A stochastic Galerkin method for the Euler equations with Roe variable transformation

Per Pettersson^{*,†}, Gianluca Iaccarino[†] and Jan Nordström[‡]

* Department of Information Technology, Uppsala University, P.O. Box 337, SE-75105 Uppsala, Sweden.

[†] Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, USA.

[‡] Department of Mathematics, Linköping University, SE-58183 Linköping, Sweden.

Abstract

The Euler equations subject to uncertainty in the initial and boundary conditions are investigated via the stochastic Galerkin approach. We present a new fully intrusive method based on a variable transformation of the continuous equations. Roe variables are employed to get quadratic dependence in the flux function and a well-defined Roe average matrix that can be determined without matrix inversion.

In previous formulations based on generalized polynomial chaos expansion of the physical variables, the need to introduce stochastic expansions of inverse quantities, or square-roots of stochastic quantities of interest, adds to the number of possible different ways to approximate the original stochastic problem. We present a method where the square roots occur in the choice of variables and no auxiliary quantities are needed, resulting in an unambiguous problem formulation.

The Roe formulation saves computational cost compared to the formulation based on expansion of conservative variables. Moreover, the Roe formulation is more robust and can handle cases of supersonic flow, for which the conservative variable formulation fails to produce a bounded solution. We use a multi-wavelet basis that can be chosen to include a large number of resolution levels to handle more extreme cases (e.g. strong discontinuities) in a robust way. For smooth cases, the order of the polynomial representation can be increased for increased accuracy.

1 Introduction

Generalized polynomial chaos (gPC) expansions are frequently used to represent uncertain quantities in numerical solutions of partial differential equations (PDEs) with uncertainty in e.g. initial data, boundary data, unknown material parameters and forcing functions due to unknown geometry. These quantities of interest are expressed as generalized Fourier series, where orthogonal polynomials (e.g. Legendre or Hermite polynomials) are commonly used as basis functions. Building on results due to Wiener [26], this is the polynomial chaos method introduced by Ghanem and Spanos [7] and generalized by [28].

Spectral convergence of the gPC expansion is observed when the solutions are sufficiently regular and continuous [28], but for general non-linear conservation laws - such as in fluid dynamics problems - the convergence is usually less favorable. Spectral expansion representations are still of interest for these problems because of the potential efficiency with respect to brute force sampling methods, but special attention must be devoted to the numerical methodology used. For some problems with steep gradients in the stochastic dimensions, gPC expansions fail entirely to capture the solution [11]. Global methods can still give superior overall performance, for instance Padé approximation methods based on rational function approximation [4], and hierarchical wavelet methods that are global methods with localized support of each resolution level [9, 10].

Intrusive gPC methods for nonlinear conservation laws have been investigated in e.g. [23], where a reduced-cost Roe solver with entropy corrector was presented, and in [22] with different localized representations of uncertainty in initial functions and problem coefficients. In previous works we investigated well-posedness and stability for an intrusive formulation of Burgers' equation with uncertainty in [14], and imposition of uncertain boundary conditions in [15].

In many nonlinear applications of the stochastic Galerkin method, truncation of the gPC expansion leads to non-unique formulations of the systems of equations. For instance, cubic products between stochastic quantities a, b and c, are represented as products of truncated approximations \tilde{a}, \tilde{b} and \tilde{c} , but the pseudo-spectral multiplication operator *, to be explicitly defined in a later section, is not associative, i.e. $(\tilde{a} * \tilde{b}) *$ $\tilde{c} \neq \tilde{a} * (\tilde{b} * \tilde{c})$. Similar problems are investigated in more detail in [8]. The need to introduce gPC expansions of inverse quantities, or square-roots of stochastic quantities of interest, adds to the number of possible different ways to approximate the original stochastic problem. This leads to ambiguity of the problem formulation. We present a method where this ambiguity is avoided since no auxiliary quantities are needed. Our formulation relies on a variable transformation where the square-root of the density is computed, but this operation can be done in a robust way in a small number of operations.

Alternative approaches to gPC methods have also been presented in the literature. Abgrall et. al. [1, 2] developed a semi-intrusive method based on a finite-volume like reconstruction technique of the discretized stochastic space. A deterministic problem is obtained by taking conditional expectations given a stochastic subcell, over which ENO constructions are used to reconstruct the fluxes in the stochastic dimensions. This makes it particularly suitable for non-smooth probability distributions. Poëtte et. al. [17] used a nonlinear projection method to bound the oscillations close to stochastic discontinuities by gPC expansion of the entropy variables obtained from a transformation of the conservative variables. Each time step is complemented by a functional minimization to obtain the entropy variables needed to update the solution vector. The method we will present here may appear similar at first sight, but it relies on a different kind of variable transformation and not on kinetic theory considerations. We do not suggest a variable transformation for general conservation laws, but a formulation that specifically targets the solution of the Euler equations with uncertainty in the variables. It is less complicated than a direct gPC expansion of the conservative variables.

In our method, the system of equations is reformulated using Roe variables so that only quadratic terms occur. No fourth-order tensors need to be approximated or calculated, resulting in increased accuracy and reduced computational cost. Moreover, there is no need to compute additional chaos expansion for inverse quantities. The Roe variable expansion provides a simple and unambiguous formulation of the Euler equations. For brevity of notation, we will refer to our expansion method as the Roe expansion, and the method based on expansion of the conservative variables as the conservative expansion.

We consider the Sod test case subject to uncertainty in the density, and uncertain diaphragm location, respectively. The uncertainty is represented with a multi-wavelet (MW) expansion in the stochastic dimension, following the framework outlined in [10]. Special cases of the MW basis include the Legendre polynomials and the piecewise constant Haar wavelets. The stochastic Galerkin system is obtained by projection of the stochastic Euler equations onto the MW basis functions.

We employ a MUSCL reconstruction in space [24] and a fourth-order Runge-Kutta method for the time integration. A Roe flux based on a stochastic Galerkin Roe average matrix will be employed. A Roe average matrix for the numerical Roe flux is derived and proven to fulfill the conditions of Roe [20] under certain circumstances.

In section 2, we present the framework for the stochastic Galerkin formulations of the Euler equations in conservative and Roe variables, which are introduced in section 3. In section 4, the numerical flux functions are introduced. We derive and prove properties of the solvers that are necessary to capture essential dynamics of hyperbolic problems. We use a Roe flux suitable for situations where the system eigenvalues can be accurately estimated. Numerical results are presented in section 5, where the previously developed techniques are evaluated. Conclusions are drawn in section 6.

2 Representation of uncertainty

Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space with event space Ω , and probability measure \mathcal{P} defined on the σ -field \mathcal{A} of subsets of Ω . Let $\xi(\omega)$ be a random variable for $\omega \in \Omega$. Consider a generalized chaos basis $\{\psi_i(\xi)\}_{i=0}^{\infty}$ spanning the space of second order (i.e. finite variance) random processes on this probability space. The basis functions are assumed to be orthonormal, i.e. they satisfy

$$\langle \psi_i \psi_j \rangle = \delta_{ij},\tag{1}$$

where the inner product between two functions $a(\xi)$ and $b(\xi)$ is defined by

$$\langle a(\xi)b(\xi)\rangle = \int_{\Omega} a(\xi)b(\xi)d\mathcal{P}(\xi).$$

Any second order random field $u(x,t,\xi)$ (i.e. $\langle u^2 \rangle < \infty$) can be expressed as

$$u(x,t,\xi) = \sum_{i=0}^{\infty} u_i(x,t)\psi_i(\xi).$$
 (2)

Independent of the choice of basis $\{\psi_i(\xi)\}_{i=0}^{\infty}$, we can express the mean and variance of $u(x, t, \xi)$ as

$$E(u(x,t,\xi)) = u_0(x,t), \quad Var(u(x,t,\xi)) = \sum_{i=1}^{\infty} u_i(x,t)^2,$$

respectively. For practical purposes, (2) is truncated to a finite number P + 1 terms, and we set

$$u(x,t,\xi) \approx \sum_{i=0}^{P} u_i(x,t)\psi_i(\xi).$$
(3)

The truncation of the gPC series leading to (3) may result in solutions that are unphysical. An extreme example is when a strictly positive quantity, say density, with uncertainty within a bounded range is represented by a polynomial expansion with infinite range, for instance Hermite polynomials of standard Gaussian variables. The Hermite series expansion converges to the true density with bounded range in the limit $P \to \infty$, but for a given order of expansion, say P = 2, the representation $\rho = \rho_0 + \rho_1 H_1(\xi)$ may result in negative density with non-zero probability since the Hermite polynomial H_1 takes arbitrarily large negative values. Similar problems may be encountered also for polynomial representations with bounded support. Furthermore, polynomial reconstruction of a discontinuity in stochastic space leads to Gibbs oscillations that may yield negative values of an approximation of a solution that is close to zero but strictly positive by definition.

For smooth problems, the gPC expansion is attractive due to its spectral convergence. For non-smooth problems such as nonlinear hyperbolic problems, discontinuities will emerge in finite time in stochastic space and gPC expansions tend to result in Gibbs oscillations. In the worst case, polynomial chaos may fail entirely to capture essential features of the solution [9].

2.1 Multi-wavelet expansion

An alternative to gPC expansions for non-smooth and oscillatory problems is generalized chaos based on a localization or discretization of the stochastic space [5, 16]. Methods based on stochastic discretization include adaptive stochastic multi-elements [25] and stochastic simplex collocation [27]. The robust properties of discretized stochastic space can also be obtained by globally defined wavelets, see [9]. In this paper, we follow the approach of [10] and use piecewise polynomial multi-wavelets (MW), defined on sub-intervals of [-1, 1]. The construction of a truncated MW basis follows the algorithm in [3].

Wavelets are defined hierarchically on different resolution levels, representing successively finer features of the solution with increasing resolution. They have nonoverlapping support within each resolution level, and in this sense they are localized. Still, the basis is global due to the overlapping support of wavelets belonging to different resolution levels. Piecewise constant wavelets, denoted Haar wavelets, do not exhibit spectral convergence, but avoid the Gibbs phenomenon in the proximity of discontinuities in the stochastic dimension.

Starting with the space \mathbf{V}_{N_p} of polynomials of degree at most N_p defined on the interval [-1, 1], the construction of multi-wavelets aims at finding a basis of piecewise polynomials for the orthogonal complement of \mathbf{V}_{N_p} in the space \mathbf{V}_{N_p+1} of polynomials of degree at most $N_p + 1$. Merging the bases of \mathbf{V}_{N_p} and that of the orthogonal complement of \mathbf{V}_{N_p} in \mathbf{V}_{N_p+1} , we obtain a piecewise polynomial basis for \mathbf{V}_{N_p+1} . Continuing the process of finding orthogonal complements in spaces of increasing degree of piecewise polynomials, leads to a basis for $L_2([-1, 1])$.

We first introduce a smooth polynomial basis on [-1, 1]. Let $\{Le_i(\xi)\}_{i=0}^{\infty}$ be the set of Legendre polynomials that are defined on [-1, 1] and orthogonal with respect to the uniform measure. The normalized Legendre polynomials are defined recursively by

$$Le_{i+1}(\xi) = \sqrt{2i+3} \left(\frac{\sqrt{2i+1}}{i+1} \xi Le_i(\xi) - \frac{i}{(i+1)\sqrt{2i-1}} Le_{i-1}(\xi) \right), i \ge 1,$$

$$Le_0(\xi) = 1, \quad Le_1(\xi) = \sqrt{3}\xi.$$

The set $\{Le_i(\xi)\}_{i=0}^{N_p}$ is an orthonormal basis for \mathbf{V}_{N_p} . Double products are readily computed from (1), and higher-order products are precomputed using numerical integration.

Following the algorithm by Alpert [3] (see Appendix A), we construct a set of mother wavelets $\{\psi_i^W(\xi)\}_{i=0}^{N_p}$ defined on the domain $\xi \in [-1, 1]$, where

$$\psi_i^W(\xi) = \begin{cases} p_i(\xi) & -1 \le \xi < 0\\ (-1)^{N_p + i + 1} p_i(\xi) & 0 \le \xi < 1\\ 0 & \text{otherwise,} \end{cases}$$
(4)

where $p_i(\xi)$ is an i^{th} order polynomial. By construction, the set of wavelets $\{\psi_i^W(\xi)\}_{i=0}^{N_p}$ are orthogonal to all polynomials of order at most N_p , hence the wavelets are orthogonal to the set $\{Le_i(\xi)\}_{i=0}^{N_p}$ of Legendre polynomials of order at most N_p . Based on translations and dilations of (4), we get the wavelet family

$$\psi_{i,j,k}^W(\xi) = 2^{j/2} \psi_i^W(2^j \xi - k), \qquad i = 0, ..., N_p, \quad j = 0, 1, ..., \quad k = 0, ..., 2^{j-1}.$$

Let $\psi_m(\xi)$ for $m = 0, ..., N_p$ be the set of Legendre polynomials up to order N_p , and concatenate the indices i, j, k into $m = (N_p + 1)(2^j + k - 1) + i$ so that $\psi_m(\xi) \equiv \psi_{i,j,k}^W(\xi)$ for $m > N_p$. With the MW basis $\{\psi_m(\xi)\}_{m=0}^{\infty}$ we can represent any random variable $u(x, t, \xi)$ with finite variance as

$$u(x,t,\xi) = \sum_{m=0}^{\infty} u_m(x,t)\psi_m(\xi),$$

which is of the form (2). In the computations, we truncate the MW series both in terms of the piecewise polynomial order N_p and the resolution level N_r . With the index $j = 0, ..., N_r$, we retain $P + 1 = (N_p + 1)2^{N_r}$ terms of the MW expansion.

The truncated MW basis is characterized by the piecewise polynomial order N_p and the number of resolution levels N_r , illustrated in Figure 1 for $N_p = 2$ and $N_r = 3$. As special cases of the MW basis, we obtain the Legendre polynomial basis for $N_r = 0$ (i = j = 0), and the Haar wavelet basis of piecewise constant functions for $N_p = 0$.

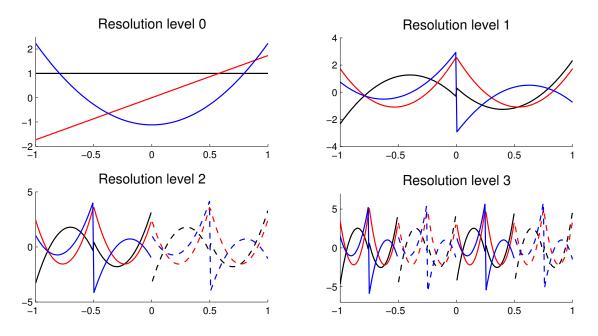


Figure 1: Multi-wavelets for $N_p = 2$, $N_r = 3$. Resolution level 0 consists of the first $N_p + 1$ Legendre polynomials. Resolution level j > 0 contains 2^{j-1} wavelets each. Each basis function is a piecewise polynomial of order N_p .

3 Euler equations with input uncertainty

Consider the 1D Euler equations, in non-dimensional form given by

$$u_t + f(u)_x = 0, \quad 0 \le x \le 1, t > 0,$$
(5)

where the solution and flux vector are given by

$$u = \begin{bmatrix} \rho \\ \rho v \\ E \end{bmatrix}, \quad f = \begin{bmatrix} \rho v \\ \rho v^2 + p \\ (E+p)v \end{bmatrix},$$

where ρ is density, v velocity, E total energy and pressure p. A perfect gas equation of state is assumed, and energy and pressure are related by

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho v^2,$$

where γ is the ratio of the specific heats. For the numerical method, we need the flux Jacobian, given by

$$\frac{\partial f}{\partial u} = \begin{bmatrix} 0 & 1 & 0\\ \frac{1}{2}(\gamma - 3)v^2 & (3 - \gamma)v & \gamma - 1\\ \frac{1}{2}(\gamma - 1)v^3 - vH & H - (\gamma - 1)v^2 & \gamma v \end{bmatrix},$$

with the total enthalpy $H = (E + p)/\rho$.

We scale the physical variables to get the dimensionless variables $\rho = \rho'/\rho'_{ref}$, $E = E'/(\gamma p'_{ref})$, $p = p'/(\gamma p'_{ref})$ and $v = v'/a'_{ref}$ where $a' = (\gamma p'/\rho')^{1/2}$ and the subscript ref denotes a reference state.

3.1 Formulation in Roe variables

Roe [20] introduced the variables

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} \rho^{1/2} \\ \rho^{1/2} v \\ \rho^{1/2} H \end{bmatrix}.$$

The flux and the conservative variables are given by

$$\hat{f}(w) = \begin{bmatrix} w_1 w_2 \\ \frac{\gamma - 1}{\gamma} w_1 w_3 + \frac{\gamma + 1}{2\gamma} w_2^2 \\ w_2 w_3 \end{bmatrix}, \quad u = \hat{g}(w) = \begin{bmatrix} w_1^2 \\ w_1 w_2 \\ \frac{w_1 w_3}{\gamma} + \frac{\gamma - 1}{2\gamma} w_2^2 \end{bmatrix}.$$

Then

$$\hat{g}(w)_t + \hat{f}_x(w) = 0$$
 (6)

is equivalent to (5). The flux Jacobian in the Roe variables is given by

$$rac{\partial \hat{f}}{\partial w} = \left[egin{array}{ccc} w_2 & w_1 & 0 \ rac{\gamma-1}{\gamma}w_3 & rac{\gamma+1}{\gamma}w_2 & rac{\gamma-1}{\gamma}w_1 \ 0 & w_3 & w_2 \end{array}
ight].$$

3.2 Stochastic Galerkin formulation of the Euler equations

Define the pseudo-spectral product u * v of order $P = P(N_p, N_r)$ by

$$(u*v)_k = \sum_{i=0}^P \sum_{j=0}^P u_i v_j \langle \psi_i \psi_j \psi_k \rangle, \qquad k = 0, ..., P,$$

where

$$\langle \psi_i \psi_j \psi_k \rangle = \int_{\Omega} \psi_i(\xi) \psi_j(\xi) \psi_k(\xi) d\mathcal{P}.$$

Alternatively, using matrix notation, we can write the spectral product as u * v = A(u)v, where

$$[A(u)]_{jk} = \sum_{i=0}^{P} u_i \langle \psi_i \psi_j \psi_k \rangle.$$
(7)

We will need the pseudo-spectral inverse q^{-*} , defined as the solution of $q * q^{-*} = 1$, and the pseudo-spectral square root, defined as the solution $q^{*/2}$ of $q^{*/2} * q^{*/2} = q$, where the spectral expansion of the quantity of interest q is assumed to be known. For more details, see [6].

Let u^P denote the vector of coefficients of the MW expansion of u of order $P = P(N_p, N_r)$. P may take the same value for two distinct pairs of (N_p, N_r) but this ambiguity in notation will not matter in the derivation of the numerical method so for brevity we use only P in the superscripts. The Euler equations represented by the conservative formulation (5) can be written as an augmented system, after stochastic Galerkin projection,

$$u_t^P + f^P (u^P)_x = 0, (8)$$

where

$$u^{P} = \begin{bmatrix} u_{1}^{P} \\ u_{2}^{P} \\ u_{3}^{P} \end{bmatrix} = \begin{bmatrix} [(u_{1})_{0}, \dots, (u_{1})_{P}]^{T} \\ [(u_{2})_{0}, \dots, (u_{2})_{P}]^{T} \\ [(u_{3})_{0}, \dots, (u_{3})_{P}]^{T} \end{bmatrix}, f^{P}(u^{P}) = \begin{bmatrix} u_{2}^{P} \\ (u_{1}^{P})^{-*} * u_{2}^{P} * u_{2}^{P} + p^{P} \\ (u_{3}^{P} + p^{P}) * u_{2}^{P} * (u_{1}^{P})^{-*} \end{bmatrix}.$$

with $p^P = (\gamma - 1)(u_3^P - (u_1^P)^{-*} * u_2^P * u_2^P/2)$. The cubic products of (8) are approximated by the application of two third-order tensors, instead of one fourth-order tensor. That is, we replace $(a * b * c)_l = \sum_{ijk} \langle \psi_i \psi_j \psi_k \psi_l a_i b_j c_k \rangle$ by the approximation $(a * b * c)_l \approx$ $((a * b) * c)_l$.

For the Roe variable formulation, the stochastic Galerkin projection of (6) gives the system

$$\hat{g}^{P}(w^{P})_{t} + \hat{f}^{P}(w^{P})_{x} = 0, \qquad (9)$$

where

$$\hat{g}^{P}(w^{P}) = \begin{bmatrix} w_{1}^{P} * w_{1}^{P} \\ w_{1}^{P} * w_{2}^{P} \\ \frac{w_{1}^{P} * w_{3}^{P}}{\gamma} + \frac{\gamma - 1}{2\gamma} w_{2}^{P} * w_{2}^{P} \end{bmatrix}, \hat{f}^{P}(w^{P}) = \begin{bmatrix} w_{1}^{P} * w_{2}^{P} \\ \frac{\gamma - 1}{\gamma} w_{1}^{P} * w_{3}^{P} + \frac{\gamma + 1}{2\gamma} w_{2}^{P} * w_{2}^{P} \\ w_{2}^{P} * w_{3}^{P} \end{bmatrix}.$$

The flux Jacobian for the stochastic Galerkin system in the Roe variables is given by

$$\frac{\partial \hat{f}^P}{\partial w^P} = \begin{bmatrix} A(w_2^P) & A(w_1^P) & 0_{(P+1)\times(P+1)} \\ \frac{\gamma-1}{\gamma}A(w_3^P) & \frac{\gamma+1}{\gamma}A(w_2^P) & \frac{\gamma-1}{\gamma}A(w_1^P) \\ 0_{(P+1)\times(P+1)} & A(w_3^P) & A(w_2^P) \end{bmatrix}.$$
 (10)

As $P \to \infty$, the formulations (8) and (9), as well as any other consistent formulation, are equivalent. However, P is assumed to be small (< 20), and truncation and conditioning of the system matrices will play an important role for the accuracy of the solution.

4 Numerical method

As our main numerical method we will use the MUSCL (Monotone Upstream-centered Schemes for Conservation Laws) scheme introduced in [24].

4.1 Expansion of conservative variables

Let *m* be the number of spatial points and $\Delta x = 1/(m-1)$ and let U^P be the spatial discretization of u^P . The semi-discretized form of (8) is given by

$$\frac{dU_j^P}{dt} + \frac{F_{j+1/2}^P - F_{j-1/2}^P}{\Delta x} = 0, \quad j = 1, ..., m,$$
(11)

where $F_{j+1/2}^P$ denotes the numerical flux function evaluated at the interface between cells j and j + 1.

For the MUSCL scheme with slope limited states U^L and U^R , we take the numerical flux

$$F_{j+\frac{1}{2}}^{P} = \frac{1}{2} \left(f^{P}(U_{j+\frac{1}{2}}^{L}) + f^{P}(U_{j+\frac{1}{2}}^{R}) \right) + \frac{1}{2} |(\tilde{J}_{c}^{P})_{j+\frac{1}{2}}| \left(U_{j+\frac{1}{2}}^{L} - U_{j+\frac{1}{2}}^{R} \right),$$
(12)

where the Roe average \tilde{J}_c^P is the pseudo-spectral generalization of the standard Roe average of the deterministic Euler equations, i.e.

$$\tilde{J}_{c}^{P}(\overline{v},\overline{H}) = \begin{bmatrix} 0_{P\times P} & I_{P\times P} & 0_{P\times P} \\ \frac{1}{2}(\gamma-3)A(\overline{v})^{2} & (3-\gamma)A(\overline{v}) & (\gamma-1)I_{P\times P} \\ \frac{1}{2}(\gamma-1)A(\overline{v})^{3} - A(\overline{v})A(\overline{H}) & A(\overline{H}) - (\gamma-1)A(\overline{v})^{2} & \gamma A(\overline{v}) \end{bmatrix}$$

where

$$\overline{v} = (\rho_L^{-*/2} + \rho_R^{-*/2}) * (\rho_L^{*/2} * v_L + \rho_R^{*/2} * v_R),$$

and

$$\overline{H} = (\rho_L^{*/2} * H_L + \rho_R^{*/2} * H_R) * (\rho_L^{-*/2} + \rho_R^{-*/2}).$$

The computation of \overline{v} and \overline{H} require the spectral square root $\rho^{*/2}$ and its inverse, that are computed solving a nonlinear and a linear system, respectively.

Further details about the formulation of the Roe average matrix are given in [22]. The scheme is a direct generalization of the deterministic MUSCL scheme. Flux limiters are applied componentwise to all MW coefficients in sharp regions. For a more detailed description of the MUSCL scheme and flux limiters, see e.g. [12], and for application to the stochastic Burgers' equation [13].

4.2 Expansion of Roe's variables

Let W^P denote the spatial discretization of w^P . The semi-discretized form of (9) is given by

$$\frac{\partial \hat{g}_{j}^{P}}{\partial t} + \frac{\hat{F}_{j+1/2}^{P} - \hat{F}_{j-1/2}^{P}}{\Delta x} = 0, \quad j = 1, ..., m,$$

with the numerical flux function

$$\hat{F}_{j+\frac{1}{2}} = \frac{1}{2} \left(\hat{f}^P(W_{j+\frac{1}{2}}^L) + \hat{f}^P(W_{j+\frac{1}{2}}^R) \right) + \frac{1}{2} |\tilde{J}_{j+\frac{1}{2}}^P| \left(W_{j+\frac{1}{2}}^L - W_{j+\frac{1}{2}}^R \right), \tag{13}$$

where $\tilde{J}^P = \tilde{J}^P(W^P)$ is the Roe matrix for the stochastic Galerkin formulation of the Euler equations in Roe's variables, to be derived below.

Each time step provides the update of the solution vector \hat{g}_j^P , j = 1, ..., m, from which we can solve for W^P to be used in the update of the numerical flux. This involves solving the nonlinear systems

$$A(W_{1,j}^P)W_{1,j}^P = \hat{g}_{1,j}^P, \quad j = 1, ..., m,$$
(14)

for $W_{1,j}^P$, and then using $W_{1,j}^P$ to solve the linear $(P+1) \times (P+1)$ -systems

$$A(W_{1,j}^P)W_{2,j}^P = \hat{g}_{2,j}^P, \quad j = 1, ..., m,$$

for $W_{2,j}^P$, and

$$A(W^P_{1,j})W^P_{3,j} = \gamma \hat{g}^P_{3,j} - \frac{\gamma - 1}{2}A(W^P_{2,j})W^P_{2,j}, \quad j = 1, ..., m,$$

for $W_{3,j}^P$.

The system (14) is solved iteratively with a trust-region-dogleg algorithm¹. Starting with the value of the previous time-step as initial guess, few iterations are required (typically 2-3). (The same method is used to solve for spectral square roots in the conservative variable formulation.)

4.3 Stochastic Galerkin Roe average matrix for Roe variables

The Roe average matrix \tilde{J}^P is given as a function of the Roe variables $w = (w_1, w_2, w_3)^T$, where each w_i is a vector of generalized chaos coefficients. It is designed to satisfy the following properties:

(i)
$$\tilde{J}^P(w^L, w^R) \to \frac{\partial \hat{f}^P}{\partial w}\Big|_{w=w'}$$
 as $w^L, w^R \to w'$.
(ii) $\tilde{J}^P(w^L, w^R) \times (w^L - w^R) = \hat{f}^P(w^L) - \hat{f}^P(w^R), \forall w^L, w^R$

(iii) \tilde{J}^P is diagonalizable with real eigenvalues and linearly independent eigenvectors.

In the standard approach introduced by Roe and commonly used for deterministic calculations, the conservative variables are mapped to the w variables, which are then averaged.

In the deterministic case, we have

$$\hat{f}^{L} - \hat{f}^{R} = \tilde{J}(w^{L}, w^{R}) \times (w^{L} - w^{R}),$$
(15)

¹This is the default algorithm for **fsolve** in Matlab. For more details, see [18].

where

$$\tilde{J}(w^L, w^R) = \begin{bmatrix} \overline{w}_2 & \overline{w}_1 & 0\\ \frac{\gamma - 1}{\gamma} \overline{w}_3 & \frac{\gamma + 1}{\gamma} \overline{w}_2 & \frac{\gamma - 1}{\gamma} \overline{w}_1\\ 0 & \overline{w}_3 & \overline{w}_2 \end{bmatrix}.$$

Overbars denote arithmetic averages of assumed left and right values of a variable, i.e.

$$\overline{w}_j = \frac{w_j^L + w_j^R}{2}, \quad j = 1, 2, 3$$

It is a straightforward extension of the analysis by Roe in [20] to show properties (i) and (ii) for the Roe variables, without mapping to the conservative variables. To prove (iii) we note that there exists an eigenvalue decomposition

$$\tilde{J} = VDV^{-1},\tag{16}$$

where

$$V = \begin{bmatrix} \frac{w_1}{w_3} & -\frac{w_1}{w_3} & -\frac{w_1}{w_3} \\ \frac{w_2 - \sqrt{w_2^2 + 8w_1 w_3 \gamma(\gamma - 1)}}{2\gamma w_3} & \frac{w_2 + \sqrt{w_2^2 + 8w_1 w_3 \gamma(\gamma - 1)}}{2\gamma w_3} & 0 \\ 1 & 1 & 1 \end{bmatrix},$$
(17)

$$D = \begin{bmatrix} \frac{w_2(1+2\gamma) - \sqrt{8w_1w_3\gamma(\gamma-1) + w_2^2}}{2\gamma} & 0 & 0\\ 0 & \frac{w_2(1+2\gamma) + \sqrt{8w_1w_3\gamma(\gamma-1) + w_2^2}}{2\gamma} & 0\\ 0 & 0 & w_2 \end{bmatrix}.$$
 (18)

The eigenvalues of \tilde{J} are real and distinct, so property (iii) is also satisfied.

Now consider the stochastic Galerkin formulation, i.e. assume that the w_i 's are vectors of generalized chaos coefficients. The stochastic Galerkin Roe average matrix \tilde{J}^P for the Roe variables formulation is a generalization of the mapping (15), i.e. of the matrix \tilde{J} . We define

$$\tilde{J}^{P}(w^{L}, w^{R}) = \tilde{J}^{P}(\overline{w}) = \begin{bmatrix} A(\overline{w}_{2}) & A(\overline{w}_{1}) & 0_{P \times P} \\ \frac{\gamma - 1}{\gamma} A(\overline{w}_{3}) & \frac{\gamma + 1}{\gamma} A(\overline{w}_{2}) & \frac{\gamma - 1}{\gamma} A(\overline{w}_{1}) \\ 0_{P \times P} & A(\overline{w}_{3}) & A(\overline{w}_{2}) \end{bmatrix},$$
(19)

where the submatrix $A(w_j)$ is given by (7) and $\overline{w} = (w^L + w^R)/2$.

Proposition 1. Property (i) is satisfied by (19).

Proof. With
$$w^L = w^R = w'$$
, $\tilde{J}^P(w^L, w^R) = \tilde{J}^P(w', w') = \left. \frac{\partial \hat{f}^P}{\partial w^P} \right|_{w=w'}$ by (10).

Proposition 2. Property (ii) is satisfied by (19).

Proof.

$$\tilde{J}^{P}(w^{L}, w^{R}) \times (w^{L} - w^{R}) = \frac{1}{2} \left(\tilde{J}^{P}(w^{L}) + \tilde{J}^{P}(w^{R}) \right) (w^{L} - w^{R}) = \frac{1}{2} \tilde{J}^{P}(w^{L}) w^{L} - \frac{1}{2} \tilde{J}^{P}(w^{R}) w^{R} = \hat{f}^{P}(w^{L}) - \hat{f}^{P}(w^{R}), \quad (20)$$

where the last equality follows from the fact that the stochastic Galerkin generalizations of the Euler equations are homogeneous of degree 1. \Box

To prove (iii), we will need the following proposition.

Lemma 1. Let $A(w_j)$ (j = 1, 2, 3) be defined by (7) and $A(w_j) = Q\Lambda_j Q^T$ be an eigenvalue decomposition with constant eigenvector matrix Q and assume that Λ_1 and Λ_3 are non-singular. Then the stochastic Galerkin Roe average matrix \tilde{J}^P has an eigenvalue decomposition $\tilde{J}^P = X \tilde{\Lambda}^P X^{-1}$ with a complete set of eigenvectors.

Proof. We will use the Kronecker product \otimes , defined for two matrices B (of size $m \times n$) and C by

$$B \otimes C = \begin{bmatrix} b_{11}C & \dots & b_{1n}C \\ \vdots & \ddots & \vdots \\ b_{m1}C & \dots & b_{mn}C \end{bmatrix}.$$

The eigenvalue decompositions of each $(P+1) \times (P+1)$ matrix block of (19) have the same eigenvector matrix Q, hence we can write

$$\tilde{J}^P = (I_3 \otimes Q) \hat{J} (I_3 \otimes Q^T) \tag{21}$$

where

$$\hat{J} = \begin{bmatrix} \Lambda_2 & \Lambda_1 & 0_{(P+1)\times(P+1)} \\ \frac{\gamma-1}{\gamma}\Lambda_3 & \frac{\gamma+1}{\gamma}\Lambda_2 & \frac{\gamma-1}{\gamma}\Lambda_1 \\ 0_{(P+1)\times(P+1)} & \Lambda_3 & \Lambda_2 \end{bmatrix}.$$

By assumption, $I_3 \otimes Q$ is non-singular, and it remains to show that \hat{J} has distinct eigenvectors. Let

$$S = \operatorname{diag}(\Lambda_1 \Lambda_3^{-1}, \sqrt{(\gamma - 1)/\gamma} \Lambda_1^{1/2} \Lambda_3^{-1/2}, I_{(P+1)\times(P+1)}).$$

By assumption, Λ_1 and Λ_3 are invertible, so S and S^{-1} exist. We have

$$J^{S} \equiv S^{-1}\hat{J}S = \begin{bmatrix} \Lambda_{2} & \left[\frac{\gamma-1}{\gamma}\Lambda_{1}\Lambda_{3}\right]^{1/2} & 0_{P\times P} \\ \left[\frac{\gamma-1}{\gamma}\Lambda_{1}\Lambda_{3}\right]^{1/2} & \frac{\gamma-1}{\gamma}\Lambda_{2} & \left[\frac{\gamma-1}{\gamma}\Lambda_{1}\Lambda_{3}\right]^{1/2} \\ 0_{P\times P} & \left[\frac{\gamma-1}{\gamma}\Lambda_{1}\Lambda_{3}\right]^{1/2} & \Lambda_{2} \end{bmatrix}.$$
 (22)

Clearly, J^S is symmetric and has the same eigenvalues as \hat{J} and \tilde{J}^P . Hence, J^S has an eigenvalue decomposition $J^S = Y \tilde{\Lambda}^P Y^T$. Then,

$$\hat{J} = SY\tilde{\Lambda}^P Y^T S^{-1} = SY\tilde{\Lambda}^P (SY)^{-1}.$$
(23)

Combining (21) and (23), we get

$$\tilde{J}^P = [(I_3 \otimes Q)SY]\tilde{\Lambda}^P [(I_3 \otimes Q)SY]^{-1}.$$

Setting $X = (I_3 \otimes Q)SY$, we get the eigenvalue decomposition $\tilde{J}^P = X\tilde{\Lambda}^P X^{-1}$. By assumption, S and Y are non-singular, and we have that

$$\det (X) = \det \left((I_3 \otimes Q) SY \right) \neq 0,$$

which proves that X is non-singular, and thus \tilde{J}^P has a complete set of eigenvectors. \Box

Proposition 3. Property (iii) is satisfied by (19).

Proof. Lemma 1 shows that since the eigenvalue matrix $\tilde{\Lambda}^P$ is also the eigenvalue matrix of the symmetric matrix J^S defined in (22), the eigenvalues are all real. Lemma 1 also shows that the eigenvectors are distinct.

The conditions in Lemma 1 are true for certain basis functions assuming moderate stochastic variation, but it can not be guaranteed for every case, and it certainly does not hold for pathological cases with e.g. negative density. The requirement of nonsingularity of Λ_1, Λ_3 is not very restrictive since it amounts to excluding unphysical behavior, for instance naturally positive quantities taking negative values with nonzero probability. The assumption of constant eigenvectors of the matrix A holds for Haar wavelets (i.e. multi-wavelets with $N_p = 0$), for all orders $P + 1 = 2^{N_r}$, with $N_r \in \mathbb{N}$. See appendix B for a proof sketch. Expressions for the first constant eigenvalue decompositions are included in appendix C for Haar wavelets and piecewise linear multi-wavelets. The eigenvectors of A for P + 1 = 1, 2, 4, 8 are shown to be constant, but we do not give a proof that this is true for piecewise linear multi-wavelets of any order P.

Remark: The Roe variable scheme has been outlined under the implicit assumption of uncertainty manifest in the variables, e.g. initial and boundary condition uncertainty. However, situations such as uncertainty in the adiabatic coefficient γ may be treated in a similar way, although it would result in additional pseudo-spectral products. Pseudo-spectral approximations of $(\gamma - 1)/\gamma$ and $(\gamma + 1)/\gamma$ could then be precomputed to sufficient accuracy.

5 Numerical results

We use the method of manufactured solutions to verify the second order convergence in space of a smooth problem using the MUSCL scheme with Roe variables. We then introduce two test cases for the non-smooth problem; case 1 with an initial function that can be exactly represented by two Legendre polynomials, and case 2 with slow initial decay of the MW coefficients in both N_p and N_r . The errors in computed quantities of interest (here variances) as functions of the order of MW are investigated. Qualitative results are then presented to indicate the behavior we can expect for the convergence of two special cases of MW, namely the Legendre polynomials and Haar wavelet basis, respectively. Robustness with respect to more extreme cases (density close to zero leading to high Mach number) is demonstrated for the Roe variable formulation for a supersonic case where the conservative variable method breaks down. Finally, we perform a comparative study of the computational time for the formulation in conservative variables and the formulation in Roe variables.

5.1 Spatial convergence

The MUSCL scheme with appropriate flux limiters is second order accurate for smooth solutions. Since the Euler solution in general becomes discontinuous in finite time,

the method of manufactured solutions [19, 21] is used to solve the Euler equations with source terms for a known smooth solution. The smooth solution is inserted into the Euler equations (5) and results in a non-zero right-hand side that is used as a source function. In order to test the capabilities of the method, we choose a solution that varies in space, time and in the stochastic dimension and with time-dependent boundary conditions. It is designed to resemble a physical solution with non-negative density and pressure. The solution is given by

$$\begin{bmatrix} \rho \\ v \\ p \end{bmatrix} = \begin{bmatrix} \rho_0 + \rho_1 \tanh(s(x_0 - x + t + \sigma\xi)) \\ \tanh(s(x_0 + v_0 - x + t + \sigma\xi)) + \tanh(-s(x_0 - v_0 - x + t + \sigma\xi)) \\ p_0 + p_1 \tanh(s(x_0 - x + t + \sigma\xi)) \end{bmatrix}.$$

The parameters are set to $\rho_0 = p_0 = 0.75$, $\rho_1 = p_1 = x_0 = 0.25$, $v_0 = 0.05$, s = 10, $\sigma = 0.1$ and $\xi \in \mathcal{U}[-1, 1]$.

We measure the error in the computed $u(x, t, \xi)$ in the $L_2(\Omega, \mathcal{P})$ norm and the discrete ℓ_2 norm,

$$\begin{aligned} \|u^{P} - u\|_{2,2} &\equiv \|u^{P} - u\|_{\ell_{2},L_{2}(\Omega,\mathcal{P})} = \left(\Delta x \sum_{i=1}^{m} \|u^{P}(x_{i},t,\xi) - u(x_{i},t,\xi)\|_{L_{2}(\Omega,\mathcal{P})}^{2}\right)^{1/2} \\ &= \left(\Delta x \sum_{i=1}^{m} \int_{\Omega} (u^{P}(x_{i},t,\xi) - u(x_{i},t,\xi))^{2} d\mathcal{P}(\xi)\right)^{1/2} = \\ &\approx \left(\Delta x \sum_{i=1}^{m} \sum_{j=1}^{q} (u^{P}(x_{i},t,\xi_{q}^{(j)}) - u(x_{i},t,\xi_{q}^{(j)}))^{2} w_{q}^{(j)}\right)^{1/2}, \quad (24) \end{aligned}$$

where a q-point quadrature rule with points $\{\xi_q^{(j)}\}_{j=1}^q$ and weights $\{w_q^{(j)}\}_{j=1}^q$ was used in the last line to approximate the integral in ξ . The Gauss-Legendre quadrature is used here since the solution is smooth in the stochastic dimension.

Figure 2 depicts the spatial convergence in the $\|.\|_{2,2}$ norm of the error in density, velocity and energy. An order $(N_p, N_r) = (10, 0)$ basis is used to represent the uncertainty. The solution dynamics is initially concentrated in the left part of the spatial domain. By the time of t = 0.4, it has moved to the right and has begun to exit the spatial domain, so the time snapshots of Figure 2 summarizes the temporal history of the spatial error decay. The theoretical optimal convergence rate for the MUSCL scheme with the van Leer flux limiter is obtained for all times and all quantities.

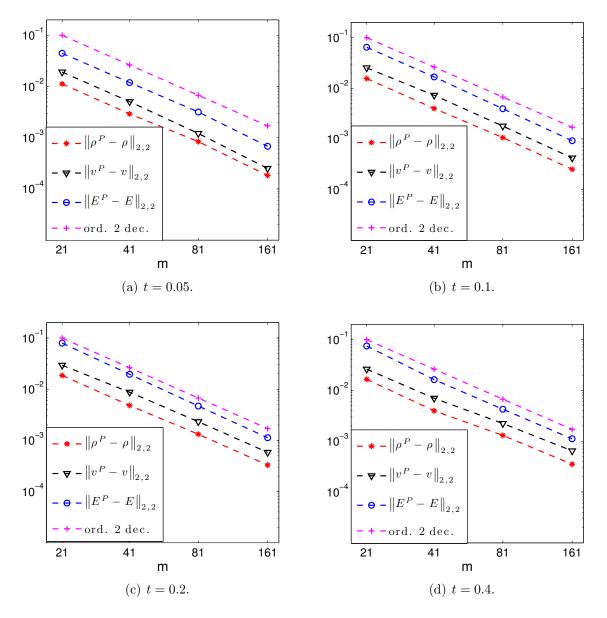


Figure 2: Convergence in space using the method of manufactured solutions, $N_p = 10, N_r = 0$ (Legendre polynomials).

5.2 Initial conditions and discontinuous solutions

We consider (5) with two different initial functions on the domain [0, 1]. Since the analytical solution of Sod's test case is known for any fixed value of the input parameters, the exact stochastic solution can be formulated as a function of the stochastic input ξ . Exact statistics can be computed by numerical integration over ξ . As case 1, assume that the density is subject to uncertainty, and all other quantities are deterministic at t = 0. The initial condition for (5) is given by

$$u(x,t=0,\xi) = \begin{cases} u_L = (1+\sigma\xi, 0, 2.5/\gamma)^T & x < 0.5\\ u_R = (0.125(1+\sigma\xi), 0, 0.25/\gamma)^T & x > 0.5 \end{cases}$$

where we assume $\xi \in \mathcal{U}[-1,1]$, $\gamma = 1.4$ and the scaling parameter $\sigma = 0.5$. This is a simple initial condition in the sense that the first two Legendre polynomials are sufficient to represent the initial function exactly. As case no 2, we consider (5) subject to uncertainty in the initial shock location. Let

$$u(x,t=0,\xi) = \begin{cases} u_L = (1, 0, 2.5/\gamma)^T & x < 0.5 + \sigma\eta \\ u_R = (0.125, 0, 0.25/\gamma)^T & x > 0.5 + \sigma\eta \end{cases}$$

where we assume $\gamma = 1.4$ and the scaling parameter $\sigma = 0.05$. Here, η takes a triangular distribution, which we parameterize as a nonlinear function in $\xi \in \mathcal{U}[-1, 1]$, i.e.

$$\eta(\xi) = (-1 + \sqrt{\xi + 1}) \mathbb{1}_{\{-1 \le \xi \le 0\}}(\xi) + (1 - \sqrt{1 - \xi}) \mathbb{1}_{\{0 < \xi \le 1\}}(\xi),$$

where the indicator function $\mathbb{1}_{\{A\}}$ of a set A is defined by $\mathbb{1}_{\{A\}}(\xi) = 1$ if $\xi \in A$ and zero otherwise. For case 2, exact representation of the initial function requires an infinite number of expansion terms in the MW basis. Figure 3 depicts the shock tube setup for the two cases, with dashed lines denoting uncertain parameters. We will also investigate another version of case 2, where the right state density is significantly reduced to obtain a strong shock.

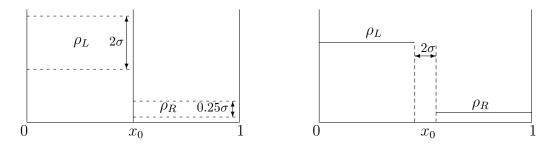


Figure 3: Schematic representation of the initial setup for case 1 (left) and case 2 (right).

5.3 Initial conditions and resolution requirements

For case 2, it should be noted that although the initial shock position can be exactly described by the first two terms of the Legendre polynomial chaos expansion, this is not the case for the initial state variables. In fact, for the the polynomial chaos expansions of the density, momentum and energy, the error decay only slowly with the number of expansion terms. Thus, unless a reasonably large number of expansion terms are retained, the stochastic Galerkin solution of case 2 will not be accurate even for small times.

The Legendre coefficients at small times display an oscillating behavior that becomes sharper with the order of the coefficients. The wavelet coefficients exhibit peaks that get sharper with the resolution level, and require a fine mesh. Figure 4 shows the initial Legendre coefficients and the initial Haar wavelets for case 2. The numerical method has a tendency to smear the chaos coefficients, resulting in under-predicting the variance. The increasing cost of using a larger number of basis functions is further increased by the need for a finer mesh to resolve the solution modes.

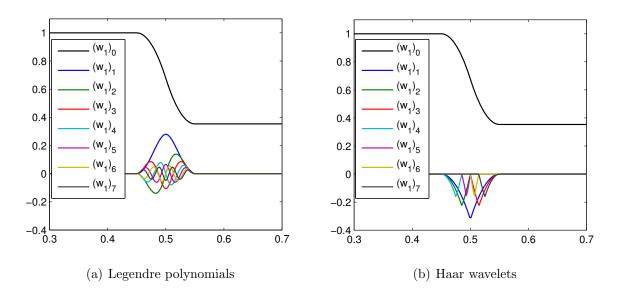


Figure 4: Initial w_1 modes for case 2, first 8 basis functions.

5.4 Convergence of multi-wavelet expansions

For moderate simulation times, the numerical solution on a sufficiently fine spatial mesh converges as the order of MW expansion increases by increasing the polynomial degree N_p or the resolution level N_r . Figure 5 shows the decay in the error of the variance of velocity and energy as a function of N_p and N_r . For well-behaved cases like these, one may freely choose between increasing N_p and N_r , in order to increase the accuracy of the solution of the quantity of interest.

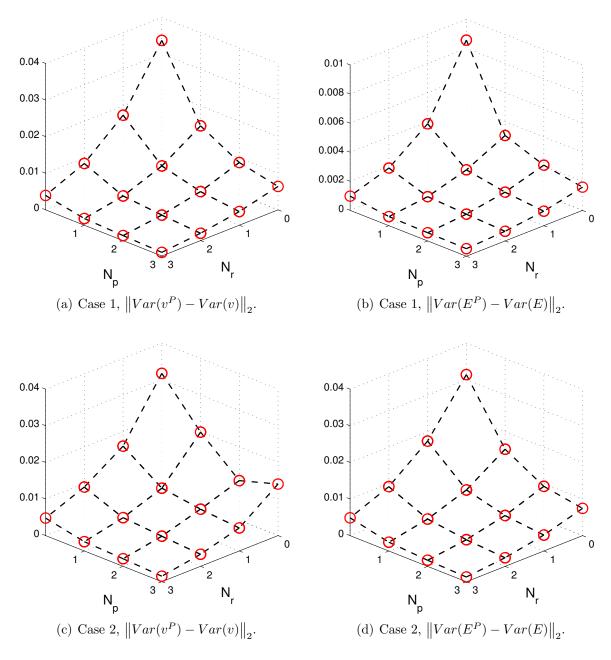


Figure 5: Decay in variance of velocity and energy as a function of the order of expansion, polynomial order N_p and resolution level N_r . Case 1, t = 0.05, 280 spatial points restricted to $x \in [0.4, 0.65]$. Solution obtained with the Roe variable scheme.

For longer simulation times or more extreme cases, e.g. supersonic flow, high-order polynomial representation (increasing N_p) may not lead to increased accuracy, but instead breakdown of the numerical method. Next, we study the qualitative properties of the MW representation of case 1 and case 2 for two extreme cases: Legendre polynomials $(N_r = 0)$ and piecewise constant Haar wavelets $(N_p = 0)$.

Figure 6 shows the density surface in the $x - \xi$ -plane of case 1 and case 2 at t = 0.15 based on exact solution evaluations, and computed with Legendre polynomials and Haar wavelets. The computed solution with Legendre polynomial reconstruction captures essential features of the exact solution, but the use of global polynomials cause oscillations downstream of the shock.

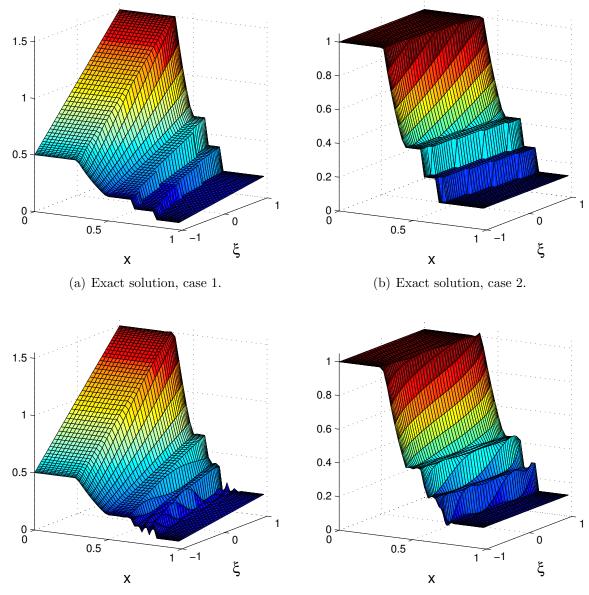
With Haar wavelets, there are no oscillations downstream, as in the Legendre polynomials case. However, the 8 'plateaus' seen in figure 6 (e) corresponds to the 8 basis functions. When the order of wavelet chaos expansion increases, the number of plateaus increases, and the solution converges to the exact solution.

From Figure 6 it is clear that the effect of the choice of multi-wavelet basis to some extent depends on the problem at hand. The Haar wavelets yield numerical solutions that are free of oscillations but converge only slowly. Oscillations around discontinuities in stochastic space should be expected when a polynomial basis is used and may lead to severe problems when variables attain unphysical values, e.g. when the oscillations downstream of the shock leads to negative density. Thus, more robust multi-wavelets are required for problems with stronger shocks, as we demonstrate below.

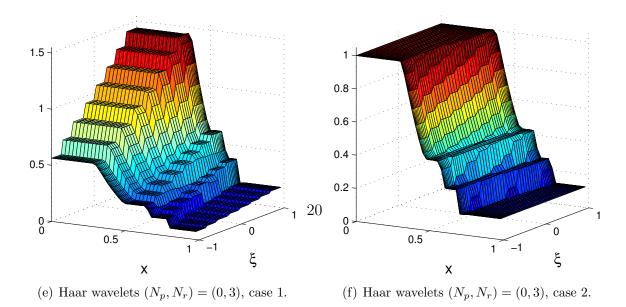
5.5 Robustness

The stochastic Galerkin method applied to the Roe variables gives a more robust method than the conservative variables formulation. Figure 7 shows the relative errors of the solution in the 2, 2 norm for modified versions of case 2 with stronger shocks, $\rho_L = 1$, and a range of right state densities, $\rho_R = 2^{-k}$, k = 3, ..., 8 for 8 basis wavelets. This corresponds to Mach numbers up to Ma = 2.0. Figure 7 also includes the relative error of the Mach number to verify that the cases solved for were reasonable close to the supersonic range they model. For this problem, the conservative variable formulation was unstable except for the original subsonic case 2 ($\rho_R = 0.125$). Thus, although no clear explanation is known to us, it seems that the Roe variable formulation is more suitable for problems where robustness is an issue.

Legendre polynomials are not suitable for this problem. As seen in Figure 6 (c) and (d), the solution is oscillatory in the right state close to the shock. If the right state density is small, as in this supersonic case, such oscillations cause the density to be very close to zero, or even negative. This leads to an unphysical solution and breakdown of the numerical method.



(c) 8 Legendre polynomials $(N_p, N_r) = (8, 0)$, case (d) 8 Legendre polynomials $(N_p, N_r) = (8, 0)$, case 1.



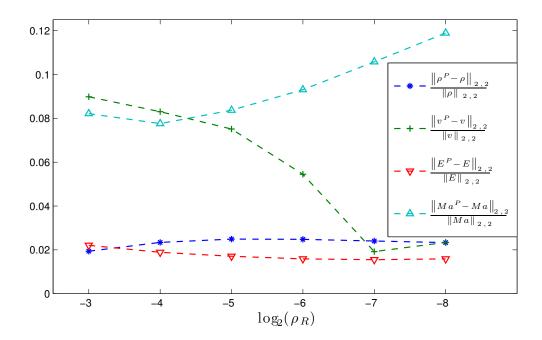


Figure 7: Relative error in density, velocity, energy and Mach number at t = 0.15 for different shock strengths. m = 300 spatial points, 8 Haar wavelets ($N_p = 0$, $N_r = 3$).

5.6 Computational cost

For two stochastic Galerkin systems of order $P = (N_p + 1)2^{N_r} - 1$ and $P' = (N'_p + 1)2^{N'_r} - 1$ where P = P' but $N_p \neq N'_p$, $N'_r \neq N_r$, the size of the problem and the computational cost is the same. Although the different bases could possibly result in properties that make them very different in e.g. the number of iterations required to solve the nonlinear matrix problems, no such tendency was observed. The numerical experiments yield very similar computational costs for the cases tested.

In order to compare the computational cost of the Roe variable expansion with that of the conservative expansion, a similar experimental setup is used for both methods. Sufficiently small test cases are run in order not to exceed the cache limit which would slow down the simulation time for fine meshes and bias the result. We used test case 1 for short simulation times.

In the experiments, the same time step has been used for the different variable expansions, since the stability limit is very similar. Table 1 displays the relative simulation time of the two different variable expansions for increasing number of Haar wavelets ($\tilde{P} = P + 1 = 2^{N_r}$, $N_p = 0$). One time unit is defined as the time for the numerical simulation of a single deterministic problem using the same numerical method with similar input conditions, discretization and time step. The higher computational cost for the conservative variable formulation is due to the need to compute inverse quantities and cubic spectral products. The Roe variable formulation only requires the solution of the nonlinear system for the square root of the density and quadratic flux

function evaluations. The relative benefit of the Roe variable expansion decreases with the order of wavelet expansion. This is due to the increasing cost of forming spectral products that dominates the total cost for high-order expansions.

Order of MW	$\tilde{P}=2$	$\tilde{P} = 4$	$\tilde{P}=8$	$\tilde{P} = 16$
Time Roe variables Time cons. variables	14 410	16 457	26 534	$60 \\ 643$

Table 1: Relative simulation time using conservative variables and Roe variables, respectively. One time unit is defined as the simulation time of a single deterministic problem with the same time-step as for the MW cases.

6 Conclusions

An intrusive formulation of the stochastic Euler equations based on Roe variables is presented. A Roe average matrix for the standard MUSCL-Roe scheme with Roe variables is derived, and we prove that it satisfies the conditions stated by Roe under certain conditions.

The Legendre polynomial basis exactly represents the input uncertainty in our first test case, but it leads to oscillations around the discontinuity in stochastic space. The Haar wavelets do not represent the input uncertainty exactly in either test case, but are more robust to discontinuities.

The Roe variable formulation is robust for supersonic problems where the conservative variable formulation fails, but only for localized basis functions of the generalized chaos representation. For global Legendre polynomials, the discontinuities in stochastic space lead to oscillations and unphysical behavior of the solution and numerical instability. Wavelet functions are more robust in this respect, and do not yield oscillations around discontinuities in stochastic space.

The Roe variable formulation leads to speedup compared to the conservative variable formulation. The relative speedup decreases with the order of generalized chaos since the total computational cost for high-order expansions is no longer dominated by spectral inversion and square root calculations. Instead, the main cost lies in the formation of spectral product matrices. However, for low order multi-wavelet expansions, the speedup is significant.

We demonstrate the need for robust flux functions by presenting cases where the standard MUSCL-Roe flux fails to capture the solution. The design of a robust numerical method is also highly dependent on the choice of the stochastic basis. The Haar wavelets are not only more robust than Legendre polynomials for representation of discontinuities in stochastic space, but also admit the proof of existence of a Roe matrix and more specifically the hyperbolicity of the stochastic Galerkin formulation. This implies that the truncated problem mimics the original problem - a desirable feature.

If the representation of the initial function has not converged, the solution at future times can not be accurate. Case 2 illustrates the need to find a representation of uncertainty with fast decay of the coefficients of the generalized chaos expansion.

Acknowledgements

This work is supported by King Abdullah University of Science and Technology (KAUST).

Appendix A Generation of multi-wavelets

Algorithm 1	Generation	of multi-wavelets
-------------	------------	-------------------

The goal is to generate the mother-wavelets $\{\psi_i^W(\xi)\}_{i=0}^{N_p}$ of (4). Start with a set of functions $\{f_k^1\}_{k=0}^{N_p}$, defined by

$$f_k^1(\xi) = \begin{cases} \xi^k, & \xi \in [-1,0], \\ -\xi^k, & \xi \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$

STEP 1: Orthogonalize with respect to the monomials $1, ..., \xi^{N_p}$ using Gram-Schmidt to obtain $\{f_k^2\}_{k=0}^{N_p}$.

STEP 2:

for
$$i \leftarrow 0$$
 to $N_p - 1$ do
Make sure $\langle f_i^{i+1} \xi^{N_p + i} \rangle \neq 0$ (otherwise reorder).
for $j = i + 1$ to N_0 do
 $w = \frac{\langle f_j^{i+2} \xi^{N_p + i} \rangle}{\langle f_i^{i+2} \xi^{N_p + i} \rangle}$
 $f_j^{i+3} \leftarrow f_j^{i+2} - w f_i^{i+2}$
end for
end for
STEP 3: Orthogonalize $\{f_i^{i+2}\}_{i=0}^{N_p}$ using G-S.
for $i \leftarrow N_p$ to 0 do
 $\psi_i^W(\xi) \leftarrow$ Apply Gram-Schmidt to f_i^{i+2} .
end for
output $\{\psi_i^W(\xi)\}_{i=0}^{N_p}$.

Appendix B Constant eigenvectors of A

Proposition 1. The matrix A defined by (7) for Haar wavelets $\{\psi_j\}_{j=0}^P$ has constant eigenvectors for all $P + 1 = 2^{N_r}$, $N_r \in \mathbb{N}$.

Sketch of proof. We will use induction on the order P of wavelet chaos to show that the matrix A has constant eigenvectors for all orders P. In order to do this, we will need

certain features of the structure of A. To facilitate the notation, denote $\tilde{P} = P + 1$. We can express $A_{2\tilde{P}}$ in terms of the matrix $A_{\tilde{P}}$. Two properties of the triple product $\langle \psi_i \psi_j \psi_k \rangle$ will be used to prove that A indeed has the matrix structure presented.

Property 1: Let $i \in \{0, ..., \tilde{P} - 1\}$, $j = k \in \{\tilde{P}, ..., 2\tilde{P} - 1\}$ and let j' and j'' be the progenies of j. Then

$$\langle \psi_i \psi_j^2 \rangle = \langle \psi_i \psi_{j'}^2 \rangle = \langle \psi_i \psi_{j''}^2 \rangle.$$

Property 2: Consider the indices $i \in \{\tilde{P}, ..., 2\tilde{P} - 1\}, j = k \in \{2\tilde{P}, ..., 4\tilde{P} - 1\}$. Then

$$\langle \psi_i \psi_j^2 \rangle = \begin{cases} \tilde{P}^{1/2} & \text{if } j \text{ first progeny of } i \\ -\tilde{P}^{1/2} & \text{if } j \text{ second progeny of } i \\ 0 & \text{otherwise} \end{cases}$$

As induction hypothesis, we assume that given $A_{\tilde{P}}$ for some $\tilde{P} = 2^{N_r}$, $N_r \in \mathbb{N}$, the next order of triple product matrix $A_{2\tilde{P}}$ can be written

$$A_{2\tilde{P}} = \left[\begin{array}{cc} A_{\tilde{P}} & Q_{\tilde{P}} M_{\tilde{P}} \\ M_{\tilde{P}} Q_{\tilde{P}}^T & \Lambda \end{array} \right]$$

where $Q_{\tilde{P}}$ is the matrix of constant eigenvectors of $A_{\tilde{P}}$ satisfying $||Q_{\tilde{P}}||_2^2 = \tilde{P}$, $M_{\tilde{P}} = \text{diag}(w_{\tilde{P}}, ..., w_{2\tilde{P}-1})$ and Λ is diagonal and contains the eigenvalues of $A_{\tilde{P}}$. Then, we have that

$$\begin{bmatrix} A_{\tilde{P}} & Q_{\tilde{P}}M_{\tilde{P}} \\ M_{\tilde{P}}Q_{\tilde{P}}^T & \Lambda \end{bmatrix} \begin{bmatrix} Q_{\tilde{P}} \\ \pm \tilde{P}^{1/2}I \end{bmatrix} = \begin{bmatrix} Q_{\tilde{P}}\Lambda \pm \tilde{P}^{1/2}QM_{\tilde{P}} \\ \tilde{P}M_{\tilde{P}} \pm \tilde{P}^{1/2}\Lambda \end{bmatrix} = \begin{bmatrix} Q_{\tilde{P}} \\ \pm \tilde{P}^{1/2}I \end{bmatrix} (\Lambda \pm \tilde{P}^{1/2}M),$$

so the eigenvalues and eigenvectors of $A_{2\tilde{P}}$ are given by $\Lambda \pm \tilde{P}^{1/2} M_{\tilde{P}}$ and $[Q_{\tilde{P}}, \pm \tilde{P}^{1/2}I]^T$, respectively. For the next order of expansion, $4\tilde{P}$, we have

$$A_{4\tilde{P}} = \begin{bmatrix} A_{\tilde{P}} & Q_{\tilde{P}}M_{\tilde{P}} \\ M_{\tilde{P}}Q_{\tilde{P}}^{T} & \Lambda \end{bmatrix}^{T} & \begin{bmatrix} Q_{\tilde{P}}\otimes[1,1] \\ \tilde{P}^{1/2}I\otimes[1,-1] \end{bmatrix} M_{2\tilde{P}} \\ M_{2\tilde{P}} \begin{bmatrix} Q_{\tilde{P}}\otimes[1,1] \\ \tilde{P}^{1/2}I\otimes[1,-1] \end{bmatrix}^{T} & \Lambda\otimes I_{2} + \tilde{P}^{1/2}M_{\tilde{P}}\otimes\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{bmatrix}$$
(25)

To see that this is indeed the structure of $A_{4\tilde{P}}$, note that any non-zero matrix entry not already present in $A_{2\tilde{P}}$, can be deduced using properties 1 and 2, and scaling the rows/columns by multiplication by the diagonal matrix $M_{2\tilde{P}}$. The structure of $A_{4\tilde{P}}$ follows from the construction of the Haar wavelet basis, but we do not give a proof here.

One can verify that $A_{4\tilde{P}}$ given by (25) has the eigenvectors and eigenvalues

$$\begin{split} Q_{4\tilde{P}} &= \left[\begin{array}{c} Q_{\tilde{P}} \otimes [1,1] \\ \tilde{P}^{1/2} I_{\tilde{P}} \otimes [1,-1] \\ \pm (2\tilde{P})^{1/2} I_{2\tilde{P}} \end{array} \right], \\ \Lambda_{4\tilde{P}} &= \Lambda \otimes I_2 + \tilde{P}^{1/2} M_{\tilde{P}} \otimes \left[\begin{array}{c} 1 & 0 \\ 0 & -1 \end{array} \right] \pm (2\tilde{P})^{1/2} M_{2\tilde{P}} \end{split}$$

so the eigenvectors are constant (but the eigenvalues are variable in the coefficients $(w_i)_j$ through $M_{\tilde{P}}$ and $M_{2\tilde{P}}$). The base cases $\tilde{P} = 1$, $\tilde{P} = 2$, can easily be verified, so by induction $A_{\tilde{P}}$ has constant eigenvectors for all $\tilde{P} = 2^{N_r}$, $N_r \in \mathbb{N}$.

Appendix C Eigenvalue decompositions of A

C.1 Piecewise constant multi-wavelets (Haar wavelets)

$$\begin{aligned} \mathbf{C.1.1} \quad N_r &= 2 \\ Q &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}, \quad \Lambda &= diag \begin{bmatrix} u_0 + u_1 + \sqrt{2}u_2 \\ u_0 + u_1 - \sqrt{2}u_2 \\ u_0 - u_1 + \sqrt{2}u_3 \\ u_0 - u_1 - \sqrt{2}u_3 \end{bmatrix} \end{aligned}$$

C.1.2 $N_r = 3$

$$\Lambda = diag \begin{bmatrix} u_0 + u_1 + \sqrt{2}u_2 + 2u_4 \\ u_0 + u_1 + \sqrt{2}u_2 - 2u_4 \\ u_0 + u_1 - \sqrt{2}u_2 + 2u_5 \\ u_0 + u_1 - \sqrt{2}u_2 - 2u_5 \\ u_0 - u_1 + \sqrt{2}u_3 + 2u_6 \\ u_0 - u_1 + \sqrt{2}u_3 - 2u_6 \\ u_0 - u_1 - \sqrt{2}u_3 - 2u_7 \\ u_0 - u_1 - \sqrt{2}u_3 - 2u_7 \end{bmatrix}$$

C.2 Piecewise linear multi-wavelets

$$\mathbf{C.2.1} \quad N_r = 1$$

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{3}+1}{4} & \frac{\sqrt{3}-1}{4} & \frac{\sqrt{3}+1}{4} & -\frac{\sqrt{3}-1}{4} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{3}-1}{4} & -\frac{\sqrt{3}+1}{4} & \frac{\sqrt{3}-1}{4} & -\frac{\sqrt{3}+1}{4} \end{bmatrix}, \quad \Lambda = diag \begin{bmatrix} u_0 - \frac{\sqrt{3}+1}{2}u_1 - u_2 - \frac{\sqrt{3}-1}{2}u_3 \\ u_0 + \frac{\sqrt{3}-1}{2}u_1 + u_2 - \frac{\sqrt{3}+1}{2}u_3 \\ u_0 + \frac{\sqrt{3}+1}{2}u_1 - u_2 + \frac{\sqrt{3}-1}{2}u_3 \\ u_0 - \frac{\sqrt{3}-1}{2}u_1 + u_2 - \frac{\sqrt{3}+1}{2}u_3 \\ u_0 - \frac{\sqrt{3}-1}{2}u_1 + u_2 - \frac{\sqrt{3}+1}{2}u_3 \end{bmatrix}$$

C.2.2 $N_r = 2$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}}$$

$$\Lambda = diag \begin{bmatrix} u_0 + \sqrt{\frac{14+3\sqrt{3}}{8}}u_1 - \frac{\sqrt{3}+1}{2}u_2 + \frac{\sqrt{3}+1}{4}u_3 - \sqrt{2}u_6 + \frac{\sqrt{3}-1}{\sqrt{2}}u_7 \\ u_0 - \sqrt{\frac{14+3\sqrt{3}}{8}}u_1 - \frac{\sqrt{3}+1}{2}u_2 - \frac{\sqrt{3}+1}{4}u_3 - \sqrt{2}u_4 - \frac{\sqrt{3}-1}{\sqrt{2}}u_5 \\ u_0 - \sqrt{\frac{14-3\sqrt{3}}{8}}u_1 - \frac{\sqrt{3}-1}{2}u_2 + \frac{\sqrt{3}-1}{4}u_3 + \sqrt{2}u_4 + \frac{\sqrt{3}+1}{\sqrt{2}}u_5 \\ u_0 + \sqrt{\frac{14-3\sqrt{3}}{8}}u_1 - \frac{\sqrt{3}-1}{2}u_2 - \frac{\sqrt{3}-1}{4}u_3 + \sqrt{2}u_6 - \frac{\sqrt{3}+1}{\sqrt{2}}u_7 \\ u_0 + \frac{\sqrt{3}+1}{4}u_1 + \frac{\sqrt{3}-1}{2}u_2 - \sqrt{\frac{14-5\sqrt{3}}{8}}u_3 + \sqrt{2}u_6 + \frac{\sqrt{3}+1}{\sqrt{2}}u_7 \\ u_0 - \frac{\sqrt{3}+1}{4}u_1 + \frac{\sqrt{3}-1}{2}u_2 + \sqrt{\frac{14-5\sqrt{3}}{8}}u_3 + \sqrt{2}u_4 - \frac{\sqrt{3}+1}{\sqrt{2}}u_5 \\ u_0 - \frac{\sqrt{3}-1}{4}u_1 + \frac{\sqrt{3}+1}{2}u_2 + \sqrt{\frac{14+5\sqrt{3}}{8}}u_3 - \sqrt{2}u_4 + \frac{\sqrt{3}-1}{\sqrt{2}}u_5 \\ u_0 + \frac{\sqrt{3}-1}{4}u_1 + \frac{\sqrt{3}+1}{2}u_2 - \sqrt{\frac{14+5\sqrt{3}}{8}}u_3 - \sqrt{2}u_6 - \frac{\sqrt{3}-1}{\sqrt{2}}u_7 \end{bmatrix}$$

References

- [1] Remi Abgrall. A simple, flexible and generic deterministic approach to uncertainty quantifications in non linear problems: application to fluid flow problems. Rapport de recherche, INRIA Bordeaux, 2008.
- [2] Remi Abgrall, Pietro Marco Congedo, Christophe Corre, and Stéphane Galera. A simple semi-intrusive method for Uncertainty Quantification of shocked flows, comparison with a non-intrusive Polynomial Chaos method. In *ECCOMAS CFD*, Lisbonne, Portugal, June 2010.
- [3] Bradley K. Alpert. A class of bases in L2 for the sparse representations of integral operators. SIAM J. Math. Anal., 24(1):246–262, January 1993.
- [4] T. Chantrasmi, A. Doostan, and G. Iaccarino. Padé-Legendre approximants for uncertainty analysis with discontinuous response surfaces. J. Comput. Phys., 228:7159–7180, October 2009.
- [5] Manas K. Deb, Ivo M. Babuska, and J. Tinsley Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Comput. Methods Appl. Mech. Eng.*, 190(48):6359 – 6372, 2001.
- [6] Bert J. Debusschere, 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.*, 26:698–719, February 2005.
- [7] Roger G. Ghanem and Pol D. Spanos. Stochastic finite elements: a spectral approach. Springer-Verlag, New York, 1991.
- [8] O. P. Le Maître and O. M. Knio. Spectral Methods for Uncertainty Quantification. Springer, first edition, 2010.
- [9] O. P. Le Maître, O. M. Knio, H. N. Najm, and R. G. Ghanem. Uncertainty propagation using Wiener-Haar expansions. J. Comput. Phys., 197:28–57, June 2004.
- [10] 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., 197:502–531, July 2004.
- [11] O. P. Le Maître, H. N. Najm, P. P. Pébay, R. G. Ghanem, and O. M. Knio. Multi-resolution-analysis scheme for uncertainty quantification in chemical systems. SIAM J. Sci. Comput., 29:864–889, March 2007.
- [12] R. J. LeVeque. Finite-Volume Methods for Hyperbolic Problems. Cambridge University Press, Cambridge, 2002.
- [13] Per Pettersson, Qaiser Abbas, Gianluca Iaccarino, and Jan Nordström. Efficiency of shock capturing schemes for Burgers' equation with boundary uncertainty. Enumath 2009, The eight European Conference on Numerical Mathematics and Advanced Applications, June 29-July 3, 2009.

- [14] Per Pettersson, Gianluca Iaccarino, and Jan Nordström. Numerical analysis of the Burgers' equation in the presence of uncertainty. J. Comput. Phys., 228:8394– 8412, December 2009.
- [15] Per Pettersson, Jan Nordström, and Gianluca Iaccarino. Boundary procedures for the time-dependent Burgers' equation under uncertainty. Acta Mathematica Scientia, 30(2):539–550, 2010.
- [16] C.L. Pettit and P.S. Beran. Spectral and multiresolution Wiener expansions of oscillatory stochastic processes. J. Sound Vib., 294:752–779, 2006.
- [17] Gaël Poëtte, Bruno Després, and Didier Lucor. Uncertainty quantification for systems of conservation laws. J. Comput. Phys., 228:2443–2467, April 2009.
- [18] M. J. D. Powell. A Fortran subroutine for solving systems of nonlinear algebraic equations. In P. Rabinowitz, editor, *Numerical Methods for Nonlinear Algebraic Equations*, chapter 7. P. Rabinowitz, 1970.
- [19] Patrick J. Roache. Verification of codes and calculations. AIAA Journal, 36(5):696–702, May 1988.
- [20] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. J. Comput. Phys., 43(2):357 – 372, 1981.
- [21] Lee Shunn, Frank E. Ham, and Parviz Moin. Verification of variable-density flow solvers using manufactured solutions. J. Comput. Physics, 231(9):3801–3827, 2012.
- [22] J. Tryoen, O. P. Le Maître, M. Ndjinga, and A. Ern. Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems. J. Comput. Phys., 229(18):6485 - 6511, 2010.
- [23] J. Tryoen, O. P. Le Maître, M. Ndjinga, and A. Ern. Roe solver with entropy corrector for uncertain hyperbolic systems. J. Comput. Appl. Math., 235:491–506, November 2010.
- [24] B. van Leer. Towards the ultimate conservative difference scheme. V A secondorder sequel to Godunov's method. J. Comput. Phys., 32:101–136, July 1979.
- [25] Xiaoliang Wan and George Em Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. J. Comput. Phys., 209:617–642, November 2005.
- [26] Norbert Wiener. The homogeneous chaos. American Journal of Mathematics, 60(4):897–936, 1938.
- [27] Jeroen A.S. Witteveen, Alex Loeven, and Hester Bijl. An adaptive stochastic finite elements approach based on Newton-Cotes quadrature in simplex elements. *Computers & Fluids*, 38(6):1270 – 1288, 2009.
- [28] Dongbin Xiu and George Em Karniadakis. The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput., 24(2):619–644, 2002.