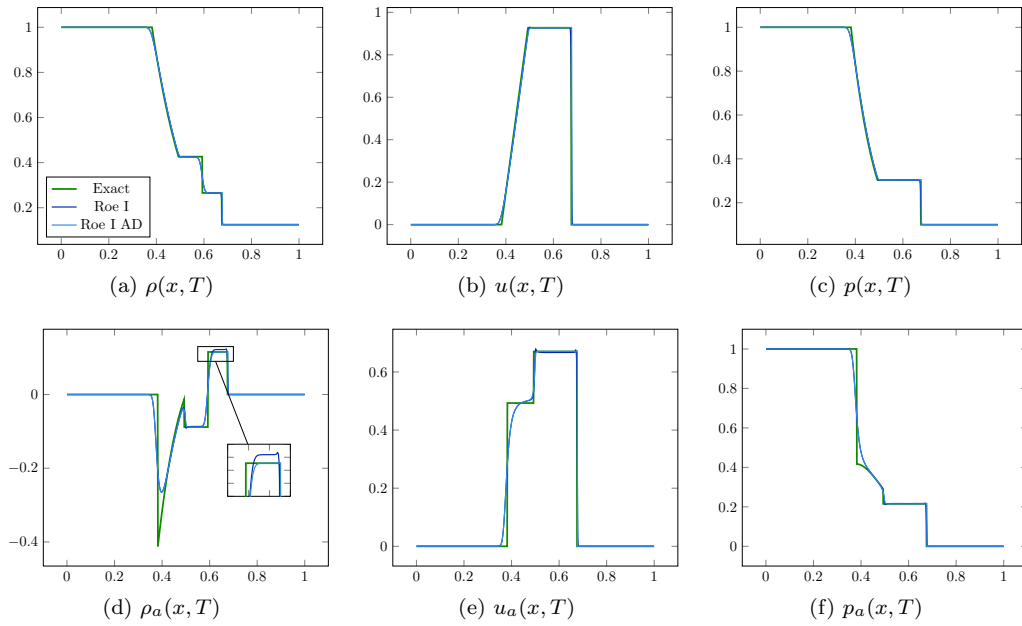


Figure 10 – First order schemes, with and without numerical diffusion. HLL-type scheme for the sensitivity.

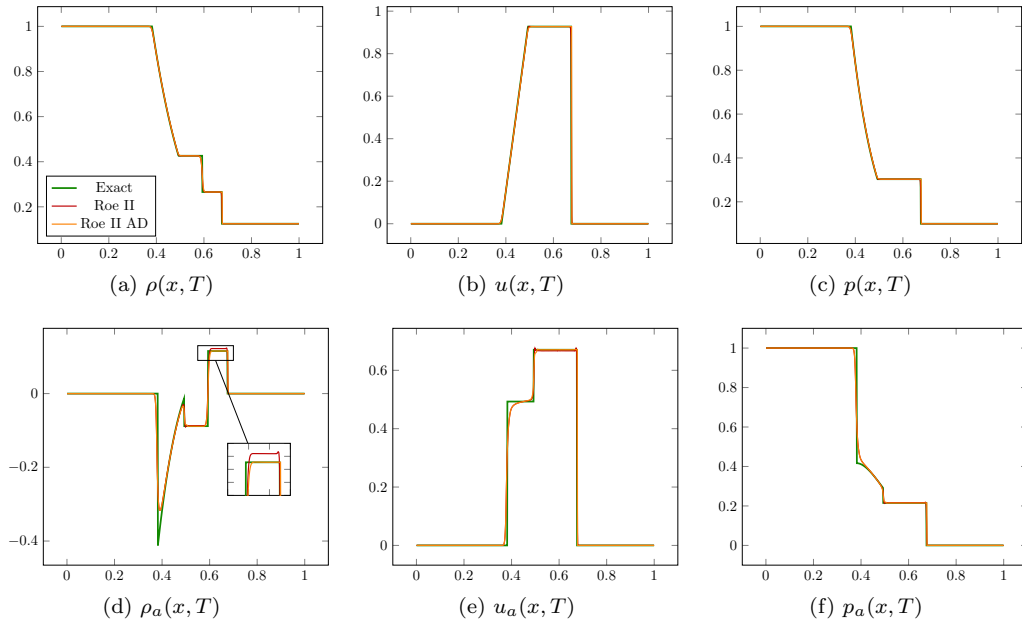


Figure 11 – Second order schemes, with and without numerical diffusion. HLL-type scheme for the sensitivity.

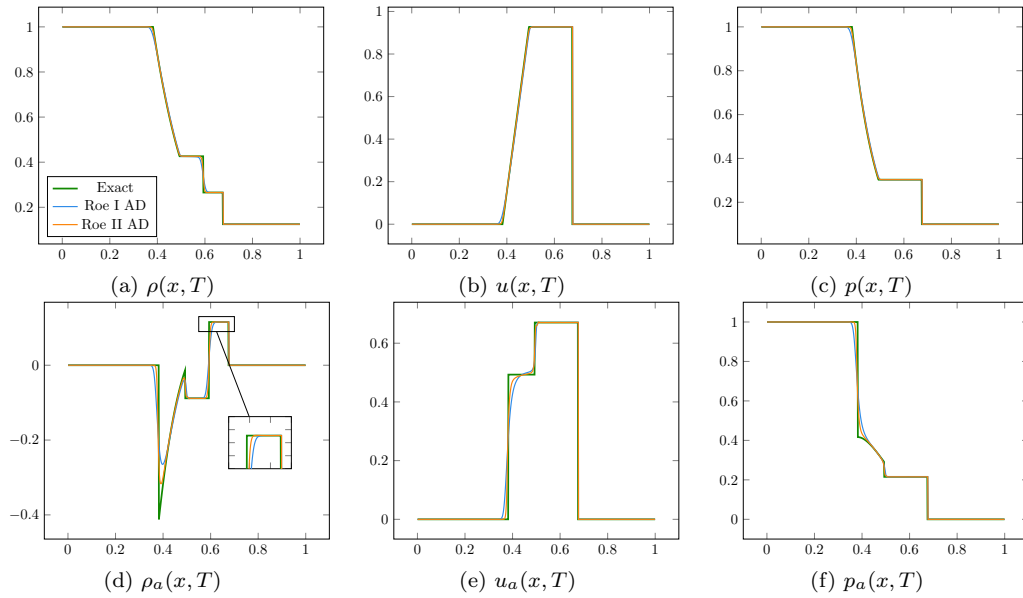


Figure 12 – First and second order schemes, without numerical diffusion. HLL-type scheme for the sensitivity.

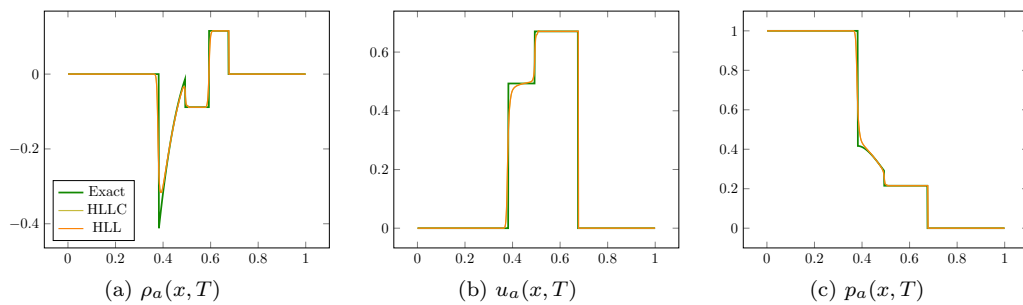


Figure 13 – Second order antidiffusive schemes: HLL and HLLC comparison.

the model. In the following,  $X$  will stand for one of the variables, i.e.  $X$  can either be  $\rho$ ,  $u$  or  $p$ , and  $X_a$  the corresponding sensitivity. We use the notation  $\mu_X$  to indicate the average of the variable  $X$  and  $\sigma_X^2$  for its variance. Once this two quantities are known, one can build a confidence interval for the variable  $X$  as:  $CI_X = [\mu_X - \kappa\sigma_X, \mu_X + \kappa\sigma_X]$ . The coefficient  $\kappa$  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  $\kappa = 2$  provides a 66% confidence interval.

**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 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:

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

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 analysis method.** SA is a deterministic approach to estimate the average  $\mu_X$  and the variance  $\sigma_X^2$  of the output  $X$ . Let  $\mu_{\mathbf{a}}$  be the average of the uncertain vector  $\mathbf{a}$ , and  $\sigma_{\mathbf{a}}$  the covariance matrix:

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

where  $M$  is the number of uncertain parameters,  $\mu_{a_i}$  the average of the  $i$ -th parameter,  $\sigma_{a_i}^2$  its variance and  $\text{cov}(\cdot, \cdot)$  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:

$$\begin{aligned} \sigma_X^2 &= E[(X(\mathbf{a}) - \mu_X)^2] = E \left[ \left( \sum_{i=1}^M X_{a_i}(\mu_{\mathbf{a}}) (a_i - \mu_{a_i}) \right)^2 \right] = \\ &= \sum_{i=1}^M X_{a_i}^2(\mu_{\mathbf{a}}) E[(a_i - \mu_{a_i})^2] + \sum_{\substack{i,j=1 \\ i \neq j}}^M X_{a_i}(\mu_{\mathbf{a}}) X_{a_j}(\mu_{\mathbf{a}}) E[(a_i - \mu_{a_i})(a_j - \mu_{a_j})]. \end{aligned}$$

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

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

Higher order estimates require higher order sensitivities [20].

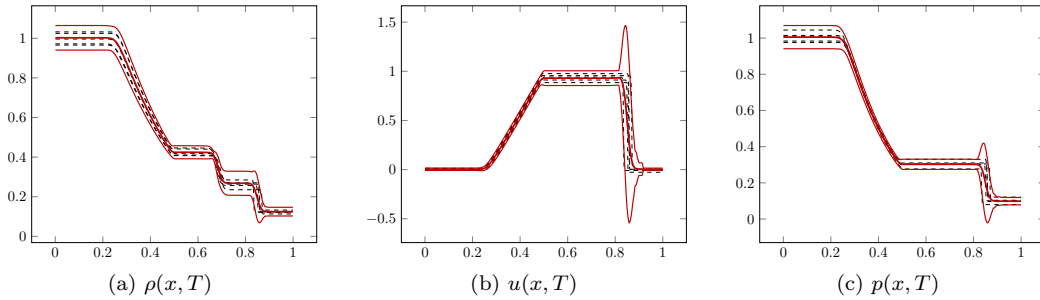


Figure 14 – Monte Carlo approach. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines.

## 6.2 Numerical results

We applied the uncertainty quantification techniques described in the previous subsection to the test case already introduced in section 3. The uncertain parameters are the left and right values of the physical variables for the state, i.e.:

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

with the following average and covariance matrix:

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

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. In Figure 14 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 Figures 15-16 we show the results of the SA 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 with respect to the ones we get with the Monte Carlo approach. 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. For these reasons, we suggest that the corrected sensitivity are more appropriate in this context. The results obtained with the corrected sensitivities are shown in Figure 17: the confidence interval obtained correspond to the ones obtained with the Monte Carlo approach, apart for the shock zone. Of course, the SA approach does not capture the uncertainty in the neighbourhood 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 SA approach, and it is the case with and without correction. However, the SA approach is less expensive: the Monte Carlo approach requires 1000 solutions of the state, whilst the SA approach only one solution of the state and as many solution of the sensitivity as the number of uncertain parameters, in this case 6. Finally, in Figure 18 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.

## 7 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

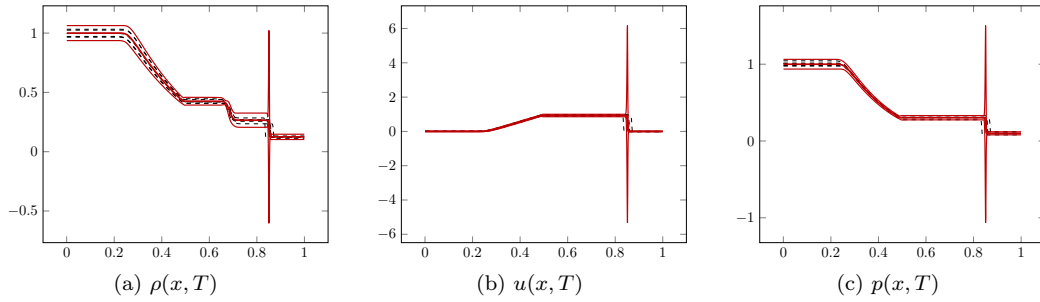


Figure 15 – SA approach without correction. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines

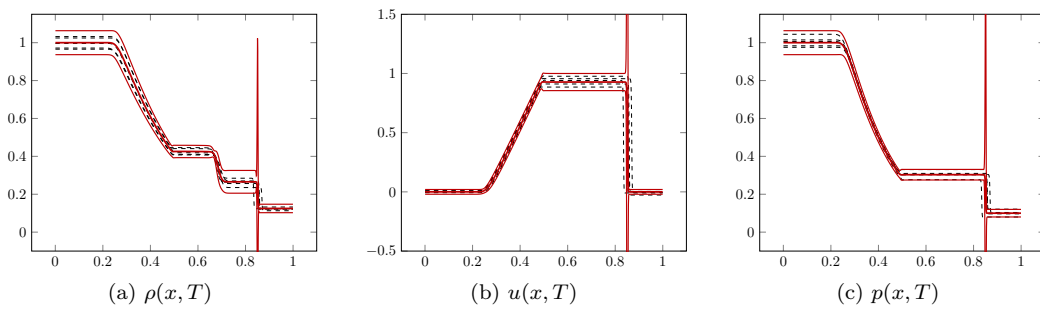


Figure 16 – SA approach without correction. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines - zoom.

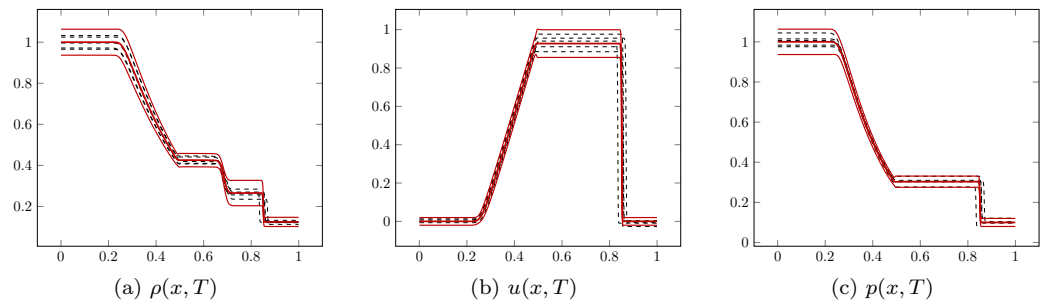


Figure 17 – SA approach with correction. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines

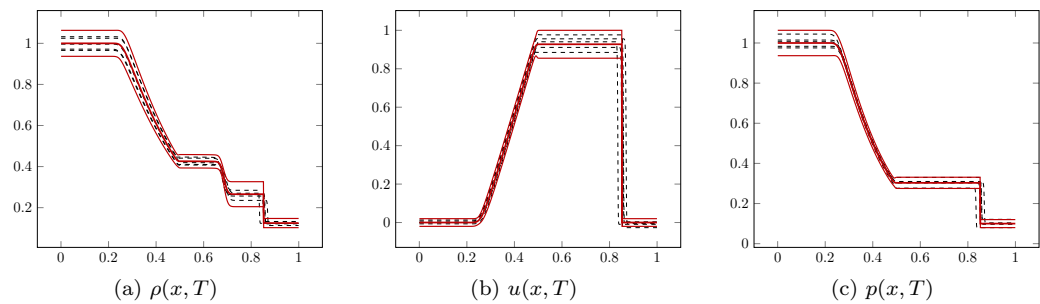


Figure 18 – SA approach with correction, anti-diffusive scheme. Average and the average plus and minus twice the standard deviation in red. Five samples in black dashed lines

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 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*, 2017. In review.
- [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] C. Delenne. Propagation de la sensibilité dans les modèles hydrodynamiques., 2014.
- [10] 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.
- [11] 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.
- [12] J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Communications on pure and applied mathematics*, 18(4):697–715, 1965.
- [13] 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.
- [14] 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.
- [15] M. D. Gunzburger. *Perspectives in flow control and optimization*, volume 5. Siam, 2003.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] O.M. Knio and O.P. Le Maître. Uncertainty propagation in CFD using polynomial chaos decomposition. *Fluid Dynamics Research*, 38(9):616–640, September 2006.

- [20] 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.
- [21] B. Mohammadi and O. Pironneau. *Applied Optimal Shape Design for Fluids*. Oxford University Press, 2001.
- [22] 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.
- [23] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of computational physics*, 43(2):357–372, 1981.
- [24] 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.
- [25] R. Walters. Towards stochastic fluid mechanics via polynomial chaos. In *41st AIAA Aerospace Sciences Meeting and Exhibit, Reno, USA*, 2003.
- [26] R. W. Walters and L. Huysse. Uncertainty analysis for fluid mechanics with applications. Technical report, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION HAMPTON VA LANGLEY RESEARCH CENTER, 2002.
- [27] D.B. Xiu and G.E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of Computational Physics*, (187):137–167, 2003.