

# Roe Solver with Entropy Corrector for Uncertain Hyperbolic Systems <sup>☆</sup>

J. Tryoen<sup>a,b,\*</sup>, O. Le Maître<sup>b</sup>, M. Ndjinga<sup>c</sup>, A. Ern<sup>a</sup>

<sup>a</sup> Université Paris Est, CERMICS, Ecole des Ponts, 77455 Marne la Vallée cedex 2, France

<sup>b</sup> LIMSI-CNRS, 91403 Orsay cedex, France

<sup>c</sup> CEA/Saclay/DEN/DM2S/SFME/LGLS, 91191 Gif-sur-Yvette cedex, France

---

## Abstract

This paper deals with intrusive Galerkin projection methods with Roe-type solver for uncertain hyperbolic systems using a finite volume discretization in physical space and a piecewise continuous representation at the stochastic level. The aim of this paper is to design a cost-effective adaptation of the deterministic Dubois and Mehlman corrector to avoid entropy-violating shocks in the presence of sonic points. The adaptation relies on an estimate of the eigenvalues and eigenvectors of the Galerkin Jacobian matrix of the deterministic system of the stochastic modes of the solution and on a correspondence between these approximate eigenvalues and eigenvectors for the intermediate states considered at the interface. Some indicators are derived to decide where a correction is needed, thereby reducing considerably the computational costs. The effectiveness of the proposed corrector is assessed on the Burgers and Euler equations including sonic points.

*Key words:* Uncertainty Quantification, hyperbolic systems, conservation laws, stochastic spectral methods, Galerkin projection, Roe solver, entropy correction

---

## 1. Introduction

To date, most computer simulations are based on deterministic mathematical models, where all input data are assumed to be perfectly known. In fact, this is seldom the case. A great challenge in recent years has been to provide effective tools for uncertainty quantification and propagation, that is, to quantify the resulting uncertainty in the numerical solution due to the uncertainty in input data (such as model parameters, initial and boundary conditions, and geometry). Stochastic spectral methods and so-called Chaos expansions originally introduced by Ghanem and Spanos [6] provide such tools by parametrizing the uncertain input data by random variables with known distribution functions and expanding the stochastic solution on a basis of (orthogonal) polynomials associated with the chosen random variables. The determination of the stochastic modes of the solution can then be achieved by non-intrusive methods based on the use of the numerical code solving the deterministic model for a sampling of the random variable or by a stochastic Galerkin projection of the model equation yielding a reformulated deterministic problem for the stochastic modes of the solution. Such methods have been successfully applied to a large variety of engineering problems governed by elliptic, parabolic or ordinary differential equations models. Two specific difficulties for the application of such methods to uncertain hyperbolic systems are the development of discontinuities in the solution in both physical and stochastic spaces and the nonlinearities in the flux functions. Different approaches have been proposed relying on multi-element probabilistic collocation methods (Lin et al. [11]), ENO-like reconstructions [1], pseudo-intrusive methods where the stochastic modes of the flux are computed

---

<sup>☆</sup>This work is partially supported by GNR MoMaS (ANDRA, BRGM, CEA, EdF, IRSN, PACEN-CNRS). O.P. Le Maître is partially supported by the French National Research Agency (Grant ANR-08-JCJC-0022).

\* *Corresponding author:* Université Paris Est, CERMICS, Ecole des Ponts, 77455 Marne la Vallée cedex 2, France.  
Phone: (33)-1 64 15 36 65

*Email addresses:* tryoenj@cermics.enpc.fr (J. Tryoen), olm@limsi.fr (O. Le Maître), michael.ndjinga@cea.fr (M. Ndjinga), ern@cermics.enpc.fr (A. Ern)

*Preprint submitted to ...*

*January 4, 2010*

by quadrature methods (Ge et al. [5], Poette et al. [12]), and fully intrusive methods using generalized Polynomial Chaos (gPC) methods in the case of the scalar linear wave equation (Gottlieb and Xiu [7]) and Multi Element-gPC (ME-gPC) methods in the case of nonlinear hyperbolic problems with relatively small fluctuations of the random quantities (Lin et al. [10]).

We proposed in [14] a fully intrusive method for nonlinear hyperbolic systems using a Finite Volume (FV) discretization in physical space and a piecewise continuous representation [2, 9, 15] at the stochastic level. To this end the stochastic space is discretized using tensor-product stochastic elements supporting local polynomial bases, and a stochastic Galerkin projection is applied to the original stochastic hyperbolic problem to derive the Galerkin system, that is, the set of deterministic equations coupling the stochastic modes of the solution. The nonlinear fluxes in the Galerkin system are computed in a pseudo-spectral way with the tools described in [3]. Furthermore, a cost-effective Roe-type solver for the Galerkin system has been designed in [14] using an upwinding determined from an approximate spectrum of the Galerkin Jacobian matrix evaluated at suitable Roe state. This matrix is referred to as the Roe Galerkin Jacobian matrix. Numerical tests on the stochastic Burgers and Euler equations (without sonic points) in one spatial dimension and, respectively, in two and one stochastic dimensions (number of stochastic variables) indicated that the method is accurate and robust while maintaining moderate computational costs.

Nevertheless, while effectively stabilizing the numerical method, the Roe scheme is known in the deterministic case to provide unphysical (entropy-violating) solutions in the presence of sonic points. In such a situation, a corrector is needed to construct numerical fluxes consistent with the physics. As in the deterministic case, non-physical expansion shocks occur in the stochastic case in the vicinity of sonic points; this is illustrated in the numerical experiments presented below. In the literature, different entropy correctors have been proposed for the Roe scheme. The most used are due to Harten [8] and Roe [13] and consist in adding a linear diffusive term in situations where the linearized Roe flux leads to an entropy-violating solution. In fact, these corrections appear more as “spreading devices” which act upon the approximate solution rather than a remedy to the fact that the approximate flux is linear in situations where a nonlinear description is crucial. The Dubois and Mehlman (DM) corrector is more general because it consists in a nonlinear modification of the flux in the vicinity of sonic points. Moreover, this corrector is non-parametric.

The aim of the present paper is to propose a cost-effective adaptation of the deterministic DM corrector to the methodology proposed in [14] derived in [14]. Direct application of the DM corrector to the Roe solver for the Galerkin system is not straightforward. First, the DM corrector needs the reconstruction of the intermediate states corresponding to each couple of left and right states at each interface cell in the physical domain and therefore the knowledge of the eigenvectors of the Roe Galerkin Jacobian matrix. The dimension of the Galerkin system can be quite large and, as for the Roe-type solver developed in [14], we want to avoid the actual decomposition of the Galerkin Jacobian matrices. Moreover, Dubois and Mehlman corrector assumes a strictly hyperbolic system, while this is generally not the case for the Galerkin system, where eigenvalues with multiplicity larger than one are common, leading to difficulties in defining the intermediate states. To overcome these issues, Theorem 3 in [14] provides us with an inexpensive estimate of the eigenvectors of the Roe Galerkin Jacobian matrix and an explicit correspondence between the approximate spectrum and the approximate eigenvectors of the Roe Galerkin Jacobian matrix for the various intermediate states. These facts play herein a central role to adapt the DM corrector to the present setting. A further important ingredient already exploited in [14] in a different context is that the Galerkin Jacobian matrix is block diagonal owing to the choice of piecewise continuous representation in the stochastic space. As a result, the DM corrector can be applied independently in each stochastic element. Finally, different indicators are derived to decide where the correction is actually needed, the purpose being to further reduce the computational costs.

The paper is organized as follows. In Section 2, we briefly recall the Galerkin projection of stochastic hyperbolic systems and the Roe-type solver introduced in [14]. The adaptation of the DM entropy corrector to the Galerkin system is described in Section 3. Simulations results are presented in Section 4. We adopt the following notation: lower case symbols represent deterministic quantities, whereas upper case symbols represent stochastic quantities.

## 2. Galerkin projection of stochastic hyperbolic systems and Roe-type solver

### 2.1. Stochastic hyperbolic systems

We are interested in stochastic nonlinear hyperbolic systems with uncertain input quantities that can be parametrized for simplicity by  $N$  independent identically distributed random variables  $\xi \equiv \{\xi_1, \dots, \xi_N\}$  uniformly distributed on  $\Xi = [0, 1]^N$ . Let us denote  $p_\xi = 1$  the density function of  $\xi$  and let  $L^2(\Xi, p_\xi)$  be the space of second-order random variables defined on the probability space  $\mathcal{P}_\xi \equiv (\Xi, \mathcal{B}_\Xi, p_\xi)$ , where  $\mathcal{B}_\Xi$  is the Borel set of  $\Xi$ . The expectation operator in  $\mathcal{P}_\xi$  is denoted for any random variable  $H(\xi)$  defined on  $\mathcal{P}_\xi$  by  $\langle H \rangle \equiv \int_\Xi H(y) p_\xi(y) dy$ . Let  $(x, t, \xi) \in \Omega \times [0, T] \times \Xi$ , where  $\Omega$  is the physical domain and  $T$  the simulation time. We seek for  $U(x, t, \xi) \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$  solving almost surely the following conservative system

$$\begin{cases} \frac{\partial}{\partial t} U(x, t, \xi) + \frac{\partial}{\partial x} F(U(x, t, \xi); \xi) = 0, \\ U(x, t = 0, \xi) = U_0(x, \xi), \end{cases} \quad (1)$$

where  $\mathcal{A}_U \subset \mathbb{R}^m$ ,  $m \geq 1$  is the set of admissible values for the solutions of the deterministic version of (1) and  $F(U; \xi) \in \mathbb{R}^m \otimes L^2(\Xi, p_\xi)$  is the stochastic flux vector. For instance, for the one-dimensional Euler system,  $m = 3$ , the components of  $U$  are the density, the impulse, and the total energy, and  $\mathcal{A}_U$  corresponds to the states for which the density and pressure are positive. The stochastic system (1) is assumed to be hyperbolic in the sense that the stochastic Jacobian matrix  $\nabla_U F \in \mathbb{R}^{m,m} \otimes L^2(\Xi, p_\xi)$  is  $\mathbb{R}$ -diagonalizable almost surely.

### 2.2. Stochastic discretization

We discretize the stochastic space  $\Xi$  using tensor-product stochastic elements, resulting from successive dyadic partitions of  $\Xi$ , and supporting local polynomial bases; we approximate the solution in the stochastic space of fully tensorized piecewise polynomial functions  $\mathcal{S}^{N_o, N_r}$ , where  $N_r \geq 0$  denotes the resolution level (controlling the size  $2^{-N_r}$  of the stochastic elements) and  $N_o \geq 0$  denotes the expansion order (controlling the degree of the piecewise polynomial approximation in each variable  $\xi_i$ ), see details in [9]. Therefore, the space  $\mathcal{S}^{N_o, N_r}$  has dimension

$$\dim \mathcal{S}^{N_o, N_r} = P_\pi P_\sigma =: P + 1, \quad (2)$$

where  $P_\pi := (N_o + 1)^N$  is the dimension of the local polynomial basis on each stochastic element, and  $P_\sigma := 2^{N_r}$  is the number of stochastic elements.

We use throughout this paper the so-called Stochastic Element (SE) basis to span  $\mathcal{S}^{N_o, N_r}$ , which corresponds to local fully tensorized (rescaled) Legendre polynomial bases such that the approximate solution in  $\mathcal{S}^P := \mathcal{S}^{N_o, N_r}$  is expanded as a series in the form

$$U(x, t, \xi) \approx U^P(x, t, \xi) = \sum_{\alpha=0}^P u_\alpha(x, t) \Psi_\alpha(\xi), \quad (3)$$

where the deterministic  $\mathbb{R}^m$ -valued fields  $u_\alpha(x, t)$  are called the stochastic modes of the solution (in  $\mathcal{S}^P$ ) and  $\{\Psi_\alpha\}_{\alpha=0, \dots, P}$  denotes the SE basis. In practice,  $\alpha$  represents a double index  $(\alpha_\sigma, \alpha_\pi)$ , where  $\alpha_\sigma$  refers to the stochastic element and  $\alpha_\pi$  to the polynomial function in the stochastic element.

### 2.3. The Galerkin system

Projecting (1) on the SE basis, we obtain the Galerkin system which couples all the stochastic modes in the form

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} f(u(x, t)) = 0, \\ u(x, t = 0) = u^0(x), \end{cases} \quad (4)$$

where  $u(x, t) = (u_0(x, t), \dots, u_P(x, t))^T \in \mathbb{R}^{m(P+1)}$ , and  $f(u(x, t)) = (f_0(u), \dots, f_P(u))^T \in \mathbb{R}^{m(P+1)}$  are respectively the vector of the stochastic modes and the Galerkin flux vector with

$$f_\alpha(u) \equiv \langle \Psi_\alpha F(U^P; \cdot) \rangle, \quad \alpha = 0, \dots, P, \quad \text{and} \quad U^P = \sum_{\beta=0}^P u_\beta \Psi_\beta(\xi). \quad (5)$$

Moreover,  $u^0 = (\langle \Psi_\alpha U^0 \rangle)_{\alpha=0, \dots, P}$ . The component vector  $u$  must belong to the admissible set  $\mathcal{A}_u \subset \mathbb{R}^{m(P+1)}$  such that  $u \in \mathcal{A}_u \Leftrightarrow U^P(\xi) = \sum_{\alpha=0}^P u_\alpha \Psi_\alpha(\xi) \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$ . The Galerkin Jacobian matrix of order  $m(P+1)$  defined as

$$(\nabla_u f(u))_{\alpha, \beta=0, \dots, P} = \langle \nabla_U F(U^P; \cdot) \Psi_\alpha \Psi_\beta \rangle_{\alpha, \beta=0, \dots, P} \quad (6)$$

has a diagonal block structure owing to the decoupling of the problem over different stochastic elements. In particular, for a given stochastic element  $\alpha_\sigma$ ,  $1 \leq \alpha_\sigma \leq P_\sigma$ , the corresponding diagonal block of size  $P_\sigma \times P_\sigma$  is denoted by  $[\nabla_u f]^{\alpha_\sigma}$ . In the sequel, all the developments are performed on each stochastic element independently. To alleviate the notation, we drop the superscript  $\alpha_\sigma$ .

In two specific cases, the Galerkin system (4) has been shown in [14] to be hyperbolic, that is, the Galerkin Jacobian matrix  $\nabla_u f$  is  $\mathbb{R}$ -diagonalizable. These correspond to stochastic problems (1) with symmetric Jacobian matrices  $\nabla_U F$  (almost surely) or with deterministic eigenvectors (*i.e.* not affected by the uncertainty). In the general case, we consider the approximate Galerkin Jacobian matrix  $\overline{\nabla_u f}$  which is also block-diagonal and such that for each stochastic element  $\alpha_\sigma$ ,  $1 \leq \alpha_\sigma \leq P_\sigma$ , the coefficient of the corresponding block  $[\overline{\nabla_u f}]^{\alpha_\sigma}$  are obtained by approaching the coefficients of  $[\nabla_u f]^{\alpha_\sigma}$  by a Gauss quadrature in the form

$$\left( [\overline{\nabla_u f}(u)]^{\alpha_\sigma} \right)_{\alpha_\pi, \beta_\pi=1, \dots, P_\pi} = \left( \sum_{\gamma=1}^{P_\pi} \varpi_\gamma \nabla_U F(U^P(\xi_\gamma); \xi_\gamma) \Psi_{\alpha_\sigma, \alpha_\pi}(\xi_\gamma) \Psi_{\alpha_\sigma, \beta_\pi}(\xi_\gamma) \right)_{\alpha_\pi, \beta_\pi=1, \dots, P_\pi}, \quad (7)$$

where  $\{\xi_\gamma\}_{\gamma=1, \dots, P_\pi}$  is the set of the Gauss points of the stochastic element  $\alpha_\sigma$ , with associated weights  $\{\varpi_\gamma\}_{\gamma=1, \dots, P_\pi}$ . This matrix has been proven in [14] to be  $\mathbb{R}$ -diagonalizable with eigenvalues  $\{\lambda'_{k\eta}\}_{k=1, \dots, m, \eta=1, \dots, P_\pi}$  and right and left eigenvectors  $\{r'_{k\eta}\}_{k=1, \dots, m, \eta=1, \dots, P_\pi}$  and  $\{l'_{k\eta}\}_{k=1, \dots, m, \eta=1, \dots, P_\pi}$  given by

$$\begin{cases} \lambda'_{k\eta} = \Lambda^k(\xi_\eta), & \forall k = 1, \dots, m, \eta = 1, \dots, P_\pi, \\ (r'_{k\eta})_{\beta=1, \dots, P_\pi} = (\varpi_\eta R^k(\xi_\eta) \Psi_\beta(\xi_\eta))_{\beta=1, \dots, P_\pi}, & \forall k = 1, \dots, m, \eta = 1, \dots, P_\pi, \\ (l'_{k\eta})_{\beta=1, \dots, P_\pi} = (\varpi_\eta L^k(\xi_\eta) \Psi_\beta(\xi_\eta))_{\beta=1, \dots, P_\pi}, & \forall k = 1, \dots, m, \eta = 1, \dots, P_\pi, \end{cases} \quad (8)$$

where  $\{\Lambda(\xi)\}_{k=1, \dots, m}$ ,  $\{R^k(\xi)\}_{k=1, \dots, m}$ , and  $\{L^k(\xi)\}_{k=1, \dots, m}$  are respectively the eigenvalues and right and left eigenvectors of  $\nabla_U F(U^P(\xi); \xi)$ . In [14], we used the spectrum of  $\overline{\nabla_u f}$  to approximate the spectrum of  $\nabla_u f$  in the framework of a procedure to compute upwind matrices for Roe-type solvers; this procedure is briefly recalled in section 2.4. In section 3, we present an original method to implement an entropy corrector using the eigenvectors of  $\overline{\nabla_u f}$  to approximate the eigenvectors of  $\nabla_u f$ .

To shorten the notation, it is convenient to use a single index for the eigenvalues and eigenvectors in the form  $\{\lambda'_\gamma\}_{\gamma=1, \dots, mP_\pi}$ ,  $\{r'_\gamma\}_{\gamma=1, \dots, mP_\pi}$ , and  $\{l'_\gamma\}_{\gamma=1, \dots, mP_\pi}$ . The prime indicates that these quantities approximate their counterpart associated with  $\nabla_u f$ .

#### 2.4. Roe-type solver without entropy correction

The Galerkin system (4) is discretized in physical space and time using a FV method in the form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (\varphi(u_i^n, u_{i+1}^n) - \varphi(u_{i-1}^n, u_i^n)), \quad (9)$$

where  $u_i^n$  is an approximation to the cell-average in physical space of the solution  $u$  in the cell of center  $i\Delta x$  with width  $\Delta x$  at the time discrete  $t^n$  and  $\varphi(\cdot, \cdot)$  is the Galerkin numerical flux.

The method presented in [14] is based on a Roe-type solver without entropy correction. As in the deterministic context, it can be viewed as an approximate Riemann solver where the Galerkin flux  $f(u)$  is replaced at each interface  $LR$  separating left and right states  $(u_L, u_R)$  by the linear approximation

$$f^{\text{Roe}}(u_L, u_R, u) = f(u_L) + a(u_L, u_R) \cdot (u - u_L), \quad (10)$$

where  $a(u_L, u_R)$  is a Roe-linearized approximation of the Galerkin Jacobian matrix. To define  $a(u_L, u_R)$ , we assume that the original stochastic problem (1) possesses a Roe state  $U_{LR}^{\text{Roe}}(\xi)$  on each interface  $LR$  separating left and right stochastic states  $(U_L(\xi), U_R(\xi))$  such that  $\nabla_U F(U_{LR}^{\text{Roe}}(\xi))$  is a Roe linearized matrix for the stochastic problem. Then, we have proven in [14] that

$$a(u_L, u_R) = \langle \nabla_U F(U_{LR}^{\text{Roe}}; \cdot) \Psi_\alpha \Psi_\beta \rangle_{\alpha, \beta=0, \dots, P} \quad (11)$$

is a Roe linearized matrix for the Galerkin problem, under the assumption that this matrix is  $\mathbb{R}$ -diagonalizable. Introducing  $u_{LR}^{\text{Roe}} = (\langle \Psi_\alpha U_{LR}^{\text{Roe}} \rangle)_{\alpha=0, \dots, P} \in \mathbb{R}^{m(P+1)}$ , we use the notation  $\nabla_u f(u_{LR}^{\text{Roe}})$  instead of  $a(u_L, u_R)$ , that is,  $u_{LR}^{\text{Roe}}$  plays formally the role of a Roe state for the Galerkin system and  $\nabla_u f(u_{LR}^{\text{Roe}})$  can be viewed as an approximate Roe linearized matrix. The Galerkin numerical flux is chosen in the form

$$\varphi(u_L, u_R) \equiv \varphi^{\text{Roe}}(u_L, u_R) = \frac{f(u_L) + f(u_R)}{2} - |\nabla_u f(u_{LR}^{\text{Roe}})| \frac{u_R - u_L}{2}. \quad (12)$$

Furthermore, so as to avoid the expensive spectral decomposition of the Roe Galerkin Jacobian matrix  $\nabla_u f(u_{LR}^{\text{Roe}})$  to compute its absolute value, we proposed in [14] to approximate  $|\nabla_u f(u_{LR}^{\text{Roe}})|$  by a polynomial transformation  $q_{d, \{\lambda'\}}$  applied to  $\nabla_u f(u_{LR}^{\text{Roe}})$ . This polynomial transformation is conveniently determined independently in each stochastic element  $\alpha_\sigma$ ,  $1 \leq \alpha_\sigma \leq P_\sigma$ , using the approximate spectrum  $\{\lambda'_\gamma\}_{\gamma=1, \dots, mP_\sigma}$  of  $[\nabla_u f(u_{LR}^{\text{Roe}})]^{\alpha_\sigma}$  and minimizing the error  $\sum_{\gamma=1}^{mP_\sigma} (|\lambda'_\gamma| - q_{d, \{\lambda'\}}(\lambda'_\gamma))^2$  (see [14] for additional details). Finally, the time-step  $\Delta^n t$  is computed using a CFL-type condition in the form

$$\frac{\Delta^n t}{\Delta x} = \frac{\text{CFL}}{\max_{LR \in \mathcal{I}, \gamma=1, \dots, mP_\sigma} |\lambda'_\gamma(u_{LR}^{\text{Roe}})|}, \quad (13)$$

where  $\mathcal{I}$  denotes the set of interfaces  $LR$  and CFL denotes a user-dependent parameter  $\leq 1$ . When working with more than one stochastic element, (13) yields a local time step  $\Delta^n t^{\alpha_\sigma}$ ,  $1 \leq \alpha_\sigma \leq P_\sigma$ , for each stochastic element, and the global time step is selected as  $\Delta^n t = \min_{1 \leq \alpha_\sigma \leq P_\sigma} \Delta^n t^{\alpha_\sigma}$ .

The above methodology has been extensively tested and analyzed in [14] on the stochastic Burgers and Euler equations without sonic points. It can be expected that an entropy corrector is needed as in the deterministic case to prevent entropy-violating shocks across sonic points. Test cases will effectively illustrate this point in section 4.

### 3. DM-type entropy correction

In this section, we propose an entropy correction for the above Roe-type solver based on an adaptation to the present context of the non-parametrized entropy corrector proposed by Dubois and Mehlman [4].

#### 3.1. General principles of the DM corrector

The methodology of Dubois and Mehlman holds for general deterministic hyperbolic systems. It consists in adding a correction to the numerical Roe flux to avoid entropy-violating shocks, the correction relying on a nonlinear modification of the numerical flux in the vicinity of sonic points. The detection and treatment of such shocks involve three steps. First, the eigenvalues corresponding to the Roe state at each physical interface separating left and right states are put in increasing order and the corresponding eigenvectors of the Roe Galerkin Jacobian matrix are ordered accordingly. Secondly, the components of the difference between the left and right states on the eigenvector basis are determined, and suitable intermediate states

are computed. Finally, a sonic expansion wave is detected between the  $(i - 1)$ -th and  $i$ -th intermediate states if the  $i$ -th eigenvalue associated to the  $(i - 1)$ -th intermediate state is negative whereas the  $i$ -th eigenvalue associated to the  $i$ -th intermediate state is positive. In this case, a nonlinear modification of the Roe numerical flux  $\varphi^{\text{Roe}}$  is added. An important aspect of the DM corrector is a correspondence between the sets of eigenvalues for all the intermediate states. This is reflected by the fact that the DM corrector has been applied to strictly hyperbolic deterministic systems where the Jacobian matrix  $\nabla_u f$  has real distinct eigenvalues and a complete set of linearly independent eigenvectors. Moreover, the left and right states at each interface are assumed to be sufficiently close, see [4, Theorem 6.1], so that the correspondence between the eigenvalues can be achieved through a common reference state.

We present now the direct application of the DM corrector to the numerical scheme of section 2.4. At each interface  $LR$  of the physical space with left and right states  $u_L$  and  $u_R$  respectively, we assume for the time being that the Roe Galerkin Jacobian matrix  $[\nabla_u f(u_{LR}^{\text{Roe}})]^{\alpha_\sigma}$  of the stochastic element  $\alpha_\sigma$ ,  $1 \leq \alpha_\sigma \leq P_\sigma$ , is  $\mathbb{R}$ -diagonalizable with distinct eigenvalues. Then, we can write  $[\nabla_u f(u_{LR}^{\text{Roe}})]^{\alpha_\sigma} = \sum_{\gamma=1}^{mP_\pi} \lambda_\gamma(u_{LR}^{\text{Roe}}) l_\gamma(u_{LR}^{\text{Roe}}) \otimes r_\gamma(u_{LR}^{\text{Roe}})$  with eigenvalues  $(\lambda_\gamma)_{\gamma=1, \dots, mP_\pi}$ , left eigenvectors  $(l_\gamma)_{\gamma=1, \dots, mP_\pi}$ , and right eigenvectors  $(r_\gamma)_{\gamma=1, \dots, mP_\pi}$ . The detection of the sonic points needs to build for all  $(u_L, u_R)$  the  $mP_\pi$  intermediate states

$$\begin{aligned} u_0 &= u_L \\ &\vdots \\ u_\gamma &= u_{\gamma-1} + \alpha_\gamma r_\gamma(u_{LR}^{\text{Roe}}) \\ &\vdots \\ u_{mP_\pi} &= u_R \end{aligned}$$

where  $(\alpha_\gamma)_{\gamma=1, \dots, mP_\pi}$  are the components of the vector  $u_L - u_R$  in the basis of right eigenvectors,

$$u_L - u_R = \sum_{\gamma=1}^{mP_\pi} \alpha_\gamma r_\gamma(u_{LR}^{\text{Roe}}). \quad (14)$$

Then, the set of sonic indices is  $S = \{\gamma, \lambda_\gamma(u_{\gamma-1}) < 0 < \lambda_\gamma(u_\gamma)\}$  and an interface is said to be sonic if  $S$  is non-empty. We modify  $f^{\text{Roe}}(u_L, u_R, u)$  only at sonic interfaces. To this end, we introduce the following modified flux function parametrized by  $u_L$  and  $u_R$ :

$$f^{\text{DM}}(u_L, u_R, u) = f(u_L) + \sum_{\gamma=1}^{mP_\pi} g_\gamma(w_\gamma(u)) r_\gamma(u_{LR}^{\text{Roe}}), \quad (15)$$

where  $(w_\gamma(u))_{\gamma=1, \dots, mP_\pi}$  are the characteristic variables defined by

$$u - u_L = \sum_{\gamma=1}^{mP_\pi} w_\gamma(u) r_\gamma(u_{LR}^{\text{Roe}}).$$

The functions  $(g_\gamma)_{\gamma=1, \dots, mP_\pi}$  are parametrized by the intermediate states  $(u_\gamma)_{\gamma=1, \dots, mP_\pi}$  and are defined according to

$$\begin{aligned} &\text{if } \gamma \notin S, \quad \forall w, \quad g_\gamma(w) = \lambda_\gamma(u_{LR}^{\text{Roe}})w, \\ &\text{if } \gamma \in S, \quad g_\gamma(w) = \begin{cases} p_\gamma(w) & \forall w \in (0, \alpha_\gamma), \\ \lambda_\gamma(u_{LR}^{\text{Roe}})w & \forall w \notin (0, \alpha_\gamma), \end{cases} \end{aligned}$$

where  $p_\gamma$  is the unique Hermite polynomial of degree 3 defined by the conditions:

$$p_\gamma(0) = 0, \quad p_\gamma(\alpha_\gamma) = \lambda_\gamma(u_{LR}^{\text{Roe}})\alpha_\gamma, \quad p'_\gamma(0) = \lambda_\gamma(u^{\gamma-1}), \quad p'_\gamma(\alpha_\gamma) = \lambda_\gamma(u_\gamma).$$

Elementary algebra shows that  $p_\gamma(w) = aw^3 + bw^2 + cw$  with

$$a = \frac{\lambda_\gamma(u_\gamma) + \lambda_\gamma(u_{\gamma-1}) - 2\lambda_\gamma(u_{LR}^{\text{Roe}})}{\alpha_\gamma^2}, \quad b = \frac{3\lambda_\gamma(u_{LR}^{\text{Roe}}) - 2\lambda_\gamma(u_{\gamma-1}) - \lambda_\gamma(u_\gamma)}{\alpha_\gamma}, \quad c = \lambda_\gamma(u_{\gamma-1}).$$

Away from sonic points, the modified flux  $f^{\text{DM}}$  coincides with the linearized Roe flux  $f^{\text{Roe}}$ . Finally, the modified Galerkin numerical flux  $\varphi^{\text{DM}}(u_L, u_R)$  has for expression

$$\varphi^{\text{DM}}(u_L, u_R) = \varphi^{\text{Roe}}(u_L, u_R) + \sum_{\gamma \in S} \max \left\{ \frac{g_\gamma(w_\gamma^*)}{\alpha_\gamma}, \frac{g_\gamma(w_\gamma^*)}{\alpha_\gamma} - \lambda_\gamma(u_{LR}^{\text{Roe}}) \right\} \alpha_\gamma r_\gamma(u_{LR}^{\text{Roe}}), \quad (16)$$

where

$$w_\gamma^* = \frac{-\lambda_\gamma(u_{\gamma-1})\alpha_\gamma}{3\lambda_\gamma(u_{LR}^{\text{Roe}}) - 2\lambda_\gamma(u_{\gamma-1}) - \lambda_\gamma(u_\gamma) + \sqrt{(3\lambda_\gamma(u_{LR}^{\text{Roe}}) - \lambda_\gamma(u_\gamma) - \lambda_\gamma(u_{\gamma-1}))^2 - \lambda_\gamma(u_{\gamma-1})\lambda_\gamma(u_\gamma)}}$$

is the value for which  $g_\gamma$  reaches its unique extremum in  $(0, \alpha_\gamma)$ .

### 3.2. Adaptation to the Galerkin system

An important point is that the Galerkin system is generally not strictly hyperbolic, i.e., some eigenvalues are multiple. Moreover, for practical reasons, we want to avoid the need of the actual spectral decomposition of the diagonal blocks of the Galerkin Jacobian matrix since the size of these blocks is still quite large. To this purpose, we consider for left and right states  $(u_L, u_R)$  and associated Roe Galerkin Jacobian matrix  $[\nabla_u f(u_{LR}^{\text{Roe}})]^{\alpha_\sigma}$ , the approximate eigenvalues  $\{\lambda'_\gamma\}_{\gamma=1, \dots, mP_\pi}$  and the corresponding approximate eigenvectors  $\{r'_\gamma\}_{\gamma=1, \dots, mP_\pi}$  defined by (8). We build the  $mP_\pi$  approximate intermediate states

$$\begin{aligned} u'_0 &= u_L \\ &\vdots \\ u'_\gamma &= u'_{\gamma-1} + \alpha'_\gamma r'_\gamma(u_{LR}^{\text{Roe}}) \\ &\vdots \\ u'_{mP_\pi} &= u_R \end{aligned}$$

where  $(\alpha'_\gamma)_{\gamma=1, \dots, mP_\pi}$  are the components of the vector  $u_L - u_R$  on the basis of approximate eigenvectors  $(r'_\gamma(u_{LR}^{\text{Roe}}))_{\gamma=1, \dots, mP_\pi}$ . A key ingredient is that the numbering of the approximate eigenvalues and eigenvectors in the Roe approximate state via the Gauss points enables us to make the correspondence between the eigenvalues of the various intermediate states, and consequently to determine the set of the sonic indices  $S = \{\gamma, \lambda'_\gamma(u'_{\gamma-1}) < 0 < \lambda'_\gamma(u'_\gamma)\}$  in a coherent way. We can then modify the Roe numerical flux  $\varphi^{\text{Roe}}(u_L, u_R)$  defined by (12) as described above.

## 4. Results

### 4.1. Test case 1: Burgers equation

The purpose of this test case is to assess the method on the stochastic scalar conservation law of Burgers in the case where a sonic point is present almost surely. We compare the numerical solution with the exact solution and study the spatial convergence of the method. A comparison with a Monte Carlo method is also presented.

#### 4.1.1. Problem definition

We consider the Burgers equation

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0, \quad F(U) = \frac{U^2}{2}, \quad (17)$$

with stochastic initial condition  $U^0(x, \xi)$  defined using two uncertain states, the first one almost surely negative and the second one almost surely positive. We take for  $x \in [0, 1]$ ,

$$U^0(x, \xi) = \begin{cases} U^-(\xi_1) & x < 1/2, \\ U^+(\xi_2) & x \geq 1/2, \end{cases} \quad (18)$$

with

$$\begin{aligned} U^-(\xi_1) &= -1 + 0.05(2\xi_1 - 1), & \xi_1 &\sim \mathcal{U}[0, 1] \rightarrow U^- \sim \mathcal{U}[-1.05, -0.95], \\ U^+(\xi_2) &= 1 + 0.1(2\xi_2 - 1), & \xi_2 &\sim \mathcal{U}[0, 1] \rightarrow U^+ \sim \mathcal{U}[0.9, 1.1]. \end{aligned} \quad (19)$$

Therefore, the problem has 2 stochastic dimensions ( $N = 2$ ). The initial condition is illustrated in Figure 1. The stochastic Burgers equation is time-integrated using the methodology described above. Unless specified,  $N_c = 250$  cells are used for space discretization with a CFL constant set to 0.95 in (13), and the stochastic parameters are  $N_o = 3$  for the expansion order and  $N_r = 3$  for the resolution level so that the dimension of the stochastic approximation space is  $\dim \mathcal{S}^{N_o, N_r} = (N_o + 1)^{N_r} 2^{N_r} = 1024$ . Since  $m = 1$  here, the Galerkin system is of size 1024. The computation of the nonlinear flux  $F(U)$  relies on spectral methods detailed in [14].

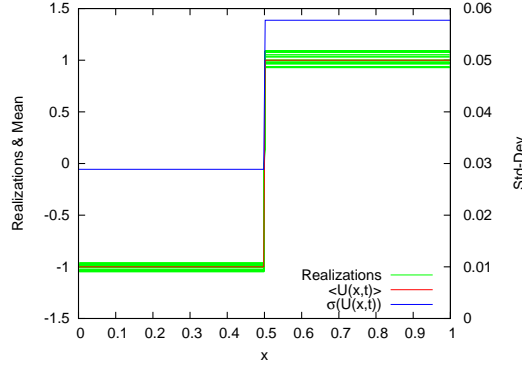


Figure 1: Random initial condition for test case 1: sample set of 20 random realizations, mean, and standard deviation. Scale on the right for the standard deviation.

For the random initial conditions specified by (19) and setting  $t_s = \frac{x-1/2}{t}$ , the (exact) stochastic entropy solution of the Burgers equation consists in an expansion wave whose expression can be easily derived for  $t > 0$  and  $\xi = (\xi_1, \xi_2) \in [0, 1]^2$ :

$$U^{ex}(x, t, (\xi_1, \xi_2)) = \begin{cases} U^-(\xi_1) & t_s \leq U^-(\xi_1), \\ t_s & U^-(\xi_1) < t_s < U^+(\xi_2), \\ U^+(\xi_2) & t_s \geq U^+(\xi_2). \end{cases} \quad (20)$$

An important remark is that the solution is independent of the uncertainty for  $(x, t)$  such that

$$\sup_{\xi_1 \in [0, 1]} U^-(\xi_1) < t_s < \inf_{\xi_2 \in [0, 1]} U^+(\xi_2), \quad (21)$$

so that the variance vanishes for  $(x, t)$  satisfying this condition. Specifically, this yields  $-0.95t + 0.5 < x < 0.9t + 0.5$ . Moreover, it is readily verified that the mean and standard deviation of the exact solution are both piecewise affine functions in physical space at all times.

#### 4.1.2. Numerical results

We compare in Figure 2 the stochastic solution at  $t = 0.3$  obtained without entropy corrector (left) and with the corrector proposed above (right). The solution expectation and standard deviation, together with a random sample set of realizations, are plotted. The realizations are reconstructed from the stochastic expansions of the solutions using a unique set of randomly generated realizations of  $\xi \in [0, 1]^2$ . As in the deterministic case, we observe in the left panel of Figure 2 that the Roe-type solver without entropy corrector does not capture the entropy solution of the problem. In fact, it captures an entropy-violating shock together with small perturbations due to the fact that the right and left initial states are almost surely non symmetric. With the additional entropy correction proposed above, the entropy solution is well-captured, as shown in the right panel of Figure 2. Moreover, we have verified that the sonic points of the Galerkin system are, as in the deterministic case, detected only for the interface located at  $x = 0.5$ . In fact, the uncertainty is not very important for this test case so that all the sonic waves have the same physical location almost surely for the times we consider. We have also noticed that the set of the sonic indices in a stochastic element corresponding to  $x = 0.5$  has cardinality  $(No + 1)^2$ , meaning that all the eigenvalues of the Galerkin Jacobian matrices change their sign at the interface and consequently all the components of the Galerkin numerical flux need to be supplemented by an entropy correction.

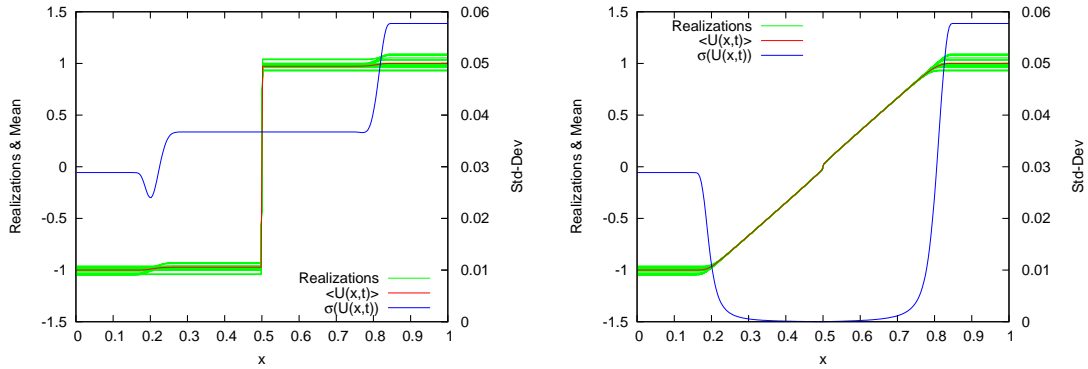


Figure 2: Stochastic solution of the Burgers equation at  $t = 0.3$  obtained without (left) and with (right) entropy corrector. The solution mean and standard deviation are plotted as a function of  $x$ , together with a reconstruction of 20 randomly generated realizations of the solution. Scale on the right for the standard deviation.

In Figure 3, the two first moments of the computed solution are compared with the two first moments of the exact solution for two different choices of the parameter CFL. It is seen that the means of the computed and exact solutions are in agreement in the two cases, although, as in the deterministic case, a slight jump remains at  $x = 0.5$  for the computed mean. The standard deviation is correctly captured by our method: it is almost zero on an interval included in  $[-0.95t + 0.5, 0.9t + 0.5]$  (where the exact standard deviation is zero). However, significant errors remain in the vicinity of  $x = 0.2$  and  $x = 0.8$  especially for the case  $CFL = 0.7$  since the Roe-type scheme causes a spreading of the numerical solution in physical space. This issue can be partly handled by increasing the CFL constant, as reflected in the right panel of Figure 3.

A further interesting observation is that the numerical solution depends almost surely only on  $\xi_1$  for  $x < 0.5$  and only on  $\xi_2$  for  $x > 0.5$ , as does the exact solution. This is shown in Figure 4 where the stochastic solution  $U(x, t, \xi)$  is plotted as a function of  $\xi = (\xi_1, \xi_2)$  for  $x = 0.2$  (left) and  $x = 0.8$  (right). In Figure 5, we represent the stochastic solution at  $t = 0.3$  as a function of  $(x, \xi_1)$  for  $0 \leq x \leq 0.25$  (left) and as a function of  $(x, \xi_2)$  for  $0.75 \leq x \leq 1$  (right). This figure confirms the validity of the stochastic expansion since the stochastic solution is smooth as expected. In addition, it corroborates the use of a polynomial degree  $No > 1$  to correctly represent the nonlinearity of the solution with respect to the stochastic variable.

The jump which remains in the vicinity of the sonic point at  $x = 0.5$  is expected as in the deterministic case, because the numerical solver is first-order. Since the standard deviation of the solution is almost surely zero at the center of the domain, the jump coincides with the value obtained in the deterministic case almost surely. The jump can consequently be represented in the left panel of Figure 6 as a function of the spatial

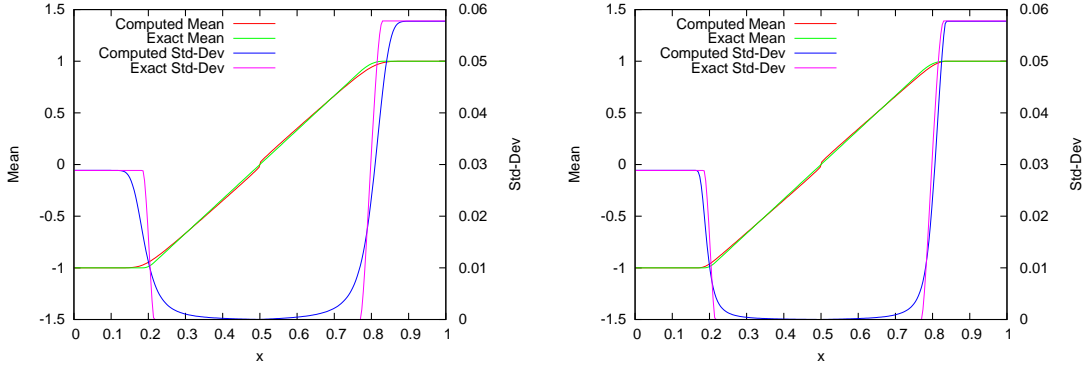


Figure 3: Stochastic Burgers equation: comparison of the mean and standard deviation of the numerical solution at  $t = 0.3$  with the corresponding mean and standard deviation of the exact solution for  $CFL = 0.7$  (left) and  $CFL = 0.99$  (right).

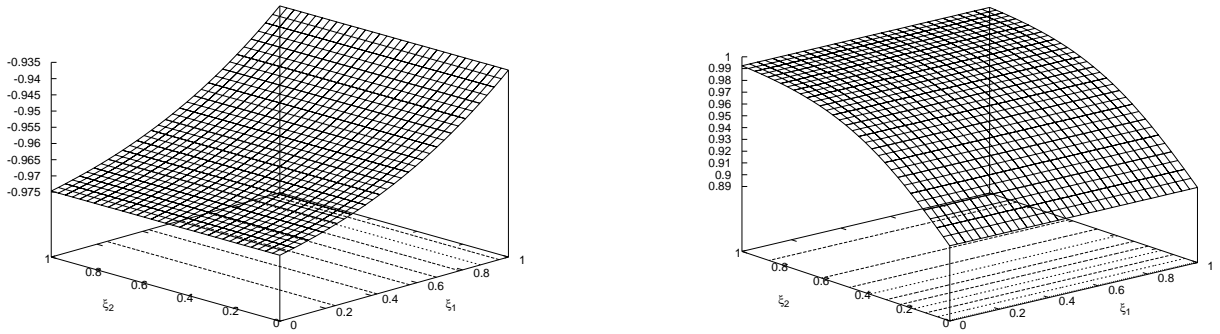


Figure 4: Stochastic solution of the Burgers equation as a function of  $(\xi_1, \xi_2)$  at  $t = 0.3$  and  $x = 0.2$  (left) or  $x = 0.8$  (right).

step by taking the difference between the value of the mean of the computed solution on the right cell and on the left cell of the interface located at  $x = 0.5$ . It decreases linearly with the spatial step; this is the same behavior as in the deterministic case.

To further analyze the spatial convergence of the method, we consider the error measure

$$\epsilon_{ex}^2(t) \equiv \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^{Nc} \Delta x \left( U_j^{\text{No,Nr}}(t, \xi^{(i)}) - U^{ex}(x_j, t, \xi^{(i)}) \right)^2, \quad (22)$$

where, for each element  $\xi^{(i)}$  in a sample set,  $U_j^{\text{No,Nr}}(t, \xi^{(i)})$  and  $U^{ex}(x_j, t, \xi^{(i)})$  are respectively evaluated from the computed solution at the cell  $j$  in physical space and from the exact solution of the corresponding deterministic problem at the center of the cell  $j$  in physical space. We use a sample set of size  $M = 100000$  uniformly drawn from  $\Xi$ . The right panel of Figure 6 reports the error  $\epsilon_{ex}^2$  at time  $t = 0.3$  as a function of the spatial step. We observe a linear decay rate of the error because the spatial error largely dominates in this test case where the solution is smooth both in physical and stochastic spaces.

#### 4.1.3. Comparison of the Galerkin method with a Monte Carlo method

We finally compare the Galerkin solution with a Monte Carlo (MC) solution obtained from the resolution of 100000 deterministic problems, each of these problems corresponding to a realization of  $\xi$  in a sample set. The two methodologies yield different results, as shown in Figure 7. In the left panel, the absolute value

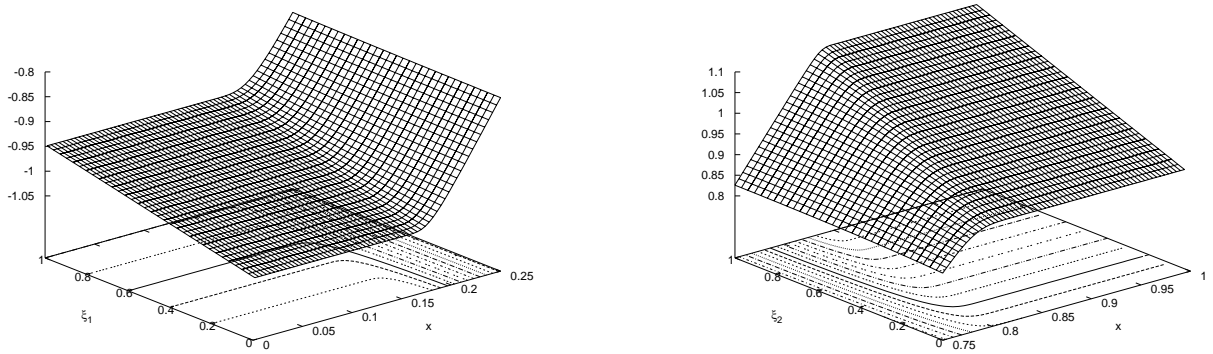


Figure 5: Stochastic solution of the Burgers equation at  $t = 0.3$  as a function of  $(x, \xi_1)$  for  $0 \leq x \leq 0.25$  (left) and  $(x, \xi_2)$  for  $0.75 \leq x \leq 1$  (right).

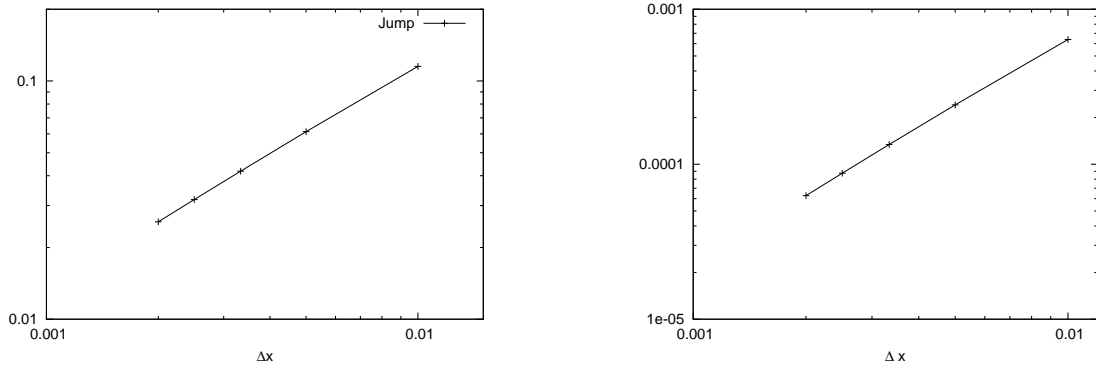


Figure 6: Left: Height of the jump of the numerical solution as a function of the spatial step  $\Delta x$ . Right: Stochastic error  $\epsilon_{ex}^2$  as a function of the spatial step  $\Delta x$ . Computations at  $t = 0.3$ .

of the difference between the standard deviation of the exact solution and that of the Galerkin solution is compared with the absolute value of the difference between the standard deviation of the exact solution and that of the MC solution. We notice that the Galerkin method yields better results than the MC method in the vicinity of the sonic point at  $x = 0.5$ . Indeed, the standard deviation does not vanish for the MC method because the latter considers a time step  $\Delta t$  specific to each realization of  $\xi$  and consequently a time step different, thereby creating a variability of the jump at  $x = 0.5$ . The Galerkin method considers the same time step  $\Delta t$  for all the stochastic elements, see discussion after (13), and therefore overcomes these difficulties by introducing some numerical diffusion. However, this diffusion creates numerical errors in the vicinity of  $x = 0.2$  and  $x = 0.8$  which are more important than for the MC method. Let us consider the MC error measure

$$\epsilon_{ex,MC}^2(t) \equiv \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^{Nc} \Delta x \left( U_j^{MC}(t, \xi^{(i)}) - U^{ex}(x_j, t, \xi^{(i)}) \right)^2, \quad (23)$$

where, for each element  $\xi^{(i)}$  in a sample set,  $U_j^{MC}(t, \xi^{(i)})$  and  $U^{ex}(x_j, t, \xi^{(i)})$  are respectively evaluated from the solution of the corresponding deterministic problem at the cell  $j$  in physical space computed with the deterministic solver and from the exact solution of the corresponding deterministic problem at the center of the cell  $j$  in physical space. Again, a sample set of size  $M = 100000$  is used. In the right panel of Figure 7, we compare the quantities  $\epsilon_{ex,MC}^2$  and  $\epsilon_{ex}^2$  as a function of the spatial step. All in all, the MC error is slightly

smaller than the Galerkin error. Lastly, we have noticed that taking a constant  $\Delta t$  for the MC method corresponding to the time step of the Galerkin method produces the same results as the Galerkin method.

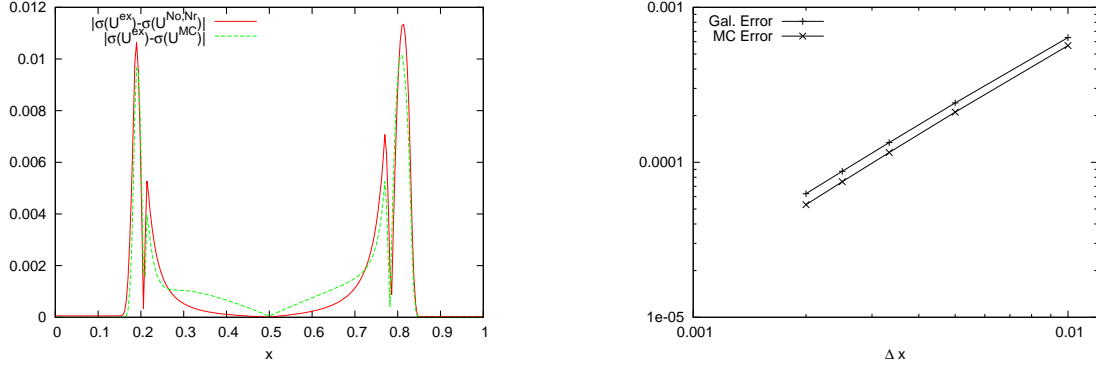


Figure 7: Left: Difference between the standard deviation of the exact solution and the Galerkin solution compared with the difference between the standard deviation of the exact solution and the MC solution. Right: Stochastic error  $\epsilon_{ex}^2$  as a function of the spatial step  $\Delta x$  for the Galerkin solution and the MC solution. Computations at  $t = 0.3$ .

#### 4.2. Test case 2: Euler equations

In this section, the method is tested on the stochastic Euler equations with one random parameter. The goal of this test case is to assess the method on a nonlinear hyperbolic system of conservation laws, so that the set of approximate eigenvectors of the Galerkin system is more complex than in the scalar case. Moreover, we propose a test case where sonic points appear just in a portion of the stochastic space. We consider the one-dimensional Sod shock-tube problem, where the flow of an ideal gas is governed by the Euler equations. Conventional thermodynamic notation is used instead of the lower/upper case convention adopted previously. The conserved quantities are the fluid density  $\rho$ , the impulse  $q = \rho v$  (with  $v$  the velocity), and the total energy  $E = 1/2 \rho v^2 + \rho e$ , where the first term is the kinetic energy and the second one the internal energy (per unit volume). The tube extends over one unit of length and is opened at  $x = 0$  and  $x = 1$ . Boundary conditions are  $\frac{\partial \rho}{\partial x} = \frac{\partial q}{\partial x} = \frac{\partial E}{\partial x} = 0$  at both extremities of the tube. The discretization uses  $N_c = 250$  cells in physical space.

##### 4.2.1. Problem definition

We consider an uncertainty on the initial Mach number parametrized by a unique random variable  $\xi$  having a uniform distribution in  $[0, 1]$ :

$$Ma^0(\xi) = \begin{cases} 0.7 + 0.5\xi & x \in [0, 1/4], \\ 2.46 \times (0.7 + 0.5\xi) & x \in ]1/4, 1]. \end{cases} \quad (24)$$

We set

$$U(x, t, \xi) = (\rho(x, t, \xi), q(x, t, \xi), E(x, t, \xi)) \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi), \quad (25)$$

where  $\mathcal{A}_U \subset \mathbb{R}^3$  is the set of admissible states such that the density and the pressure are positive and

$$F(U; \xi) = (F_\rho(U; \xi), F_q(U; \xi), F_E(U; \xi)) = (q(\xi), (q^2/\rho + p)(\xi), (v(E + p))(\xi)) \in \mathbb{R}^3 \otimes L^2(\Xi, p_\xi), \quad (26)$$

with the pressure  $p$  given by the ideal gas law

$$p(\rho, q, E) = (\gamma - 1) \left( E - \frac{1}{2} \rho v^2 \right), \quad (27)$$

and  $\gamma = 1.4$  is the adiabatic coefficient. The other initial conditions are

$$p^0(x) = \begin{cases} 0.05 & x \in [0, 1/4], \\ 0.008 \times 0.05 & x \in ]1/4, 1], \end{cases} \quad \rho^0(x) = \begin{cases} 1.4 & x \in [0, 1/4], \\ 1.4 \times 0.03 & x \in ]1/4, 1]. \end{cases} \quad (28)$$

#### 4.2.2. Numerical solver

At each physical interface, the numerical flux  $\varphi^{\text{DM}}(u_L, u_R)$  defined by (16) has to be evaluated. We approximate the stochastic quantities present in the numerical flux by their projection on  $\mathcal{S}^P$  that we compute in a pseudo-spectral way. To this end, we rely on [3], which describes tools for accurate evaluations of polynomial and non-polynomial functions of variables represented by stochastic expansions. For the Euler equations, the nonlinearities can be handled by the use of three pseudo-spectral stochastic operations, namely the pseudo-spectral product  $*$ , the pseudo-spectral inverse  $^{-*}$  obtained from the resolution of a linear system, and the pseudo-spectral square root  $^{*/2}$  obtained from the resolution of a nonlinear system (see [3, 14]). The first component of  $\varphi^{\text{DM}}(u_L, u_R)$  is the Roe-type numerical flux  $\varphi^{\text{Roe}}(u_L, u_R)$  whose construction is detailed in [14]. We describe here the evaluation of the correction part of  $\varphi^{\text{DM}}(u_L, u_R)$  for the case of the Euler equations. To this end, we need the approximate eigenvalues and right eigenvectors of  $\nabla_u f(u_{LR}^{\text{Roe}})$ . We recall that the Galerkin Jacobian matrix  $\nabla_u f(u) \in \mathbb{R}^{3(P+1), 3(P+1)}$  is approximated as

$$\begin{aligned} \nabla_u f(u) &= \langle \nabla_U F(U^P; \cdot) \Psi_\alpha \Psi_\beta \rangle_{\alpha\beta} \\ &\approx \left( \sum_{\delta=0}^P (\nabla_U F^*(U^P; \cdot))_\delta \mathcal{M}_{\alpha\beta\delta} \right)_{\alpha\beta}, \end{aligned} \quad (29)$$

where  $\nabla_U F^*(U^P; \xi)$  is the pseudo-spectral approximation of the projection of  $\nabla_U F(\cdot; \xi)$  on  $\mathcal{S}^P$  [14].

Denoting by  $H^* \equiv (E^P + p^*) * \rho^{-*}$  the pseudo-spectral approximation of the enthalpy on  $\mathcal{S}^P$ ,  $\nabla_U F^*(U^P; \cdot)$  is defined as

$$\nabla_U F^*(U^P; \cdot) = \begin{pmatrix} 0 & 1 & 0 \\ 1/2(\gamma - 3)(v^* * v^*) & -(\gamma - 3)v^* & \gamma - 1 \\ 1/2(\gamma - 1)(v^* * (v^* * v^*)) - v^* * H^* & H^* - (\gamma - 1)(v^* * v^*) & \gamma v^* \end{pmatrix}. \quad (30)$$

Consequently, the eigenvalues of  $\nabla_u f(u_{LR}^{\text{Roe}})$  can be approximated by the approximation in  $\mathcal{S}^P$  of the stochastic eigenvalues of  $\nabla_U F^*(U_{LR}^{\text{Roe}, P}(\xi); \xi)$  at Gauss points in each stochastic element, that is,

$$(v_{LR}^{\text{Roe},*} \pm c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta=0, \dots, N_0}, \text{ and } v_{LR}^{\text{Roe},*}(\xi_\eta)_{\eta=0, \dots, N_0}, \quad (31)$$

where  $v^{\text{Roe}}(\xi)$  and  $c^{\text{Roe}}(\xi)$  are respectively the stochastic Roe velocity and corresponding sound velocity. Moreover, the eigenvectors of  $\nabla_u f(u_{LR}^{\text{Roe}})$  can be approximated by  $(\omega_\eta R^{k,*}(\xi_\eta) \Psi_\beta)_{\beta=0, \dots, N_0, \eta=0, \dots, N_0, k=1, 2, 3}$ , where  $(R^{k,*}(\xi))_{k=1, 2, 3}$  are the approximation in  $\mathcal{S}^P$  of the stochastic eigenvectors of  $\nabla_U F^*(U_{LR}^{\text{Roe}, P}(\xi); \xi)$ , that is,

$$\begin{aligned} R^{1,*} &= (1, v^{\text{Roe},*} + c^{\text{Roe},*}, H^{\text{Roe},*} + v^{\text{Roe},*} * c^{\text{Roe},*})^T, \\ R^{2,*} &= (1, v^{\text{Roe},*} - c^{\text{Roe},*}, H^{\text{Roe},*} + v^{\text{Roe},*} * c^{\text{Roe},*})^T, \\ R^{3,*} &= (1, v^{\text{Roe},*}, H^{\text{Roe},*} - (c^{\text{Roe},*} * c^{\text{Roe},*})/(\gamma - 1))^T, \end{aligned} \quad (32)$$

where  $H^{\text{Roe},*}(\xi)$  is the pseudo-spectral approximation in  $\mathcal{S}^P$  of the stochastic Roe enthalpy. Thus, we have all the tools necessary to compute the intermediate states for all the left and right states  $(u_L, u_R)$ . To determine the set of the sonic indices  $S'$ , we use the approximate left and right eigenvalues  $(v_L^* \pm c_L^*)(\xi_\eta)_{\eta=0, \dots, N_0}$  and  $v_L^*(\xi_\eta)_{\eta=0, \dots, N_0}$ , that we compare with  $(v_R^* \pm c_R^*)(\xi_\eta)_{\eta=0, \dots, N_0}$  and  $v_R^*(\xi_\eta)_{\eta=0, \dots, N_0}$ .

#### 4.2.3. Results analysis

In this section we present and analyze the results for the shock tube problem with uncertainty in the initial Mach number. We begin with a general analysis of the results, taking  $N_0 = 2$  and  $N_r = 3$  as stochastic discretization parameters, so that the dimension of the stochastic space is 24.

In the deterministic case and for the initial condition (28) for a certain realization of  $Ma^0(\xi)$ , a shock wave generated at the discontinuity travels to the right with velocity  $v + c$ , while a rarefaction fan travels to the left with velocity  $v - c$ , and a contact discontinuity wave travels to the right with velocity  $v$ . Here, the

uncertain initial Mach number will affect the propagation velocity of the shock, contact discontinuity, and rarefaction fan. Solutions for different realizations of  $Ma^0(\xi)$  exhibit similar patterns as in the deterministic case, but with different slopes for the shock, contact discontinuity, and rarefaction fan. Moreover, the uncertainty in the initial Mach number is such that sonic points are generated only for  $\xi \in [0, 0.6]$ . To assess the validity of the stochastic expansion and the effectiveness of the DM-type entropy corrector, we show in Figure 8 a reconstruction of the stochastic density  $\rho(x, t, \xi)$  at  $t = 1$  obtained without (left) and with (right) entropy corrector. On the left panel, we observe that an entropy-violating shock is generated at the vicinity of  $x = 0.2$  and for  $\xi \in [0, 0.6]$ , while for  $\xi \in [0.6, 1]$  the expansion wave is well-captured. When using the DM-type entropy corrector in the right panel of Figure 8, the sonic expansion wave is also well-captured, even if a small jump remains due to the first-order nature of the numerical scheme. We also mention that the solver captures well the discontinuity in physical as well as in stochastic space in the vicinity of the shock wave; this point has already been investigated in [14].

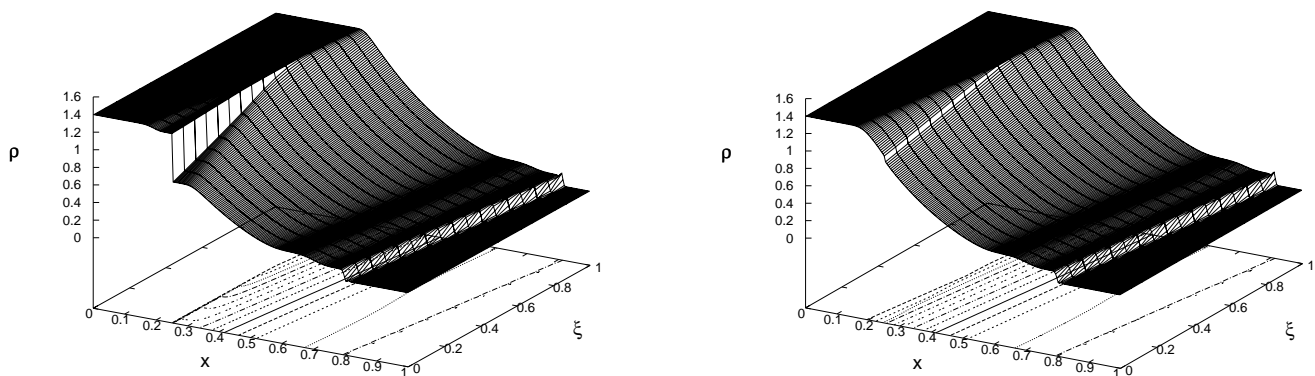


Figure 8: Reconstruction of the stochastic density  $\rho(x, t, \xi)$  at  $t = 1$  obtained without (left) and with (right) the entropy corrector. Computations with  $N_r = 3$  and  $N_o = 2$ .

We compare in Figure 9 at  $t = 1$  the mean and standard deviation of the density computed with the Galerkin method with that computed with a MC method where deterministic (discrete) Euler problems corresponding to a sample set of size 10000 are solved. Following the remarks in section 4.1.3, we compute the solution for the two methods with a fixed time step  $\Delta t$ ; this time step is for simplicity the minimum of the time steps obtained in the Galerkin solver. We notice that the means and standard deviations are in excellent agreement.

We will not study the stochastic convergence with respect to the stochastic parameters  $N_o, N_r$ , since a study has already been provided in [14]. Nevertheless, we can mention that for this test case, a low resolution level is sufficient to capture the sonic expansion wave, while a higher resolution level is needed in the vicinity of the shock wave to capture the discontinuity in the stochastic space. This fact corroborates the need for stochastic adaptivity, which is the focus of ongoing efforts.

#### 4.2.4. CPU improvements

We describe in this section a way to improve cost-effectiveness in the detection of sonic points. To illustrate our idea, we show in Figure 10 at  $t = 1$  and for selected interfaces at various positions in physical space, the approximate eigenvalues  $(v_{LR}^{\text{Roe},*} \pm c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta=0,\dots,N_o}$  and  $v_{LR}^{\text{Roe},*}(\xi_\eta)_{\eta=0,\dots,N_o}$  corresponding to each stochastic element and the associated density functions. This figure deserves several comments. First, we notice that the density functions have completely different patterns for the different interfaces considered. For  $x = 0.15$ ,  $x = 0.2$ , and  $x = 0.25$ , the sets of eigenvalues corresponding to  $v - c$ ,  $v$ , and  $v + c$  have distinct supports, while these sets have overlapping supports for  $x = 0.9$  (and are therefore represented in three

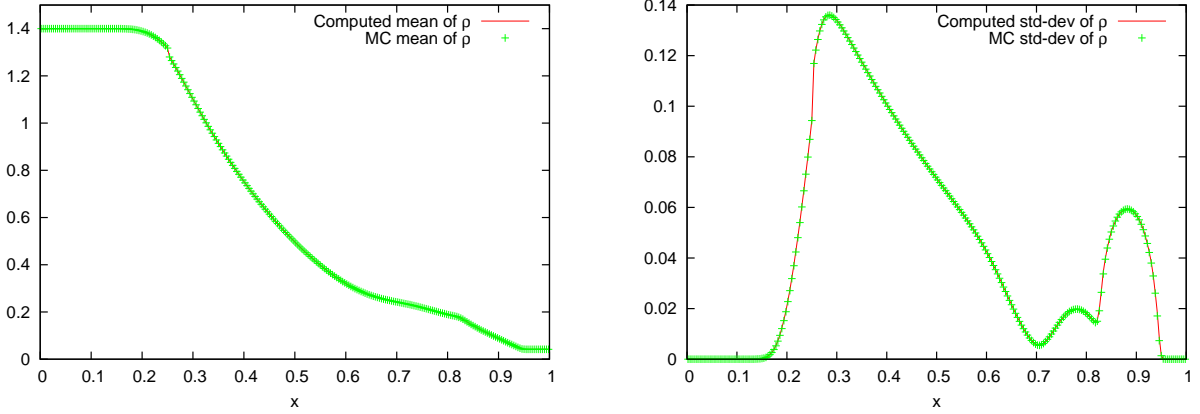


Figure 9: Comparison of the mean and standard deviation of the numerical density at  $t = 1$ , computed with a Galerkin method (using  $N_r = 3$  and  $N_o = 2$ ) and a MC method (using a sample set of size 10000).

different panels in the bottom row of Figure 10). For the interface  $x = 0.9$ , which is affected by the shock wave, it is interesting to notice that each set of eigenvalues can be split into two different subsets, except for a few eigenvalues which actually correspond to the discontinuity in the stochastic space. On the contrary, for the three panels in the upper row, the eigenvalues of each stochastic element are very close and the density functions hit their maximum values where the eigenvalues are practically equal, that is, for the stochastic elements supporting a low variability. This phenomenon is particularly visible at  $x = 0.2$  and  $x = 0.25$ . Moreover, the plateaus in the density function correspond to stochastic elements where eigenvalues have locally uniform distributions. A further important remark is that the first three panels correspond to positions located in the vicinity of the sonic points, and only the set of eigenvalues  $(v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta=0,\dots,N_o}$  can take negative values. Therefore, the key point to reduce computational time is to test the detection of sonic points only using the eigenvalues  $(v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta=0,\dots,N_o}$ . To further reduce the CPU time, we only test on each stochastic element  $\alpha_\sigma$ ,  $1 \leq \alpha_\sigma \leq P_\sigma$ , if the mean  $E^{\alpha_\sigma}[v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*}]$  changes its sign at the interface, in other words, if

$$(E^{\alpha_\sigma}[(v_L^* - c_L^*)] - \text{ctol} * v_{ref} < 0 < E^{\alpha_\sigma}[(v_R^* - c_R^*)] + \text{ctol} * v_{ref}), \quad 1 \leq \alpha_\sigma \leq P_\sigma, \quad (33)$$

where  $\text{ctol}$  is a fixed parameter and  $v_{ref}$  is a reference velocity here taken equal to one as reflected in the size of eigenvalues reported in Figure 11 below.

We show in Figure 11 the portion of the domain  $(x, \xi)$  which is actually selected for the entropy correction and the number of sonic points detected ( $\text{card } S'$ ) for two different values of the parameter  $\text{ctol}$ . For  $\text{ctol} = 1e^{-2}$  (left), we observe a rectangular portion of the domain on the left which is positive to the test (33) but does not actually correspond to sonic points. The flagged interfaces correspond in fact to the portion of the domain where the initial condition passes from the sonic case to the subsonic case. For both cases, three sonic points are detected in the four stochastic elements corresponding to  $\xi \in [0, 0.5]$ , and two sonic points are detected in the fifth stochastic element.

We close our discussion by an estimate of the CPU times reported in Table 1. We observe that the above ideas enable us to save a considerable amount of CPU time since the CPU time can be divided by a factor of two. Moreover, we indicate the fraction of stochastic cells which are actually tested for the correction. This quantity does not change significantly when decreasing  $\text{ctol}$  below  $1e^{-3}$ . Finally, we define the error measure on the density as

$$\epsilon_h^2(t) \equiv \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^{N_c} \Delta x \left( \rho_j^{\text{No}, \text{Nr}}(t, \xi^{(i)}) - \rho_j^{\text{MC}}(t, \xi^{(i)}) \right)^2, \quad (34)$$

where, for each element  $\xi^{(i)}$  in a sample set,  $\rho_j^{\text{No}, \text{Nr}}(t, \xi^{(i)})$  and  $\rho_j^{\text{MC}}(t, \xi^{(i)})$  are respectively evaluated from

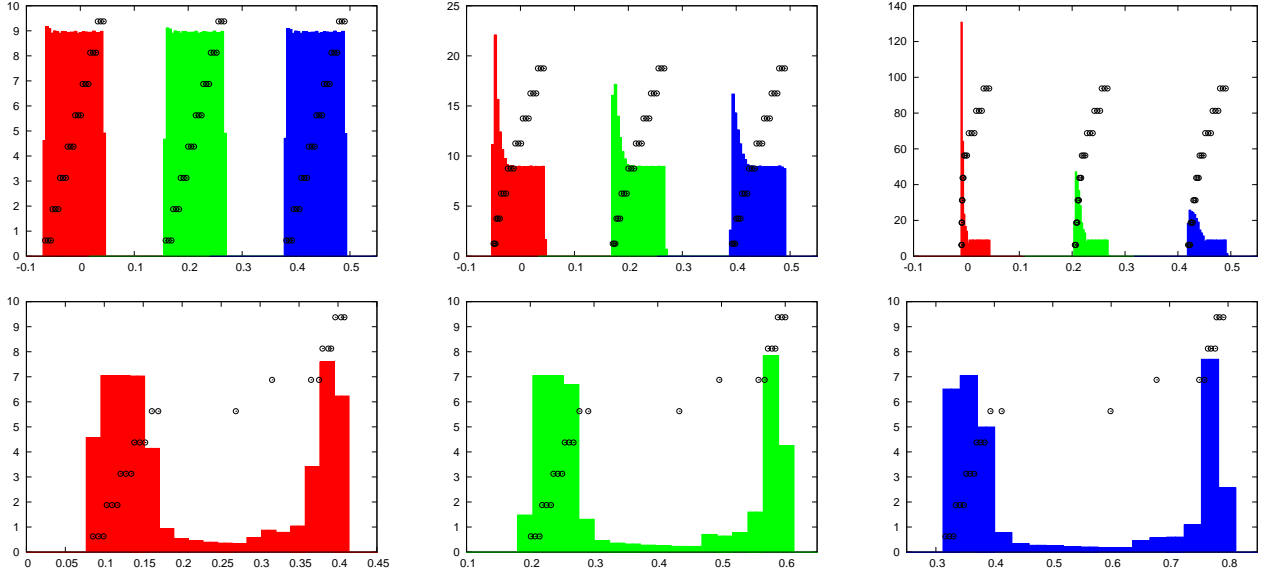


Figure 10: Approximate eigenvalues  $(v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta=0,\dots,N_o}$  (red/gray),  $v_{LR}^{\text{Roe},*}(\xi_\eta)_{\eta=0,\dots,N_o}$  (green/light gray), and  $(v_{LR}^{\text{Roe},*} + c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta=0,\dots,N_o}$  (blue/dark gray) corresponding to each stochastic element together with their density functions for  $x = 0.15$ ,  $x = 0.2$ ,  $x = 0.25$  (top), and  $x = 0.9$  (bottom). Computations at  $t = 1$  with  $N_r = 3$  and  $N_o = 2$ .

the computed solution at the cell  $j$  in physical space and from the solution of the corresponding deterministic problem at the cell  $j$  in physical space computed with the deterministic solver. We use a sample set of size  $M = 100000$  uniformly drawn from  $\Xi$ . In this specific case, we observe that the value of  $ctol$  does not affect the error demonstrating that a criteria based on the averaged velocity over the considered stochastic element is sufficient.

### 5. Conclusion and discussion

We have proposed an extension of the entropy corrector of Dubois and Mehlman [4] to the computation of the numerical Roe flux for the Galerkin system arising from the projection of stochastic hyperbolic equations. The proposed method relies on approximate eigenvalues and eigenvectors of the Roe Galerkin Jacobian matrix using the underlying Gauss quadrature points and weights, thereby avoiding the need for the costly decomposition of the Roe Galerkin Jacobian matrix. In addition, such a procedure can be

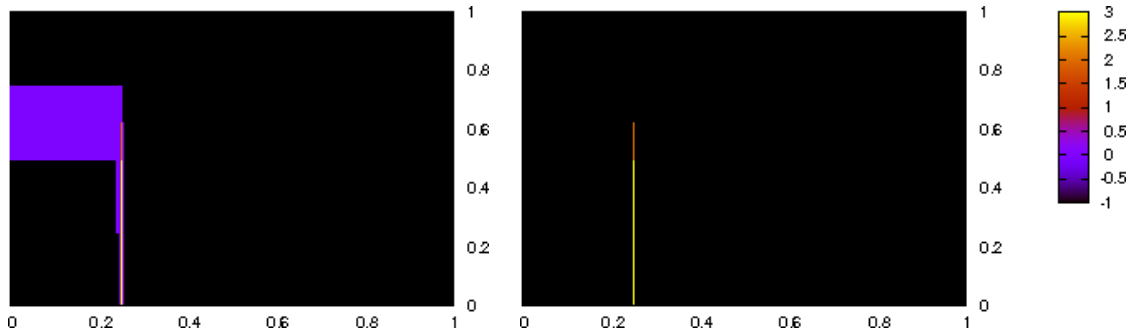


Figure 11: Portion of the domain  $(x, \xi)$  selected for the entropy correction and card  $S'$ . Value=-1 if no correction. Value=0 if test. Card  $S'$  else. Computations with  $ctol = 1e^{-2}$  (left) and  $ctol = 1e^{-5}$  (right).

$\dim \mathcal{S}^{Nr, No}$	No = 1, Nr = 3 16		No = 2, Nr = 3 24		No = 3, Nr = 3 32	
	$T_{CPU}$	$factor$	$T_{CPU}$	$factor$	$T_{CPU}$	$factor$
$ctol = +\infty$	11.7	1.0e-0	16.1	1.0e-0	21.6	1.0e-0
$ctol = 1e^{-1}$	8.2	3.7e-1	11.8	3.7e-1	16.4	3.7e-1
$ctol = 1e^{-2}$	6.5	7.1e-2	9.8	7.1e-2	13.9	7.1e-2
$ctol = 1e^{-3}$	6.1	2.8e-3	9.3	2.8e-3	13.5	2.8e-3
$ctol = 0$	6	2.5e-3	9.2	2.5e-3	13.4	2.5e-3
$\epsilon_h$	1.32e-3		7.17e-4		2.88e-4	

Table 1: Normalized computational times  $T_{CPU}$ , fraction  $factor$  of stochastic cells actually tested, and error measure  $\epsilon_h$  for different stochastic discretization parameters Nr and No. computations with  $N_c = 250$ .

applied to the case of non-strictly hyperbolic systems since it provides an unambiguous relation between the eigenvalues, eigenvectors and reconstructed intermediate states. Numerical tests for the Burgers and Euler equations have demonstrated the effectiveness of the proposed corrector in providing entropy solutions. A robust criterion has also been presented for the *a priori* localization of the physical cell interfaces requiring a correction, yielding a significant reduction of the corrector computational overhead.

The present corrector still involves some limitations that will have to be overcome to handle general situations. Firstly, the corrector, as well as the Roe Galerkin solver, assumes the availability of the explicit expression of the stochastic eigenvalues and eigenvectors of the stochastic hyperbolic system. This is not a limitation for the Burgers and Euler equations. However, for general systems of conservation laws, *a priori* knowledge of the solution characteristics may not be possible. Alternative strategies to deal with such general problems are still an open question. Secondly, the approximate eigenvalues and eigenvectors of the Roe Galerkin Jacobian matrix assume a fully tensorized polynomial basis on each stochastic element to construct the corresponding set of Gauss points and weights. For problems involving a large number of random variables, sparse polynomial bases are mandatory to temper the curse of dimensionality. Such bases prevent the immediate construction of approximate eigenvalues and eigenvectors as above. In [14], we have shown that for partially tensorized stochastic bases, the computation of the upwind matrix can still be performed using the set of fully tensorized Gauss points. A similar approach could be followed for the correction: whenever a sonic point is detected on a physical interface, the left and right Galerkin states could be locally lifted to the (higher dimensional) fully tensorized stochastic approximation space in which the flux correction can be determined by means of the present method, before being projected back to the original sparse stochastic space. Owing to the pertinent criteria proposed for the detection of sonic points, these lifting/projection procedures would not be used frequently, so that this approach would not incur significant computational overheads. Such a procedure however remains to be implemented and tested.

To close the discussion, we observe that the proposed Roe solver and entropy corrector require sufficient stochastic resolution to properly capture solution discontinuities in the stochastic space. Results presented here and in [14] use a uniform stochastic refinement, while in fact the solutions are essentially smooth everywhere except in the vicinity of localized shocks and discontinuities. This observation calls for the use of adaptive methods where the stochastic resolution (and possibly also the spatial one) is selected with regard to the local smoothness of the solution in order to optimize the computational effort. In doing so, the tensorized structure of the deterministic (spatial) and stochastic approximation spaces will be lost. Current research focuses on the implementation of such adaptive strategies, with particular emphasis on the derivation of robust error estimators to drive the refinement (and coarsening) of the discretization spaces.

## References

- [1] R. Abgrall. A simple, flexible and generic deterministic approach to uncertainty quantifications in non linear problems: application to fluid flow problems. *Rapport de recherche INRIA*, 00325315, 2008.
- [2] Manas K. Deb, Ivo M. Babuška, and J. Tinsley Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Comput. Methods Appl. Mech. Engrg.*, 190(48):6359–6372, 2001.

- [3] Bert J. Debuschere, Habib N. Najm, Philippe P. Pébay, Omar M. Knio, Roger G. Ghanem, and Olivier P. Le Maître. Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM J. Sci. Comput.*, 26(2):698–719, 2004.
- [4] François Dubois and Guillaume Mehlman. A non-parameterized entropy correction for Roe’s approximate Riemann solver. *Numer. Math.*, 73(2):169–208, 1996.
- [5] L. Ge, K.F. Cheung, and M.H. Kobayashi. Stochastic Solution for Uncertainty Propagation in Nonlinear Shallow-Water Equations. *Journal of Hydraulic Engineering*, Decembre 2008:1732–1743, 2008.
- [6] Roger G. Ghanem and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Dover, 2003.
- [7] David Gottlieb and Dongbin Xiu. Galerkin method for wave equations with uncertain coefficients. *Commun. Comput. Phys.*, 3(2):505–518, 2008.
- [8] Ami Harten. High resolution schemes for hyperbolic conservation laws. *J. Comput. Phys.*, 49(3):357–393, 1983.
- [9] O. P. Le Maître, H. N. Najm, R. G. Ghanem, and O. M. Knio. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *J. Comput. Phys.*, 197(2):502–531, 2004.
- [10] G. Lin, C.-H. Su, and G. E. Karniadakis. Predicting shock dynamics in the presence of uncertainties. *J. Comput. Phys.*, 217(1):260–276, 2006.
- [11] G. Lin, C.-H. Su, and G. E. Karniadakis. Stochastic modeling of random roughness in shock scattering problems: theory and simulations. *Comput. Methods Appl. Mech. Engrg.*, 197(43-44):3420–3434, 2008.
- [12] G. Poette, B. Després, and D. Lucor. Uncertainty quantification for systems of conservation laws. *J. Comput. Phys.*, 228(7):2443–2467, 2009.
- [13] Philip L. Roe. Some contributions to the modelling of discontinuous flows. In *Large-scale computations in fluid mechanics, Part 2 (La Jolla, Calif., 1983)*, volume 22 of *Lectures in Appl. Math.*, pages 163–193. Amer. Math. Soc., Providence, RI, 1985.
- [14] J. Tryoen, O. Le Maître, M. Ndjinga, and A. Ern. Intrusive Projection Methods with Upwinding for Uncertain Nonlinear Hyperbolic Systems. *J. Comput. Phys.*, 2009, Under Review.
- [15] Xiaoliang Wan and George Em Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.*, 28(3):901–928 (electronic), 2006.