



**HAL**  
open science

# Méthodes de Galerkin stochastiques adaptatives pour la propagation d'incertitudes paramétriques dans les modèles hyperboliques

Julie Tryoen

► **To cite this version:**

Julie Tryoen. Méthodes de Galerkin stochastiques adaptatives pour la propagation d'incertitudes paramétriques dans les modèles hyperboliques. Mathématiques générales [math.GM]. Université Paris-Est, 2011. Français. NNT : 2011PEST1054 . pastel-00795322

**HAL Id: pastel-00795322**

**<https://pastel.hal.science/pastel-00795322>**

Submitted on 27 Feb 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ÉCOLE DOCTORALE MSTIC  
MATHÉMATIQUES, SCIENCES ET TECHNOLOGIES  
DE L'INFORMATION ET DE LA COMMUNICATION

## THÈSE

pour obtenir le titre de

**Docteur ès Sciences**

de l'Université Paris-Est

**Mention : Mathématiques Appliquées**

Présentée et soutenue le 21 novembre 2011 par

**Julie Tryoen**

# Méthodes de Galerkin stochastiques adaptatives pour la propagation d'incertitudes paramétriques dans les systèmes hyperboliques

Thèse dirigée par Alexandre ERN et Olivier LE MAÎTRE  
préparée au CERMICS, Ecole des Ponts ParisTech et au LIMSI, CNRS

### Jury :

<i>Président :</i>	Denis TALAY	-	INRIA, Sophia Antipolis
<i>Rapporteurs :</i>	Bruno DESPRÉS	-	LJLL, Université Pierre et Marie Curie, Paris 6
	Fabio NOBILE	-	Ecole Polytechnique Fédérale de Lausanne
<i>Examineurs :</i>	Rémi ABGRALL	-	INRIA, Bordeaux
	Thierry GALLOUËT	-	LATP, CMI, Université Aix-Marseille
	Marie POSTEL	-	LJLL, Université Pierre et Marie Curie, Paris 6
<i>Directeurs :</i>	Alexandre ERN	-	CERMICS, Université Paris-Est
	Olivier LE MAÎTRE	-	LIMSI, CNRS



# Méthodes de Galerkin stochastiques adaptatives pour la propagation d'incertitudes paramétriques dans les systèmes hyperboliques

---

## Résumé

---

On considère des méthodes de Galerkin stochastiques pour des systèmes hyperboliques faisant intervenir des données en entrée incertaines de lois de distribution connues paramétrées par des variables aléatoires. On s'intéresse à des problèmes où un choc apparaît presque sûrement en temps fini. Dans ce cas, la solution peut développer des discontinuités dans les domaines spatial et stochastique. On utilise un schéma de Volumes Finis pour la discrétisation spatiale et une projection de Galerkin basée sur une approximation polynomiale par morceaux pour la discrétisation stochastique. On propose un solveur de type Roe avec correcteur entropique pour le système de Galerkin, utilisant une technique originale pour approcher la valeur absolue de la matrice de Roe et une adaptation du correcteur entropique de Dubois et Mehlmann. La méthode proposée reste coûteuse car une discrétisation stochastique très fine est nécessaire pour représenter la solution au voisinage des discontinuités. Il est donc nécessaire de faire appel à des stratégies adaptatives. Comme les discontinuités sont localisées en espace et évoluent en temps, on propose des représentations stochastiques dépendant de l'espace et du temps. On formule cette méthodologie dans un contexte multi-résolution basé sur le concept d'arbres binaires pour décrire la discrétisation stochastique. Les étapes d'enrichissement et d'élagage adaptatifs sont réalisées en utilisant des critères d'analyse multi-résolution. Dans le cas multidimensionnel, une anisotropie de la procédure adaptative est proposée. La méthodologie est évaluée sur le système des équations d'Euler dans un tube à choc et sur l'équation de Burgers en une et deux dimensions stochastiques.

---

## Mots clés

---

Quantification d'incertitudes, systèmes hyperboliques, lois de conservation, méthodes spectrales stochastiques, projection de Galerkin, Volumes Finis, solveur de Roe, correction entropique, analyse multi-résolution, multi-ondelettes, discrétisation stochastique adaptative, adaptation anisotrope



# Adaptive stochastic Galerkin methods for parametric uncertainty propagation in hyperbolic systems

---

## Abstract

---

This work is concerned with stochastic Galerkin methods for hyperbolic systems involving uncertain data with known distribution functions parametrized by random variables. We are interested in problems where a shock appears almost surely in finite time. In this case, the solution exhibits discontinuities in the spatial and in the stochastic domains. A Finite Volume scheme is used for the spatial discretization and a Galerkin projection based on piecewise polynomial approximation is used for the stochastic discretization. A Roe-type solver with an entropy correction is proposed for the Galerkin system, using an original technique to approximate the absolute value of the Roe matrix and an adaptation of the Dubois and Mehlman entropy corrector. Although this method deals with complex situations, it remains costly because a very fine stochastic discretization is needed to represent the solution in the vicinity of discontinuities. This fact calls for adaptive strategies. As discontinuities are localized in space and time, stochastic representations depending on space and time are proposed. This methodology is formulated in a multiresolution context based on the concept of binary trees for the stochastic discretization. The adaptive enrichment and coarsening steps are based on multiresolution analysis criteria. In the multidimensional case, an anisotropy of the adaptive procedure is proposed. The method is tested on the Euler equations in a shock tube and on the Burgers equation in one and two stochastic dimensions.

---

## Key words

---

Uncertainty Quantification, hyperbolic systems, conservation laws, stochastic spectral methods, Galerkin projection, Finite Volumes, Roe solver, entropy correction, multiresolution analysis, multiwavelets, adaptive stochastic discretization, anisotropic adaptation



A mes parents.





# Remerciements

Mes premiers remerciements vont naturellement à mes deux directeurs de thèse, Alexandre Ern et Olivier Le Maître, qui m'ont permis de poursuivre le travail initié lors de mon stage de master au CEA Saclay par une thèse en codirection à l'Ecole des Ponts et au LIMSI. Je les remercie pour l'aide compétente qu'ils m'ont apportée, pour leur disponibilité, leur patience et leurs encouragements. Merci à Alexandre, qui par ses vastes connaissances mathématiques et son esprit de rigueur m'a été d'une aide précieuse dans toute la partie théorique de cette thèse. Merci à Olivier, de m'avoir fait découvrir l'univers fascinant des incertitudes, ainsi que pour son appui dans toute la partie numérique et programmation de cette thèse.

Je remercie ensuite Michaël Ndjinga du CEA Saclay, mon co-encadrant de stage de master, qui par ses connaissances en systèmes hyperboliques, a grandement participé aux premiers résultats de cette thèse. Merci également à Jean-Marc Martinez du CEA Saclay, à l'initiative de ce stage, en concertation avec Olivier et Michaël.

Je suis très reconnaissante à Bruno Després et à Fabio Nobile du temps qu'ils m'ont consacré en tant que rapporteurs sur cette thèse. Mes remerciements s'adressent également à Rémi Abgrall, Thierry Gallouët, Marie Postel et Denis Talay pour leur présence dans le jury, et à l'ensemble des membres du jury pour l'intérêt qu'ils ont porté à ce travail.

Je salue aussi mes collègues du CERMICS et du LIMSI, et les remercie pour leur bonne humeur, leur soutien et leurs conseils. En particulier, je remercie Christian Tenaud du LIMSI pour nos multiples discussions sur l'analyse multi-résolution.

Enfin, mes pensées vont à mes parents, à mon frère et ma sœur, qui m'ont toujours soutenue et encouragée, ainsi qu'à ma famille et mes amis, pour tout ce qu'ils m'apportent. Et merci à Florent, qui m'a considérablement épaulée par ses conseils, son écoute et sa patience de chaque jour.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Propagation et quantification des incertitudes . . . . .	2
1.1.1	Les incertitudes en simulation numérique . . . . .	2
1.1.2	Cadre probabiliste et paramétrisation . . . . .	3
1.1.3	Propagation d'incertitudes . . . . .	5
1.2	Méthodes spectrales stochastiques . . . . .	7
1.2.1	Chaos Polynomial de dimension finie . . . . .	7
1.2.2	Application à la propagation d'incertitudes . . . . .	13
1.2.3	Méthodes non-intrusives . . . . .	16
1.2.4	Méthodes de Galerkin (intrusives) . . . . .	20
1.3	Systèmes hyperboliques . . . . .	22
1.3.1	Systèmes hyperboliques déterministes . . . . .	22
1.3.2	Systèmes hyperboliques stochastiques . . . . .	25
1.3.3	Méthodes spectrales stochastiques : principales difficultés et état de l'art . . . . .	27
1.4	Description des travaux . . . . .	29
<b>I</b>	<b>Solveur de Roe stochastique</b>	<b>31</b>
<b>2</b>	<b>Conception et analyse du solveur</b>	<b>33</b>
2.1	Projection de Galerkin . . . . .	33
2.1.1	Rappel du cadre probabiliste pour les incertitudes paramétriques . . . . .	34
2.1.2	Discretisation stochastique . . . . .	34
2.1.3	Le système de Galerkin . . . . .	36
2.1.4	Calcul des non-linéarités . . . . .	37
2.2	Hyperbolicité du système de Galerkin . . . . .	38
2.2.1	Systèmes hyperboliques stochastiques symétriques . . . . .	39
2.2.2	Vecteurs propres indépendants de l'incertitude . . . . .	39
2.2.3	Cas général . . . . .	40
2.3	Approximation en espace et en temps . . . . .	43
2.3.1	Matrice de Roe et état de Roe . . . . .	44
2.3.2	Une méthode efficace pour approcher la valeur absolue d'une matrice . . . . .	45
2.3.3	Le schéma de décentrement . . . . .	48
2.4	Correcteur entropique . . . . .	49
2.4.1	Principes généraux . . . . .	50
2.4.2	Application directe au système de Galerkin . . . . .	50
2.4.3	Adaptation au système de Galerkin . . . . .	52

<b>3</b>	<b>Résultats numériques</b>	<b>55</b>
3.1	Burgers . . . . .	55
3.1.1	Cas test 1 : vitesses des ondes positives . . . . .	55
3.1.2	Cas test 2 : vitesses des ondes positives et négatives . . . . .	62
3.1.3	Cas test 3 : point sonique . . . . .	72
3.2	Euler . . . . .	78
3.2.1	Cas test 4 : tube à choc sans point sonique . . . . .	78
3.2.2	Cas test 5 : tube à choc avec point sonique . . . . .	87
3.3	Conclusion . . . . .	94
<b>II</b>	<b>Adaptation stochastique</b>	<b>95</b>
<b>4</b>	<b>Outils d'Analyse Multi-Résolution 1D</b>	<b>97</b>
4.1	Analyse Multi-Résolution Multi-Ondelettes . . . . .	98
4.1.1	Ingrédients de base . . . . .	98
4.1.2	Système orthogonal MW d'Alpert . . . . .	100
4.1.3	Seuillage . . . . .	108
4.1.4	Complétion à la Harten . . . . .	109
4.2	Arbres binaires . . . . .	111
4.2.1	Principales définitions . . . . .	111
4.2.2	Grille associée à un arbre binaire . . . . .	112
4.2.3	Espace d'approximation stochastique associé à un arbre binaire	113
4.2.4	Inclusion d'arbres . . . . .	114
4.3	Adaptation à une fonction connue . . . . .	114
4.3.1	Principe . . . . .	114
4.3.2	Aspects algorithmiques . . . . .	117
4.3.3	Résultats . . . . .	119
4.4	Synthèse . . . . .	124
<b>5</b>	<b>Solveur de Roe adaptatif 1D</b>	<b>125</b>
5.1	Aspects algorithmiques . . . . .	126
5.1.1	Présentation générale du solveur de Roe adaptatif . . . . .	126
5.1.2	Union d'arbres binaires . . . . .	128
5.1.3	Opérateurs de restriction et de prédiction . . . . .	128
5.1.4	Élagage et enrichissement d'arbres . . . . .	130
5.2	Application aux équations d'Euler . . . . .	132
<b>6</b>	<b>Extension au cas multidimensionnel</b>	<b>137</b>
6.1	Arbres binaires multiD . . . . .	137
6.1.1	Indicateur directionnel de partition . . . . .	137
6.1.2	Arbres équivalents . . . . .	139
6.1.3	Union d'arbres binaires multiD . . . . .	139
6.1.4	Opérateurs de restriction et de prédiction . . . . .	140

6.2	Elagage et enrichissement . . . . .	140
6.2.1	Elagage anisotrope . . . . .	141
6.2.2	Enrichissement isotrope . . . . .	142
6.3	Application à l'équation de Burgers . . . . .	142
6.3.1	Cas test 1 : vitesse des ondes positives . . . . .	142
6.3.2	Cas test 2 : vitesses des ondes positives et négatives . . . . .	143
<b>7</b>	<b>Conclusion et perspectives</b>	<b>153</b>
	<b>Bibliographie</b>	<b>155</b>



# Introduction

*" If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts, he shall end in certainties. "*  
*F. Bacon - 1605.*

## Sommaire

<b>1.1 Propagation et quantification des incertitudes . . . . .</b>	<b>2</b>
1.1.1 Les incertitudes en simulation numérique . . . . .	2
1.1.2 Cadre probabiliste et paramétrisation . . . . .	3
1.1.3 Propagation d'incertitudes . . . . .	5
<b>1.2 Méthodes spectrales stochastiques . . . . .</b>	<b>7</b>
1.2.1 Chaos Polynomial de dimension finie . . . . .	7
1.2.2 Application à la propagation d'incertitudes . . . . .	13
1.2.3 Méthodes non-intrusives . . . . .	16
1.2.4 Méthodes de Galerkin (intrusives) . . . . .	20
<b>1.3 Systèmes hyperboliques . . . . .</b>	<b>22</b>
1.3.1 Systèmes hyperboliques déterministes . . . . .	22
1.3.2 Systèmes hyperboliques stochastiques . . . . .	25
1.3.3 Méthodes spectrales stochastiques : principales difficultés et état de l'art . . . . .	27
<b>1.4 Description des travaux . . . . .</b>	<b>29</b>

De nos jours, la simulation numérique de systèmes complexes est utilisée dans de nombreux domaines de l'ingénierie. L'essor de la simulation numérique repose en grande partie sur les progrès constants des techniques numériques et la puissance croissante des calculateurs. La simulation numérique est un outil essentiel pour les ingénieurs tout au long du processus de conception. Elle réduit le besoin en études expérimentales qui sont souvent longues, coûteuses et délicates à mettre en œuvre. Cependant, les simulations numériques doivent être conçues, réalisées et vérifiées avec soin, afin de fournir des informations utiles et fiables sur le système à étudier. De fait, la confiance que l'on peut avoir dans un calcul est un aspect clé de l'interprétation et de l'analyse des résultats. Les simulations impliquent différents types d'erreurs, dont la compréhension et la quantification est cruciale pour appréhender les différences entre les prédictions numériques et le comportement réel du système.

Ce travail porte sur la propagation et la quantification des incertitudes sur les données en entrée dans des modèles gouvernés par des systèmes hyperboliques



stochastiques. Après une introduction générale sur le cadre probabiliste, on présente une classe de méthodes appelées méthodes spectrales stochastiques qui permet de propager les incertitudes sur les données dans un modèle mathématique. En particulier, on décrit les méthodes de Galerkin, qu'on appliquera aux systèmes hyperboliques stochastiques. On rappelle ensuite quelques définitions et propriétés des systèmes hyperboliques déterministes et on définit la notion de systèmes hyperboliques stochastiques, pour lesquels on évoque les difficultés rencontrées lors de l'application des méthodes spectrales stochastiques et l'état de l'art correspondant. Enfin, on présente les contributions développées dans ce manuscrit.

## 1.1 Propagation et quantification des incertitudes

Dans cette section, on passe d'abord en revue les différents types d'incertitudes en simulation numérique. On se concentre ensuite sur les incertitudes sur les données en adoptant une approche paramétrique pour les décrire. Enfin, on passe en revue différentes méthodes pour propager les incertitudes paramétriques dans un modèle mathématique.

### 1.1.1 Les incertitudes en simulation numérique

Classiquement, les erreurs intervenant dans les simulations peuvent être regroupées en trois grandes catégories (Soize [76]) :

- Les erreurs de **modélisation** : les simulations sont basées sur la résolution de modèles mathématiques censés restituer les caractéristiques essentielles du système à étudier. Souvent, des simplifications physiques sont utilisées pour faciliter la résolution du modèle mathématique, basées sur l'analyse du problème ou sur certaines hypothèses. Le modèle obtenu n'est par conséquent pas capable de reproduire exactement le comportement du système réel.
- Les erreurs **numériques** : le modèle mathématique sélectionné est résolu en utilisant des techniques de discrétisation et des algorithmes appropriés. La résolution numérique du modèle fournit en général une approximation de la solution exacte du modèle et introduit donc une erreur numérique. Celle-ci peut être contrôlée et réduite à un niveau arbitrairement bas, au moins en théorie, en utilisant des discrétisations plus fines ou des ressources de calcul plus importantes.
- Les erreurs sur les **données** : le modèle mathématique fait intervenir un ensemble de données qui spécifient les caractéristiques du système simulé : conditions aux limites, conditions initiales, constantes physiques, géométrie, forçages extérieurs. Dans beaucoup de cas, les données ne peuvent pas être spécifiées avec une grande précision, soit pour des raisons expérimentales (erreurs de mesure), soit à cause de certaines limitations dans la connaissance du système ou du fait d'une variabilité intrinsèque irréductible.

On s'est intéressé dans ce travail à la dernière catégorie d'incertitudes. En particulier, on s'est attaché à développer des méthodes permettant de caractériser l'impact

d'une connaissance imprécise des données en entrée sur la réponse du modèle. Ainsi, en notant les données par  $D$ , la réponse  $U$  du système est régie par un opérateur  $\mathcal{M}$  (le modèle), que l'on peut écrire sous la forme abstraite

$$\mathcal{M}(U, D) = 0. \quad (1.1)$$

Par la suite, on fait toujours l'hypothèse de modèles bien posés, *i.e.*, on suppose que pour tout  $D$  fixé, il existe une et une seule solution  $U$  satisfaisant (1.1). Dans le domaine de la simulation numérique, la prise en compte de ce type d'incertitudes est un point clé, notamment en analyse de risque, en sûreté et en conception. Les objectifs sont, entre autres, de fournir des barres d'erreur numérique facilitant la comparaison avec des observations expérimentales et ainsi de mieux juger de la qualité des modèles physiques employés ; d'identifier les données incertaines ayant le plus grand impact sur la simulation et devant donc être mesurées ou contrôlées avec le plus de précision ; de mener une analyse de sûreté (probabilité de dépassement de valeurs critiques) et de jauger le degré de confiance que l'on peut accorder aux calculs, lors de la prise de décisions de conception par exemple.

### 1.1.2 Cadre probabiliste et paramétrisation

Pour atteindre les objectifs énoncés ci-dessus, un cadre probabiliste se révèle adapté : puisque les données d'entrée ne peuvent être déterminées exactement, il apparaît naturel de les considérer comme des quantités aléatoires. Dans ce manuscrit, on adopte la convention de notation suivante : les symboles en minuscules représentent des quantités déterministes, tandis que les symboles en majuscule représentent des quantités aléatoires.

On définit les données incertaines par le vecteur aléatoire  $D$  sur un espace de probabilité  $\mathcal{P} := (\Theta, \Sigma, d\mu)$ , dont on suppose implicitement l'existence, où  $\Theta$  est l'ensemble des événements aléatoires,  $\Sigma$  la  $\sigma$ -algèbre associée, et  $d\mu$  la mesure de probabilité. On réécrit (1.1) sous la forme

$$\mathcal{M}(U(\theta), D(\theta)) = 0, \quad \text{p.s. en } \theta. \quad (1.2)$$

L'opérateur d'espérance est défini par

$$E[U] := \int_{\Theta} U(\theta) d\mu(\theta), \quad \forall U \text{ intégrable sur } \mathcal{P}. \quad (1.3)$$

On se placera dans l'espace  $L^2(\Theta, d\mu)$  engendré par les variables aléatoires du second-ordre sur  $\mathcal{P}$ . Cet espace est équipé du produit scalaire défini grâce à l'opérateur d'espérance  $E[\cdot]$  et de la norme associée  $\|\cdot\|_{\Theta}$  :

$$E[UV] := \int_{\Theta} U(\theta)V(\theta) d\mu(\theta) \quad \forall U, V \in L^2(\Theta, d\mu), \quad (1.4)$$

$$U \in L^2(\Theta, d\mu) \mapsto \|U\|_{\Theta}^2 := E[U^2] < +\infty. \quad (1.5)$$

On se place ici dans le cadre des méthodes paramétriques, qui sont conditionnées par une paramétrisation des données incertaines à l'aide d'un ensemble fini de  $N$  variables aléatoires (VAs) indépendantes définies sur  $\mathcal{P}$ , parfois appelé **germe**, noté

$$\xi(\theta) = \{\xi_1(\theta), \dots, \xi_N(\theta)\}, \quad (1.6)$$

de distribution connue  $p_\xi$  et à valeurs dans un ensemble  $\Xi \subset \mathbb{R}^N$  appelé **domaine stochastique**. On note  $\mathcal{P}_\xi := (\Xi, \mathcal{B}_\Xi, p_\xi d\xi)$  l'espace de probabilité image, où  $\mathcal{B}_\Xi$  est l'ensemble de Borel de  $\Xi$  et  $d\xi$  la mesure de Lebesgue dans  $\mathbb{R}^N$ . On réécrit (1.2) sous la forme

$$\mathcal{M}(U(\xi), D(\xi)) = 0, \quad \text{p.s. en } \xi. \quad (1.7)$$

On suppose que le modèle (1.7) a une solution unique presque sûrement et on appelle  $U \equiv U(\xi)$  la **solution stochastique** du modèle. On peut exprimer l'espérance de  $U$  aussi bien dans l'espace de départ, que dans l'espace associé aux  $\{\xi_i\}_{i=1}^N$ ,

$$E[U] = \int_{\Theta} U(\xi(\theta)) d\mu(\theta) = \int_{\Xi} U(y) p_\xi(y) dy = \langle U \rangle. \quad (1.8)$$

On utilisera les crochets  $\langle \cdot \rangle$  pour rendre clair le fait que l'espérance est mesurée par rapport à la distribution de probabilité des variables aléatoires  $\xi$ . On définit alors l'espace des variables aléatoires du second ordre sur  $\mathcal{P}_\xi$ , noté  $L^2(\Xi, p_\xi)$ , équipé du produit scalaire, noté également à l'aide de crochets sous la forme  $\langle \cdot, \cdot \rangle$ , et de la norme associée  $\|\cdot\|_\Xi$  :

$$\langle U, V \rangle := \int_{\Xi} U(\xi) V(\xi) p_\xi(\xi) d\xi \quad \forall U, V \in L^2(\Xi, p_\xi), \quad (1.9)$$

$$U \in L^2(\Xi, p_\xi) \mapsto \|U\|_\Xi^2 := \langle U, U \rangle < +\infty. \quad (1.10)$$

Par la suite, les quantités aléatoires seront toutes supposées du second ordre, *i.e.* dans  $L^2(\Xi, p_\xi)$  (et donc également dans  $L^2(\Theta, d\mu)$ ).

Pour effectuer la paramétrisation de  $D$ , on dispose de plusieurs techniques (Ditlevsen *et al.* [26]) : les décompositions de Karhunen–Loève pour les processus stochastiques [56] ou, plus généralement, les décompositions en polynômes de chaos (*cf* section 1.2). Le type d'information dont l'on dispose sur une quantité aléatoire en entrée peut être très différent d'une application à une autre. Il peut s'agir de sa loi de probabilité complète ou simplement d'un échantillon de réalisations. Généralement, pour construire la représentation probabiliste d'une donnée d'entrée, on peut considérer un problème d'optimisation, dépendant de l'information disponible. La question de la convergence de cette représentation peut s'avérer délicate, particulièrement lorsque l'on a seulement accès à un échantillon de réalisations ; on a alors recours à des techniques d'optimisation particulières basées par exemple sur l'inférence Bayésienne ou sur des principes de maximum d'entropie (Das *et al.* [15], Desceliers *et al.* [22, 23], Soize et Ghanem [77, 78]). Dans ce travail, on fait l'hypothèse que l'on connaît toujours la loi de probabilité des données incertaines  $D$  et que celles-ci sont facilement paramétrisables par un vecteur  $\xi$  de  $N$  variables aléatoires indépendantes.

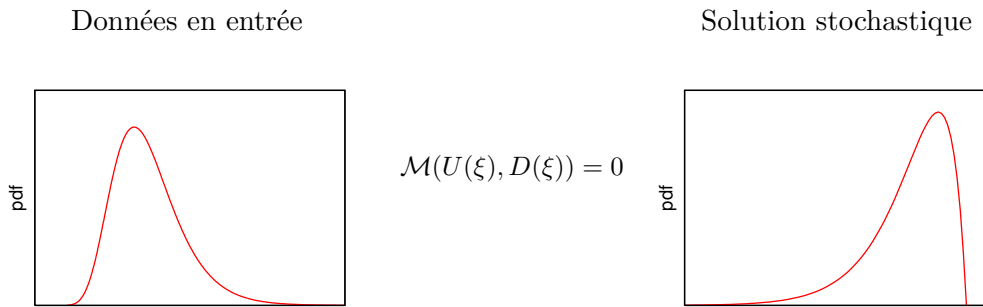


FIGURE 1.1 – Synoptique d’une propagation d’incertitudes : connaissant la loi de probabilité des données incertaines  $D$  (à gauche), on réalise une paramétrisation à partir d’un ensemble de VAs (le germe) suivant une loi de probabilité donnée ; l’étape de propagation consiste ensuite à déterminer la loi de probabilité de la solution  $U$  induite par le germe (à droite) au travers du modèle mathématique  $\mathcal{M}$ .

### 1.1.3 Propagation d’incertitudes

La figure 1.1 illustre la propagation d’incertitudes à travers le modèle mathématique (1.7). Formellement, on cherche à déterminer la densité de probabilité de la solution  $U$  du modèle (1.7) induite par les données aléatoires paramétrées  $D \equiv D(\xi)$ . Cela revient à chercher la dépendance fonctionnelle de  $U \equiv U(\xi)$  vis-à-vis des VAs paramétrant les données incertaines. Il existe plusieurs méthodes pour la propagation d’incertitudes, dont on dresse ci-dessous un bref panorama.

#### 1.1.3.1 Méthodes déterministes

Bien qu’on ne s’intéresse pas à cette classe de méthodes dans ce manuscrit, on présente ici une courte revue de quelques techniques déterministes pour la propagation d’incertitudes.

**Méthodes de perturbation** Les méthodes de perturbation, abondamment utilisées dans de nombreux domaines de l’ingénierie (Kleiber et Hien [43]), sont basées sur un développement tronqué des processus aléatoires *via* des séries de Taylor autour de leur valeur moyenne. Typiquement, des développements du second ordre au plus sont utilisés car le système d’équations résultant devient extrêmement complexe pour des degrés plus élevés. Une limitation inhérente à ces méthodes est que l’ampleur des incertitudes, aussi bien en entrée qu’en sortie, ne peut être trop large, car autrement ces méthodes ne se comportent pas bien.

**Méthode des moments** Dans cette approche, on tente de calculer les moments statistiques de la solution stochastique directement. Les inconnues sont les moments statistiques de la solution, et leurs équations sont directement dérivées en intégrant les équations du modèle. Par exemple, l’espérance est déterminée en intégrant ces

équations sur le domaine stochastique. La difficulté essentielle réside dans la fermeture des termes couplant les équations sur les différents moments statistiques de la solution.

**Méthodes basées sur les opérateurs** Ces méthodes sont basées sur la manipulation des opérateurs stochastiques dans les équations du modèle. Elles incluent le développement de Neumann, qui exprime l'inverse de l'opérateur stochastique par série de Neumann (Shinozuka et Deodatis [74], Yamazaki *et al.* [90]) et la méthode d'intégration à poids (Deodatis et Shinozuka [20, 21]). De façon similaire aux méthodes de perturbation, ces méthodes sont également limitées à des petites incertitudes. Leur application est souvent fortement dépendante de l'opérateur sous-jacent et est typiquement limitée aux problèmes linéaires ou faiblement non-linéaires.

### 1.1.3.2 Méthodes d'échantillonnage Monte Carlo

L'une des méthodes les plus communément utilisées pour la propagation d'incertitudes est la méthode d'échantillonnage Monte Carlo (MC) ou l'une de ses variantes. L'idée fondamentale de cette méthode est de générer un ensemble de réalisations du vecteur aléatoire  $\xi$  pour construire un ensemble de  $M$  réalisations des données d'entrée  $\{D^1, \dots, D^M\}$  générées selon la loi de probabilité  $p_\xi$ . À chacune de ces réalisations correspond une réalisation déterministe du problème et une solution unique, notée  $U^{(j)} \equiv U(\xi^{(j)})$ , telle que

$$\mathcal{M}(U(\xi^{(j)}), D(\xi^{(j)})) = 0, \quad j \in \{1, \dots, M\}. \quad (1.11)$$

À partir de la collection des  $\{U^1, \dots, U^M\}$ , appelée échantillon de la solution, il est possible d'estimer les statistiques de  $U$ , les corrélations entre les composantes de la solution, des densités de probabilité, etc. Par exemple, l'espérance mathématique de  $U$  peut être estimée par

$$\langle U \rangle \approx \frac{1}{M} \sum_{j=1}^M U^{(j)}. \quad (1.12)$$

L'avantage de la méthode MC est qu'elle est très facilement implémentable puisqu'elle nécessite uniquement l'exécution répétitive de simulations déterministes. Cette méthode est d'autre part très robuste. Néanmoins, l'une de ses principales limitations concerne le faible taux de convergence des statistiques avec le nombre  $M$  de réalisations. Par exemple, l'espérance converge typiquement en  $O(1/\sqrt{M})$ . De nombreuses techniques ont été développées pour accélérer la convergence statistique des estimateurs : Importance Sampling, Variance Reduction, Latin Hyper-Cube (Grigoriu [35], Liu [54], McKay *et al.* [60]), Multi-Level MC (Barth *et al.* [7], Mishra et Schwab [61]), ...

### 1.1.3.3 Méthodes spectrales stochastiques

Les méthodes spectrales stochastiques sont basées sur une approche radicalement différente qui consiste en la construction (ou la reconstruction) de la dépen-

dance fonctionnelle de la solution par rapport au germe. Cette dépendance fonctionnelle est typiquement exprimée par un développement en série du type

$$U(\xi) = \sum_{\alpha} u_{\alpha} \Phi_{\alpha}(\xi), \quad (1.13)$$

où les  $\Phi_{\alpha}$  sont des fonctionnelles des VAs, choisies de manière adéquate, et les  $u_{\alpha}$  sont les coefficients déterministes du développement spectral de la solution. Une fois ce développement en série obtenu, celui-ci peut être directement exploité pour déterminer les statistiques de  $U$ , analytiquement ou via un échantillonnage de  $\xi$ . Une description détaillée de ces méthodes peut être trouvée, entre autres, dans les livres de Ghanem et Spanos [32], Le Maître et Knio [45] et Xiu [87]. Elles font l'objet de la section suivante.

## 1.2 Méthodes spectrales stochastiques

Cette section décrit les différents outils nécessaires à une approche spectrale stochastique : développements spectraux en polynômes de Chaos et Chaos généralisé, calculs statistiques à partir de développements spectraux stochastiques, ainsi que les différentes approches non-intrusives et intrusives pour le calcul des coefficients  $u_{\alpha}$  dans (1.13).

### 1.2.1 Chaos Polynomial de dimension finie

Le développement en Polynômes de Chaos (PC) est basé sur la théorie du Chaos Homogène de Wiener [86] et sur les travaux de Cameron et Martin [10]. Soit  $\{\xi_i\}_{i=1}^N$  un ensemble de  $N$  variables gaussiennes centrées, normalisées et mutuellement orthogonales (ce qui implique qu'elles sont aussi indépendantes), définies sur l'espace de probabilité  $\mathcal{P}$ . Pour tout entier  $N_0 \geq 0$ , soit  $\hat{\Gamma}_{N_0}$  l'espace de tous les polynômes en  $\{\xi_i\}_{i=1}^N$  de degré total au plus égal à  $N_0$ , et pour  $N_0 \geq 1$ ,  $\tilde{\Gamma}_{N_0}$  le supplémentaire orthogonal de  $\hat{\Gamma}_{N_0-1}$  dans  $\hat{\Gamma}_{N_0}$ , *i.e.*

$$\hat{\Gamma}_{N_0} = \hat{\Gamma}_{N_0-1} \oplus \tilde{\Gamma}_{N_0}, \quad (1.14)$$

et  $\mathcal{B}_{N_0}$  une base orthogonale de  $\tilde{\Gamma}_{N_0}$ . Par densité dans  $L^2(\Xi, p_{\xi})$ , on a

$$L^2(\Xi, p_{\xi}) = \bigoplus_{i=0}^{+\infty} \tilde{\Gamma}_i, \quad (1.15)$$

avec la convention  $\tilde{\Gamma}_0 = \hat{\Gamma}_0$ . Le sous-espace  $\tilde{\Gamma}_{N_0}$  est appelé le  $N_0$ -ième Chaos Homogène de dimension  $N$  et la base  $\mathcal{B}_{N_0}$  est appelée Chaos Polynomial d'ordre  $N_0$ . Le sous-espace  $\tilde{\Gamma}_{N_0}$  est constitué de tous les polynômes de degré égal à  $N_0$ , faisant intervenir toutes les combinaisons possibles des variables aléatoires  $\{\xi_i\}_{i=1}^N$ .

Toute variable aléatoire  $U \in L^2(\Xi, p_\xi)$  admet une représentation en fonction des chaos polynomiaux (Cameron et Martin [10]) :

$$\begin{aligned}
U(\xi) &= u_0\Gamma_0 + \sum_{i_1=1}^N u_{i_1}\Gamma_1(\xi_{i_1}) + \sum_{i_1=1}^N \sum_{i_2=1}^{i_1} u_{i_1 i_2}\Gamma_2(\xi_{i_1}, \xi_{i_2}) \\
&+ \sum_{i_1=1}^N \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} u_{i_1 i_2 i_3}\Gamma_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) \\
&+ \sum_{i_1=1}^N \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} \sum_{i_4=1}^{i_3} u_{i_1 i_2 i_3 i_4}\Gamma_4(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}, \xi_{i_4}) + \dots, \tag{1.16}
\end{aligned}$$

où  $\Gamma_{N_0}(\xi)$  s'entend comme un polynôme de la base  $\mathcal{B}_{N_0}$ , chaque polynôme pouvant changer pour chacun des termes des sommations multiples. Par exemple, pour un développement bidimensionnel ( $N = 2$ ), l'équation (1.16) devient

$$\begin{aligned}
U(\xi) &= u_0\Gamma_0 + \sum_{i_1=1}^2 u_{i_1}\Gamma_1(\xi_{i_1}) + \sum_{i_1=1}^2 \sum_{i_2=1}^{i_1} u_{i_1 i_2}\Gamma_2(\xi_{i_1}, \xi_{i_2}) \\
&+ \sum_{i_1=1}^2 \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} u_{i_1 i_2 i_3}\Gamma_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) \\
&+ \sum_{i_1=1}^2 \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} \sum_{i_4=1}^{i_3} u_{i_1 i_2 i_3 i_4}\Gamma_4(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}, \xi_{i_4}) + \dots \tag{1.17}
\end{aligned}$$

ou encore

$$\begin{aligned}
U(\xi) &= u_0\Gamma_0 + u_1\Gamma_1(\xi_1) + u_2\Gamma_1(\xi_2) + u_{11}\Gamma_2(\xi_1, \xi_1) + u_{21}\Gamma_2(\xi_2, \xi_1) + u_{22}\Gamma_2(\xi_2, \xi_2) \\
&+ u_{111}\Gamma_3(\xi_1, \xi_1, \xi_1) + u_{211}\Gamma_3(\xi_2, \xi_1, \xi_1) + u_{221}\Gamma_3(\xi_2, \xi_2, \xi_1) \\
&+ u_{222}\Gamma_3(\xi_2, \xi_2, \xi_2) + u_{1111}\Gamma_4(\xi_1, \xi_1, \xi_1, \xi_1) + \dots \tag{1.18}
\end{aligned}$$

Cette représentation est convergente au sens de la moyenne quadratique (théorème de Cameron–Martin) :

$$\lim_{N_0 \rightarrow +\infty} \left\langle \left( u_0\Gamma_0 + \dots + \sum_{i_1=1}^N \dots \sum_{i_{N_0}=1}^{i_{N_0-1}} u_{i_1, \dots, i_{N_0}}\Gamma_{N_0}(\xi_{i_1}, \dots, \xi_{i_{N_0}}) - U(\xi) \right)^2 \right\rangle = 0. \tag{1.19}$$

Par construction, les chaos polynomiaux d'ordre supérieur à 0 ont tous une espérance nulle et les polynômes sont tous mutuellement orthogonaux vis à vis de la mesure de probabilité gaussienne.

Pour faciliter la manipulation des développements en PC, on définit une relation univoque entre les fonctionnelles  $\Gamma()$  et de nouvelles fonctionnelles  $\Phi()$  dans le but d'obtenir une expression plus compacte du développement de la variable aléatoire :

$$U(\xi) = \sum_{\alpha=1}^{+\infty} u_\alpha \Phi_\alpha(\xi), \quad \xi = \{\xi_1, \dots, \xi_N\}, \tag{1.20}$$

où les coefficients déterministes du développement  $u_\alpha$  sont appelés **modes stochastiques**. On adopte la convention  $\Phi_1 = 1$  comme fonction de base de  $\mathcal{B}_0$  et on suppose de plus que les  $\Phi_\alpha$  sont ordonnées par ordre polynomial croissant.

**Remarque 1.1.** *Classiquement, la sommation commence à l'indice  $\alpha = 0$ . Cependant on choisira de la commencer à  $\alpha = 1$  pour simplifier les notations dans les chapitres suivants.*

### 1.2.1.1 Base PC unidimensionnelle

Une façon simple de construire les PC N-dimensionnels est de procéder par tensorisation partielle de polynômes unidimensionnels. On s'intéresse donc dans un premier temps aux polynômes d'une unique variable aléatoire  $\xi$ . Ces polynômes de chaos sont orthogonaux et la densité de probabilité de  $\xi$  est donnée par

$$p_\xi(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2). \quad (1.21)$$

Soit  $\Phi_\alpha(\xi)$  le polynôme de chaos unidimensionnel de degré  $\alpha$ . Le polynôme de degré 0 est  $\Phi_1 = 1$ . La condition d'orthogonalité s'exprime par

$$\langle \Phi_i, \Phi_j \rangle = \delta_{ij} \langle \Phi_i^2 \rangle. \quad (1.22)$$

Conventionnellement, on adopte la normalisation  $\langle \Phi_\alpha^2 \rangle = (\alpha - 1)!$ . Les polynômes unidimensionnels ainsi définis sont orthogonaux pour la mesure gaussienne et constituent la famille bien connue des polynômes de Hermite (*cf* par exemple le livre de Abramowitz et Stegun [2]). Dans la figure 1.2, les polynômes de Hermite  $\Phi_\alpha(\xi)$  sont tracés pour  $\xi \in [-3, 3]$  et  $\alpha \in \{1, \dots, 5\}$ .

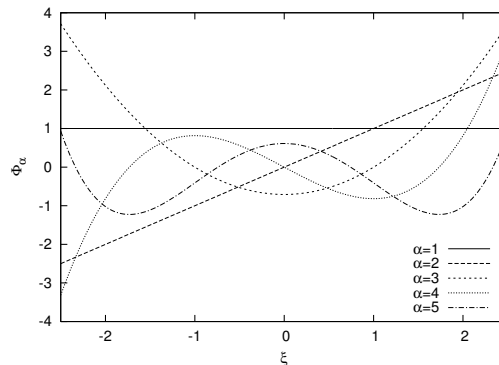


FIGURE 1.2 – Polynômes unidimensionnels de Hermite pour  $\alpha \in \{1, \dots, 5\}$ .

### 1.2.1.2 Base PC multidimensionnelle

On procède maintenant au cas N-dimensionnel et on cherche à construire  $\tilde{\Gamma}_{\text{No}}$  à partir des polynômes unidimensionnels de Hermite  $\{H_\alpha\}_{\alpha \in \mathbb{N}}$ . Puisque les VAs sont



indépendantes, la densité de probabilité de  $\xi = \{\xi_1, \dots, \xi_N\}$  est donnée par

$$p_\xi(y) = \prod_{i=1}^N p_{\xi_i}(y_i), \quad (1.23)$$

où  $p_{\xi_i}(y_i)$  est la densité de probabilité commune à tous les  $\xi_i$  donnée par (1.21). Soit  $\lambda(\text{No})$  l'ensemble de multi-indices défini par

$$\lambda(\text{No}) = \left\{ \gamma = \{\gamma_1, \dots, \gamma_N\} \in \mathbb{N}^N \mid \sum_{i=1}^N \gamma_i = \text{No} \right\}. \quad (1.24)$$

Avec ces définitions, on construit le No-ième Chaos Homogène de dimension N selon

$$\tilde{\Gamma}_{\text{No}} = \text{span}_{\gamma \in \lambda(\text{No})} \left\{ \prod_{i=1}^N H_{\gamma_i}(\xi_i) \right\}. \quad (1.25)$$

Ainsi, pour le cas 2D, le développement de Hermite a pour expression

$$\begin{aligned} U(\xi) &= u_0 H_0 + u_1 H_1(\xi_1) + u_2 H_1(\xi_2) \\ &+ u_{11} H_2(\xi_1) + u_{21} H_1(\xi_2) H_1(\xi_1) + u_{22} H_2(\xi_2) \\ &+ u_{111} H_3(\xi_1) + u_{211} H_1(\xi_2) H_2(\xi_1) + u_{221} H_2(\xi_2) H_1(\xi_1) \\ &+ u_{222} H_3(\xi_2) + u_{1111} H_4(\xi_1) + \dots \end{aligned} \quad (1.26)$$

L'expression ci-dessus est réécrite sous la forme plus compacte :

$$U(\xi) = \sum_{\alpha=1}^{+\infty} u_\alpha \Phi_\alpha(\xi_1, \xi_2). \quad (1.27)$$

Les premiers polynômes de chaos bidimensionnels sont tracés dans la figure 1.3.

**Remarque 1.2.** *L'ensemble de multi-indices (1.24) correspond à la tensorisation partielle de la base PC unidimensionnelle, classiquement utilisée. Dans ce travail, on aura également recours à la tensorisation complète de la base PC, pour laquelle l'ensemble de multi-indices est*

$$\Lambda(\text{No}) = \left\{ \gamma = \{\gamma_1, \dots, \gamma_N\} \in \mathbb{N}^N \mid \gamma_i \leq \text{No}, \forall i \in \{1, \dots, N\} \right\}. \quad (1.28)$$

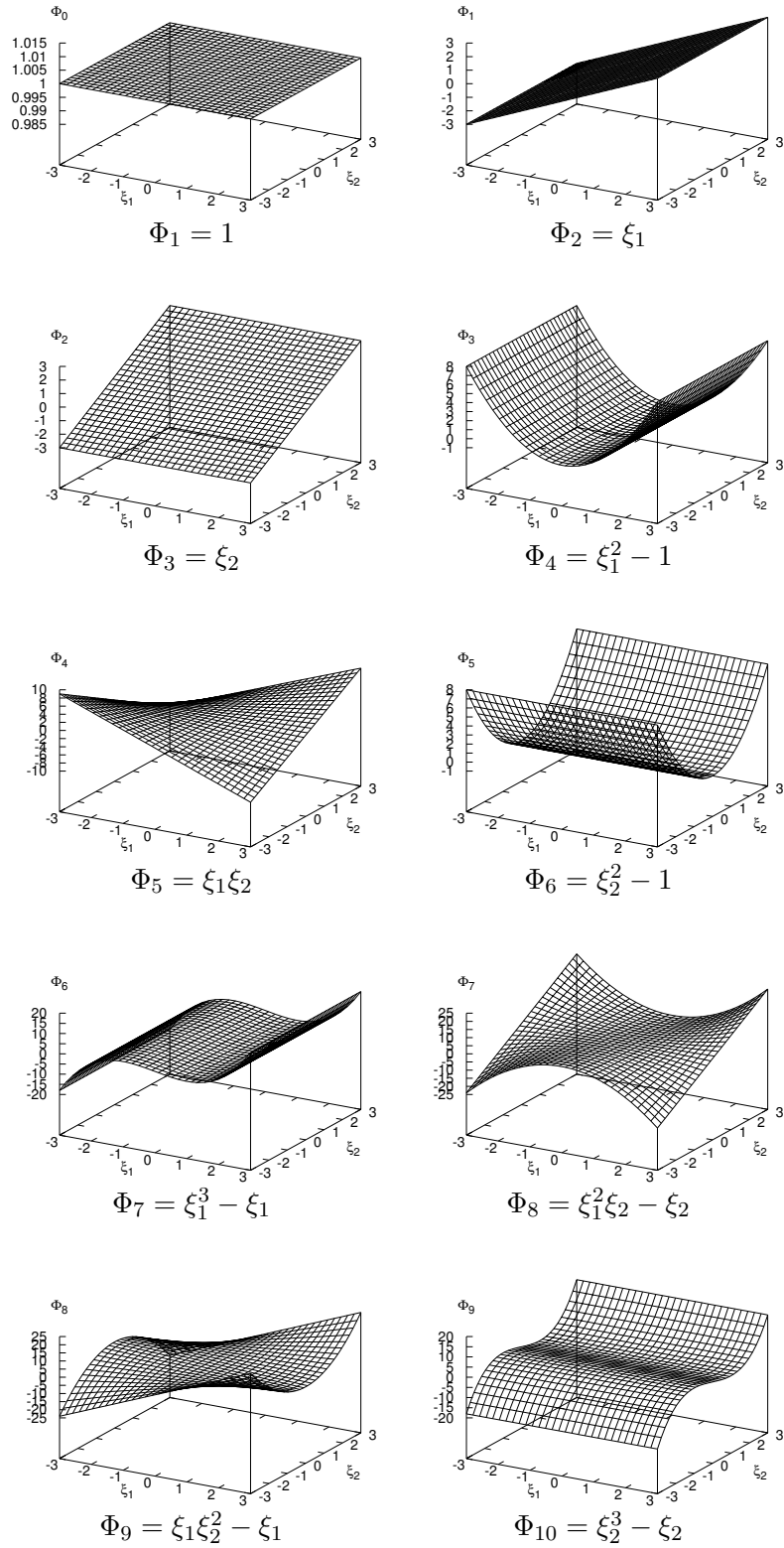
### 1.2.1.3 Développement en PC tronqué

On adopte la notation condensée du développement en PC de la variable aléatoire  $U$  :

$$U(\xi) = \sum_{\alpha=1}^{+\infty} u_\alpha \Phi_\alpha(\xi). \quad (1.29)$$

Il faut à présent tronquer le développement à un ordre polynomial No pour que le développement soit fini. Le nombre total de termes dans la série, après la troncature du développement à l'ordre No, est donné par (Ghanem et Spanos [32])

$$P = \frac{(N + \text{No})!}{N! \text{No}!}. \quad (1.30)$$

FIGURE 1.3 – Polynômes de chaos bidimensionnels ( $N = 2$ ) d'ordre 0, 1, 2 et 3.

Le développement en PC tronqué peut par conséquent être exprimé par

$$U(\xi) = \sum_{\alpha=1}^P u_{\alpha} \Phi_{\alpha}(\xi) + \epsilon_{N_o}(\xi), \quad (1.31)$$

où l'erreur de troncature  $\epsilon_{N_o}$  dépend de  $N_o$  et est elle-même une variable aléatoire. De par (1.19), le développement tronqué converge au sens de la moyenne quadratique pour  $N_o$  tendant vers l'infini :

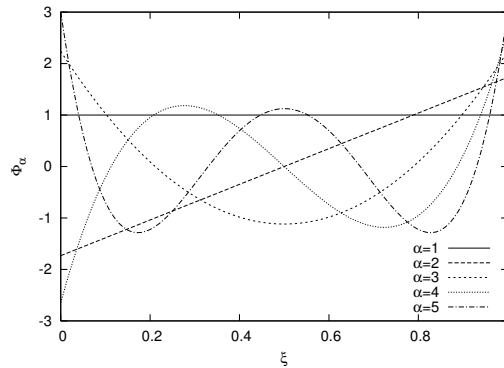
$$\lim_{N_o \rightarrow +\infty} \langle \epsilon_{N_o}^2 \rangle = 0. \quad (1.32)$$

Compte tenu de la dépendance du nombre de termes  $P$  intervenant dans le développement en PC de  $U$  avec  $N$  et  $N_o$ , la représentation PC sera efficace d'un point de vue numérique si des valeurs modérées de  $N_o$  sont suffisantes pour approcher précisément  $U$  avec un germe de taille  $N$  pas trop grande.

#### 1.2.1.4 Chaos Polynomial généralisé (gPC)

On peut se demander s'il n'existe pas d'autres familles de polynômes orthogonaux conduisant à une erreur de représentation plus faible pour le même nombre de termes dans le développement en PC. On peut ainsi remarquer que si  $U$  est une variable aléatoire gaussienne, alors la base PC de Hermite est optimale, puisque le développement avec  $N_o = 1$  fournit une représentation exacte. Cette remarque peut être généralisée : le développement polynomial optimal sera celui construit en utilisant la mesure correspondant à la loi de probabilité de la variable aléatoire que l'on cherche à représenter. Cependant, la loi de probabilité de  $U$  n'est généralement pas connue *a priori*, c'est pourquoi on choisit classiquement de générer la base PC à partir de la mesure associée aux données incertaines que l'on veut propager dans le modèle.

Suivant cette idée, si l'incertitude d'entrée correspond à une variable aléatoire uniformément distribuée sur un intervalle fini donné, on choisira la base générée par des variables aléatoires uniformément distribuées sur  $[0, 1]$ , ce qui conduit aux polynômes de Legendre, définis par convention par rapport à l'intervalle de référence  $[0, 1]$  (*cf* par exemple le livre de Abramowitz et Stegun [2]). Les cinq premiers polynômes de Legendre sont tracés sur la figure 1.4. Plus généralement, Xiu et Karniadakis [89] ont montré que pour de nombreuses lois de probabilité, les familles de polynômes correspondantes, appelées Chaos Polynomiaux généralisé (gPC), sont déterminées par le schéma de Askey [5]. La correspondance entre les principales lois de probabilité (mesure) et leur famille de polynômes orthogonaux est donnée dans le Tableau 1.1. Dans le cas de mesures pour lesquelles on ne dispose pas d'une famille orthogonale de polynômes, il est en général possible de recourir à une construction numérique de la base PC selon une procédure d'orthogonalisation de Gram–Schmidt (*cf* par exemple Stoer et Bulirsch [79]). Enfin, on notera que l'expression des PC multidimensionnels en termes de produits de polynômes unidimensionnels présente l'avantage de s'étendre naturellement au cas où les variables aléatoires  $\xi_i$  sont indépendantes et associées à des mesures différentes.

FIGURE 1.4 – Polynômes unidimensionnels de Legendre pour  $\alpha \in \{1, \dots, 5\}$ .**Tableau 1.1** Familles de loi de probabilité et familles de polynômes orthogonaux construits pour les mesures correspondantes.

	Distribution $\xi$	Polynômes $\Phi_\alpha(\xi)$	Support
VA continue	Gaussienne	Hermite	$(-\infty, +\infty)$
	$\gamma$	Laguerre	$[0, +\infty)$
	$\beta$	Jacobi	$[a, b]$
	Uniforme	Legendre	$[a, b]$
VA discrète	Poisson	Charlier	$\{0, 1, 2, \dots\}$
	Binomiale	Krawtchouk	$\{0, 1, 2, \dots, n\}$
	Binomiale négative	Meixner	$\{0, 1, 2, \dots\}$
	Hypergéométrique	Hahn	$\{0, 1, 2, \dots, n\}$

Une remarque importante est que le développement en PC usuel (Hermite) et le développement en PC généralisé sont basés sur des polynômes globaux. Or, cette représentation polynomiale continue s'avère inappropriée dans certaines situations, notamment lorsque la variable aléatoire  $U$  ne dépend pas continûment de  $\xi$ . Pour surmonter cette difficulté, des méthodes utilisant des éléments finis stochastiques (Deb *et al.* [16]), des multi-éléments gPC (Wan et Karniadakis [85]) et des développements en Multi-Ondelettes (Le Maître *et al.* [47, 48, 49]) peuvent être utilisées afin de rendre le développement spectral plus local grâce à la décomposition du domaine stochastique en différentes régions ou différentes échelles. On utilisera ce type de représentation pour traiter les problèmes hyperboliques stochastiques dès le chapitre 2.

### 1.2.2 Application à la propagation d'incertitudes

On revient à présent à la propagation *via* le modèle mathématique des incertitudes sur les données d'entrée  $D$ , de loi de probabilité prescrite et supposées paramétrées à l'aide d'un nombre fini  $N$  de variables aléatoires indépendantes  $\xi = \{\xi_1, \dots, \xi_N\}$  de loi de probabilité  $p_\xi$ . On considère les polynômes  $\Phi_\alpha$   $N$ -

dimensionnels orthogonaux pour  $p_\xi$  et on cherche à déterminer le développement de la solution stochastique induite par les données incertaines  $D$  sur le sous-espace

$$S^P := \text{span}_{\alpha \in \{1, \dots, P\}} \{\Phi_\alpha\} \subset L^2(\Xi, p_\xi), \quad (1.33)$$

appelé **espace d'approximation stochastique**, soit

$$U(\xi) \approx U^P(\xi) = \sum_{\alpha=1}^P u_\alpha \Phi_\alpha(\xi). \quad (1.34)$$

On choisit donc la même base pour  $U$  que celle utilisée pour paramétrer  $D$ . Les coefficients  $u_\alpha$  sont appelés les **modes stochastiques** de  $U$ . Lorsque  $U$  est connue, un choix naturel est  $u_\alpha = \langle U, \Phi_\alpha \rangle / \langle \Phi_\alpha^2 \rangle$ , si bien que  $U^P$  est la projection  $L^2(\Xi, p_\xi)$ -orthogonale de  $U$  sur l'espace d'approximation stochastique  $S^P$ . Le facteur de normalisation  $\langle \Phi_\alpha^2 \rangle$  dépend de la base utilisée et est connu analytiquement pour les bases PC. Dans le cas général, on dispose de deux grandes classes de méthodes pour déterminer les modes stochastiques : les méthodes non-intrusives et les méthodes de Galerkin, détaillées respectivement dans les sections 1.2.3 et 1.2.4.

Le développement spectral stochastique d'une quantité aléatoire fournit une représentation pratique en vue de sa caractérisation par des quantités statistiques. Ces informations peuvent être obtenues avec un coût de calcul faible pourvu que les modes stochastiques  $u_\alpha$  de la quantité d'intérêt soient connus. Quelques exemples sont donnés ci-dessous.

### 1.2.2.1 Variables aléatoires

Soit  $U \in L^2(\Xi, p_\xi)$  et

$$U^P(\xi) = \sum_{\alpha=1}^P u_\alpha \Phi_\alpha(\xi), \quad (1.35)$$

son approximation sur la base PC orthogonale tronquée  $\{\Phi_1, \dots, \Phi_P\}$  (on rappelle que  $\langle \Phi_\alpha, \Phi_\beta \rangle = 0, \alpha \neq \beta$ , et  $\Phi_1 = 1$ ). L'espérance de  $U^P$  est donnée par

$$\langle U^P \rangle = \langle \Phi_1, U^P \rangle = \sum_{\alpha=1}^P u_\alpha \langle \Phi_1, \Phi_\alpha \rangle = u_1, \quad (1.36)$$

en raison de l'orthogonalité de la base. Le coefficient  $u_1$  est par conséquent l'espérance de la variable aléatoire  $U^P$ . De plus, en utilisant la définition de la variance,  $\sigma^2(U^P)$ , on obtient

$$\sigma^2(U^P) = \left\langle (U^P - \langle U^P \rangle)^2 \right\rangle = \left\langle \left( \sum_{\alpha=2}^P u_\alpha \Phi_\alpha \right)^2 \right\rangle \quad (1.37)$$

$$= \sum_{\alpha, \beta=2}^P u_\alpha u_\beta \langle \Phi_\alpha, \Phi_\beta \rangle = \sum_{\alpha=2}^P u_\alpha^2 \langle \Phi_\alpha^2 \rangle. \quad (1.38)$$

Des expressions similaires peuvent être obtenues pour les moments d'ordre plus élevé de  $U^P$  en fonction de ses modes stochastiques. De plus, les statistiques d'une variable aléatoire peuvent être estimées à partir de stratégies d'échantillonnage : des réalisations  $U^P(\xi)$  peuvent être obtenues en échantillonnant  $\xi$  suivant sa densité  $p_\xi$  et en évaluant le développement en PC aux points  $\xi$  de l'échantillon. On peut également estimer des densités de probabilité, des fonctions de répartition, des probabilités d'événements, réaliser une analyse de sensibilité locale ou globale, etc... Il est intéressant de remarquer que dans le contexte de l'analyse d'incertitudes, la connaissance du développement en PC de la solution stochastique permet d'accéder plus simplement et plus efficacement à des quantités statistiques que dans le cas des méthodes MC.

### 1.2.2.2 Vecteurs aléatoires

Le développement en PC d'une variable aléatoire peut immédiatement être généralisé à la représentation de vecteurs aléatoires  $U$  du second ordre à valeurs dans  $\mathbb{R}^m$  ( $U : \Xi \mapsto \mathbb{R}^m$ ). Soit  $U_i$  la  $i$ -ème composante d'un vecteur aléatoire. Le développement tronqué en PC s'écrit sous la forme

$$U_i^P(\xi) = \sum_{\alpha=1}^P (u_i)_\alpha \Phi_\alpha(\xi), \quad \forall i \in \{1, \dots, m\}, \quad (1.39)$$

ou encore, sous forme vectorielle,

$$U^P(\xi) = \sum_{\alpha=1}^P u_\alpha \Phi_\alpha(\xi), \quad (1.40)$$

où le vecteur  $u_\alpha = ((u_1)_\alpha, \dots, (u_m)_\alpha)^T \in \mathbb{R}^m$  est le  $\alpha$ -ème mode stochastique de  $U^P$ . À nouveau,  $u_1$  est l'espérance du vecteur aléatoire  $U^P$ . De plus, les matrices de corrélation et de covariance de  $U^P$  ont pour expressions

$$r_{U^P} = \sum_{\alpha=1}^P u_\alpha u_\alpha^T \langle \Phi_\alpha^2 \rangle, \quad c_{U^P} = \sum_{\alpha=2}^P u_\alpha u_\alpha^T \langle \Phi_\alpha^2 \rangle. \quad (1.41)$$

### 1.2.2.3 Processus stochastiques

Le développement en PC peut immédiatement être généralisé aux processus stochastiques  $U : \Omega \times \Xi \mapsto \mathbb{R}$ , où  $\Omega$  est un domaine de l'espace physique, en faisant dépendre les modes stochastiques  $u_\alpha$  de l'indice  $x \in \Omega$ . Ainsi,

$$U^P(x, \xi) = \sum_{\alpha=1}^P u_\alpha(x) \Phi_\alpha(\xi). \quad (1.42)$$

La fonction déterministe  $u_\alpha(x)$  est le  $\alpha$ -ème mode stochastique du processus  $U^P$ . On a  $u_1(x) = \langle U^P(x, \cdot) \rangle$ . De plus, la fonction de corrélation de  $U^P$  est donnée par

$$R_{U^P}(x, x') = \langle U^P(x) U^P(x') \rangle = \sum_{\alpha=1}^P u_\alpha(x) u_\alpha(x') \langle \Phi_\alpha^2 \rangle, \quad (1.43)$$

et la fonction de covariance est donnée de façon similaire par

$$C_{UP}(x, x') = \sum_{\alpha=2}^P u_{\alpha}(x)u_{\alpha}(x') \langle \Phi_{\alpha}^2 \rangle. \quad (1.44)$$

### 1.2.3 Méthodes non-intrusives

On appelle méthodes non-intrusives les méthodes basées sur un ensemble de résolutions du modèle déterministe, correspondant à des réalisations de  $\xi$ , afin d'approcher les modes stochastiques de la solution. Le code de calcul du modèle déterministe peut ainsi être utilisé en boîte noire, en associant à chaque réalisation des paramètres la solution du modèle correspondant. L'avantage principal de ces méthodes est que les codes existants ne nécessitent aucune adaptation particulière. De plus, il est en général possible de planifier les simulations de modèles déterministes, de sorte que l'exécution peut être menée en parallèle. Ces caractéristiques rendent les méthodes non-intrusives très attrayantes pour la propagation d'incertitudes paramétriques dans les modèles complexes, les applications industrielles et les situations où l'on a uniquement accès à des codes déterministes.

Le coût numérique des méthodes non-intrusives est essentiellement proportionnel au nombre de résolutions du modèle déterministe nécessaires pour construire l'approximation. Ce nombre peut s'avérer très important, en particulier lorsque le nombre  $N$  de variables aléatoires dans la paramétrisation est grand (problème de la dimensionnalité). Ce problème se pose aussi pour les méthodes de Galerkin (intrusives) décrites à la section 1.2.4. Par conséquent, de nombreux travaux de recherche visent à réduire la complexité de ces méthodes. On dresse ci-après un bref panorama des différentes stratégies non-intrusives disponibles dans la littérature. Pour simplifier, on se restreint au cas de variables aléatoires à valeurs scalaires ; l'extension à des cas plus généraux (vecteurs et processus) ne pose pas de difficultés.

#### 1.2.3.1 Projection Spectrale Non-Intrusive

La projection spectrale non-intrusive (NISP, de l'anglais "Non Intrusive Spectral Projection") consiste à chercher  $U^P$  comme la projection orthogonale de  $U$  sur  $S^P$  et à calculer les modes stochastiques de  $U^P$  par intégration numérique. Grâce à l'orthogonalité de la base PC, les modes stochastiques sont donnés par

$$u_{\alpha} = \frac{\langle U, \Phi_{\alpha} \rangle}{\langle \Phi_{\alpha}^2 \rangle}, \quad \forall \alpha \in \{1, \dots, P\}. \quad (1.45)$$

Différentes techniques ont été proposées pour estimer numériquement l'intégrale  $\langle U, \Phi_{\alpha} \rangle$ , basées sur des stratégies d'échantillonnage pseudo-aléatoires ou des méthodes de cubature.

La méthode Monte Carlo (MC) est la technique de simulation la plus simple : un échantillon de réalisations indépendantes de  $\xi$  est généré à partir de  $p_{\xi}$  en utilisant un générateur pseudo-aléatoire. Soit  $\xi^j$  le  $j$ -ème élément de l'échantillon et  $U^{(j)} \equiv$

$U(\xi^{(j)})$  la solution du modèle correspondante. Le  $\alpha$ -ème mode stochastique de  $U^P$  a pour expression

$$u_\alpha = \frac{1}{M} \sum_{j=1}^M U^{(j)} \Phi_\alpha(\xi^{(j)}) + \epsilon_M, \quad (1.46)$$

où  $M$  est la dimension de l'échantillon et  $\epsilon_M$  l'erreur d'échantillonnage. Cet estimateur est non biaisé ( $\langle \epsilon_M \rangle = 0$ ) et la variance de  $\epsilon_M$  tend vers zéro asymptotiquement en  $O(1/\sqrt{M})$  pour  $M$  suffisamment grand. Ce faible taux de convergence, bien qu'indépendant de  $N$ , reste la limitation principale de l'utilisation de MC pour NISP. Des erreurs d'échantillonnage plus faibles, à  $M$  fixé, peuvent néanmoins être atteintes en utilisant des stratégies plus efficaces, comme l'échantillonnage Latin Hypercube (LHS) (McKay *et al.* [60]) ou Quasi Monte Carlo (QMC) (Morokoff et Caflisch [62]). L'idée de base, dans le cas des mesures uniformes, est de forcer l'échantillonneur à choisir des points qui couvrent le domaine stochastique de manière plus uniforme que pour MC standard.

Les méthodes de cubature constituent une alternative aux méthodes MC. La définition des modes stochastiques (1.45) montre que leur évaluation équivaut au calcul de l'intégrale multidimensionnelle de la fonction  $F_\alpha(\xi) := U(\xi)\Phi_\alpha(\xi)$  sur  $\Xi \subset \mathbb{R}^N$  avec un poids non négatif  $p_\xi$  sous forme produit. En fait, on doit calculer pour chaque  $\alpha \in \{1, \dots, P\}$  l'intégrale

$$\mathcal{I}^{(N)} F_\alpha = \int_{\Xi} F_\alpha(y) p_\xi(y) dy. \quad (1.47)$$

Une cubature est une approximation de l'intégrale multidimensionnelle  $\mathcal{I}^{(N)} F_\alpha$  en une somme discrète,

$$\mathcal{I}^{(N)} F_\alpha \approx \mathcal{Q}^{(N)} F_\alpha = \sum_{j=1}^{N_Q} F_\alpha(\xi^j) w^j, \quad (1.48)$$

où  $\xi^j \in \Xi$  et  $w^j \in \mathbb{R}$ ,  $j \in \{1, \dots, N_Q\}$ , sont respectivement les nœuds et les poids de la cubature à  $N_Q$  nœuds. Classiquement, les cubatures sont construites à partir de formules de quadrature unidimensionnelles. Soient  $\mathcal{I}^{(1)} F_\alpha$  l'intégrale unidimensionnelle et  $\mathcal{Q}^{(1)} F_\alpha$  son approximation par quadrature :

$$\mathcal{I}^{(1)} F_\alpha = \int_a^b F_\alpha(y) p(y) dy \approx \mathcal{Q}^{(1)} F_\alpha \equiv \sum_{j=1}^n F_\alpha(\xi^j) w^j, \quad (1.49)$$

où les bornes de l'intégrale  $a$  et  $b$  peuvent être infinies. Les formules de quadrature de Gauss à  $n$  points offrent un degré d'exactitude optimal pour les intégrales unidimensionnelles puisqu'elles intègrent exactement des fonctions polynomiales de degré  $\leq 2n - 1$  (Stoer et Bulirsch [79]). Cependant, elles présentent l'inconvénient d'avoir une distribution des nœuds qui dépend de  $n$ . De ce fait, si l'on souhaite augmenter progressivement le nombre de nœuds, on adopte souvent des formules de quadrature imbriquées, où les nœuds d'une quadrature font partie des nœuds des formules d'ordres supérieurs. On peut citer les formules de quadrature de Newton–Cotes (Stoer



et Bulirsch [79]), de Clenshaw–Curtis [11] et de Fejér (Waldvogel [84]). Les cubatures  $N$ -dimensionnelles peuvent être construites par tensorisation de cubatures unidimensionnelles. Par exemple, si les paramètres aléatoires sont identiquement distribués, de sorte que la même formule de quadrature peut être utilisée dans toutes les directions, on obtient

$$\mathcal{I}^{(N)} F_\alpha \approx \mathcal{Q}^{(N)} F_\alpha = \left( \mathcal{Q}^{(1)} \otimes \dots \otimes \mathcal{Q}^{(1)} \right) F_\alpha. \quad (1.50)$$

Le nombre total de points de cubature vaut alors  $N_Q = n^N$  et augmente de façon exponentielle avec  $N$ . Ce problème de la dimensionnalité peut être tempéré par l’usage de formules de quadrature partiellement tensorisées, qui permettent de réduire considérablement le nombre de nœuds de cubature nécessaires pour un niveau de précision donné. Ces formules de quadrature sont basées sur des méthodes de tensorisation partielle, dont la première méthode de construction fut proposée par Smolyak [75]. Elle consiste en un algorithme général basé sur la construction de produits tensorisés creux. Son application à NISP permet une réduction considérable des coûts du calcul, en comparaison aux formules de cubature sous forme produit (*cf* par exemple Keese et Matthies [41] et Reagan *et al.* [72]). Néanmoins, cette approche reste encore trop coûteuse pour les problèmes stochastiques faisant intervenir un grand nombre de variables aléatoires ; une piste actuellement explorée est l’emploi de grilles creuses adaptatives (*cf* par exemple Gerstner et Griebel [31] et Crestaux [14]).

### 1.2.3.2 Moindres carrés

L’estimation des coefficients du développement de la solution d’un modèle peut être formulée alternativement par la résolution d’un problème de minimisation au sens des moindres carrés. L’estimation au sens des moindres carrés est un outil fréquemment utilisé en statistiques où l’on cherche à déterminer les paramètres d’un modèle à partir d’un ensemble de mesures. La différence est que dans ce cas, les  $P$  coefficients du développement sont déterminés à partir d’un ensemble de mesures (calculs ou observations). Soit  $\{(\xi^j, U^j)\}_{j \in \{1, \dots, N_{\text{obs}}\}}$  l’ensemble des observations, où  $U^j \equiv U(\xi^j)$  est la solution du modèle pour les paramètres  $D(\xi^j)$  donnés par  $\xi^j$ . Les modes stochastiques  $(u_\alpha)_{\alpha \in \{1, \dots, P\}}$  peuvent être approchés par la solution d’un problème d’optimisation pour la somme des carrés des résidus

$$R((u_\alpha)_{\alpha \in \{1, \dots, P\}}) \equiv \sum_{j=1}^{N_{\text{obs}}} (R^j)^2 = \sum_{i=1}^{N_{\text{obs}}} \left( U^j - U^P(\xi^j) \right)^2, \quad (1.51)$$

où les résidus  $R^j$  sont simplement les distances entre les observations et les prédictions du modèle de substitution donné par l’expression de  $U^P$ . Les modes stochastiques approchés  $u_\alpha$  sont obtenus par minimisation de la fonctionnelle  $R$ . Se pose alors la question de la sélection des points de minimisation  $\xi^j$ . Dans cette optique, les méthodes de plan d’expériences et de construction optimale sont des

domaines de recherche qui se concentrent sur la construction d'échantillons fournissant des propriétés optimales pour les problèmes aux moindres carrés (*cf* par exemple Pukelsheim [71] et Berveiller *et al.* [9]).

### 1.2.3.3 Méthodes de collocation

Contrairement aux deux méthodes évoquées ci-dessus (projection et moindres carrés), les méthodes de collocation sont basées sur l'interpolation de la solution stochastique. En particulier, on cherche une approximation  $\tilde{U}(\xi)$  de  $U(\xi)$  telle que

$$\tilde{U}(\xi^j) = U(\xi^j), \quad \forall j \in \{1, \dots, N_L\}, \quad (1.52)$$

pour un ensemble de  $N_L$  points d'interpolation  $\xi^j$ . Par définition, l'approximation est exacte aux  $N_L$  points de collocation. On peut alors construire une approximation de la solution stochastique dans un espace vectoriel  $S^{N_L}$  de dimension  $N_L$  qui satisfait les  $N_L$  contraintes (1.52). Cette approximation peut se développer sur une base  $\{\tilde{\Phi}_j\}_{j \in \{1, \dots, N_L\}}$  sous la forme

$$\tilde{U}(\xi) = \sum_{j=1}^{N_L} \tilde{u}_j \tilde{\Phi}_j(\xi), \quad (1.53)$$

où les coefficients du développement  $\tilde{u}_j$  sont déterminés pour satisfaire la contrainte (1.52). Usuellement, on utilise des fonctions de base qui vérifient

$$\tilde{\Phi}_j(\xi^{j'}) = \begin{cases} 1, & \text{si } j = j', \\ 0, & \text{sinon,} \end{cases} \quad \forall j, j' \in \{1, \dots, N_L\}, \quad (1.54)$$

de sorte que

$$\tilde{u}_j = U(\xi^j), \quad \forall j \in \{1, \dots, N_L\}. \quad (1.55)$$

Le choix le plus courant consiste en l'interpolation polynomiale spectrale où le support des fonctions  $\tilde{\Phi}_j$  est le domaine entier  $\Xi$ . En une dimension stochastique, cela revient à prendre pour fonctions de base les  $N_L$  polynômes de Lagrange unidimensionnel définis par

$$\mathcal{L}_j(\xi) = \prod_{j'=1, j' \neq j}^{N_L} \frac{\xi - \xi^{j'}}{\xi^j - \xi^{j'}}, \quad \forall j \in \{1, \dots, N_L\}. \quad (1.56)$$

Le choix des points de collocation s'avère très important pour la qualité de l'approximation et de ses propriétés de convergence lorsque  $N_L$  augmente (*cf* par exemple Agarwal et Wong [3]). En particulier, l'usage des points de Gauss et de Chebychev est préféré à celui de points équidistants car il offre de meilleures propriétés d'interpolation. Les points de Chebychev présentent également l'avantage d'être des ensembles imbriqués (en multipliant le nombre de points par un entier pour passer d'un ensemble de points à l'autre), caractéristique intéressante pour l'interpolation adaptative. L'extension des formules d'interpolation unidimensionnelles

au cas N-dimensionnel est immédiate en utilisant les grilles tensorisées, de façon similaire aux formules d'intégration. Pour cette construction, la grille de points d'interpolation et les polynômes d'interpolation sont tous les deux des produits des objets unidimensionnels associés. La complexité de ces méthodes, coûteuse car proportionnelle au nombre de points d'interpolation, peut à nouveau être réduite en utilisant des méthodes creuses. En fait, les techniques de grilles creuses, telles l'algorithme de Smolyak, peuvent être réutilisées pour construire des méthodes d'interpolation creuse et également des schémas d'interpolation adaptatifs. Des exemples de méthodes de collocation à grilles creuses pour les problèmes stochastiques peuvent être trouvés dans les travaux de Mathelin et Hussaini [58], Xiu et Hesthaven [88], Babuska *et al.* [6], Nobile *et al.* [68]. Des techniques adaptatives ont été récemment considérées par Nobile *et al.* [67], Ganapathysubramanian et Zabaras [29] et Ma et Zabaras [57].

#### 1.2.4 Méthodes de Galerkin (intrusives)

Contrairement aux méthodes non-intrusives basées sur des réalisations individuelles pour déterminer la réponse du modèle stochastique aux entrées aléatoires, les méthodes de Galerkin sont basées sur un formalisme de résidus pour construire des systèmes d'équations gouvernant les modes stochastiques  $\{u_\alpha\}_{\alpha \in \{1, \dots, P\}}$ . De telles méthodes sont dites intrusives puisqu'elles nécessitent dans une certaine mesure la réécriture du code de calcul déterministe. La projection de Galerkin est un outil classique pour la résolution des problèmes spectraux et dans la formulation des méthodes d'éléments finis (voir par exemple le livre de Ern et Guermond [28]). Dans le contexte stochastique, elle fut proposée par Ghanem et Spanos [32] comme méthode de calcul pour déterminer le développement en PC de la solution d'équations linéaires stochastiques.

On se place dans le cadre général d'un problème de propagation d'incertitudes décrit par la relation (1.7). On rappelle qu'on a fait l'hypothèse que ce problème est bien posé, en d'autres termes, que la loi de probabilité des données est telle que le problème mathématique  $\mathcal{M}(U(\xi), D(\xi)) = 0$  a presque sûrement une unique solution dans un espace fonctionnel adéquat que l'on note  $V$ , supposé indépendant de l'événement aléatoire. En notant  $S \equiv L^2(\Xi, p_\xi)$  l'espace stochastique de second ordre, l'espace fonctionnel dans lequel chercher la solution stochastique  $U$  de (1.7) est alors l'espace produit  $V \otimes S$ .

La solution stochastique  $U$  est approchée dans un sous-espace  $V \otimes S^P$  sous forme d'un développement stochastique en PC  $U^P(\xi)$  donné par (1.34). La projection de Galerkin nécessite deux étapes : la première est l'introduction du développement tronqué de la solution dans le problème stochastique, la deuxième est la projection de l'équation stochastique ainsi obtenue sur la base du développement en PC afin d'obtenir un ensemble d'équations que les modes stochastiques doivent satisfaire. On appellera  $U^P$  la **solution de Galerkin** du système couplant les modes stochas-

tiques. Celle-ci est obtenue en introduisant (1.34) dans le problème stochastique :

$$\mathcal{M} \left( \sum_{\alpha=1}^P u_{\alpha} \Phi_{\alpha}(\xi), D(\xi) \right) \equiv \mathcal{R}^P((u_{\alpha})_{\alpha \in \{1, \dots, P\}}). \quad (1.57)$$

En général, le terme de droite de (1.57) ne vaut pas zéro, sauf si  $P \rightarrow +\infty$ . En fait, pourvu que le résidu aléatoire  $\mathcal{R}^P$  soit une quantité aléatoire du second-ordre, celui-ci peut être développé formellement sous la forme

$$\mathcal{R}^P = \sum_{\alpha=1}^{+\infty} r_{\alpha} \Phi_{\alpha}(\xi). \quad (1.58)$$

La formulation faible du problème consiste à trouver les modes stochastiques  $u_{\alpha} \in V$ ,  $\alpha \in \{1, \dots, P\}$ , tels que

$$\left\langle \mathcal{M} \left( \sum_{\beta=1}^P u_{\beta} \Phi_{\beta}(\xi), D(\xi) \right), \Phi_{\alpha}(\xi) \right\rangle = 0, \quad \forall \alpha \in \{1, \dots, P\}. \quad (1.59)$$

La construction du problème spectral (1.59) conduit à diverses difficultés que l'on rencontre classiquement lorsque l'on applique la méthode de Galerkin à des modèles mathématiques complexes. La première difficulté est relative à la dimension du problème spectral : la forme produit de l'espace de discrétisation de la solution  $V \otimes S^P$  montre que, comparé au problème déterministe, la taille de l'ensemble des équations à résoudre va être  $P$  fois plus grand que pour le problème déterministe. De plus, le problème (1.59) peut exhiber une structure différente de son équivalent déterministe, ce qui implique qu'on peut avoir besoin d'une nouvelle classe de solveurs ou de codes numériques pour le discrétiser. Par conséquent, les bénéfices de l'approche intrusive vont dépendre de compromis entre les coûts de calcul supplémentaires associés à la résolution numérique de (1.59) et des gains en précision dus à la convergence spectrale de la représentation en PC. Ainsi, le développement de stratégies numériques efficaces est une préoccupation centrale des méthodes de Galerkin. Différentes approches peuvent être envisagées pour maintenir un coût de calcul aussi faible que possible. Certaines de ces approches sont développées par exemple dans le livre de Le Maître et Knio [45]. Une seconde difficulté résultant de la méthode de Galerkin concerne le traitement des non-linéarités, difficulté que l'on rencontrera dans le cas des systèmes hyperboliques et sur laquelle on reviendra.

Les méthodes de Galerkin appliquées aux problèmes elliptiques et paraboliques stochastiques sont relativement bien comprises. De telles méthodes ont été appliquées dans de nombreux domaines (*cf* Ghanem et Spanos [32] et ses références). Concernant les modèles de fluides visqueux, les travaux antérieurs concernent en particulier les équations de Navier–Stokes (Le Maître *et al.* [46, 50]), les fluides à bas nombre de Mach et les applications microfluides électrochimiques (Debusschere *et al.* [17]). Des travaux récents sur les écoulements de fluide incertains peuvent être trouvés dans Knio et Le Maître [44] et dans Najm *et al.* [64, 65]. L'application des méthodes spectrales stochastiques (et en particulier les méthodes de Galerkin) aux lois de conservation et aux systèmes hyperboliques (en particulier les fluides non visqueux) posent des problèmes supplémentaires, qui sont décrits ci-après.

### 1.3 Systèmes hyperboliques

On commence cette section par une brève présentation sur les systèmes hyperboliques de lois de conservation dans un cadre déterministe. Une description plus détaillée de ces systèmes et des schémas utilisés pour les approcher peut être trouvée, entre autres, dans les livres de Godlewski et Raviart [33], de Toro [80] et de Després [24]. On présente ensuite les difficultés rencontrées lors de l'application de méthodes spectrales stochastiques aux systèmes hyperboliques stochastiques ainsi que l'état de l'art correspondant.

#### 1.3.1 Systèmes hyperboliques déterministes

Pour simplifier, on se limite ici au cas d'une dimension d'espace. On considère un système de lois de conservation sous la forme conservative

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} f(u(x, t)) = 0, \\ u(x, t = 0) = u^0(x). \end{cases} \quad (1.60)$$

Le problème (1.60) est posé sur un intervalle  $\Omega \subset \mathbb{R}$  et un intervalle de temps  $[0, T]$ . On ne discutera pas ici de la question des conditions aux limites ; on suppose par exemple que  $u(x, t)$  est à support compact dans  $\Omega$  pour tout  $t \in [0, T]$ . La fonction

$$u : (x, t) \in \Omega \times [0, T] \mapsto u(x, t) \in \mathbb{R}^m \quad (1.61)$$

est le vecteur d'état des variables conservatives, la fonction

$$f : u \in \mathbb{R}^m \mapsto f(u) \in \mathbb{R}^m \quad (1.62)$$

la fonction flux, et  $u^0$  la condition initiale. Par ailleurs, la solution  $u$  est souvent à valeurs dans un ensemble d'états admissibles  $\mathcal{A}_u \subset \mathbb{R}^m$ . Par exemple, pour l'équation de Burgers,  $m = 1$ ,  $u$  est la vitesse et on peut prendre  $\mathcal{A}_u = \mathbb{R}$ , alors que pour le système des équations d'Euler unidimensionnel,  $m = 3$ , les composantes de  $u$  sont la densité, la quantité de mouvement (par unité de volume) et l'énergie totale (par unité de volume) et  $\mathcal{A}_u$  correspond à l'ensemble des états pour lesquels la densité et la pression sont positives. Enfin, une propriété importante de (1.60) est que l'intégrale en espace de toutes les composantes de  $u$  est conservée au cours du temps.

Pour une solution  $u$  suffisamment régulière, le système (1.60) peut également s'écrire sous la forme non conservative

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) + \nabla_u f(u(x, t)) \frac{\partial}{\partial x} u(x, t) = 0, \\ u(x, t = 0) = u^0(x). \end{cases} \quad (1.63)$$

où  $\nabla_u f : u \in \mathcal{A}_u \mapsto \nabla_u f(u) \in \mathbb{R}^{m,m}$  est la matrice Jacobienne de la fonction flux.

**Définition 1.1.** *Le système (1.60) est dit hyperbolique si et seulement si pour tout  $u \in \mathcal{A}_u$ , la matrice  $\nabla_u f(u)$  est diagonalisable dans  $\mathbb{R}$ . Le système est dit strictement hyperbolique si toutes les valeurs propres de cette matrice sont distinctes.*

On rappelle qu'une matrice  $A \in \mathbb{R}^{m,m}$  est diagonalisable dans  $\mathbb{R}$  si il existe

1.  $m$  valeurs propres  $\lambda^1, \dots, \lambda^m$  (comptées avec leur multiplicité),
2. une base de  $m$  vecteurs propres à droite  $r^1, \dots, r^m$ ,
3. une base de  $m$  vecteurs propres à gauche  $l^1, \dots, l^m$ ,

tels que pour tout  $k \in \{1, \dots, m\}$ ,

$$Ar^k = \lambda^k r^k, \quad A^T l^k = \lambda^k l^k, \quad \langle r^k, l^{k'} \rangle_{\mathbb{R}^m} = \delta_{k,k'}. \quad (1.64)$$

En posant

$$R = \begin{pmatrix} r^1 & \dots & r^m \end{pmatrix}, \quad D = \text{diag}(\lambda^k)_{k \in \{1, \dots, m\}} \quad \text{et} \quad L = \begin{pmatrix} l^1 & \dots & l^m \end{pmatrix}, \quad (1.65)$$

les relations ci-dessus deviennent

$$A = RDL^T, \quad R^{-1} = L^T. \quad (1.66)$$

On peut également écrire

$$A = \sum_{k=1}^m \lambda^k r^k \otimes l^k. \quad (1.67)$$

### 1.3.1.1 Solutions faibles : la condition de Rankine-Hugoniot

Lorsque la fonction flux  $f$  est non linéaire, des discontinuités peuvent apparaître en temps fini, même pour des données initiales très régulières. Il faut donc comprendre la solution de (1.60) dans un sens faible. On désigne par  $\cdot$  le produit scalaire dans  $\mathbb{R}^m$ .

**Définition 1.2.** *Une fonction mesurable et localement bornée  $u : \Omega \times [0, T] \mapsto \mathcal{A}_u$  est solution faible de (1.60) si pour toute fonction  $\psi$  de classe  $C^1$  à support compact dans  $\Omega \times [0, T)$ , à valeurs dans  $\mathbb{R}^m$ , on a*

$$\int_{\Omega} \int_0^T \left( u(x, t) \cdot \frac{\partial}{\partial t} \psi(x, t) + f(u(x, t)) \cdot \frac{\partial}{\partial x} \psi(x, t) \right) dx dt = - \int_{\Omega} u^0(x) \cdot \psi(x, 0) dx. \quad (1.68)$$

Dans le cas où la solution  $u$  est régulière par morceaux, la relation de Rankine-Hugoniot doit être satisfaite le long des courbes où  $u$  est discontinue. En particulier, si on considère une solution faible de (1.60) qui a la structure suivante :

$$u(x, t) = \begin{cases} u_L(x, t), & x < \sigma(t), \\ u_R(x, t), & x > \sigma(t), \end{cases} \quad (1.69)$$

où  $\Sigma = \{(\sigma(t), t)\}$  est une courbe régulière dans  $\Omega \times [0, T]$ , alors les états gauche et droit  $u_L$  et  $u_R$  satisfont localement (1.60) au sens fort en dehors de  $\Sigma$  et la condition de Rankine–Hugoniot stipule que, le long de  $\Sigma$ ,

$$f(u_R) - f(u_L) = \sigma'(t)(u_R - u_L). \quad (1.70)$$

### 1.3.1.2 Solutions entropiques

La classe des solutions faibles est trop large car on perd l'unicité de la solution. Il faut donc introduire un critère afin de sélectionner parmi les solutions faibles de (1.60) l'unique solution qui est physiquement admissible.

La solution d'un problème hyperbolique peut être obtenue comme la limite de la solution visqueuse d'un problème parabolique où le terme de viscosité tend vers zéro avec  $\epsilon$  :

$$\frac{\partial}{\partial t} u_\epsilon(x, t) + \frac{\partial}{\partial x} f(u_\epsilon(x, t)) = \epsilon \frac{\partial}{\partial x} \left( \mathcal{D} \frac{\partial}{\partial x} u_\epsilon \right), \quad (1.71)$$

où  $\mathcal{D}$  est une certaine matrice de diffusion. Le système (1.60) que l'on considère est invariant par le changement de coordonnées espace-temps ( $\hat{t} = -t, \hat{x} = -x$ ) : il est dit réversible, alors que le système (1.71) lui ne l'est pas, les termes dissipatifs le rendant irréversible. Il s'agit alors de garder la mémoire de l'irréversibilité dans le passage à la limite  $\epsilon \rightarrow 0$ . À cette fin, on introduit la notion d'entropie mathématique et de solution entropique.

**Définition 1.3.** Une fonction  $s : u \in \mathcal{A}_u \mapsto s(u) \in \mathbb{R}$  est une entropie du système (1.60) si et seulement si  $s$  est convexe et il existe une fonction  $g : u \in \mathcal{A}_u \mapsto g(u) \in \mathbb{R}$  dite flux d'entropie telle que

$$\nabla_u f^T \nabla_u s = \nabla_u g. \quad (1.72)$$

**Définition 1.4.** Une fonction mesurable et localement bornée  $u : \Omega \times [0, T] \mapsto \mathcal{A}_u$  est solution (faible) entropique de (1.60) si elle est solution faible et si pour toute fonction  $\psi$  de classe  $C^1$  à support compact dans  $\Omega \times [0, T)$ , à valeurs dans  $\mathbb{R}$ , et pour toute entropie  $s$  de flux d'entropie  $g$ , on a

$$\int_{\Omega} \int_0^T \left( s(u) \frac{\partial}{\partial t} \psi(x, t) + g(u) \frac{\partial}{\partial x} \psi(x, t) \right) dx dt \geq - \int_{\Omega} s(u_0(x)) \psi(x, 0) dx. \quad (1.73)$$

Dans le cas des lois de conservation scalaires ( $m = 1$ ), on a le théorème d'unicité suivant :

**Théorème 1.1.** (Kruzhkov) On suppose  $f \in C^1(\mathbb{R})$  et  $u_0 \in L^\infty(\mathbb{R})$ . Alors, le problème (1.60) admet une et une seule solution faible entropique bornée. De plus, cette solution satisfait le principe du maximum suivant :

$$(\forall x \in \Omega, \alpha \leq u_0(x) \leq \beta) \Rightarrow (\forall (x, t) \in \Omega \times [0, T], \alpha \leq u(x, t) \leq \beta). \quad (1.74)$$

### 1.3.1.3 Schémas numériques

Les schémas numériques utilisés pour la résolution des systèmes hyperboliques se doivent d'assurer la propriété de conservation des variables conservatives (les composantes du vecteur  $u$ ) au cours du temps, d'où l'appellation de schémas conservatifs. Dans ce manuscrit, on se concentre sur les méthodes de Volumes Finis (VF).

Les méthodes VF sont construites grâce à des bilans de flux entre cellules. Elles respectent les principales propriétés mathématiques du modèle, capturent correctement les ondes se propageant dans le système en limitant (autant que possible) la diffusion numérique et sont relativement précises et robustes même en présence de forts gradients. Par la suite, on utilise une discrétisation numérique de type Euler explicite en temps et une discrétisation VF d'ordre un en espace. Soit un pas d'espace  $\Delta x$  (uniforme pour simplifier) et des temps discrets  $t^n$  avec un pas de temps  $\Delta t^n = t^{n+1} - t^n$ . On considère un schéma VF de la forme suivante

$$u_i^{n+1} = u_i^n - \frac{\Delta t^n}{\Delta x} (\varphi(u_i^n, u_{i+1}^n) - \varphi(u_{i-1}^n, u_i^n)), \quad (1.75)$$

où  $u_i^n$  est une approximation de la valeur moyenne en espace de la solution  $u$  dans la cellule de centre  $i\Delta x$  au temps  $t^n$  et  $\varphi(\cdot, \cdot)$  est un flux numérique à deux points. En particulier, on s'intéresse au flux numérique de type Roe [73], qui assure une faible diffusion numérique et propage les ondes à la bonne vitesse (sauf dans le cas de points soniques sur lequel on reviendra). Enfin, du fait de la nature explicite de (1.75), le pas de temps doit vérifier une condition CFL de la forme

$$\Delta t \leq \frac{\Delta x}{|\lambda|}, \quad (1.76)$$

où  $\lambda$  est la vitesse maximale des ondes sur toutes les interfaces.

Des méthodes d'ordre plus élevé sont disponibles dans la littérature mais ne seront pas utilisées dans ce manuscrit : on peut citer, entre autres, les méthodes WENO introduites par Liu *et al.* [55], les méthodes de Galerkin discontinues, cf par exemple les livres de Hesthaven et Warburton [39] et de Di Pietro et Ern [25], ...

### 1.3.2 Systèmes hyperboliques stochastiques

On considère à présent des incertitudes sur les coefficients ou les conditions initiales d'un système hyperbolique. Pour simplifier, on se place toujours dans le cas de domaines spatiaux unidimensionnels. L'extension aux dimensions spatiales supérieures est directe du point de vue stochastique. On cherche alors  $U(x, t, \xi)$  résolvant presque sûrement en  $\xi$  le système conservatif suivant :

$$\begin{cases} \frac{\partial}{\partial t} U(x, t, \xi) + \frac{\partial}{\partial x} F(U(x, t, \xi); \xi) = 0, \\ U(x, t = 0, \xi) = U^0(x, \xi). \end{cases} \quad (1.77)$$

Le problème (1.77) est posé sur un intervalle borné  $\Omega \subset \mathbb{R}$  (le **domaine spatial**). On rappelle que  $\xi \in \Xi$ , le **domaine stochastique**. L'ensemble  $\Omega$  étant borné, le



problème (1.77) doit être complété par des conditions aux bords. Celles-ci seront spécifiées dans le chapitre 3 lorsque les cas tests seront présentés. Ainsi,

$$U : (x, t, \xi) \in \Omega \times [0, T] \times \Xi \mapsto U(x, t, \xi) \in \mathcal{A}_U \quad (1.78)$$

est le vecteur d'état incertain des variables conservatives paramétré par  $\xi$ , appelé **solution stochastique** du problème (1.77), où  $\mathcal{A}_U \subset \mathbb{R}^m$ ,  $m \geq 1$ , est l'ensemble des états admissibles, que l'on suppose indépendant de l'événement aléatoire. Par ailleurs,  $U^0(x, \xi)$  est une paramétrisation par  $\xi$  de la condition initiale incertaine et

$$F : (U; \xi) \in \mathcal{A}_U \times \Xi \mapsto F(U; \xi) \in \mathbb{R}^m \quad (1.79)$$

est la fonction flux. La dépendance explicite de  $F$  en  $\xi$  indique que  $F$  peut faire intervenir des coefficients incertains, également paramétrés par  $\xi$ .

Pour une solution  $U$  suffisamment régulière, le système (1.77) peut également s'écrire sous la forme non conservative

$$\begin{cases} \frac{\partial}{\partial t} U(x, t, \xi) + \nabla_U F(U(x, t, \xi); \xi) \frac{\partial}{\partial x} U(x, t, \xi) = 0, \\ U(x, t = 0, \xi) = U^0(x, \xi). \end{cases} \quad (1.80)$$

On fait l'hypothèse que ce système est presque sûrement hyperbolique au sens où la **matrice Jacobienne stochastique**

$$\nabla_U F : (U; \xi) \in \mathcal{A}_U \times \Xi \mapsto \nabla_U F(U; \xi) \in \mathbb{R}^{m,m} \quad (1.81)$$

est diagonalisable dans  $\mathbb{R}$  presque sûrement, *i.e.*, pour presque tout  $\xi \in \Xi$ , il existe

1.  $m$  valeurs propres  $\Lambda^1(U; \xi), \dots, \Lambda^m(U; \xi)$  (comptées avec leur multiplicité),
2. une base de  $m$  vecteurs propres à droite  $R^1(U; \xi), \dots, R^m(U; \xi)$ ,
3. une base de  $m$  vecteurs propres à gauche  $L^1(U; \xi), \dots, L^m(U; \xi)$ ,

tels que

$$\nabla_U F(U; \xi) = R(U; \xi) D(U; \xi) L(U; \xi)^T, \quad (1.82)$$

avec

$$\begin{aligned} R(U; \xi) &= \begin{pmatrix} R^1(U; \xi) & \dots & R^m(U; \xi) \end{pmatrix}, \quad D(U; \xi) = \text{diag}(\Lambda^k(U; \xi))_{k \in \{1, \dots, m\}}, \\ \text{et } L(U; \xi) &= \begin{pmatrix} L^1(U; \xi) & \dots & L^m(U; \xi) \end{pmatrix}. \end{aligned} \quad (1.83)$$

Les matrices  $R(U; \xi)$ ,  $D(U; \xi)$  et  $L(U; \xi)$  sont dans  $\mathbb{R}^{m,m}$ . Pour alléger les notations, la dépendance des valeurs propres et des vecteurs propres en  $U$  est omise par la suite.

**Remarque 1.3.** *Mishra et Schwab ont récemment prouvé dans [61], sous des hypothèses relativement générales, l'existence et l'unicité des solutions stochastiques entropiques pour les lois de conservation scalaires stochastiques où l'incertitude porte sur la condition initiale, ainsi que l'existence de moments statistiques d'ordre quelconque pour ces solutions stochastiques entropiques.*

### 1.3.3 Méthodes spectrales stochastiques : principales difficultés et état de l'art

L'objectif de cette section est de présenter les principales difficultés et l'état de l'art concernant l'application des méthodes spectrales stochastiques aux systèmes hyperboliques stochastiques.

#### 1.3.3.1 Principales difficultés

La principale difficulté est que les solutions de systèmes hyperboliques stochastiques peuvent présenter des discontinuités (dans le domaine spatial) en temps fini dues au développement d'ondes de choc et de discontinuités de contact (dans les variables spatiales). Bien que ces discontinuités concernent les variables spatiales, leur vitesse de propagation peut être affectée par les incertitudes, et ainsi engendrer des discontinuités de la solution par rapport aux variables stochastiques. Pour illustrer cette difficulté, on représente sur la figure 1.5 la densité stochastique d'un gaz parfait dans un tube à choc unidimensionnel avec une incertitude sur la vitesse du son. La densité est représentée en fonction de  $x$ , la variable spatiale, et  $\xi$ , la variable stochastique qui paramétrise l'aléa, à un temps fixé après plusieurs réflexions des ondes aux extrémités du tube. On observe que pour toute réalisation de  $\xi$ , la solution présente une discontinuité dans le domaine spatial, et que pour certaines positions dans le domaine spatial, la solution va également être discontinue dans le domaine stochastique. Par conséquent, des bases de polynômes continus en  $\xi$  ne sont plus appropriées pour la mise en œuvre des méthodes spectrales stochastiques, en raison de phénomènes de type Gibbs (Le Maître *et al.* [47]). Pour surmonter cette difficulté, on utilisera des développements spectraux plus locaux, similaires à ceux cités précédemment en fin de section 1.2.1.4.

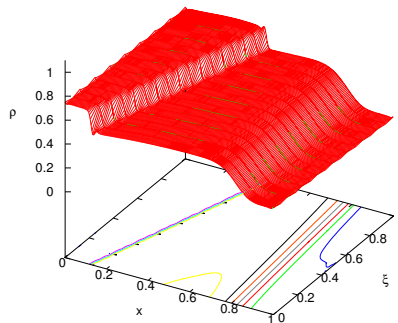


FIGURE 1.5 – Densité stochastique dans un tube à choc avec vitesses des ondes incertaines en fonction de  $x$  et  $\xi$  à  $t$  fixé.

Une autre difficulté a pour origine les non-linéarités dans les flux physiques du système hyperbolique stochastique soulevant le problème du calcul de ces flux dans le contexte des méthodes de Galerkin que l'on souhaite utiliser. En effet, toutes les opérations mathématiques doivent être appliquées aux développements stochastiques qui représentent les variables. Une approche intéressante est d'utiliser les

techniques pseudo-spectrales proposées par Debusschere *et al.* [18]. On y reviendra dans le chapitre 2.

Enfin, une question importante concerne l'hyperbolicité du système de Galerkin obtenu à partir du système (1.77), d'autant plus que l'on souhaite concevoir une méthode VF pour le discrétiser en espace et en temps. En effet, la projection de Galerkin de (1.77) fournit un système conservatif de lois de conservation couplant les modes stochastiques de la solution. Il est donc légitime de se demander si ce système est hyperbolique, *i.e.*, si la matrice Jacobienne associée à ce système est diagonalisable dans  $\mathbb{R}$  dans une base complète de vecteurs propres de  $\mathbb{R}^{mP}$ . Ce point sera traité dans le chapitre 2.

### 1.3.3.2 Etat de l'art

Concernant les approches non-intrusives, Mathelin *et al.* [59] ont appliqué les méthodes de collocation polynomiales aux équations d'Euler dans le cas de solutions continues (sans choc). Plus récemment, Abgrall [1] a proposé de décomposer le domaine stochastique en sous-domaines et d'effectuer des reconstructions non oscillantes basées sur des polynômes par morceaux de type ENO/WENO afin d'obtenir les solutions stochastiques discontinues des équations de convection, de Burgers et d'Euler. Barth [8] a ensuite considéré le cas où la solution est cherchée dans un espace produit composé de polynômes par morceaux en  $x$  et  $\xi$  reconstruit par une méthode de type WENO. Par ailleurs, Lin *et al.* [53] ont utilisé une approche basée sur les méthodes de collocation probabiliste multi-éléments pour un écoulement de fluide supersonique venant impacter un obstacle de rugosité aléatoire. Le domaine stochastique est également découpé en sous-domaines et dans chacun d'eux, une base de polynômes orthogonaux est construite à partir de la fonction de densité conditionnelle (au sens de se trouver ou non dans le sous-domaine), base sur laquelle le développement stochastique local de la solution est cherché.

À notre connaissance, très peu de méthodes spectrales stochastiques intrusives ont été développées pour les systèmes hyperboliques stochastiques. L'équation des ondes a été traitée dans le cas scalaire avec des méthodes gPC par Gottlieb et Xiu [34]. Le cas de systèmes hyperboliques non-linéaires est évidemment plus compliqué. Dans le contexte des méthodes intrusives, une question cruciale est la construction d'un schéma adéquat pour approcher en espace et en temps le problème d'évolution associé à la projection de Galerkin. Typiquement, on souhaite utiliser un schéma VF avec un décentrement adéquat. Ge *et al.* [30] ont appliqué la méthode gPC à des solutions continues (sans choc) des équations des eaux peu profondes en utilisant un schéma de Godunov mais en calculant les flux du système de Galerkin de manière non-intrusive par des méthodes de quadrature.

Une méthode originale a été proposée récemment par Poette *et al.* [69, 70] pour l'équation de Burgers et les équations d'Euler (avec une position d'interface incertaine entre deux fluides). Une caractéristique intéressante de l'approche est que le développement polynomial porte sur des variables entropiques et non sur les variables conservatives d'origine, de sorte qu'il peut être prouvé que la projection de

Galerkin conduit à un système hyperbolique. L'algorithme numérique nécessite une procédure de minimisation pour retrouver le développement de la solution. Comme dans [30], les flux numériques du schéma VF utilisés pour discrétiser le système de Galerkin sont calculés par des formules de quadrature.

Enfin, dans le cas où le décentrement du schéma VF est calculé de manière intrusive, des écoulements de fluides supersoniques contre un coin avec des fluctuations en entrée aléatoires ou des oscillations aléatoires du coin autour de son sommet ont été étudiés par Lin *et al.* [52] en utilisant des méthodes multi-éléments gPC et en considérant un décentrement basé sur les valeurs moyennes (dans le domaine stochastique) des vecteurs propres de la matrice Jacobienne du système de Galerkin. Comme mentionné dans [52], cette approche n'est justifiée que dans le cas de fluctuations relativement petites des quantités aléatoires. Notre travail améliore ce point aussi bien théoriquement que numériquement.

## 1.4 Description des travaux

L'objectif de ce travail est de concevoir, analyser et mettre en œuvre des méthodes de Galerkin pour les systèmes hyperboliques non linéaires stochastiques.

La première partie de ce mémoire consiste à construire et tester numériquement un solveur de type Roe pour le système de Galerkin associé au problème stochastique d'origine, en utilisant une discrétisation VF dans l'espace physique et une représentation polynomiale continue par morceaux dans le domaine stochastique [16, 48, 85]. Un point saillant est le développement d'une méthode efficace pour le calcul de la matrice de décentrement. Cette première partie du mémoire a fait l'objet de deux publications en collaboration avec O. Le Maître, M. Ndjinga et A. Ern [82, 83]. Dans la première, la méthodologie est étudiée théoriquement et numériquement. Dans la deuxième, un correcteur entropique est développé. Ce correcteur est basé sur une adaptation au cas stochastique du correcteur de Dubois et Mehlmann [27] proposé dans le contexte déterministe. Cette adaptation est loin d'être directe en raison de la possibilité d'apparition de valeurs propres multiples dans la matrice Jacobienne du système de Galerkin. La description du solveur de Roe avec correcteur entropique pour les systèmes hyperboliques non linéaires fait l'objet du chapitre 2. Son application à l'équation de Burgers et aux équations d'Euler est présentée au chapitre 3, dans le cas de deux dimensions stochastiques et une dimension stochastique respectivement.

La méthode proposée, bien que capable de s'appliquer à des situations complexes, reste coûteuse car une discrétisation stochastique très fine est nécessaire pour représenter la solution au voisinage des discontinuités. Par conséquent, la deuxième partie de ce mémoire consiste à élaborer des stratégies adaptatives dans le domaine stochastique. Puisque les discontinuités sont localisées en espace et évoluent en temps, on propose des représentations stochastiques dépendant de l'espace et du temps. Ainsi, chaque cellule VF supporte sa propre discrétisation stochastique et celle-ci évolue en temps. La discrétisation n'est donc plus basée sur une tensorisation

des espaces d'approximation stochastique et déterministe, ce qui, à notre connaissance, constitue une contribution originale. Pour cela, on reformule la méthodologie dans un contexte multi-résolution basé sur le concept d'arbres binaires afin de décrire la discrétisation du domaine stochastique, de façon similaire aux travaux initiés par Harten dans le cas déterministe [38]. L'enrichissement et l'élagage adaptatifs des arbres binaires définissant les espaces d'approximation stochastiques sont alors réalisés en utilisant des critères d'analyse multi-résolution. De plus, une adaptation anisotrope est proposée pour les domaines stochastiques multidimensionnels. Le concept d'analyse multi-résolution et ses différents outils sont introduits dans le chapitre 4 dans le cas unidimensionnel et testés sur des exemples simples. Dans le chapitre 5, on utilise ces techniques pour construire un solveur de Roe adaptatif, que l'on applique aux équations d'Euler dans le cas d'une dimension stochastique. Enfin, l'extension de cette méthode au cas multidimensionnel est développée dans le chapitre 6 et son application à l'équation de Burgers est présentée dans le cas de deux dimensions stochastiques. Cette partie a fait l'objet d'un proceeding [81] et fait l'objet d'une publication actuellement en cours de rédaction.

Première partie

Solveur de Roe stochastique



# Conception et analyse du solveur

## Sommaire

<b>2.1</b>	<b>Projection de Galerkin</b>	<b>33</b>
2.1.1	Rappel du cadre probabiliste pour les incertitudes paramétriques	34
2.1.2	Discrétisation stochastique	34
2.1.3	Le système de Galerkin	36
2.1.4	Calcul des non-linéarités	37
<b>2.2</b>	<b>Hyperbolicité du système de Galerkin</b>	<b>38</b>
2.2.1	Systèmes hyperboliques stochastiques symétriques	39
2.2.2	Vecteurs propres indépendants de l'incertitude	39
2.2.3	Cas général	40
<b>2.3</b>	<b>Approximation en espace et en temps</b>	<b>43</b>
2.3.1	Matrice de Roe et état de Roe	44
2.3.2	Une méthode efficace pour approcher la valeur absolue d'une matrice	45
2.3.3	Le schéma de décentrement	48
<b>2.4</b>	<b>Correcteur entropique</b>	<b>49</b>
2.4.1	Principes généraux	50
2.4.2	Application directe au système de Galerkin	50
2.4.3	Adaptation au système de Galerkin	52

On s'intéresse à des systèmes hyperboliques stochastiques non linéaires de la forme (1.77). Les incertitudes peuvent résulter d'une variabilité sur la condition initiale et/ou de coefficients dans le modèle, et sont paramétrées par un ensemble de  $N$  variables aléatoires indépendantes. On décrit dans ce chapitre une méthode de Galerkin stochastique pour la propagation de ces incertitudes, basée sur la construction d'un solveur de Roe stochastique pour le système de Galerkin régissant l'évolution des modes stochastiques de la solution.

## 2.1 Projection de Galerkin

Après un rappel du cadre probabiliste, on définit la discrétisation stochastique utilisée. On applique ensuite une projection de Galerkin pour construire un système déterministe couplant les modes stochastiques de la solution.



### 2.1.1 Rappel du cadre probabiliste pour les incertitudes paramétriques

On utilise un espace de probabilité abstrait  $\mathcal{P} = (\Theta, \Sigma, d\mu)$ , où  $\Theta$  est l'ensemble des événements aléatoires,  $\Sigma$  la  $\sigma$ -algèbre associée et  $d\mu$  la mesure de probabilité. Pour toute variable aléatoire  $U(\theta)$  définie sur  $\mathcal{P}$ , l'espérance de  $U$  est

$$E[U] := \int_{\Theta} U(\theta) d\mu(\theta). \quad (2.1)$$

Soit  $L^2(\Theta, d\mu)$  l'espace des variables aléatoires du second ordre sur  $\mathcal{P}$ . Par la suite, les quantités aléatoires sont supposées du second ordre.

En vue de la discrétisation stochastique, on introduit un ensemble fini de  $N$  variables aléatoires  $\xi(\theta) := \{\xi_1(\theta), \dots, \xi_N(\theta)\}$  définies sur  $\mathcal{P}$  de distributions connues. Ces variables aléatoires vont être utilisées pour décrire des incertitudes sur les coefficients ou les conditions initiales du système hyperbolique. Pour simplifier, les variables aléatoires  $\xi_i(\theta)$  sont supposées indépendantes identiquement distribuées à valeurs dans  $\mathbb{R}$ , de sorte que la fonction de densité jointe de  $\xi(\theta)$  se factorise sous la forme

$$p_{\xi}(y) = \prod_{i=1}^N p(y_i), \quad (2.2)$$

où  $p(y_i)$  est la fonction densité de probabilité de  $\xi_i(\theta)$ . De plus, on note  $\Xi \subset \mathbb{R}^N$  l'ensemble des valeurs prises par  $\xi$ , que l'on appelle **domaine stochastique**, et  $\mathcal{P}_{\xi}$  l'espace de probabilité image,  $\mathcal{P}_{\xi} := (\Xi, \mathcal{B}_{\Xi}, p_{\xi} d\xi)$ , où  $\mathcal{B}_{\Xi}$  est l'ensemble de Borel de  $\Xi$  et  $d\xi$  la mesure de Lebesgue dans  $\mathbb{R}^N$ . Soit  $L^2(\Xi, p_{\xi})$  l'espace des variables aléatoires du second ordre sur  $\mathcal{P}_{\xi}$ . L'opérateur espérance sur l'espace image  $\mathcal{P}_{\xi}$  relatif à l'espérance sur  $\mathcal{P}$ , noté par des crochets, est obtenue par l'identité

$$E[U] = \int_{\Theta} U(\xi(\theta)) d\mu(\theta) = \int_{\Xi} U(y) p_{\xi}(y) dy =: \langle U \rangle. \quad (2.3)$$

### 2.1.2 Discrétisation stochastique

Pour approcher la solution  $U$  d'un système hyperbolique stochastique de la forme (1.77) dans  $L^2(\Xi, p_{\xi})$ , une discrétisation stochastique du problème est nécessaire. Celle-ci est obtenue en considérant une base Hilbertienne de fonctionnelles en  $\xi$  engendrant  $L^2(\Xi, p_{\xi})$ ,

$$L^2(\Xi, p_{\xi}) = \overline{\text{span}\{\Phi_1(\xi), \Phi_2(\xi), \dots\}}, \quad \langle \Phi_{\alpha} \Phi_{\beta} \rangle = \delta_{\alpha\beta}, \quad (2.4)$$

où  $\delta_{\alpha\beta}$  est le symbole de Kronecker.

La solution discrète est cherchée dans un espace de dimension finie  $S^P$ , appelé **espace d'approximation stochastique**, construit par troncature de la base Hilbertienne :

$$S^P := \text{span}\{\Phi_1(\xi), \Phi_2(\xi), \dots, \Phi_P(\xi)\} \subset L^2(\Xi, p_{\xi}), \quad \dim(S^P) =: P. \quad (2.5)$$

Pour simplifier, on suppose que  $\xi$  est un vecteur aléatoire uniforme sur  $\Xi := [0, 1]^N$ , *i.e.*  $\xi \sim \mathcal{U}([0, 1]^N)$  (une transformation isoprobabiliste peut être utilisée pour effectuer une projection des variables aléatoires indépendantes d'origine sur ce vecteur aléatoire [48, 49]). L'espace de probabilité image est alors  $\mathcal{P}_\xi := ([0, 1]^N, \mathcal{B}_{[0,1]^N}, d\xi)$ , où  $\mathcal{B}_{[0,1]^N}$  est l'ensemble de Borel de  $[0, 1]^N$  et  $d\xi$  la mesure de Lebesgue dans  $\mathbb{R}^N$ .

Le domaine stochastique  $\Xi$  est décomposé dyadiquement et la solution stochastique est approchée par des fonctions polynomiales par morceaux. Outre le nombre  $N$  de variables aléatoires  $\xi_i$  dans la paramétrisation, cette approximation dépend

- du niveau de résolution  $Nr \geq 0$ , qui contrôle la taille minimale des éléments stochastiques (en abrégé, SE pour “stochastic element”), *i.e.*, les cellules de discrétisation dans le domaine stochastique,
- de l'ordre polynomial  $No \geq 0$  (qui contrôle le degré de l'approximation polynomiale dans chaque SE). On se restreint ici à un ordre polynomial unique dans tout le domaine stochastique.

Le cas  $Nr = 0$  correspond à l'approximation continue classique (développement de Wiener-Legendre), alors que le choix  $Nr > 0$  et  $No = 0$  conduit au développement de Wiener-Haar (approximation constante par morceaux). Soit  $\mathbf{i} = (i_1, \dots, i_N) \in \{1, \dots, 2^{Nr}\}^N$  un multi-indice et soit

$$K_{\mathbf{i}} = \{\xi \in [0, 1]^N, \forall j \in \{1, \dots, N\}, \xi_j \in [2^{-Nr}(i_j - 1), 2^{-Nr}i_j]\} \quad (2.6)$$

l'élément stochastique associé. La discrétisation stochastique de  $\Xi$  est représentée pour différents  $Nr$  dans le cas de deux dimensions stochastiques ( $N = 2$ ) sur la figure 2.1. L'espace d'approximation stochastique  $S^P$  est alors défini comme l'espace des fonctions polynomiales par morceaux

$$S^P := S^{No, Nr} := \{f : [0, 1]^N \rightarrow \mathbb{R}, \forall \mathbf{i} \in \{1, \dots, 2^{Nr}\}^N, f|_{K_{\mathbf{i}}} \in \mathbb{Q}_{No}^N[\xi]\}, \quad (2.7)$$

où  $\mathbb{Q}_{No}^N[\xi]$  est l'espace vectoriel des polynômes réels dans  $\mathbb{R}^N$  de degré  $\leq No$  en chaque variable  $\xi_i$ . Il est également possible de travailler avec des espaces d'approximation stochastique plus petits que ceux définis dans (2.7), par exemple des espaces de polynômes de degré total  $\leq No$  obtenus en utilisant des tensorisations polynomiales partielles à la place des tensorisations polynomiales complètes. En outre, les espaces  $S^{No, Nr}$  forment une famille hiérarchique d'espaces d'approximation stochastiques puisque  $S^{No, Nr} \subset S^{No', Nr}$  pour  $No \leq No'$  et  $S^{No, Nr} \subset S^{No, Nr'}$  pour  $Nr \leq Nr'$ .

Deux bases différentes peuvent être utilisées pour engendrer l'espace  $S^{No, Nr}$ . La première est le système hiérarchique de Multi-Ondelettes (MW) d'ordre  $No$  et de niveau de résolution  $Nr$  introduit dans [48]. La deuxième est le système de bases polynomiales de Legendre locales que l'on nommera base SE, où chaque fonction de  $S^{No, Nr}$  est développée sur chaque élément stochastique de volume  $2^{-NNr}$  sur une base locale complètement tensorisée de taille  $(No+1)^N$  de polynômes de Legendre, définis par convention par rapport à l'intervalle de référence  $[0, 1]$ . En vue d'algorithmes adaptatifs, la base MW fournit un cadre de travail naturel, comme on le verra dans les chapitres 4 et 5. La base SE est plus pratique pour l'analyse théorique et

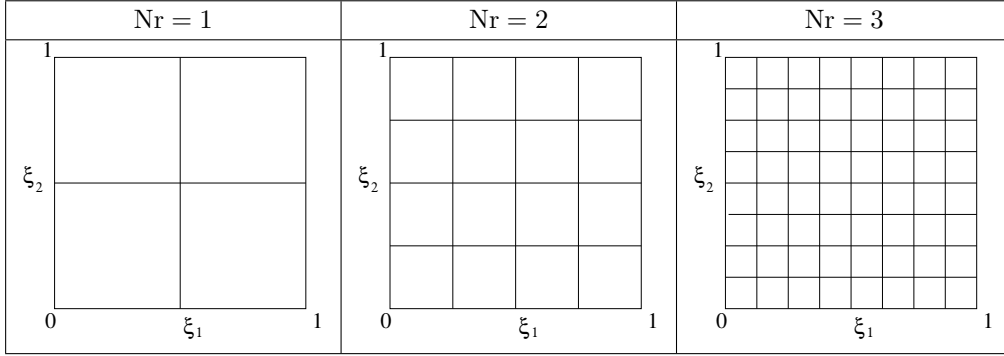


FIGURE 2.1 – Discretisation du domaine stochastique dans le cas de deux dimensions stochastiques ( $N = 2$ ) pour différents niveaux de résolution.

l'implémentation. Par conséquent, on utilisera dans les chapitres 2 et 3 la base SE, que l'on notera  $\{\Phi_\alpha(\xi)\}_{\alpha \in \{1, \dots, P\}}$ . En pratique,  $\alpha$  est un double indice,  $\alpha = (\alpha_\sigma, \alpha_\pi)$ , le premier indice faisant référence à l'élément stochastique et le second à la fonction polynomiale dans l'élément stochastique. Par la suite, on pose  $P_\pi := (N_o + 1)^N$ ,  $P_\sigma := 2^{NN_r}$  et  $P := P_\pi P_\sigma$ , si bien que l'espace d'approximation stochastique  $S^P := S^{N_o, N_r}$  est de dimension  $P$ . Enfin, on définit les ensembles

$$\mathcal{P}_\pi := \{1, \dots, P_\pi\}, \quad \mathcal{P}_\sigma := \{1, \dots, P_\sigma\} \quad \text{et} \quad \mathcal{P} := \{1, \dots, P\}. \quad (2.8)$$

### 2.1.3 Le système de Galerkin

On approche la solution stochastique  $U$  par une fonction  $U^P \in S^P$ , appelée **solution de Galerkin**, sous la forme

$$U(x, t, \xi) \approx U^P(x, t, \xi) = \sum_{\alpha \in \mathcal{P}} u_\alpha(x, t) \Phi_\alpha(\xi). \quad (2.9)$$

Les champs déterministes  $u_\alpha(x, t)$  à valeurs dans  $\mathbb{R}^m$  sont appelés les **modes stochastiques** de la solution (dans  $S^P$ ). La connaissance des modes stochastiques permet ensuite de calculer des quantités statistiques intéressantes sur la solution  $U^P$ , comme l'espérance, la variance, des moments d'ordre plus élevés, les fonctions de densité, les matrices de corrélation, etc. . . , à partir d'expressions analytiques ou d'échantillonnages de  $\Xi$  (*cf* section 1.2.2).

Le calcul des modes stochastiques  $u_\alpha(x, t)$  est basé sur une interprétation faible, ou projection de Galerkin, de (1.77), où l'on injecte le développement (2.9) de la solution. Après projection de (1.77) sur la base de  $S^P$  et prise en compte de l'orthonormalité de la base, on obtient

$$\begin{cases} \frac{\partial}{\partial t} u_\alpha(x, t) + \frac{\partial}{\partial x} \langle \Phi_\alpha, F(U^P; \cdot) \rangle = 0, & \forall \alpha \in \mathcal{P}, \\ u_\alpha(x, t = 0) = \langle \Phi_\alpha, U^0 \rangle, & \forall \alpha \in \mathcal{P}. \end{cases} \quad (2.10)$$

L'équation (2.10) montre que le  $\alpha$ -ème mode stochastique de la solution de Galerkin est régi par une équation qui, en général, couple tous les modes stochastiques au

travers du terme  $\langle \Phi_\alpha, F(U^P; \cdot) \rangle$ . On définit les vecteurs des modes stochastiques et le vecteur flux de Galerkin par

$$u(x, t) = \begin{pmatrix} u_1(x, t) \\ \vdots \\ u_P(x, t) \end{pmatrix}, \quad f(u(x, t)) = \begin{pmatrix} f_1(u) \\ \vdots \\ f_P(u) \end{pmatrix}, \quad (2.11)$$

avec

$$f_\alpha(u) := \langle \Phi_\alpha, F(U^P; \cdot) \rangle, \quad \alpha \in \mathcal{P}, \quad \text{et} \quad U^P = \sum_{\beta \in \mathcal{P}} u_\beta \Phi_\beta(\xi). \quad (2.12)$$

Le système de Galerkin déterministe se met alors simplement sous la forme

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} f(u(x, t)) = 0, \\ u(x, t = 0) = u^0(x), \end{cases} \quad (2.13)$$

où la condition initiale s'obtient par projection orthogonale de  $U^0$  sur  $S^P$  en posant  $u^0 = (\langle \Phi_\alpha, U^0 \rangle)_{\alpha \in \mathcal{P}}$ . Ainsi, le problème (2.13) a la même forme que le problème stochastique d'origine (1.77), à ceci près que le vecteur d'état est de taille  $mP$ . En pratique, le vecteur  $u$  doit appartenir à l'ensemble admissible  $\mathcal{A}_u \subset \mathbb{R}^{mP}$  défini de sorte que  $u \in \mathcal{A}_u \Leftrightarrow U^P(\xi) = \sum_{\alpha \in \mathcal{P}} u_\alpha \Phi_\alpha(\xi) \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$ , où  $\mathcal{A}_U \in \mathbb{R}^m$  est l'ensemble admissible associé au problème stochastique d'origine (1.77).

#### 2.1.4 Calcul des non-linéarités

On peut difficilement envisager des projections de Galerkin exactes des flux stochastiques puisque celles-ci engendreraient des coûts très importants. Une approche raisonnable est de considérer une approximation de la projection de Galerkin du flux  $F(U^P; \xi)$  sur  $S^P$ . Comme motivé par Debusschere *et al* dans [18], des calculs pseudo-spectraux permettent des gains de calcul significatifs, en introduisant des erreurs numériques négligeables pourvu que la résolution stochastique soit suffisamment fine. Il est également possible de calculer le flux de Galerkin de façon non intrusive par des formules de quadrature, *cf* par exemple Ge *et al*. [30] et Poette *et al*. [69, 70].

Des outils pour des évaluations précises de fonctions (polynomiales et non polynomiales) des variables représentées par des développements stochastiques sont décrits dans [18]. Soient  $a(\xi) = \sum_{\alpha \in \mathcal{P}} a_\alpha \Phi_\alpha(\xi) \in S^P$  et  $b(\xi) = \sum_{\beta \in \mathcal{P}} b_\beta \Phi_\beta(\xi) \in S^P$ . Le produit  $ab$  peut être développé sous la forme

$$(ab)(\xi) = \left( \sum_{\alpha \in \mathcal{P}} a_\alpha \Phi_\alpha(\xi) \right) \left( \sum_{\beta \in \mathcal{P}} b_\beta \Phi_\beta(\xi) \right) = \sum_{\alpha, \beta \in \mathcal{P}} a_\alpha b_\beta \Phi_\alpha \Phi_\beta(\xi). \quad (2.14)$$

En général,  $(ab) \notin S^P$  puisque cette fonction possède des termes de degré  $> \text{No}$ . On appelle produit de Galerkin, et on note  $(a * b)$ , la projection  $L^2$ -orthogonale de

( $ab$ ) sur  $S^P$ . Ce produit est donné par

$$(a * b) := \sum_{\alpha \in \mathcal{P}} (a * b)_\alpha \Phi_\alpha, \quad (a * b)_\alpha = \sum_{\beta, \delta \in \mathcal{P}} a_\beta b_\delta \mathcal{M}_{\alpha\beta\delta}, \quad (2.15)$$

où on a introduit le tenseur de multiplication d'ordre trois

$$\mathcal{M}_{\alpha\beta\delta} := \langle \Phi_\alpha \Phi_\beta \Phi_\delta \rangle. \quad (2.16)$$

Ce tenseur d'ordre trois dépend uniquement de la base stochastique, peut être calculé une fois pour toutes au début de la simulation et son caractère creux dans l'espace stochastique peut être exploité pour son stockage. Le produit de Galerkin est l'opération de base intervenant dans l'évaluation de la projection dans  $S^P$  de non-linéarités plus générales. Par exemple, l'inverse de Galerkin, noté  $a^{-*}$ , et la racine carré de Galerkin, notée  $a^{*/2}$ , sont définis respectivement par l'approximation de la projection  $L^2$ -orthogonale  $1/a$  sur  $S^P$  obtenue à partir de la résolution du système linéaire  $a * a^{-*} = 1$ , et l'approximation de la projection  $L^2$ -orthogonale de  $\sqrt{a}$  sur  $S^P$  obtenue à partir de la résolution (par une méthode de Newton) du système non linéaire  $a^{*/2} * a^{*/2} = a$ .

## 2.2 Hyperbolicité du système de Galerkin

Avant de détailler la construction de la méthode numérique pour approcher le système de Galerkin (2.13), on étudie l'hyperbolicité de ce système. On introduit la **matrice Jacobienne de Galerkin** de dimension  $mP$  :

$$(\nabla_u f(u))_{\alpha, \beta \in \mathcal{P}} = \langle \nabla_U F(U^P; \cdot) \Phi_\alpha \Phi_\beta \rangle_{\alpha, \beta \in \mathcal{P}}, \quad (2.17)$$

et on cherche à savoir si cette matrice est diagonalisable dans  $\mathbb{R}$ .

Il est intéressant de remarquer que les deux matrices Jacobiennes de Galerkin utilisant une représentation sur la base MW et sur la base SE pour la discrétisation stochastique sont équivalentes du point de vue de la diagonalisation dans  $\mathbb{R}$ . En effet, soit  $\{\Psi_\alpha(\xi)\}_{\alpha \in \mathcal{P}}$  la base MW et soit  $B \in \mathbb{R}^{P, P}$  la matrice de transition entre les deux bases, telle que  $\Psi_\alpha(\xi) = \sum_{\gamma \in \mathcal{P}} B_{\alpha\gamma} \Phi_\gamma(\xi)$ , pour tout  $\alpha \in \mathcal{P}$ , *i.e.*,  $B_{\alpha\gamma} = \langle \Psi_\alpha \Phi_\gamma \rangle$ . Soit  $\nabla_u f^{MW} \in \mathbb{R}^{mP, mP}$  la représentation de la matrice Jacobienne de Galerkin dans la base MW. Alors, pour tout  $\alpha, \beta \in \mathcal{P}$ ,

$$(\nabla_u f^{MW})_{\alpha\beta} = \langle \nabla_U F(U^P; \cdot) \Psi_\alpha \Psi_\beta \rangle = \sum_{\gamma, \delta \in \mathcal{P}} \langle \nabla_U F(U^P; \cdot) B_{\alpha\gamma} \Phi_\gamma B_{\beta\delta} \Phi_\delta \rangle \quad (2.18)$$

$$= \sum_{\gamma, \delta \in \mathcal{P}} B_{\alpha\gamma} (\nabla_u f)_{\gamma\delta} B_{\beta\delta} = (B \nabla_u f B^T)_{\alpha\beta}. \quad (2.19)$$

De plus,  $B$  est orthogonale de par l'orthonormalité des deux bases, ce qui implique que  $\nabla_u f^{MW}$  et  $\nabla_u f$  sont semblables et prouve l'équivalence des deux représentations par rapport à la diagonalisation dans  $\mathbb{R}$ .

L'avantage d'utiliser la base SE plutôt que la base MW est que la matrice Jacobienne de Galerkin  $\nabla_u f$  a une structure bloc diagonale grâce au découplage du

problème sur les différents éléments stochastiques. En effet,  $(\nabla_u f)_{\alpha\beta} = 0$  lorsque  $\text{supp}(\Phi_\alpha) \cap \text{supp}(\Phi_\beta)$  est de mesure nulle. Pour un élément stochastique donné  $\alpha_\sigma \in \mathcal{P}_\sigma$ , le bloc diagonal correspondant de taille  $m\mathcal{P}_\pi \times m\mathcal{P}_\pi$  est noté  $[\nabla_u f]^{\alpha_\sigma}$  et a pour expression :

$$([\nabla_u f(u)]^{\alpha_\sigma})_{\alpha_\pi, \beta_\pi \in \mathcal{P}_\pi} = \left\langle \nabla_U F(U^P; \cdot) \Phi_{\alpha_\sigma, \alpha_\pi} \Phi_{\alpha_\sigma, \beta_\pi} \right\rangle_{\alpha_\pi, \beta_\pi \in \mathcal{P}_\pi}. \quad (2.20)$$

Par conséquent,  $\nabla_u f$  est diagonalisable si et seulement si chaque bloc  $[\nabla_u f]^{\alpha_\sigma}$  est diagonalisable. Il suffit donc d'étudier la question de l'hyperbolicité du système de Galerkin pour le cas  $\text{Nr} = 0$ , *i.e.*, on se limite à un élément stochastique et on supprime l'indice  $\alpha_\sigma$  dans cette étude.

### 2.2.1 Systèmes hyperboliques stochastiques symétriques

**Théorème 2.1.** *On considère indifféremment une tensorisation polynomiale complète ou partielle pour l'espace d'approximation stochastique  $S^P$ . Si la matrice Jacobienne stochastique  $\nabla_U F(\cdot; \xi)$  est symétrique, alors la matrice Jacobienne de Galerkin  $\nabla_u f$  est diagonalisable dans  $\mathbb{R}$ . En particulier, la projection de Galerkin d'une loi de conservation scalaire conduit toujours à un système hyperbolique.*

*Démonstration.* Si  $\nabla_U F(\cdot; \xi)$  est symétrique, alors la matrice Jacobienne de Galerkin  $\nabla_u f$  définie par (2.17) est également symétrique et par conséquent diagonalisable dans  $\mathbb{R}$ .  $\square$

### 2.2.2 Vecteurs propres indépendants de l'incertitude

**Théorème 2.2.** *On considère indifféremment une tensorisation polynomiale complète ou partielle pour l'espace d'approximation stochastique  $S^P$ . Si les vecteurs propres de la matrice Jacobienne stochastique  $\nabla_U F(\cdot; \xi)$  sont indépendants de l'incertitude, alors la matrice Jacobienne de Galerkin  $\nabla_u f$  est diagonalisable dans  $\mathbb{R}$ .*

*Démonstration.* Si les vecteurs propres de  $\nabla_U F(\cdot; \xi)$  sont indépendants de  $\xi$ , alors la décomposition spectrale (1.82) devient

$$\nabla_U F(\cdot; \xi) = p_0 D(\xi) p_0^{-1}, \quad \text{avec} \quad p_0 = \begin{pmatrix} r_0^1 & \dots & r_0^m \end{pmatrix},$$

où  $r_0^1, \dots, r_0^m$  sont des vecteurs de  $\mathbb{R}^m$  indépendants de  $\xi$ . Un élément générique dans  $\nabla_u f$  peut être identifié pour le multi-indice  $(\alpha_i, \beta_j)$  avec  $i, j \in \{1, \dots, m\}$  et  $\alpha, \beta \in \mathcal{P}_\pi$ , de sorte que

$$\begin{aligned}
(\nabla_u f(u))_{\alpha i, \beta j} &= \left\langle \left( \nabla_U F(U^P; \cdot) \right)_{ij} \Phi_\alpha \Phi_\beta \right\rangle = \sum_k \left\langle (p_0)_{ik} \Lambda^k (p_0^{-1})_{kj} \Phi_\alpha \Phi_\beta \right\rangle \\
&= \sum_k (p_0)_{ik} \left\langle \Lambda^k \Phi_\alpha \Phi_\beta \right\rangle (p_0^{-1})_{kj} \\
&= \sum_{k, k'} \sum_{\gamma, \gamma'} \{ \delta_{\alpha\gamma} (p_0)_{ik} \} \{ \delta_{kk'} \left\langle \Lambda^k \Phi_\gamma \Phi_{\gamma'} \right\rangle \} \{ \delta_{\gamma'\beta} (p_0^{-1})_{k'j} \} \\
&= \sum_{k, k'} \sum_{\gamma, \gamma'} (r)_{\alpha i, \gamma k} (d)_{\gamma k, \gamma' k'} (q)_{\gamma' k', \beta j} = (q d r)_{\alpha i, \beta j},
\end{aligned}$$

où  $d$  est la matrice bloc diagonale de taille  $mP_\pi \times mP_\pi$  telle que

$$(d)_{\gamma k, \gamma' k'} = \delta_{kk'} \left\langle \Lambda^k \Phi_\gamma \Phi_{\gamma'} \right\rangle,$$

et  $r$  et  $q$  sont des matrices  $mP_\pi \times mP_\pi$  telles que

$$(r)_{\alpha i, \gamma k} = \delta_{\alpha\gamma} (p_0)_{ik}, \quad (q)_{\gamma k, \beta j} = \delta_{\gamma\beta} (p_0^{-1})_{kj}.$$

Chaque bloc de la diagonale de  $d$  est symétrique et par conséquent diagonalisable dans  $\mathbb{R}$  de sorte que  $d$  est diagonalisable dans  $\mathbb{R}$ . De plus,

$$(rq)_{\alpha i, \beta j} = \sum_{\gamma, k} (r)_{\alpha i, \gamma k} (q)_{\gamma k, \beta j} = \sum_{\gamma, k} (p_0)_{ik} \delta_{\alpha\gamma} \delta_{\gamma\beta} (p_0^{-1})_{kj} = \delta_{\alpha\beta} \delta_{ij},$$

ce qui signifie que  $q = r^{-1}$ . Ceci conclut la preuve.  $\square$

**Remarque 2.1.** *Le théorème 2.2 fournit une autre preuve de l'hyperbolicité du système de Galerkin pour une loi de conservation scalaire. En effet,  $W(\xi) = 1$  est le vecteur propre de  $\nabla_U F \in \mathbb{R}$ . Une application (plus) intéressante du théorème 2.2 est l'équation des ondes scalaires avec vitesse du son incertaine. Le théorème 2.2 peut également être appliqué aux systèmes hyperboliques linéaires où l'incertitude porte uniquement sur les conditions initiales ou les conditions aux limites.*

### 2.2.3 Cas général

Dans le cas général, on ne peut pas établir que la matrice Jacobienne de Galerkin est diagonalisable dans  $\mathbb{R}$ . Cependant, on peut identifier une approximation de  $\nabla_u f$  qui soit diagonalisable dans  $\mathbb{R}$  et pour laquelle les expressions des valeurs propres et vecteurs propres sont explicitement connues. Ainsi, considérons la matrice  $\overline{\nabla_u f}$ , également bloc diagonale, obtenue en approchant les coefficients de la matrice Jacobienne de Galerkin  $\nabla_u f$  par une quadrature de Gauss. Pour un élément stochastique donné  $\alpha_\sigma \in \mathcal{P}_\sigma$ , le bloc diagonal correspondant de taille  $mP_\pi \times mP_\pi$  est noté  $[\overline{\nabla_u f}]^{\alpha_\sigma}$  et a pour expression

$$\left( [\overline{\nabla_u f}(u)]^{\alpha_\sigma} \right)_{\alpha_\pi, \beta_\pi \in \mathcal{P}_\pi} = \left( \sum_{\gamma \in \mathcal{P}_\pi} \varpi_\gamma \nabla_U F(U^P(\xi_\gamma); \xi_\gamma) \Phi_{\alpha_\sigma, \alpha_\pi}(\xi_\gamma) \Phi_{\alpha_\sigma, \beta_\pi}(\xi_\gamma) \right)_{\alpha_\pi, \beta_\pi \in \mathcal{P}_\pi}, \quad (2.21)$$

où les  $\{\xi_\gamma\}_{\gamma \in \mathcal{P}_\pi}$  sont les  $P_\pi$  points de Gauss multidimensionnels dans l'élément stochastique  $\alpha_\sigma$  et  $\{\varpi_\gamma\}_{\gamma \in \mathcal{P}_\pi}$  les poids associés. Pour étudier la diagonalisation dans  $\mathbb{R}$  de  $\overline{\nabla_u f}$ , on peut faire l'hypothèse que  $N_r = 0$  puisque l'extension à  $N_r \geq 1$  est directe grâce à la structure bloc diagonale de la matrice Jacobienne. On commence par traiter le cas d'une dimension stochastique ( $N = 1$ ), cf remarque 2.2 pour le cas multidimensionnel ( $N > 1$ ). Soient  $\{\xi_\gamma\}_{\gamma \in \mathcal{P}_\pi}$  les  $P = P_\pi = (N_o + 1)$  points de Gauss points dans  $[0, 1]$ , i.e., les  $P_\pi$  zéros du polynôme de Legendre de degré  $P_\pi$ , et soient  $\{\varpi_\gamma\}_{\gamma \in \mathcal{P}_\pi}$  les poids de quadrature associés.

**Théorème 2.3.** *On considère le cas unidimensionnel  $N = 1$ . On suppose que la matrice Jacobienne stochastique  $\nabla_U F(\cdot; \xi)$  est définie aux  $P_\pi = (N_o + 1)$  points de Gauss dans  $\Xi = [0, 1]$ . On considère la matrice  $\overline{\nabla_u f}$  obtenue en approchant les coefficients de la matrice Jacobienne de Galerkin  $\nabla_u f$  par quadrature de Gauss, sous la forme*

$$\left(\overline{\nabla_u f}(u)\right)_{\alpha, \beta \in \mathcal{P}_\pi} = \left( \sum_{\gamma \in \mathcal{P}_\pi} \varpi_\gamma \nabla_U F(U^P(\xi_\gamma); \xi_\gamma) \Phi_\alpha(\xi_\gamma) \Phi_\beta(\xi_\gamma) \right)_{\alpha, \beta \in \mathcal{P}_\pi}. \quad (2.22)$$

Alors,  $\overline{\nabla_u f}$  est diagonalisable dans  $\mathbb{R}$  et ses valeurs propres et vecteurs propres à droite et à gauche sont donnés, pour tout  $k \in \{1, \dots, m\}$ , et tout  $\eta \in \mathcal{P}_\pi$ , par

$$\begin{cases} \lambda'_{k\eta} = \Lambda^k(\xi_\eta), \\ \left(r'_{k\eta}\right)_{\beta \in \mathcal{P}_\pi} = \left(\varpi_\eta R^k(\xi_\eta) \Phi_\beta(\xi_\eta)\right)_{\beta \in \mathcal{P}_\pi}, \\ \left(l'_{k\eta}\right)_{\beta \in \mathcal{P}_\pi} = \left(\varpi_\eta L^k(\xi_\eta) \Phi_\beta(\xi_\eta)\right)_{\beta \in \mathcal{P}_\pi}, \end{cases} \quad (2.23)$$

où  $\{\Lambda^k(\xi)\}_{k \in \{1, \dots, m\}}$ ,  $\{R^k(\xi)\}_{k \in \{1, \dots, m\}}$ , et  $\{L^k(\xi)\}_{k \in \{1, \dots, m\}}$  sont respectivement les valeurs propres et les vecteurs propres à droite et à gauche de la matrice Jacobienne stochastique  $\nabla_U F(U^P(\xi); \xi)$  définis par (1.83).

*Démonstration.* On considère uniquement les vecteurs propres à droite; la preuve pour les vecteurs propres à gauche est similaire. Soient  $k \in \{1, \dots, m\}$  et  $\eta \in \mathcal{P}_\pi$ . Soit  $V_{k\eta}(\xi)$  le polynôme de degré inférieur ou égal à  $N_o$  à valeurs dans  $\mathbb{R}^m$  tel que pour tout  $\gamma \in \mathcal{P}_\pi$ ,

$$V_{k\eta}(\xi_\gamma) = \delta_{\eta\gamma} R^k(\xi_\gamma),$$

où  $\delta_{\eta\gamma}$  est le symbole de Kronecker. En observant que la quadrature est exacte pour les polynômes de degré inférieur ou égal à  $(2P_\pi - 1)$ , on obtient, pour tout  $\beta \in \mathcal{P}_\pi$ ,

$$\langle \Phi_\beta, V_{k\eta} \rangle = \sum_{\gamma \in \mathcal{P}_\pi} \varpi_\gamma \Phi_\beta(\xi_\gamma) V_{k\eta}(\xi_\gamma) = \varpi_\eta \Phi_\beta(\xi_\eta) R^k(\xi_\eta) = (r'_{k\eta})_\beta.$$

Puisque la famille  $(\Phi_\beta)_{\beta \in \mathcal{P}_\pi}$  est orthogonale dans  $L^2(\Xi, p_\xi)$ , on a pour tout  $\xi \in \Xi$ ,

$$V_{k\eta}(\xi) = \sum_{\beta \in \mathcal{P}_\pi} (r'_{k\eta})_\beta \Phi_\beta(\xi). \quad (2.24)$$



Soit  $\alpha \in \mathcal{P}_\pi$ . On observe que

$$\begin{aligned} (\overline{\nabla_u f} r'_{k\eta})_\alpha &= \sum_{\beta \in \mathcal{P}_\pi} \left( \sum_{\gamma \in \mathcal{P}_\pi} \varpi_\gamma \nabla_U F(U^P(\xi_\gamma); \xi_\gamma) \Phi_\alpha(\xi_\gamma) \Phi_\beta(\xi_\gamma) \right) (r'_{k\eta})_\beta \\ &= \sum_{\gamma \in \mathcal{P}_\pi} \varpi_\gamma \nabla_U F(U^P(\xi_\gamma); \xi_\gamma) \Phi_\alpha(\xi_\gamma) \left( \sum_{\beta \in \mathcal{P}_\pi} (r'_{k\eta})_\beta \Phi_\beta(\xi_\gamma) \right) \\ &= \sum_{\gamma \in \mathcal{P}_\pi} \varpi_\gamma \nabla_U F(U^P(\xi_\gamma); \xi_\gamma) \Phi_\alpha(\xi_\gamma) V_{k\eta}(\xi_\gamma), \end{aligned}$$

où on a utilisé (2.24) en  $\xi_\gamma$ . Ainsi, puisque  $V_{k\eta}(\xi_\gamma) = \delta_{\gamma\eta} R^k(\xi_\eta)$  et  $R^k(\xi_\eta)$  est un vecteur propre à droite de  $\nabla_U F(U^P(\xi_\eta); \xi_\eta)$ , on en déduit

$$\begin{aligned} (\overline{\nabla_u f} r'_{k\eta})_\alpha &= \varpi_\eta \nabla_U F(U^P(\xi_\eta); \xi_\eta) R^k(\xi_\eta) \Phi_\alpha(\xi_\eta) \\ &= \varpi_\eta \Lambda^k(\xi_\eta) R^k(\xi_\eta) \Phi_\alpha(\xi_\eta) = \lambda_{k\eta} (r'_{k\eta})_\alpha. \end{aligned}$$

Pour clore la preuve, il reste à vérifier que la famille  $(r'_{k\eta})_{k \in \{1, \dots, m\}, \eta \in \mathcal{P}_\pi}$  forme une base de  $\mathbb{R}^{m\mathcal{P}_\pi}$ . Ainsi, soient  $m\mathcal{P}_\pi$  nombres réels  $(a_{k\eta})_{k \in \{1, \dots, m\}, \eta \in \mathcal{P}_\pi}$  tels que  $\sum_{k=1}^m \sum_{\eta \in \mathcal{P}_\pi} a_{k\eta} r'_{k\eta} = 0$ . On considère le vecteur stochastique

$$A(\xi) = \sum_{k=1}^m \sum_{\eta \in \mathcal{P}_\pi} \sum_{\beta \in \mathcal{P}_\pi} a_{k\eta} (r'_{k\eta})_\beta \Phi_\beta(\xi).$$

En évaluant ce vecteur en  $\xi_\gamma$  pour chaque  $\gamma \in \mathcal{P}_\pi$ , on obtient

$$A(\xi_\gamma) = \sum_{k=1}^m \sum_{\eta \in \mathcal{P}_\pi} a_{k\eta} \left( \sum_{\beta \in \mathcal{P}_\pi} (r'_{k\eta})_\beta \Phi_\beta(\xi_\gamma) \right) = \sum_{k=1}^m \sum_{\eta \in \mathcal{P}_\pi} a_{k\eta} V_{k\eta}(\xi_\gamma),$$

grâce à (2.24) de sorte que, par définition de  $V_{k\eta}$ ,

$$A(\xi_\gamma) = \sum_{k=1}^m a_{k\gamma} R^k(\xi_\gamma).$$

Puisque, par hypothèse, le vecteur  $A(\xi_\gamma)$  vaut zéro et que les vecteurs  $(R^k(\xi_\gamma))_{k \in \{1, \dots, m\}}$  forment une base complète de  $\mathbb{R}^m$ , on obtient  $a_{k\gamma} = 0$  pour tout  $k \in \{1, \dots, m\}$ . Puisque  $\gamma$  est arbitraire, cela complète la preuve.  $\square$

La première application de ce résultat (cf section 2.3) sera d'utiliser le spectre de  $\overline{\nabla_u f}$  comme approximation du spectre de  $\nabla_u f$  pour construire un fit polynomial qui, appliqué à  $\nabla_u f$ , permettra d'approcher sa valeur absolue dans le contexte des matrices de décentrement pour les solveurs de type Roe. Le calcul direct de la valeur absolue de  $\nabla_u f$  par l'intermédiaire de la diagonalisation sera ainsi évité, méthode qui pourrait s'avérer d'autant plus coûteuse que la dimension stochastique est grande. Ensuite, l'utilisation des vecteurs propres de  $\overline{\nabla_u f}$  comme vecteurs approchés de  $\nabla_u f$  permettra d'étendre le correcteur entropique de Dubois et Mehlmann au solveur de Roe (cf section 2.4).

**Remarque 2.2.** *L'extension du théorème 2.3 au cas multidimensionnel ( $N > 1$ ) est directe dans le cas d'une tensorisation polynomiale complète grâce à la structure tensorisée de la base polynomiale, de sorte que les points de Gauss multidimensionnels sont simplement obtenus par tensorisation des points de Gauss unidimensionnels. Lorsque l'on travaille avec une tensorisation polynomiale partielle, le théorème 2.3 peut néanmoins être utilisé. Dans ce cas, les valeurs propres approchées servent uniquement de données pour construire le fit polynomial en vue d'évaluer la matrice de décentrement dans le solveur de type Roe. Ce point sera discuté plus loin dans la section 2.3.3 et illustré numériquement dans la section 3.1.2.3.*

## 2.3 Approximation en espace et en temps

Le système de Galerkin (2.13) est discrétisé en espace physique et en temps en utilisant une méthode Volumes Finis (VF), cf par exemple les livres de Godlewski et Raviart [33], de Toro [80] et de Després [24]. On considère un pas d'espace  $\Delta x$  (uniforme pour simplifier) et des temps discrets  $t^n$  avec un pas de temps  $\Delta t^n = t^{n+1} - t^n$  vérifiant une condition CFL spécifiée plus bas. Le schéma VF est de la forme

$$u_i^{n+1} = u_i^n - \frac{\Delta t^n}{\Delta x} (\varphi(u_i^n, u_{i+1}^n) - \varphi(u_{i-1}^n, u_i^n)), \quad (2.25)$$

où  $u_i^n$  est une approximation de la valeur moyenne en espace de la solution  $u$  dans la cellule de centre  $i\Delta x$  de taille  $\Delta x$  au temps  $t^n$  et  $\varphi(\cdot, \cdot)$  est le flux numérique de Galerkin. À une interface donnée  $LR$  séparant les états gauche et droit indexés respectivement par  $L$  et  $R$ , le flux numérique est choisi de la forme

$$\varphi(u_L, u_R) = \frac{f(u_L) + f(u_R)}{2} - a \frac{u_R - u_L}{2}, \quad (2.26)$$

où  $\frac{1}{2}(f(u_L) + f(u_R))$  est la partie centrée du flux et  $a \in \mathbb{R}^{m^P, m^P}$  est une matrice de décentrement (non négative). Il s'agit donc d'un schéma du premier ordre pour la discrétisation en espace.

On souhaite construire un solveur de type Roe pour le système de Galerkin (2.13), qui sera appelé **solveur de Roe stochastique**. Comme dans le cas déterministe, on considère le solveur de Riemann approché où le flux de Galerkin  $f(u)$  est remplacé à chaque interface  $LR$  séparant les états gauche et droit  $(u_L, u_R)$  par l'approximation linéaire

$$f^{\text{Roe}}(u_L, u_R, u) = f(u_L) + a^{\text{Roe}}(u_L, u_R) \cdot (u - u_L), \quad (2.27)$$

où  $a^{\text{Roe}}(u_L, u_R)$  est une approximation de Roe linéarisée de la matrice Jacobienne de Galerkin. Dans la section 2.3.1, on prouve qu'une telle matrice existe sous certaines hypothèses. Le flux numérique de Galerkin est alors choisi de la forme

$$\varphi(u_L, u_R) \equiv \varphi^{\text{Roe}}(u_L, u_R) := \frac{f(u_L) + f(u_R)}{2} - |a^{\text{Roe}}(u_L, u_R)| \frac{u_R - u_L}{2}. \quad (2.28)$$

Une méthode efficace pour approcher la valeur absolue d'une matrice quelconque est proposée en section 2.3.2 et la construction de la matrice de décentrement  $|a^{\text{Roe}}(u_L, u_R)|$  est détaillée dans la section 2.3.3.

### 2.3.1 Matrice de Roe et état de Roe

On fait l'hypothèse que le système hyperbolique stochastique de départ (1.77) possède une matrice de Roe  $A^{\text{Roe}}(U_L, U_R; \xi) \in \mathbb{R}^{m,m}$  presque sûrement. Rappelons que  $A^{\text{Roe}}(U_L, U_R; \xi)$  vérifie les propriétés suivantes :

- $A^{\text{Roe}}(U_L, U_R; \xi)$  est diagonalisable dans  $\mathbb{R}$ ,  $\forall U_L, U_R \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$ .
- Consistance avec la matrice Jacobienne stochastique  $\nabla_U F$ ,

$$A^{\text{Roe}}(U, U; \xi) = \nabla_U F(U; \xi), \quad \forall U \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi). \quad (2.29)$$

- Conservativité au travers des chocs,

$$F(U_R; \xi) - F(U_L; \xi) = A^{\text{Roe}}(U_L, U_R; \xi)(U_R - U_L), \quad \forall U_L, U_R \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi). \quad (2.30)$$

Le résultat suivant identifie une matrice qui joue le rôle de matrice de Roe pour le système de Galerkin.

**Théorème 2.4.** *On considère indifféremment une tensorisation polynomiale complète ou partielle pour l'espace stochastique  $S^P$ . Sous l'hypothèse ci-dessus sur l'existence de la matrice  $A^{\text{Roe}}(\cdot, \cdot; \xi)$ , la matrice  $a(u_L, u_R) \in \mathbb{R}^{m^P, m^P}$  définie pour tout  $u_L, u_R \in \mathcal{A}_u$  par*

$$a(u_L, u_R) := \left( \left\langle A^{\text{Roe}}(U_L^P, U_R^P; \cdot) \Phi_\alpha \Phi_\beta \right\rangle \right)_{\alpha, \beta \in \mathcal{P}} \quad (2.31)$$

avec  $U_L^P(\xi) = \sum_{\alpha \in \mathcal{P}} (u_L)_\alpha \Phi_\alpha(\xi)$  et  $U_R^P(\xi) = \sum_{\alpha \in \mathcal{P}} (u_R)_\alpha \Phi_\alpha(\xi)$ , vérifie les propriétés suivantes :

- Consistance avec la matrice Jacobienne de Galerkin  $\nabla_u f$ ,

$$a(u, u) = \nabla_u f(u), \quad \forall u \in \mathcal{A}_u. \quad (2.32)$$

- Conservativité au travers des chocs,

$$f(u_R) - f(u_L) = a(u_L, u_R)(u_R - u_L), \quad \forall u_L, u_R \in \mathcal{A}_u. \quad (2.33)$$

*Démonstration.* Pour prouver la consistance avec la matrice Jacobienne de Galerkin, on observe que pour tout  $u \in \mathcal{A}_u$ , en posant  $U^P = \sum_{\alpha \in \mathcal{P}} u_\alpha \Phi_\alpha(\xi)$ , il vient

$$\begin{aligned} a(u, u) &= \left( \left\langle A^{\text{Roe}}(U^P, U^P; \cdot) \Phi_\alpha \Phi_\beta \right\rangle \right)_{\alpha, \beta \in \mathcal{P}} \\ &= \left( \left\langle \nabla_U F(U^P; \cdot) \Phi_\alpha \Phi_\beta \right\rangle \right)_{\alpha, \beta \in \mathcal{P}} = \nabla_u f(u). \end{aligned}$$

Pour prouver la conservativité au travers des chocs, on observe que pour tout  $u_L, u_R \in \mathcal{A}_u$  et tout  $\alpha \in \mathcal{P}$ , en posant  $U_L^P = \sum_{\beta \in \mathcal{P}} (u_L)_\beta \Phi_\beta(\xi)$  et  $U_R^P =$

$\sum_{\beta \in \mathcal{P}} (u_R)_\beta \Phi_\beta(\xi)$ , il vient

$$\begin{aligned}
(f(u_R) - f(u_L))_\alpha &= \left\langle \left( F(U_R^P; \cdot) - F(U_L^P; \cdot) \right) \Phi_\alpha \right\rangle \\
&= \left\langle A^{\text{Roe}}(U_L^P, U_R^P; \cdot) (U_R^P - U_L^P) \Phi_\alpha \right\rangle \\
&= \left\langle A^{\text{Roe}}(U_L^P, U_R^P; \cdot) \sum_{\beta \in \mathcal{P}} \left( \langle \Phi_\beta U_R^P \rangle - \langle \Phi_\beta U_L^P \rangle \right) \Phi_\alpha \Phi_\beta \right\rangle \\
&= \sum_{\beta \in \mathcal{P}} \left\langle A^{\text{Roe}}(U_L^P, U_R^P; \cdot) \Phi_\alpha \Phi_\beta \right\rangle \left( \langle \Phi_\beta U_R^P \rangle - \langle \Phi_\beta U_L^P \rangle \right) \\
&= \sum_{\beta \in \mathcal{P}} (a(u_L, u_R))_{\alpha\beta} \left( (u_R)_\beta - (u_L)_\beta \right).
\end{aligned}$$

Ceci complète la preuve.  $\square$

On suppose enfin que pour tout  $U_L, U_R \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$ , il existe un état de Roe  $U_{LR}^{\text{Roe}} \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$  presque sûrement tel que

$$A^{\text{Roe}}(U_L, U_R; \xi) = \nabla_U F(U_{LR}^{\text{Roe}}; \xi). \quad (2.34)$$

Dans ces conditions, pour tout  $U_L^P, U_R^P \in \mathcal{A}_U \otimes S^P$ , on considère  $U_{LR}^{\text{Roe}} \in \mathcal{A}_U \otimes L^2(\Xi, p_\xi)$  tel que  $A^{\text{Roe}}(U_L^P, U_R^P; \xi) = \nabla_U F(U_{LR}^{\text{Roe}}; \xi)$  et on définit

$$a_{LR}^{\text{Roe}} := a(u_L, u_R) = \left( \left\langle \nabla_U F(U_{LR}^{\text{Roe}}; \cdot) \Phi_\alpha \Phi_\beta \right\rangle \right)_{\alpha, \beta \in \mathcal{P}}. \quad (2.35)$$

Alors, si  $a_{LR}^{\text{Roe}}$  est diagonalisable dans  $\mathbb{R}$ , cette matrice est une matrice linéarisée de Roe pour le système de Galerkin (2.13).

### 2.3.2 Une méthode efficace pour approcher la valeur absolue d'une matrice

La méthode décrite dans cette section a une portée plus large que l'approximation du système de Galerkin par un solveur de Roe. On la présente donc dans un cadre général; l'application au solveur de Roe est détaillée dans la section 2.3.3. Soit  $A$  une matrice diagonalisable dans  $\mathbb{R}$  de taille  $N_A$ . Par définition,  $|A|$  est la matrice co-diagonalisable avec  $A$  dont les valeurs propres sont les valeurs absolues des valeurs propres de  $A$ ,

$$|A| = \sum_{\gamma=1}^{N_A} |\lambda_\gamma| r_\gamma \otimes l_\gamma, \quad (2.36)$$

où  $\{\lambda_\gamma\}_{\gamma \in \{1, \dots, N_A\}}$  sont les valeurs propres réelles de  $A$ ,  $\{l_\gamma\}_{\gamma=1, \dots, N_A}$  les vecteurs propres à gauche, et  $\{r_\gamma\}_{\gamma \in \{1, \dots, N_A\}}$  les vecteurs propres à droite. Il est possible de diagonaliser  $A$  et de calculer  $|A|$  en utilisant (2.36), mais en pratique cette méthode est extrêmement coûteuse lorsque  $N_A$  est grand.

Une méthode intéressante, proposée par Ndjinga dans [66], consiste à calculer une séquence d'itérations polynomiales basées sur la connaissance exacte des valeurs

propres (ou au moins une borne explicite sur celles-ci) et convergeant vers la matrice signe si les valeurs propres sont réelles. Cependant, cette méthode devient également coûteuse quand  $N_A$  devient grand. Une autre méthode, proposée par Degond et. al dans [19], est basée sur le calcul d'un polynôme qui interpole certaines valeurs absolues des valeurs propres de  $A$ . On propose ici une nouvelle méthode basée sur le calcul d'un unique polynôme de bas degré. Par rapport aux méthodes ci-dessus, cette nouvelle méthode est à la fois moins coûteuse et mieux adaptée aux situations où seulement des approximations des valeurs propres sont connues.

On note  $\{\lambda'_\gamma\}_{\gamma \in \{1, \dots, N_A\}}$  les valeurs propres approchées de  $A$ . La méthode consiste à trouver un polynôme  $q_{d_{\text{fit}}, \{\lambda'\}}$  de degré  $d_{\text{fit}}$  ( $d_{\text{fit}}$  est fixé *a priori*) qui minimise l'erreur au sens des moindres carrés entre  $|\lambda'_\gamma|$  et  $q_{d_{\text{fit}}, \{\lambda'\}}(\lambda'_\gamma)$ , et ensuite à appliquer ce polynôme à la matrice  $A$  pour approcher  $|A|$ . Soit  $\Pi_{d_{\text{fit}}}$  l'espace vectoriel des polynômes sur  $\mathbb{R}$  de degré  $\leq d_{\text{fit}}$ . On cherche  $q_{d_{\text{fit}}, \{\lambda'\}} \in \Pi_{d_{\text{fit}}}$  tel que

$$q_{d_{\text{fit}}, \{\lambda'\}} = \arg \min_{q \in \Pi_{d_{\text{fit}}}} \sum_{\gamma=1}^{N_A} \left( |\lambda'_\gamma| - q(\lambda'_\gamma) \right)^2. \quad (2.37)$$

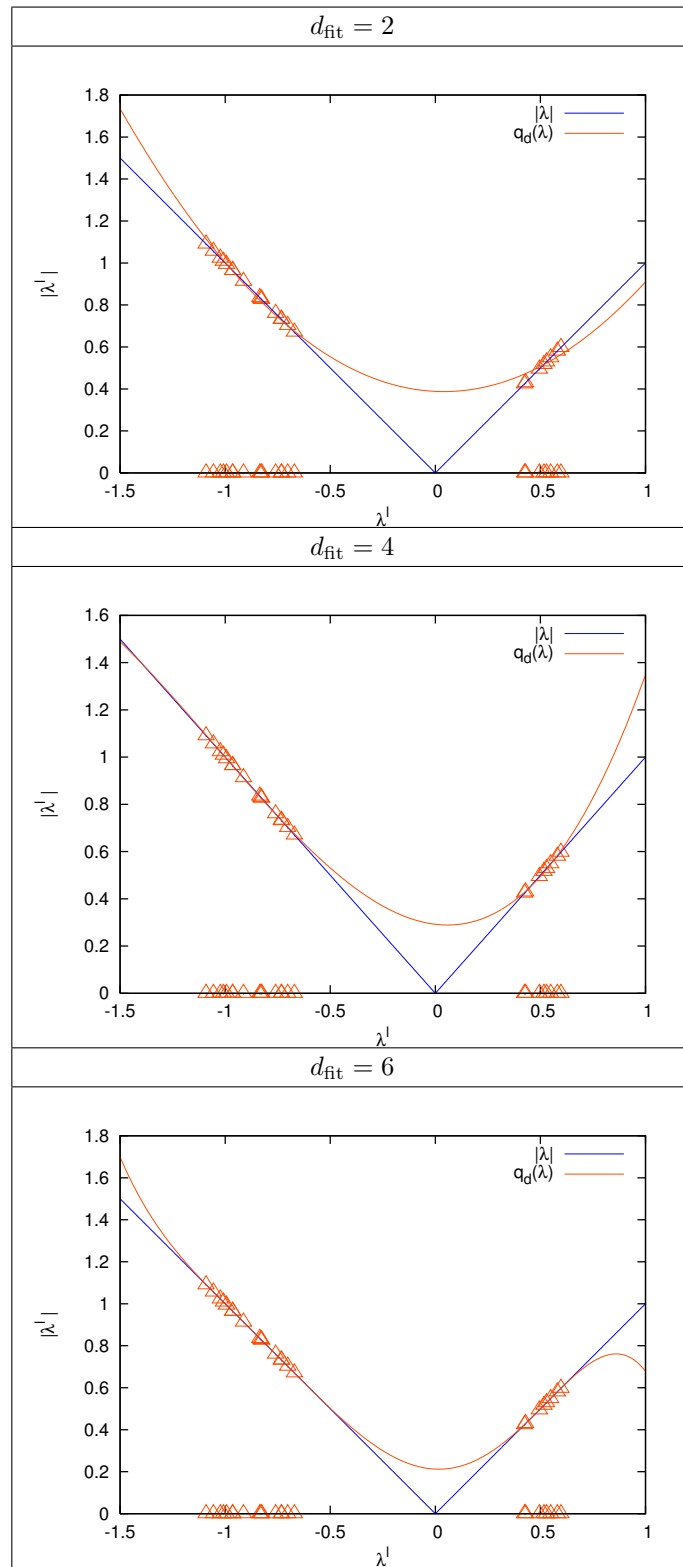
En écrivant  $q(X) = \sum_{j=0}^{d_{\text{fit}}} c_j X^j \in \Pi_{d_{\text{fit}}}$ , ce problème de minimisation est équivalent à la résolution d'un système linéaire avec les coefficients polynomiaux  $(c_j)_{j \in \{0, \dots, d_{\text{fit}}\}}$  comme inconnues. Ce système de taille  $(d_{\text{fit}} + 1) \times (d_{\text{fit}} + 1)$  s'écrit

$$\begin{pmatrix} \sum_{\gamma=1}^{N_A} \lambda'^0_\gamma \lambda'^0_\gamma & \dots & \sum_{\gamma=1}^{N_A} \lambda'^0_\gamma \lambda'^{d_{\text{fit}}}_\gamma \\ \vdots & \ddots & \vdots \\ \sum_{\gamma=1}^{N_A} \lambda'^{d_{\text{fit}}}_\gamma \lambda'^0_\gamma & \dots & \sum_{\gamma=1}^{N_A} \lambda'^{d_{\text{fit}}}_\gamma \lambda'^{d_{\text{fit}}}_\gamma \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_{d_{\text{fit}}} \end{pmatrix} = \begin{pmatrix} \sum_{\gamma=1}^{N_A} |\lambda'_\gamma| \lambda'^0_\gamma \\ \vdots \\ \sum_{\gamma=1}^{N_A} |\lambda'_\gamma| \lambda'^d_\gamma \end{pmatrix}. \quad (2.38)$$

La matrice du système (2.38) est symétrique et elle est définie positive dès que  $d \leq N_A$  lorsque les valeurs propres approchées  $\lambda'_\gamma$  sont distinctes, ce qui garantit l'existence d'une solution unique au système. Lorsque ce n'est pas le cas, on utilise une méthode de décomposition en valeurs singulières (*cf* remarque 2.3 en section 2.3.3 pour un exemple). La résolution du système linéaire (2.38) fournit les coefficients  $(c_j)_{j \in \{0, \dots, d\}}$  qui définissent le polynôme  $q_{d_{\text{fit}}, \{\lambda'\}}$ . On applique ensuite ce polynôme à la matrice  $A$  et on obtient une approximation de la matrice  $|A|$ . Pour plus d'efficacité, la méthode de Hörner [79, p. 44] peut être utilisée :  $q_{d_{\text{fit}}, \{\lambda'\}}(A)$  est réécrit sous la forme suivante

$$q_{d_{\text{fit}}, \{\lambda'\}}(A) = c_0 I + (c_1 I + (c_2 I + \dots + (c_{d_{\text{fit}}-1} I + c_{d_{\text{fit}}} A) \dots A) A). \quad (2.39)$$

Le nombre de produits matrice-matrice est ainsi réduit à  $d_{\text{fit}}$  au lieu de  $d_{\text{fit}}(d_{\text{fit}} - 1)/2$  si toutes les puissances de la matrice étaient calculées séparément. En outre, on peut encore réduire le coût de calcul dans le cas du solveur de Roe puisqu'on cherche uniquement à évaluer le produit de la matrice  $|A|$  par un vecteur  $x$  donné. En calculant directement  $q_{d_{\text{fit}}, \{\lambda'\}}(A)x$ , le coût est réduit à  $d_{\text{fit}}$  produits matrice-vecteur au lieu de  $d_{\text{fit}}$  produits matrice-matrice. Des exemples de polynômes obtenus par la méthode ci-dessus sont présentés sur la figure 2.2 pour différents degrés  $d_{\text{fit}}$  pour des valeurs propres générés aléatoirement (et réparties dans deux sous-ensembles distincts).

FIGURE 2.2 – Exemples de polynômes  $q_{d_{\text{fit}},\{\lambda^l\}}$  pour différents degrés  $d_{\text{fit}}$ .

### 2.3.3 Le schéma de décentrement

On applique la méthode présentée dans la section précédente pour approcher la valeur absolue de  $a_{LR}^{\text{Roe}} \in \mathbb{R}^{mP, mP}$  définie par (2.35) à chaque interface  $LR$  du domaine spatial. On exploite la structure bloc diagonale de la matrice  $a_{LR}^{\text{Roe}}$  pour évaluer un fit polynomial sur chaque élément stochastique  $\alpha_\sigma \in \mathcal{P}_\sigma$  indépendamment. L'avantage est qu'ainsi, le polynôme associé doit approcher moins de points, de sorte que les calculs sont à la fois plus efficaces et plus précis. Par conséquent, le flux numérique (2.28) est évalué séparément sur chaque élément stochastique  $\alpha_\sigma \in \mathcal{P}_\sigma$  en utilisant un polynôme spécifique noté  $q_{d_{\text{fit}}, \{\lambda'_\gamma\}}^\sigma$ . En supposant que l'on dispose bien des expressions explicites des valeurs propres  $\Lambda^1(\cdot; \xi), \dots, \Lambda^m(\cdot; \xi)$  de la matrice Jacobienne stochastique  $\nabla_U F$  (c'est le cas pour les lois de conservation scalaires et les équations d'Euler par exemple), on évalue ces valeurs propres en  $U_{LR}^{\text{Roe}}(\xi)$  et aux points de Gauss de chaque élément stochastique. On obtient ainsi les valeurs propres approchées  $\{\lambda'_\gamma\}_{\gamma \in \{1, \dots, mP_\pi\}}$  de  $[a_{LR}^{\text{Roe}}]^{\alpha_\sigma}$  dans chaque élément stochastique. En d'autres termes, on utilise les valeurs propres de la matrice  $\overline{\nabla_{uf}}$  définie au théorème 2.3 (une remarque très importante est que cette matrice n'est jamais construite ou évaluée).

Dans le cas d'une tensorisation polynomiale complète, le nombre de valeurs propres approchées et exactes est le même. Dans le cas d'une tensorisation polynomiale partielle, il y a plus de valeurs propres approchées que de valeurs propres exactes. Il est néanmoins raisonnable de penser que le fit polynomial obtenu approche relativement bien les valeurs propres exactes (pourvu que la matrice Jacobienne de Galerkin exacte soit diagonalisable dans  $\mathbb{R}$ ). En effet, grâce à la localisation sur chaque élément stochastique, on s'attend à ce que les valeurs propres soient groupées autour des valeurs propres de la matrice Jacobienne stochastique  $\nabla_U F$  dans chaque élément stochastique. On renvoie à la fin de la section 3.1.2.3 pour un exemple. À l'inverse, puisque l'évaluation du fit polynomial ne requiert que la connaissance de valeurs propres approchées de la matrice Jacobienne de Galerkin, on pourrait utiliser les valeurs propres de la matrice Jacobienne stochastique évaluées sur une grille creuse ou sur un échantillon de réalisations du paramètre aléatoire dans chaque élément stochastique. Cela permettrait ainsi de réduire le coût de l'évaluation du spectre approché pour le calcul du fit polynomial dans le cas de grande dimension stochastique; cette idée reste néanmoins à explorer.

**Remarque 2.3.** *Le système linéaire (2.38) peut être singulier si le nombre de valeurs propres approchées distinctes est plus petit que  $d_{\text{fit}}$ . En particulier, ceci a lieu dans le cas où l'interface  $LR$  n'a pas encore été atteinte par les ondes en provenance des zones d'où se propagent les incertitudes.*

Le flux numérique dans le schéma de volumes finis (2.25) est donc approché par

$$\varphi^{\text{Roe}}(u_L, u_R) \approx \frac{f(u_L) + f(u_R)}{2} - \text{diag}_\sigma(q_{d_{\text{fit}}, \{\lambda'_\gamma\}}^\sigma([a_{LR}^{\text{Roe}}]^\sigma)) \frac{u_R - u_L}{2}. \quad (2.40)$$

On observera que (2.40) fournit bien un flux numérique (décentré) associé au système de Galerkin (2.13). On utilise certes des données obtenues par collocation pour

calculer les polynômes  $q_{d_{\text{fit}},\{\lambda'_\gamma\}}^\sigma$ , mais ces polynômes sont appliqués aux matrices Jacobiennes de Galerkin de Roe  $[a_{LR}^{\text{Roe}}]^\alpha$ .

Le pas de temps  $\Delta t^n$  est sélectionné grâce à une condition CFL basée sur la plus grande vitesse caractéristique sur les cellules de discrétisation spatiales et stochastiques. En pratique,  $\Delta t^n$  est calculé tel que

$$\frac{\Delta t^n}{\Delta x} = \frac{C_{\text{CFL}}}{\max_{LR \in \mathcal{I}, \gamma=1, \dots, mP} |\lambda'_\gamma|}, \quad (2.41)$$

où  $\mathcal{I}$  est l'ensemble des interfaces  $LR$  et  $\{\lambda'_\gamma\}_{\gamma \in \{1, \dots, mP\}}$  sont les valeurs propres approchées identifiées ci-dessus. En d'autres termes, le maximum des valeurs propres sur tout le domaine stochastique est pris en compte pour déterminer le pas de temps et ce maximum est évalué en considérant les valeurs propres aux points de Gauss sur tous les éléments stochastiques. Par la suite, on fixe la constante  $C_{\text{CFL}}$  à 0.95.

**Remarque 2.4.** *On observe que la matrice  $q_{d_{\text{fit}},\{\lambda'_\gamma\}}^\sigma$  ( $[a_{LR}^{\text{Roe}}]^\sigma$ ) ne garantit pas le contrôle des valeurs propres de la matrice Jacobienne de Galerkin. En effet, des valeurs propres approchées ont été utilisées pour construire  $q_{d_{\text{fit}},\{\lambda'_\gamma\}}^\sigma$  et, de plus, ce polynôme ne fournit qu'un fit au sens des moindres carrés des valeurs propres. Ce problème peut éventuellement être résolu en augmentant la résolution stochastique ou en augmentant le degré polynomial  $d_{\text{fit}}$ . Ces aspects seront explorés numériquement dans la section 3.1.2.3. De plus, il est également possible de diminuer la constante  $C_{\text{CFL}}$  pour se prémunir du cas où les valeurs propres sont sous-estimées.*

## 2.4 Correcteur entropique

Bien que stabilisant correctement la méthode numérique, le schéma de Roe est connu dans le cas déterministe pour créer des solutions non physiques (violant la condition d'entropie) en présence de points soniques. Cette situation se présente lorsque certaines valeurs propres de la matrice Jacobienne du système changent de signe à une interface. Dans une telle situation, un correcteur est nécessaire pour construire des flux numériques en adéquation avec la physique. Comme dans le cas déterministe, des chocs non physiques peuvent se produire dans le cas stochastique au voisinage de points soniques dans les détentes. Ce point sera illustré dans la section 3.1.3. Dans la littérature, différents correcteurs entropiques ont été proposés pour le schéma de Roe. Les plus utilisés sont ceux de Harten [36] et Roe [73] et consistent à ajouter un terme diffusif linéaire dans les situations où le flux de Roe linéarisé produit une solution qui viole la condition d'entropie. En fait, ces corrections agissent comme des outils qui permettent "d'étaler" la solution approchée, plutôt que comme un remède au fait que le flux approché est linéaire dans des situations où une description non-linéaire du flux est nécessaire pour mieux respecter la physique. Le correcteur de Dubois et Mehlman (DM) [27] est plus général car il consiste en une modification non-linéaire du flux au voisinage des points soniques. De plus, ce correcteur est non paramétrique. Dans cette section, on propose un



correcteur entropique pour le solveur de Roe stochastique basé sur l'adaptation du correcteur DM au contexte du système de Galerkin.

### 2.4.1 Principes généraux

Le correcteur DM s'applique aux systèmes hyperboliques déterministes pour lesquels un solveur de Roe peut être utilisé. L'idée consiste à ajouter une correction au flux numérique de Roe  $\varphi_{\text{Roe}}(\cdot, \cdot)$  afin d'éviter les chocs non entropiques, la correction étant basée sur une modification non linéaire du flux numérique  $\varphi_{\text{Roe}}(\cdot, \cdot)$  au voisinage des points soniques. La détection et le traitement de tels chocs impliquent trois étapes :

1. les valeurs propres correspondant à la matrice Jacobienne en l'état de Roe à chaque interface physique séparant les états gauche et droit sont rangées par ordre croissant, et les vecteurs propres correspondants sont ordonnés en conséquence ;
2. les composantes des différences entre les états gauche et droit sur la base des vecteurs propres sont déterminées et les états intermédiaires associés sont calculés ;
3. une onde de détente sonique est détectée entre les  $(i - 1)$ -ème et  $i$ -ème états intermédiaires si la  $i$ -ème valeur propre correspondant au  $(i - 1)$ -ème état intermédiaire est négative alors que la  $i$ -ème valeur propre correspondant au  $i$ -ème état intermédiaire est positive. Dans ce cas, on ajoute une modification non linéaire au flux numérique  $\varphi^{\text{Roe}}(\cdot, \cdot)$ .

Un aspect important du correcteur DM est la correspondance entre les ensembles de valeurs propres pour tous les états intermédiaires. Cela est reflété par le fait que le correcteur DM a été appliqué à des systèmes déterministes strictement hyperboliques où la matrice Jacobienne a des valeurs propres réelles distinctes et un ensemble complet de vecteurs propres indépendants. De plus, les états gauche et droit à chaque interface sont supposés suffisamment proches, *cf* [27, théorème 6.1], de sorte que la correspondance entre les valeurs propres peut être obtenue par le biais d'un état commun de référence.

### 2.4.2 Application directe au système de Galerkin

On présente tout d'abord une application directe du correcteur DM au schéma numérique de la section 2.3 ; des adaptations seront introduites à la section suivante. Comme précédemment, on travaille sur chaque élément stochastique indépendamment. A chaque interface  $LR$  de l'espace physique avec des états gauche et droit  $u_L$  et  $u_R$  respectivement, on fait l'hypothèse pour l'instant que la matrice Jacobienne de Galerkin de Roe  $[a_{LR}^{\text{Roe}}]^{\alpha_\sigma}$  associée à l'élément stochastique  $\alpha_\sigma \in \mathcal{P}_\sigma$  est diagonalisable dans  $\mathbb{R}$  avec des valeurs propres **distinctes**. Ensuite, on écrit

$$[a_{LR}^{\text{Roe}}]^{\alpha_\sigma} = \sum_{\gamma=1}^{mP_\sigma} \lambda_\gamma(u_{LR}^{\text{Roe}}) r_\gamma(u_{LR}^{\text{Roe}}) \otimes l_\gamma(u_{LR}^{\text{Roe}}), \quad (2.42)$$

avec les valeurs propres  $(\lambda_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$ , les vecteurs propres à droite  $(r_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$  et les vecteurs propres à gauche  $(l_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$  (l'indice supérieur  $\alpha_\sigma$  est omis pour simplifier). La détection des points soniques nécessite la construction des  $(mP_\pi + 1)$  états intermédiaires  $(u_\gamma)_{\gamma \in \{0, \dots, mP_\pi\}}$  tels que

$$u_0 = u_L, \quad (2.43)$$

$$u_\gamma = u_{\gamma-1} + \beta_\gamma r_\gamma(u_{LR}^{\text{Roe}}), \quad \forall \gamma \in \{1, \dots, mP_\pi\}, \quad (2.44)$$

$$u_R = u_{mP_\pi}, \quad (2.45)$$

où  $(\beta_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$  sont les composantes du vecteur  $u_R - u_L$  dans la base des vecteurs propres à droite,

$$u_R - u_L = \sum_{\gamma=1}^{mP_\pi} \beta_\gamma r_\gamma(u_{LR}^{\text{Roe}}). \quad (2.46)$$

Ensuite, l'ensemble des indices soniques est défini comme

$$S = \{\gamma \in \{1, \dots, mP_\pi\}, \lambda_\gamma(u_{\gamma-1}) < 0 < \lambda_\gamma(u_\gamma)\}, \quad (2.47)$$

et une interface est dite sonique si  $S$  est non vide. On modifie le flux linéarisé de Roe  $f^{\text{Roe}}(u_L, u_R, u)$  aux interfaces soniques. Pour cela, on introduit la fonction de flux modifiée suivante paramétrée par  $u_L$  et  $u_R$  :

$$f^{\text{DM}}(u_L, u_R, u) = f(u_L) + \sum_{\gamma=1}^{mP_\pi} g_\gamma(w_\gamma(u)) r_\gamma(u_{LR}^{\text{Roe}}), \quad (2.48)$$

où  $(w_\gamma(u))_{\gamma \in \{1, \dots, mP_\pi\}}$  sont les variables caractéristiques définies par

$$u - u_L = \sum_{\gamma=1}^{mP_\pi} w_\gamma(u) r_\gamma(u_{LR}^{\text{Roe}}). \quad (2.49)$$

Les fonctions  $(g_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$  sont paramétrées par les états intermédiaires  $(u_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$  et sont définies de la façon suivante :

$$\text{si } \gamma \notin S, \quad g_\gamma(w) = \lambda_\gamma(u_{LR}^{\text{Roe}})w, \quad \forall w, \quad (2.50)$$

$$\text{si } \gamma \in S, \quad g_\gamma(w) = \begin{cases} p_\gamma(w) & \forall w \in (0, \beta_\gamma), \\ \lambda_\gamma(u_{LR}^{\text{Roe}})w & \forall w \notin (0, \beta_\gamma), \end{cases} \quad (2.51)$$

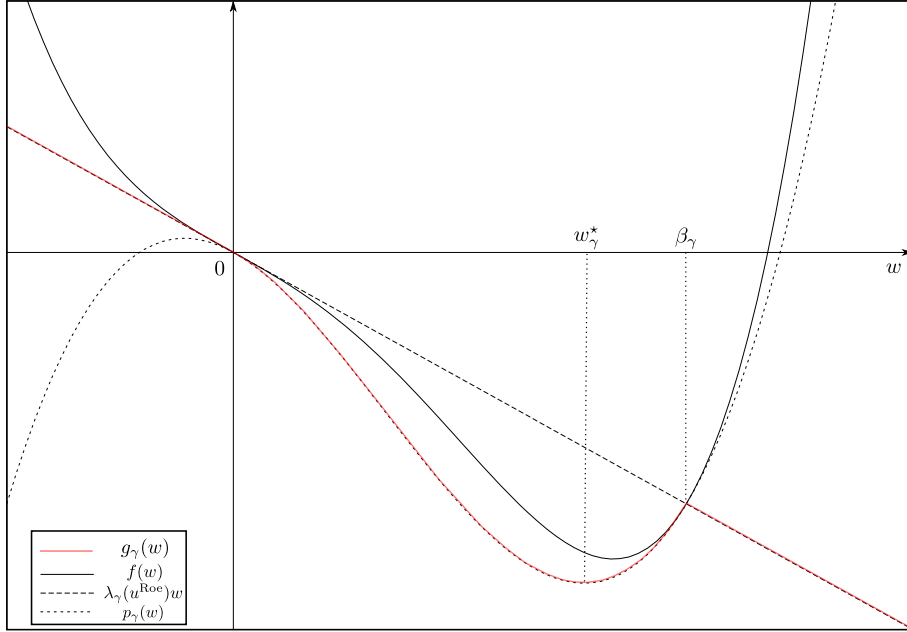
où  $p_\gamma$  est l'unique polynôme de Hermite de degré 3 défini par les conditions :

$$p_\gamma(0) = 0, \quad p_\gamma(\beta_\gamma) = \lambda_\gamma(u_{LR}^{\text{Roe}})\beta_\gamma, \quad p'_\gamma(0) = \lambda_\gamma(u_{\gamma-1}), \quad p'_\gamma(\beta_\gamma) = \lambda_\gamma(u_\gamma). \quad (2.52)$$

On vérifie que  $p_\gamma(w) = aw^3 + bw^2 + cw$  avec

$$a = \frac{\lambda_\gamma(u_\gamma) + \lambda_\gamma(u_{\gamma-1}) - 2\lambda_\gamma(u_{LR}^{\text{Roe}})}{\beta_\gamma^2}, \quad (2.53)$$

$$b = \frac{3\lambda_\gamma(u_{LR}^{\text{Roe}}) - 2\lambda_\gamma(u_{\gamma-1}) - \lambda_\gamma(u_\gamma)}{\beta_\gamma}, \quad c = \lambda_\gamma(u_{\gamma-1}). \quad (2.54)$$

FIGURE 2.3 – Interprétation géométrique de  $g_\gamma$ 

Une interprétation géométrique du polynôme  $g_\gamma$  est illustrée sur la figure 2.3. En l'absence de points soniques, le flux modifié  $f^{\text{DM}}$  coïncide avec le flux linéarisé de Roe  $f^{\text{Roe}}$ . Finalement, le flux numérique de Galerkin modifié  $\varphi^{\text{DM}}(u_L, u_R)$  a pour expression

$$\varphi^{\text{DM}}(u_L, u_R) = \varphi^{\text{Roe}}(u_L, u_R) + \sum_{\gamma \in S} \max \left\{ \frac{g_\gamma(w_\gamma^*)}{\beta_\gamma}, \frac{g_\gamma(w_\gamma^*)}{\beta_\gamma} - \lambda_\gamma(u_{LR}^{\text{Roe}}) \right\} \beta_\gamma r^\gamma(u_{LR}^{\text{Roe}}), \quad (2.55)$$

où  $w_\gamma^*$  est la valeur pour laquelle  $g_\gamma$  atteint son unique extremum sur  $(0, \beta_\gamma)$ . Un calcul direct donne

$$w_\gamma^* = \frac{-\lambda_\gamma(u_{\gamma-1})\beta_\gamma}{\lambda_\gamma^* - \lambda_\gamma(u_{\gamma-1}) + \sqrt{(\lambda_\gamma^*)^2 - \lambda_\gamma(u_{\gamma-1})\lambda_\gamma(u_\gamma)}}, \quad (2.56)$$

où on a posé  $\lambda_\gamma^* = 3\lambda_\gamma(u_{LR}^{\text{Roe}}) - \lambda_\gamma(u_\gamma) - \lambda_\gamma(u_{\gamma-1})$ .

### 2.4.3 Adaptation au système de Galerkin

Un premier point important est que le système de Galerkin n'est généralement pas strictement hyperbolique, *i.e.*, certaines valeurs propres sont multiples. De plus, pour des raisons pratiques, on veut éviter la décomposition spectrale des blocs diagonaux de la matrice Jacobienne de Galerkin car la taille de ces blocs reste grande même si on travaille sur chaque élément stochastique indépendamment. Etant donnés les états gauche et droit  $(u_L, u_R)$  et la matrice Jacobienne de Galerkin de Roe associée  $[a_{LR}^{\text{Roe}}]^{\alpha\sigma}$ , on considère les valeurs propres approchées  $\{\lambda_\gamma\}_{\gamma \in \{1, \dots, mP_\pi\}}$  et

les vecteurs propres approchés  $\{r'_\gamma\}_{\gamma \in \{1, \dots, mP_\pi\}}$  définis par (2.23). On construit les  $(mP_\pi + 1)$  états intermédiaires approchés  $(u'_\gamma)_{\gamma \in \{0, \dots, mP_\pi\}}$  tels que

$$u'_0 = u_L, \quad (2.57)$$

$$u'_\gamma = u'_{\gamma-1} + \beta'_\gamma r'_\gamma(u_{LR}^{\text{Roe}}), \quad \forall \gamma \in \{1, \dots, mP_\pi\}, \quad (2.58)$$

$$u_R = u'_{mP_\pi}, \quad (2.59)$$

où  $(\beta'_\gamma)_{\gamma \in \{1, \dots, mP_\pi\}}$  sont les composantes du vecteur  $u_R - u_L$  sur la base des vecteurs propres approchés  $(r'_\gamma(u_{LR}^{\text{Roe}}))_{\gamma \in \{1, \dots, mP_\pi\}}$ . Un ingrédient clé est que l'énumération des valeurs et vecteurs propres approchés en l'état de Roe approché via les points de Gauss permet de faire la correspondance entre les valeurs propres des différents états intermédiaires, et par conséquent de déterminer l'ensemble des indices soniques

$$S' = \{\gamma \in \{1, \dots, mP_\pi\}, \lambda'_\gamma(u'_{\gamma-1}) < 0 < \lambda'_\gamma(u'_\gamma)\} \quad (2.60)$$

de manière cohérente. On peut ensuite modifier le flux numérique de Roe  $\varphi^{\text{Roe}}(u_L, u_R)$  comme décrit à la section 2.3.3.

Le correcteur entropique ci-dessus requiert l'hypothèse d'une base polynomiale complètement tensorisée pour chaque élément stochastique afin de construire un ensemble de points et poids de Gauss correspondant aux valeurs et vecteurs propres approchés. Pour des problèmes faisant intervenir un grand nombre de variables aléatoires, des bases polynomiales creuses sont nécessaires pour réduire le problème de la dimensionnalité. De telles bases empêchent la construction immédiate de valeurs et vecteurs propres approchés. En utilisant des bases stochastiques partiellement tensorisées, on a proposé de calculer la matrice de décentrement avec l'ensemble des points de Gauss complètement tensorisés (*cf* section 2.3.3). Une approche similaire pourrait être envisagée pour le correcteur entropique : à chaque fois qu'un point sonique est détecté à une interface physique, les états gauche et droit de Galerkin pourraient être localement considérés dans l'espace d'approximation stochastique complètement tensorisé (de dimension plus grande) sur lequel le flux correctionnel pourrait être déterminé grâce à la méthode proposée ci-dessus, avant d'être re-projeté sur l'espace stochastique creux d'origine. Cette méthode est intéressante lorsque le nombre d'interfaces physiques et d'éléments stochastiques où se trouve un point sonique reste (très) modéré.



# Résultats numériques

---

## Sommaire

---

<b>3.1 Burgers</b> . . . . .	<b>55</b>
3.1.1 Cas test 1 : vitesses des ondes positives . . . . .	55
3.1.2 Cas test 2 : vitesses des ondes positives et négatives . . . . .	62
3.1.3 Cas test 3 : point sonique . . . . .	72
<b>3.2 Euler</b> . . . . .	<b>78</b>
3.2.1 Cas test 4 : tube à choc sans point sonique . . . . .	78
3.2.2 Cas test 5 : tube à choc avec point sonique . . . . .	87
<b>3.3 Conclusion</b> . . . . .	<b>94</b>

---

Dans ce chapitre, le solveur de Roe stochastique développé au chapitre précédent est appliqué à cinq cas tests. Les trois premiers utilisent l'équation de Burgers et les deux derniers les équations d'Euler. Les deux premiers cas tests Burgers et le premier cas test Euler ne présentent pas de points soniques, de sorte qu'il n'est pas nécessaire d'utiliser le correcteur entropique de la section 2.4. L'efficacité du correcteur entropique est étudiée dans les sections 3.1.3 et 3.2.2 respectivement pour des cas tests Burgers et Euler présentant des points soniques. Enfin, sauf mention explicite, on utilise une tensorisation polynomiale complète pour engendrer l'espace d'approximation stochastique.

## 3.1 Burgers

Cette section présente les résultats numériques pour l'équation de Burgers.

### 3.1.1 Cas test 1 : vitesses des ondes positives

Le but de ce premier cas test est de valider le solveur de Roe stochastique sur une loi de conservation scalaire (l'équation de Burgers) de sorte que l'hyperbolicité du système de Garlerkin est garantie d'après les théorèmes 2.1 ou 2.2. De plus, on considère un système faisant intervenir uniquement des vitesses d'ondes positives de sorte que le calcul de  $|a_{LR}^{\text{Roe}}|$  est trivial. Enfin, le recours au correcteur entropique de la section 2.4 est évité (pas de point sonique).

### 3.1.1.1 Définition du problème

On considère un domaine spatial unidimensionnel  $\Omega = [0, 1]$  avec des conditions aux bords périodiques. L'équation qui gouverne le système est donnée sous forme conservative par

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0, \quad F(U) = \frac{U^2}{2}, \quad (3.1)$$

et on considère une condition initiale incertaine  $U^0(x, \xi)$  composés de trois états déterministes constants par morceaux en  $x$ . Plus précisément, les trois états sont  $\bar{u}^1 = 1$ ,  $\bar{u}^2 = 1/2$  et  $\bar{u}^3 = 1/6$ , et la position de certaines interfaces entre ces états est incertaine : l'interface entre les états  $\bar{u}^1$  et  $\bar{u}^2$  est située à une position incertaine  $X_{1,2}$  de distribution uniforme dans  $[0.1, 0.2]$  et l'interface entre les états  $\bar{u}^2$  et  $\bar{u}^3$  est située à une position incertaine  $X_{2,3}$  de distribution uniforme dans  $[0.3, 0.4]$ . Enfin, l'interface entre les états  $\bar{u}^3$  et  $\bar{u}^1$  est située en  $x_{3,1} = 0.6$ . Les positions aléatoires  $X_{1,2}$  et  $X_{2,3}$  sont indépendantes et paramétrées en utilisant deux variables aléatoires indépendantes  $\xi_1$  et  $\xi_2$ , de distribution uniforme dans  $[0, 1]$  :

$$X_{1,2} = 0.1 + 0.1\xi_1, \quad X_{2,3} = 0.3 + 0.1\xi_2, \quad \xi_1, \xi_2 \sim \mathcal{U}[0, 1]. \quad (3.2)$$

Par conséquent, le problème a deux dimensions stochastiques ( $N = 2$  et  $\Xi = [0, 1]^2$ ) et la dimension de l'espace d'approximation stochastique d'ordre polynomial  $No$  et de niveau de résolution  $Nr$  est

$$P = 2^{2Nr}(No + 1)^2 = P_\sigma P_\pi, \quad (3.3)$$

avec  $P_\sigma = 2^{2Nr}$  et  $P_\pi = (No + 1)^2$ .

La condition initiale est discrétisée en prenant comme valeurs initiales sa projection  $L^2(\Xi, p_\xi)$ -orthogonale sur  $S^P$  moyennée sur chaque cellule spatiale. Au niveau stochastique, la discrétisation de la condition initiale utilise des approximations bilinéaires continues par morceaux sur les  $P_\sigma$  éléments stochastiques pour  $No \geq 1$  ou l'état moyen sur l'élément stochastique pour  $No = 0$ . L'approximation bilinéaire utilise une interpolation nodale aux sommets des éléments stochastiques, de sorte que les états initiaux discrets sont continus dans le domaine stochastique. Cette procédure évite la présence d'overshoots artificiels dans la donnée initiale. Cependant, aucun traitement particulier n'est appliqué pour forcer la continuité stochastique ou éviter l'apparition d'overshoots durant l'intégration en temps. La Figure 3.1 donne une illustration de la condition initiale aléatoire pour une discrétisation spatiale utilisant  $Nc = 200$  cellules uniformes dans l'espace physique. La figure montre un échantillon de 20 réalisations de la condition initiale aléatoire  $U^0(x, \xi)$ , avec son espérance et son écart type. On peut observer que les réalisations présentent des chocs (très) légèrement inclinés ; cet effet est causé par la représentation par valeur moyenne sur les cellules spatiales et peut être réduit en utilisant un maillage spatial plus fin.

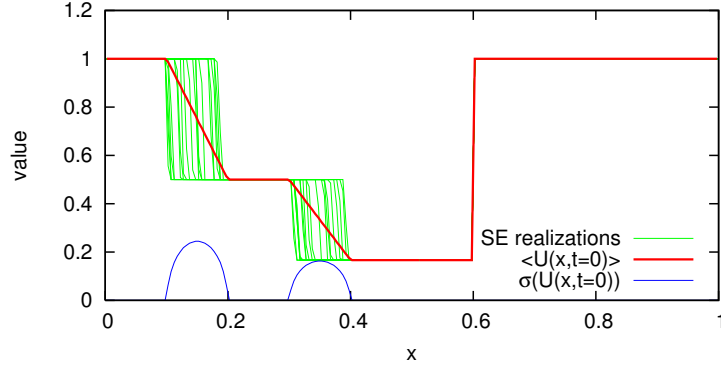


FIGURE 3.1 – Cas test 1 (Burgers) : Échantillon de 20 réalisations de la condition initiale, espérance et écart type.

### 3.1.1.2 Intégration en temps

L'équation de Burgers stochastique est intégrée en temps en utilisant le solveur de Roe sans correcteur entropique décrit dans la section 2.3.3. On rappelle que, puisque  $U$  est à valeurs scalaires, le problème de Galerkin est hyperbolique. L'évaluation du développement stochastique du flux non linéaire  $F(U)$  est basé sur une projection de Galerkin exact utilisant le tenseur de multiplication d'ordre trois  $\mathcal{M}_{\alpha\beta\delta}$  défini par l'équation (2.16). De même, la matrice Jacobienne de Galerkin est exactement évaluée en utilisant (2.17), qui devient

$$(\nabla_u f(u))_{\alpha, \beta \in \mathcal{P}} = \left( \sum_{\delta \in \mathcal{P}} u_\delta \mathcal{M}_{\alpha\beta\delta} \right)_{\alpha, \beta \in \mathcal{P}}. \quad (3.4)$$

Puisque la condition initiale est presque sûrement positive pour tout  $x$ , la solution stochastique  $U$  est presque sûrement positive pour tout  $(x, t)$  d'après le principe du maximum. Par conséquent, on s'attend à ce que le spectre de la matrice Jacobienne de Galerkin soit positif pour une discrétisation stochastique suffisante, de sorte que la matrice de décentrement se réduit à la matrice Jacobienne de Galerkin (la transformation polynomiale est en fait l'identité).

La Figure 3.2 montre la solution de Galerkin aux instants  $t = 0.2, 0.4, 0.6$  et  $0.8$ . Le calcul utilise  $N_r = 3$  et  $N_o = 3$ , la dimension de l'espace d'approximation stochastique est donc  $4^2 \times 8^2 = 1024$ . L'espérance de la solution de Galerkin et l'écart type sont tracés ainsi qu'un échantillon de réalisations. Ces réalisations sont reconstruites à partir du développement stochastique de la solution de Galerkin, en utilisant un échantillon unique de réalisations aléatoires de  $\xi \in [0, 1]^2$ .

On considère d'abord la solution de Galerkin. On observe que la méthode proposée capture correctement la dynamique de l'équation de Burgers. Les chocs sont transportés avec la bonne vitesse et les discontinuités restent raides avec le temps. À l'instant  $t = 0.2$ , le premier choc qui a pour vitesse  $0.75$  n'a pas encore atteint le second choc qui a pour vitesse  $1/3$ . À l'instant  $t = 0.4$ , une fraction des réalisations



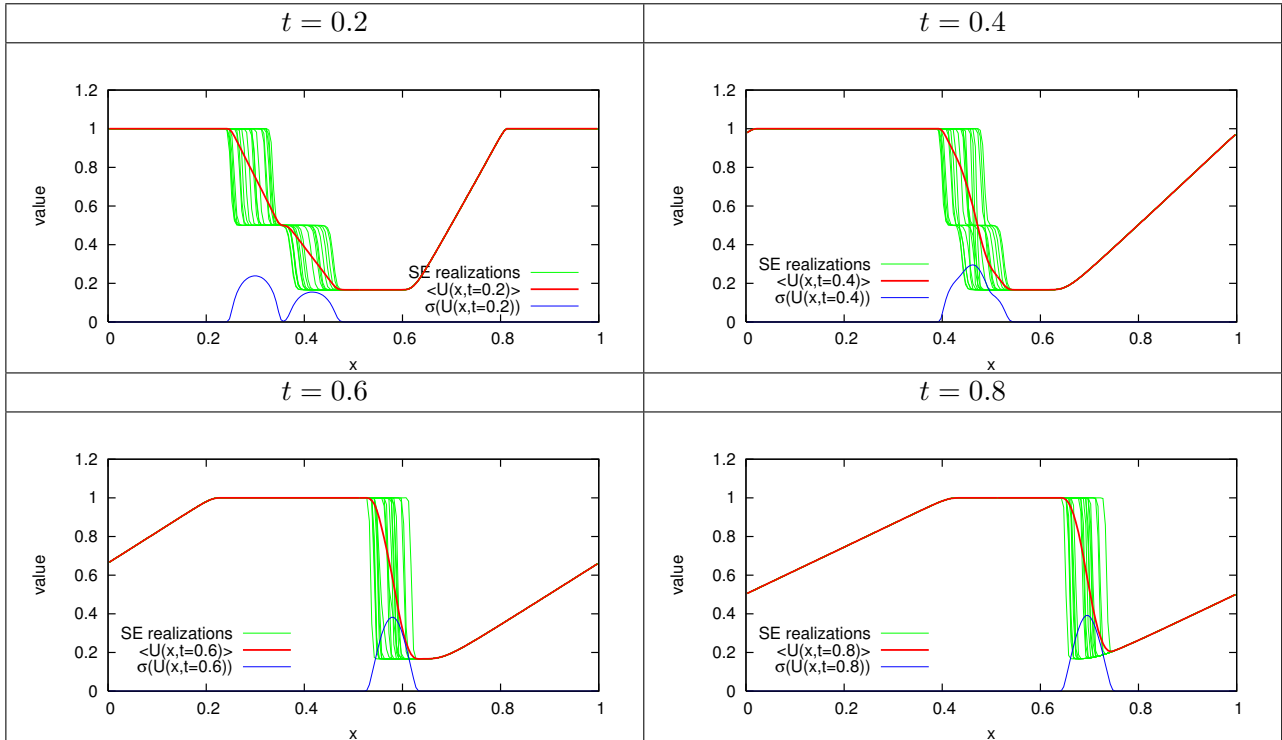


FIGURE 3.2 – Cas test 1 (Burgers) : Solution de Galerkin à différents instants. L'espérance de la solution (en rouge) et l'écart type (en bleu) sont tracés en fonction de  $x$ , ainsi qu'une reconstruction de 20 réalisations (en vert). Calculs avec  $N_r = 3$  et  $N_o = 3$ .

correspond à la situation où le premier et le second chocs ont fusionné. À l'instant  $t = 0.6$ , les chocs ont fusionné pour presque toutes les réalisations, situation qui est intégralement réalisée à l'instant  $t = 0.8$ . On peut observer que les réalisations, bien que correspondant au même échantillon de  $\xi$  dans toutes les figures, présentent une distribution différente avant et après que les chocs aient fusionné. En effet, puisque la fusion a lieu à différents instants dépendant des positions initiales des deux chocs et que les vitesses des chocs sont différentes avant et après la fusion, on ne s'attend pas à ce que la position du choc à des temps plus longs suive une loi uniforme.

La dynamique des chocs incertains peut également être analysée à partir des écarts type de la solution de Galerkin : l'écart type maximum est non seulement plus grand à l'instant  $t = 0.8$ , mais les profils sont également différents au fil du temps. Les tracés de l'espérance  $\langle U^P \rangle$  confirment les observations précédentes. Alors que l'incertitude dans la position du choc induit une évolution affine de  $\langle U^P \rangle$  avec  $x$  lorsque les deux chocs sont distincts, on observe une pente variable de  $\langle U^P \rangle$  avec  $x$  après que les chocs ont fusionné : cela indique une distribution non uniforme de la position du choc après la fusion. De même, la dynamique de l'onde de détente (déterministe) est bien capturée.

En outre, la figure 3.2 montre que le solveur de Roe stochastique ne crée pas d'in-

certitudes parasites dans la solution, dues par exemple à de la diffusion numérique. Ce fait peut être davantage apprécié sur la figure 3.3 où les diagrammes espace-temps de l'espérance de la solution et de son écart type sont tracés sur une large période de temps  $t \in [0, 2]$ . Pour des temps  $t > 0.7$ , dans un repère mobile attaché au choc subsistant, l'écart type atteint son maximum à  $t \approx 1$  avec  $\sigma(U^P) \approx 0.42$ , puis celui-ci décroît lentement. Cette décroissance n'est pas un artéfact numérique, mais est induite par l'onde de détente déterministe qui s'étale progressivement jusqu'à occuper le domaine en entier, comme on peut l'observer sur le tracé de l'espérance où le plateau  $U^P = 1$  a disparu pour  $t > 1$ .

Pour poursuivre l'analyse des résultats, on définit un point d'observation mobile  $x_o(t) = 0.25 + 0.5t$ . Le point d'observation est initialement localisé entre les deux chocs stochastiques. Puisque la vitesse de  $x_o$  est plus petite que 0.75,  $x_o$  va être rattrapé par le premier choc. De plus, puisque  $x_o$  avance plus vite que le second choc, il y a un intervalle de temps pour lequel la solution de Galerkin en  $x_o$  correspond à un ensemble d'événements  $\xi$  avec différentes configurations des chocs. Ceci peut être observé sur la figure 3.4 où la solution de Galerkin  $U^P(x_o(t), t, \xi)$  est tracée en fonction de  $\xi = (\xi_1, \xi_2)$  pour différents instants  $t \in [0.2, 0.7]$ . À l'instant  $t = 0.2$ , le point d'observation commence à être affecté par certains événements correspondant aux réalisations les plus grandes de  $X_{1,2}$  : la solution est fonction de  $\xi_1$  uniquement. À l'instant  $t = 0.3$ , pour une fraction plus large d'événements (environ 1/4), le premier choc a rattrapé le point d'observation et la solution de Galerkin exhibe deux plateaux. À l'instant  $t = 0.4$ , le point d'observation commence à atteindre le deuxième choc, introduisant une dépendance en  $\xi_2$ , alors qu'une fraction d'événements donne lieu à des chocs ayant déjà fusionné. Cela crée une solution de Galerkin avec trois plateaux distincts de valeurs respectives 1, 1/2 et 1/6, dont la configuration évolue au cours du temps. À l'instant  $t = 0.7$ , la solution au point d'observation est essentiellement constante et égale à 1, avec seulement une petite fraction d'événements pour lesquels elle vaut 1/6.

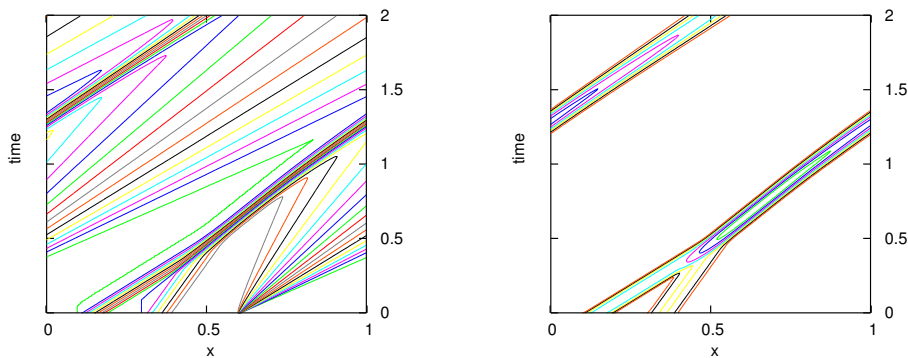


FIGURE 3.3 – Cas test 1 (Burgers) : Diagrammes espace-temps de l'espérance (à gauche) et de l'écart type (à droite) de la solution de Galerkin. Les contours sont dans l'intervalle  $[0, 1]$  avec un pas constant de 0.05. Calculs avec  $N_r = 3$  et  $N_o = 3$ .

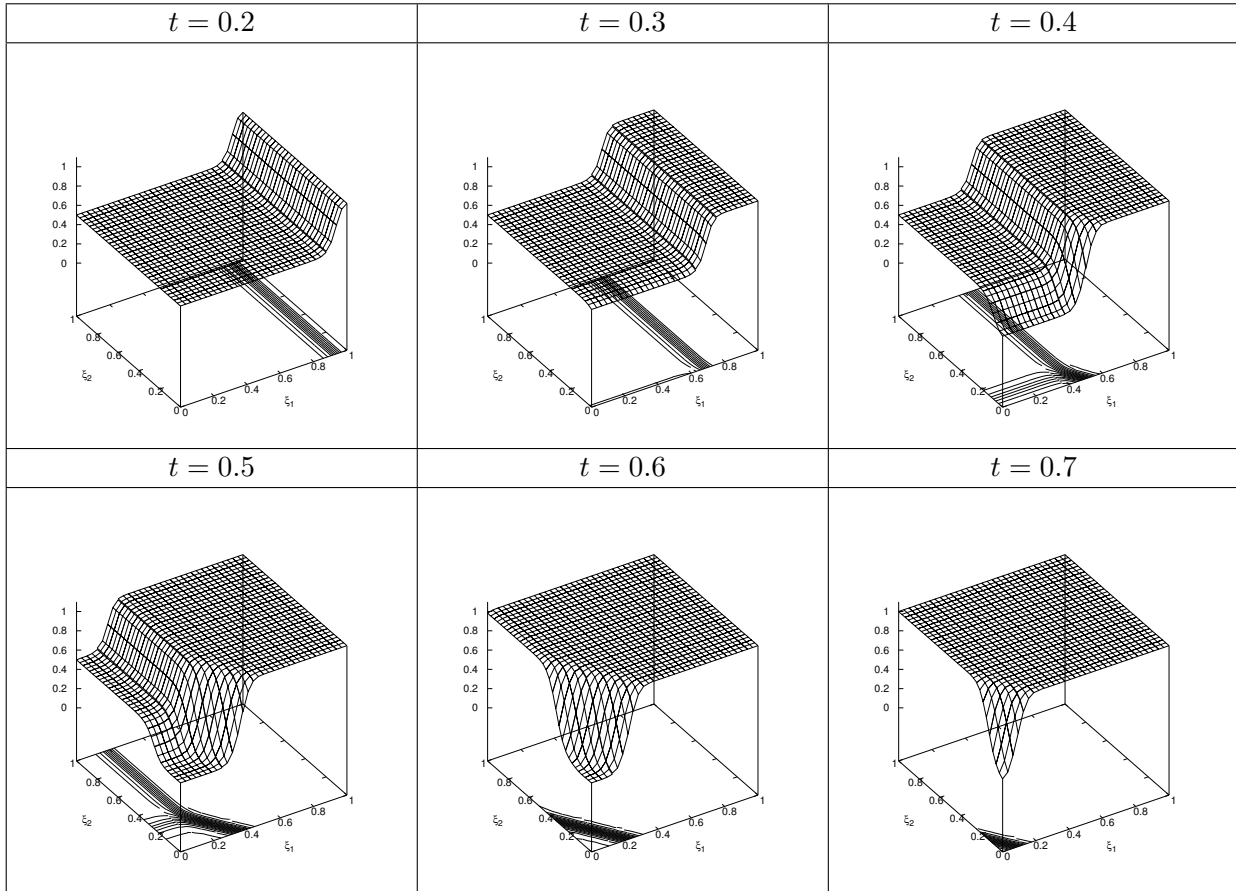


FIGURE 3.4 – Cas test 1 (Burgers) : Solution de Galerkin au point d’observation  $x_o(t)$  en fonction de  $(\xi_1, \xi_2)$  et pour différents instants. Calculs avec  $No = 3$  et  $Nr = 3$ .

Ces résultats démontrent l’efficacité de la méthode proposée pour prendre en compte la dynamique non linéaire et les interactions entre les chocs aléatoires dans le cadre de l’équation de Burgers. Cependant, les courbes de la figure 3.4 méritent quelques commentaires supplémentaires. Premièrement, bien que le schéma numérique autorise des discontinuités entre les éléments stochastiques, les solutions présentées sur la figure 3.4 apparaissent essentiellement continues. La procédure d’initialisation assure une continuité stochastique de la condition initiale, puis le schéma numérique maintient de façon satisfaisante cette propriété lorsque le temps avance pourvu que la résolution soit suffisamment fine. Deuxièmement, les transitions entre les états sont régulières. Cela est dû à la diffusion numérique de la méthode de Roe qui est connue pour étaler les chocs sur quelques cellules spatiales. La régularité de la solution de Galerkin en la variable stochastique reflète cette diffusion numérique spatiale. Ce point sera confirmé ci-dessous en observant que les transitions dans le domaine stochastique ont une largeur caractéristique indépendante de la résolution stochastique. En outre, on peut observer que les transitions

sont plus fines selon la deuxième direction stochastique ( $\xi_2$ ) que selon la première ( $\xi_1$ ). Cela est dû aux différentes vitesses des chocs (effets des différentes CFLs locales).

### 3.1.1.3 Analyse de convergence

On présente dans la figure 3.5 les solutions stochastiques au point d'observation  $x = 0.5$  et à l'instant  $t = 0.5$  pour différentes discrétisations stochastiques. Les tracés de la première ligne illustrent la convergence de l'approximation avec l'ordre polynomial  $No$ , alors que ceux de la seconde ligne mettent en évidence la convergence avec le niveau de résolution  $Nr$ . On voit que lorsque la discrétisation stochastique est trop grossière, la solution exhibe des discontinuités significatives entre les éléments stochastiques. De plus, comme mentionné ci-dessus, les largeurs des transitions entre états dans le domaine stochastique deviennent indépendantes de  $No$  et  $Nr$  lorsque ces paramètres augmentent.

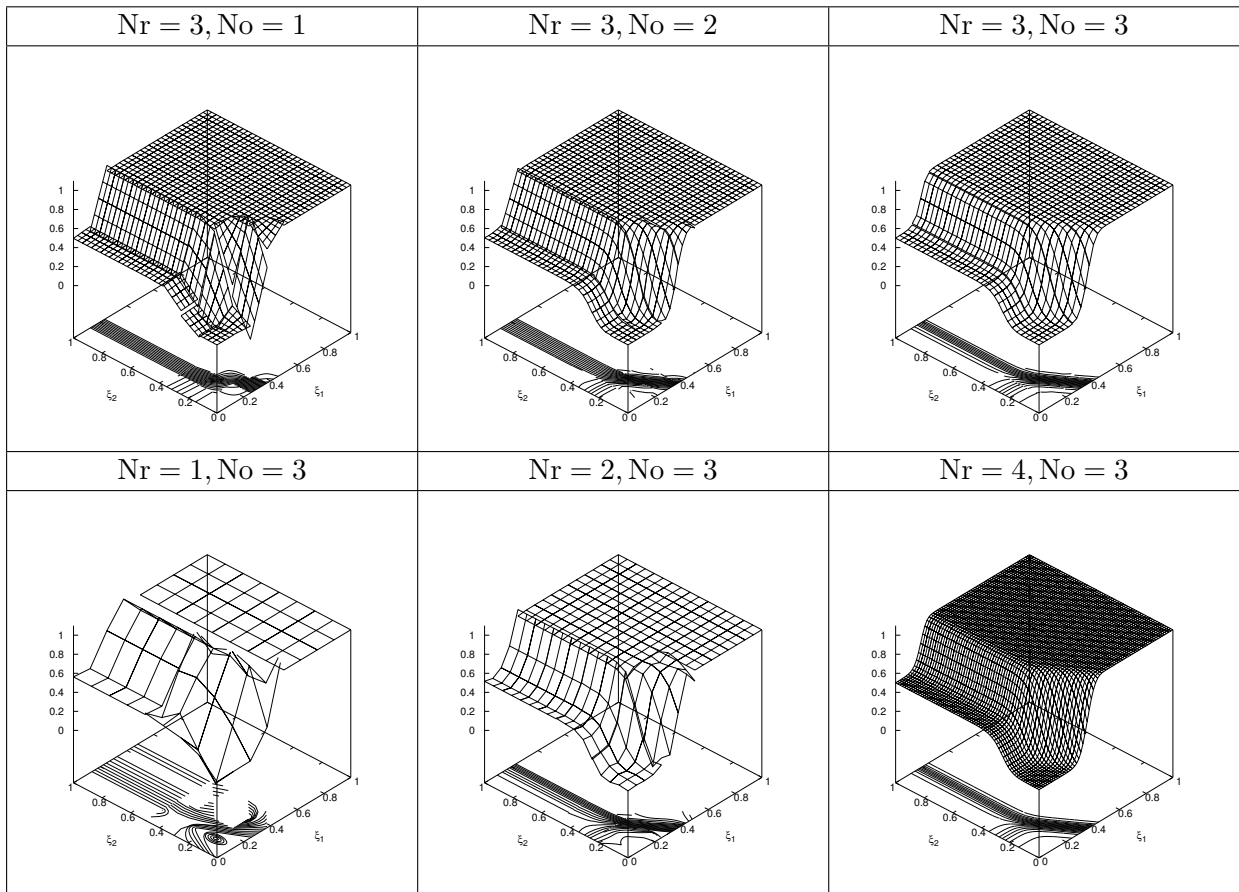


FIGURE 3.5 – Cas test 1 (Burgers) : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  en  $x = 0.5$  et à l'instant  $t = 0.5$  pour différents paramètres de discrétisation stochastique  $Nr$  et  $No$ .

### 3.1.2 Cas test 2 : vitesses des ondes positives et négatives

Le but de ce cas test est de valider la méthode à nouveau sur l'équation de Burgers (de sorte que l'hyperbolicité du système de Galerkin est garantie), mais dans une situation faisant intervenir des vitesses d'onde positives et négatives nécessitant donc le calcul de  $|a_{LR}^{\text{Roe}}|$  comme décrit dans les sections 2.3.2 and 2.3.3. On étudie à nouveau un système sans point sonique, de sorte que l'utilisation du correcteur entropique n'est pas nécessaire.

#### 3.1.2.1 Définition du problème

On considère à nouveau l'équation de Burgers mais avec une condition initiale incertaine  $U^0(x, \xi)$  définie en utilisant deux états incertains,  $U^+(\xi_1)$  et  $U^-(\xi_2)$ , le premier presque sûrement positif et le second presque sûrement négatif. On prend pour  $x \in [0, 1]$ ,

$$U^0(x, \xi) = \begin{cases} U^+(\xi_1) & x < 1/3, \\ U^-(\xi_2) & x > 2/3, \\ U^+(\xi_1)(2 - 3x) + U^-(\xi_2)(3x - 1) & 1/3 \leq x \leq 2/3, \end{cases} \quad (3.5)$$

de sorte que  $U^0(x, \xi)$  est continue en  $x$  pour tout  $\xi \in [0, 1]^2$ . Les états stochastiques sont définis par

$$\begin{aligned} U^+(\xi_1) &= 1 - 0.1(2\xi_1 - 1), & \xi_1 &\sim \mathcal{U}[0, 1] \rightarrow U^+ \sim \mathcal{U}[0.9, 1.1], \\ U^-(\xi_2) &= -1 - 0.05(2\xi_2 - 1), & \xi_2 &\sim \mathcal{U}[0, 1] \rightarrow U^- \sim \mathcal{U}[-1.05, -0.95], \end{aligned} \quad (3.6)$$

et on résout l'équation de Burgers stochastique avec des conditions aux bords de Dirichlet,  $U = U^+$  en  $x = 0$  et  $U = U^-$  en  $x = 1$ . La condition initiale est illustrée sur la figure 3.6. On utilise  $N_c = 201$  cellules pour la discrétisation spatiale.

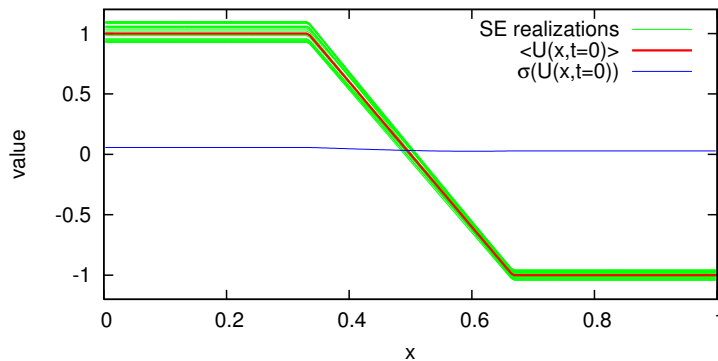


FIGURE 3.6 – Cas test 2 (Burgers) : Échantillon de 20 réalisations de la condition initiale, espérance et écart type.

### 3.1.2.2 Intégration en temps

Bien qu'initialement continue, la solution de Galerkin va développer en temps fini une discontinuité avec un saut stochastique  $|U^+ - U^-|$  et une vitesse de propagation stochastique  $(U^+ + U^-)/2$ . Le caractère stochastique de l'amplitude du choc et de la vitesse contraste avec le cas test précédent, où les sauts et la vitesse du choc étaient certains. Cela conduit à une situation plus complexe comme illustré sur la figure 3.7 où la solution de Galerkin est tracée à différents instants pour les paramètres de discrétisation stochastique  $No = 3$  et  $Nr = 3$  si bien que  $\dim \mathcal{S}^{No, Nr} = 1024$ .

Sur les réalisations de la figure 3.7, on observe l'apparition d'overshoots relatifs au phénomène de Gibbs. Pour vérifier la stabilité de la méthode, on compare dans la figure 3.8 des réalisations de la solution obtenues par le solveur de Roe stochastique (utilisant la projection de Galerkin) avec des réalisations obtenues par une méthode spectrale de projection non intrusive (fournissant une solution approchée qu'on appellera solution NISP). En particulier, on se base sur des quadratures tensorisées avec  $N_Q^2$  points sur chaque élément stochastique pour calculer les modes de la solution aux instants sélectionnés. En chaque point de quadrature, la solution NISP est calculée en utilisant un code déterministe également basé sur un solveur de Roe. Puisque la solution stochastique n'est pas un polynôme en  $\xi$ , le nombre de points de quadrature ne peut pas être sélectionné *a priori*; pour tous les résultats présentés, on utilise  $N_Q = 16$ , une valeur pour laquelle les quadratures sont suffisamment précises pour tous les ordres de développement  $No$  et niveaux de résolution  $Nr$  utilisés. Le calcul de la solution NISP requiert la résolution d'un très grand nombre ( $16^2 \times 2^{2Nr} = 256 \times 4^{Nr}$ ) d'équations de Burgers déterministes indépendantes. Sur la figure 3.8, on observe que les overshoots sont également présents sur la solution NISP. De plus, pour les deux solutions (Galerkin et NISP), les overshoots peuvent être réduits en augmentant la résolution stochastique. En outre, on compare dans la figure 3.9 la valeur maximale des solutions de Galerkin et NISP en fonction du temps pour  $Nr = 4$  et  $No = 3$ . On observe que les deux valeurs maximales augmentent avec le temps jusqu'à atteindre un plateau vers l'instant  $t \approx 2.35$ . Cela confirme le fait que des instabilités ne se développent pas en temps. D'autres cas tests (omis ici) montrent que la taille des overshoots peut être réduite en augmentant  $Nr$  à  $No$  fixé ou *vice versa*. On reviendra sur cette propriété lors des cas tests pour le solveur de Roe adaptatif (en  $Nr$  à  $No$  fixé), développé dans les chapitres 5 et 6.

Finalement, on présente dans la figure 3.10 l'évolution de la solution de Galerkin en un point d'observation fixe  $x = 0.5$  et à différents instants. Les tracés montrent l'évolution de la solution initialement continue en une solution comportant un choc séparant les états  $U^+$  ou  $U^-$  selon le signe de  $2(\xi_1 - 1/2) - (\xi_2 - 1/2)$ . De plus, on observe que les overshoots ont lieu seulement dans un voisinage de la discontinuité, *i.e.* dans les éléments stochastiques au voisinage de la discontinuité qui se développe. En utilisant une discrétisation stochastique plus fine (en augmentant  $Nr$ ), on retarde l'émergence des overshoots et on réduit la portion du domaine stochastique affectée par ceux-ci.

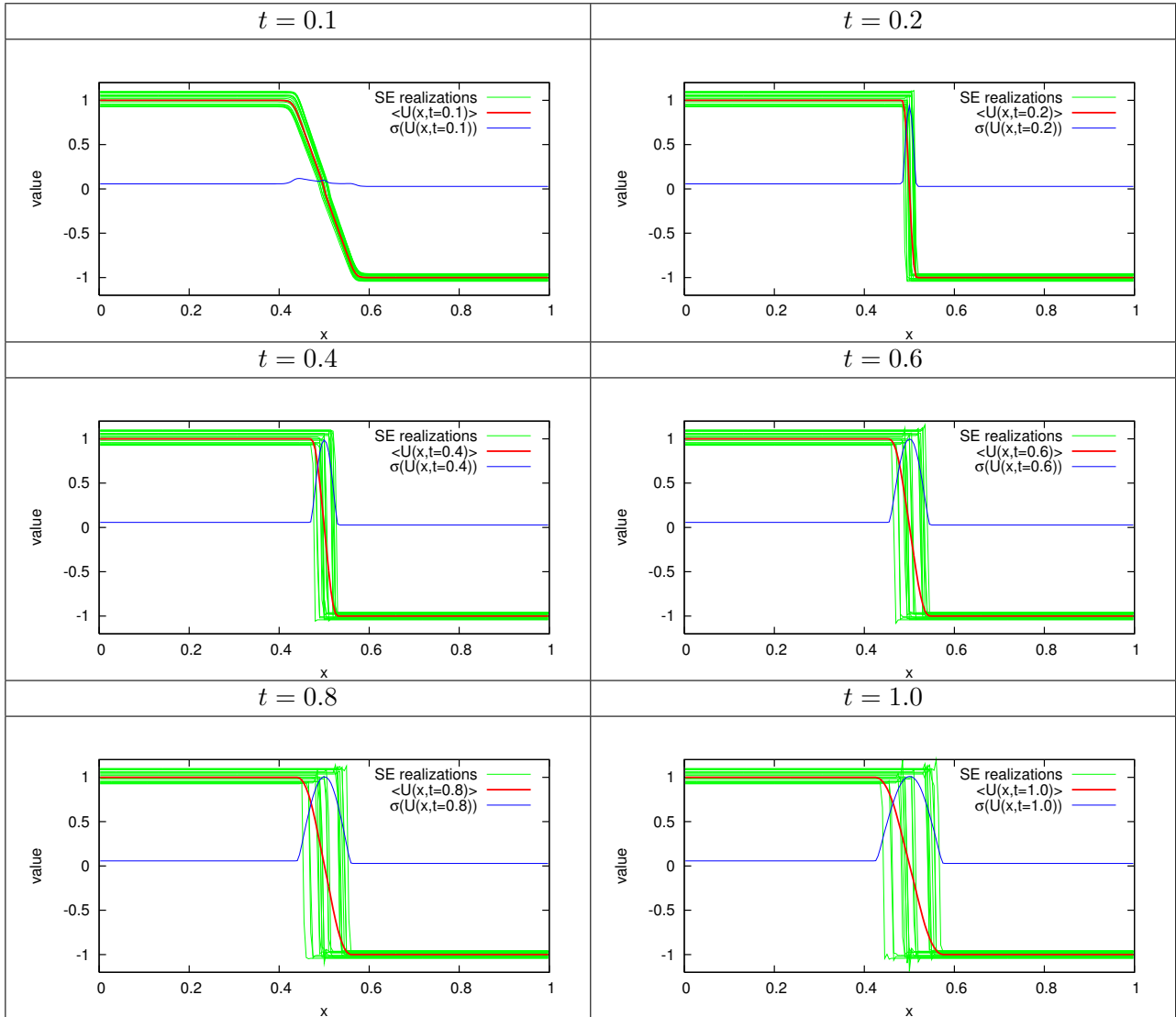


FIGURE 3.7 – Cas test 2 (Burgers) : Solution de Galerkin à différents instants. L'espérance de la solution (en rouge) et l'écart type (en bleu) sont tracés en fonction de  $x$ , ainsi qu'une reconstruction de 20 réalisations (en vert). Calculs avec  $N_r = 3$  et  $N_o = 3$ .

### 3.1.2.3 Validation de la méthode utilisée pour évaluer la matrice de décentrement

Une propriété intéressante du cas test 2 est que contrairement au précédent, il existe des cellules spatiales où la solution de Galerkin peut prendre des valeurs positives et négatives en fonction de  $\xi$ . Par conséquent, les valeurs propres de  $a_{LR}^{\text{Roe}}$  ne sont plus strictement positives, et le polynôme  $q_{d_{\text{fit}},\{\lambda'\}}$  utilisé pour approcher la valeur absolue de  $a_{LR}^{\text{Roe}}$  n'est plus trivial comme dans le cas précédent. On analyse maintenant l'impact du degré polynomial  $d_{\text{fit}}$  sur la solution de Galerkin. Dans les

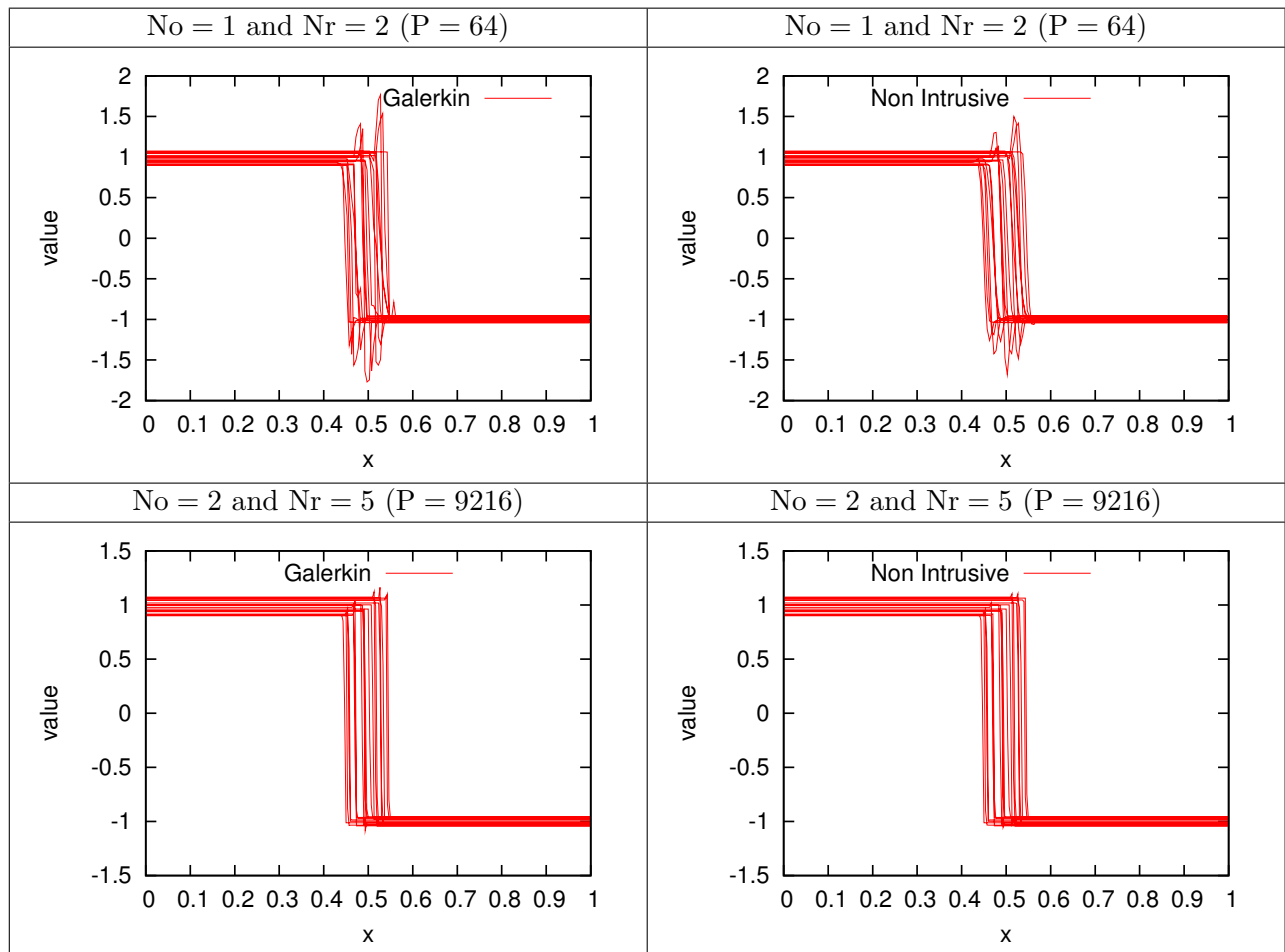


FIGURE 3.8 – Cas test 2 (Burgers) : Reconstruction de 20 réalisations de la solution de Galerkin (à gauche) et de la solution NISP (à droite) à l'instant  $t = 1.0$ . Calculs avec différents  $N_r$  et  $N_o$ .

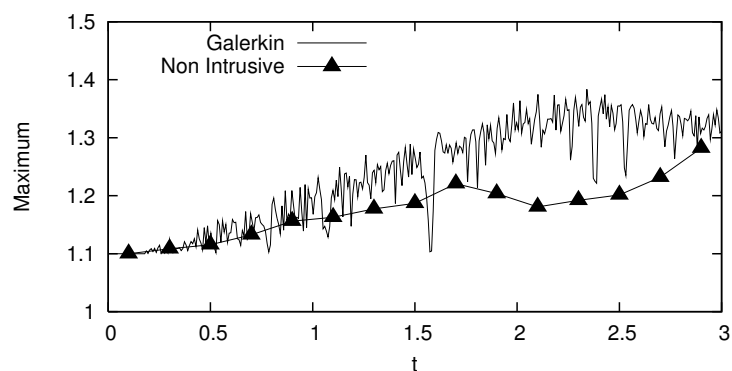


FIGURE 3.9 – Cas test 2 (Burgers) : Valeur maximale des solutions de Galerkin et NISP en fonction du temps pour  $N_r = 4$  et  $N_o = 3$ .



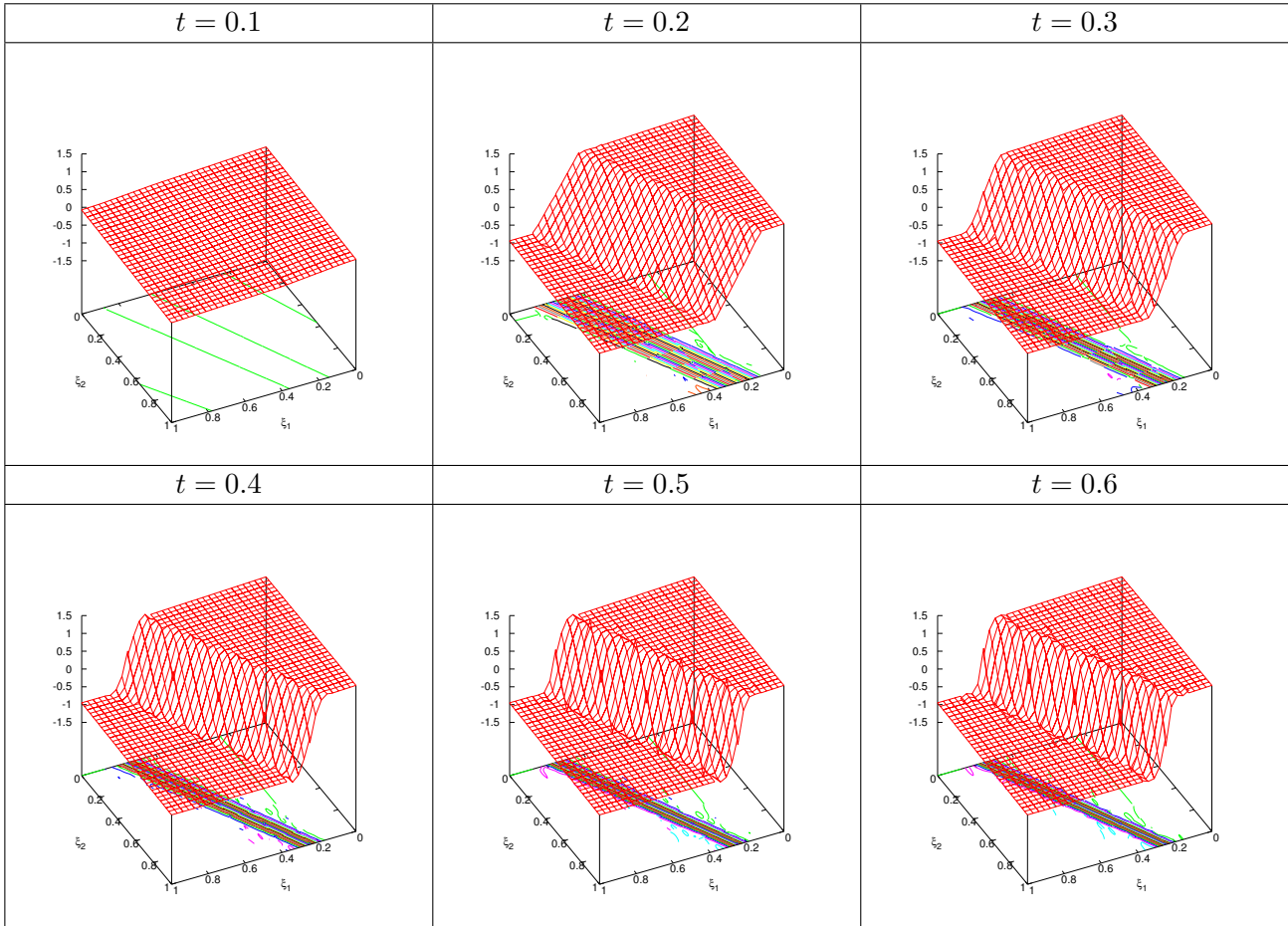


FIGURE 3.10 – Cas test 2 (Burgers) : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  en  $x = 0.5$  et à différents instants. Calculs avec  $N_r = 3$  et  $N_o = 3$ .

calculs présentés ci-dessus, on a utilisé des polynômes de degré  $d_{\text{fit}} = 3$ . Dans la figure 3.11, on reporte la solution de Galerkin en  $x = 0.5$  et à l'instant  $t = 0.5$  calculée en utilisant un degré polynomial  $d_{\text{fit}}$  croissant. On voit que pour  $d_{\text{fit}} = 1$ , la solution exhibe des discontinuités parasites et des overshoots au travers des cellules de discrétisation stochastique contenant le choc qui se développe (où la solution change de signe) ce qui signifie que les valeurs propres de la matrice de décentrement ne sont pas approchées avec suffisamment de précision. Lorsque  $d_{\text{fit}} = 2$ , les overshoots et discontinuités sont largement réduits par rapport au cas  $d_{\text{fit}} = 1$ . Augmenter ensuite  $d_{\text{fit}}$  n'apporte pas d'amélioration significative dans la solution. En fait, à ce stade l'erreur dans la solution de Galerkin est essentiellement dominée par l'erreur de discrétisation stochastique et spatiale, si bien que l'erreur d'approximation dans le calcul de  $|a_{LR}^{\text{Roe}}|$  pour  $d_{\text{fit}} > 3$  devient négligeable.

Pour mesurer plus précisément l'erreur sur l'approximation de  $|a_{LR}^{\text{Roe}}|$ , on calcule l'ensemble des valeurs propres exactes  $\{\lambda_\alpha\}_{\alpha \in \mathcal{P}}$  de  $a_{LR}^{\text{Roe}}$ . On compare ensuite les quantités  $|\lambda_\alpha|$  avec leurs approximations polynomiales  $q_{d_{\text{fit}}, \{\lambda\}}(\lambda_\alpha)$ . L'erreur est

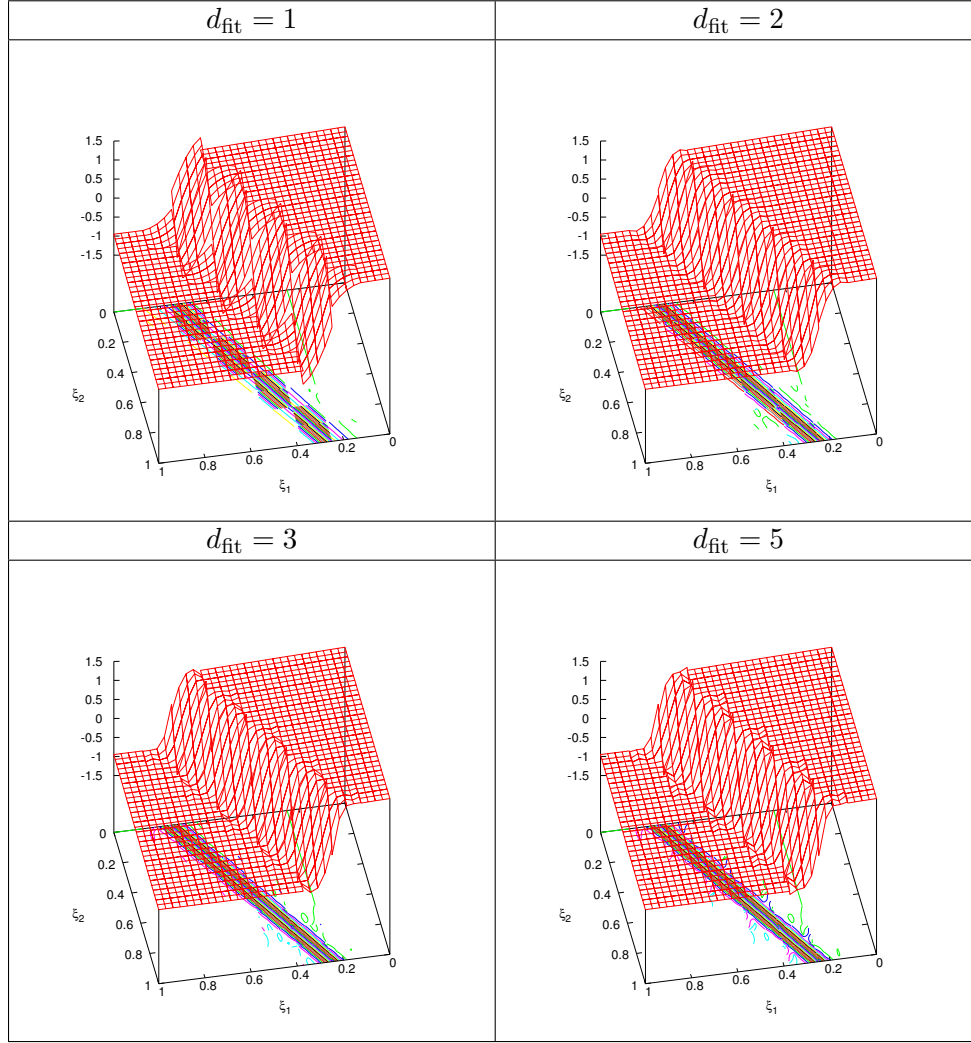


FIGURE 3.11 – Cas test 2 (Burgers) : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  à  $x = 0.5$  et  $t = 0.5$  pour différents degrés  $d_{\text{fit}}$  du polynôme  $q_{d_{\text{fit}}, \{\lambda'\}}$  utilisé dans l'approximation de la valeur absolue de  $a_{LR}^{\text{Roe}}$ . Calculs avec  $N_r = 3$  et  $N_o = 3$ .

quantifiée en utilisant les mesures suivantes :

$$\epsilon_2^2 = \frac{1}{P} \sum_{\alpha \in \mathcal{P}} (|\lambda_\alpha| - q_{d_{\text{fit}}, \{\lambda'\}}(\lambda_\alpha))^2 \quad \text{et} \quad \epsilon_\infty = \max_{\alpha \in \mathcal{P}} ||\lambda_\alpha| - q_{d_{\text{fit}}, \{\lambda'\}}(\lambda_\alpha)|. \quad (3.7)$$

On rappelle que le fit polynomial  $q_{d_{\text{fit}}, \{\lambda'\}}$  est en fait différent sur chaque élément stochastique. Dans la figure 3.12 on représente les mesures (3.7) à l'instant  $t = 0.4$  en fonction de  $x$ . On remarque tout d'abord que l'erreur se limite à la portion du domaine spatial où le choc stochastique peut être présent, et elle diminue lorsque  $d_{\text{fit}}$  augmente. Ensuite, les deux mesures d'erreur stagnent lorsque  $d_{\text{fit}}$  augmente au-dessus de 5; ceci résulte du fait que les valeurs propres estimées aux points de Gauss tensorisés et utilisées pour la détermination de  $q_{d_{\text{fit}}, \{\lambda'\}}$  ne sont pas les réelles valeurs propres exactes de  $a_{LR}^{\text{Roe}}$ .

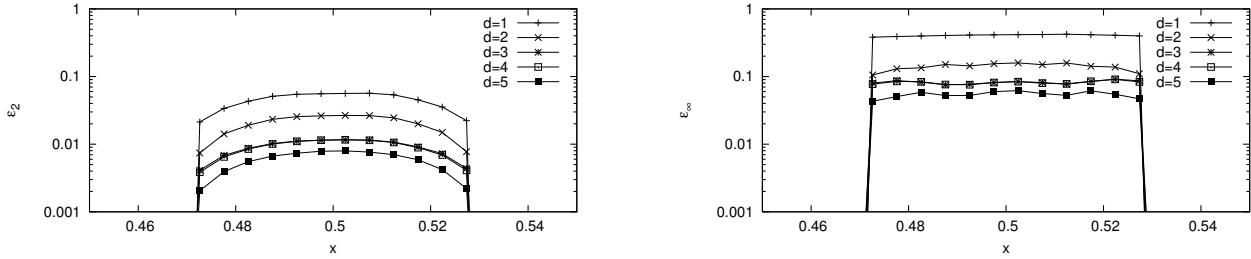


FIGURE 3.12 – Cas test 2 (Burgers) : Mesures  $\epsilon_2$  (à gauche) et  $\epsilon_\infty$  (à droite) des erreurs sur les valeurs propres de la valeur absolue de  $a_{LR}^{\text{Roe}}$  à l'instant  $t = 0.4$  et pour différents degrés  $d_{\text{fit}}$  du polynôme  $q_{d_{\text{fit}},\{\lambda\}}$ . Calculs avec  $\text{Nr} = 3$  et  $\text{No} = 3$ .

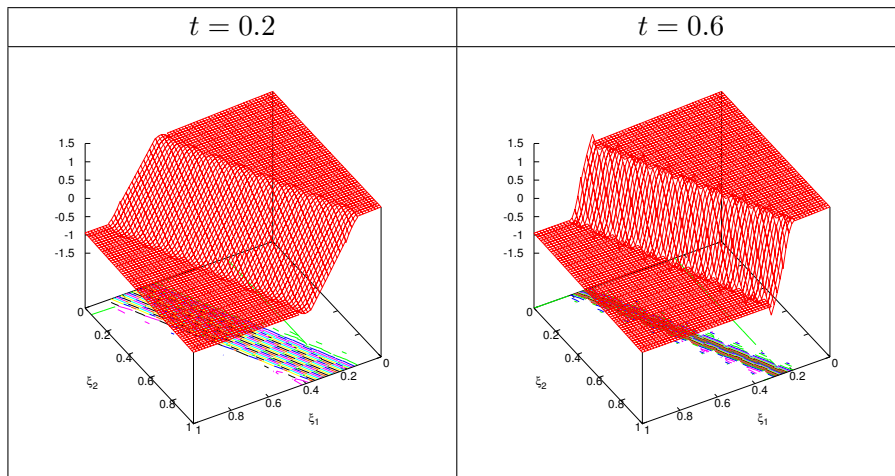


FIGURE 3.13 – Cas test 2 (Burgers) : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  en  $x = 0.5$  aux instants  $t = 0.2$  et  $t = 0.6$  dans le cas d'une tensorisation polynomiale partielle. Calculs avec  $\text{Nr} = 4$  et  $\text{No} = 2$ .

Finalement, on vérifie que la procédure utilisée pour calculer la matrice de décentrement approchée peut être appliquée lorsque l'on travaille avec une tensorisation polynomiale partielle. Pour cela, on procède comme décrit dans la section 2.3.3. La solution de Galerkin est tracée en fonction de  $(\xi_1, \xi_2)$  dans la figure 3.13 en  $x = 0.5$  aux instants  $t = 0.2$  et  $t = 0.6$  pour les paramètres stochastiques  $\text{Nr} = 4$  et  $\text{No} = 2$ . On observe que les profils sont similaires à ceux de la figure 3.10, indiquant que la matrice de décentrement approchée suffit à assurer la stabilité des calculs. On observe cependant l'apparition d'overshoots à  $t = 0.6$  au voisinage de la discontinuité qui se forme. Cela s'explique par le fait que la dimension stochastique est plus petite dans le cas de la base polynomiale partiellement tensorisée. En augmentant la discrétisation stochastique, par exemple en augmentant  $\text{Nr}$ , on réussit à faire disparaître les overshoots, cf figure 3.14 où on compare la solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  en  $x = 0.5$  à l'instant  $t = 0.5$  pour  $\text{No} = 2$  et  $\text{Nr} = 4$  et pour  $\text{No} = 2$  et  $\text{Nr} = 5$ . On observe, comme attendu, que pour capturer les discontinuités, on a besoin d'une résolution stochastique (un peu) plus fine pour la base polynomi-

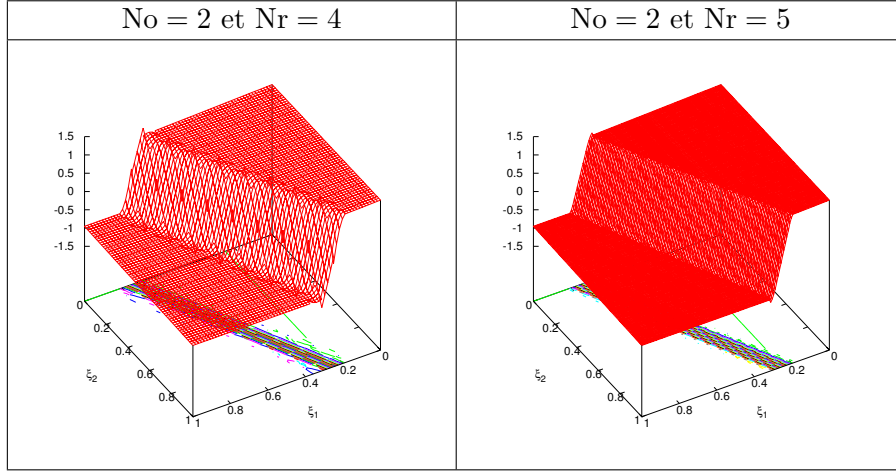


FIGURE 3.14 – Cas test 2 (Burgers) : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  en  $x = 0.5$  à l’instant  $t = 0.5$  dans le cas d’une tensorisation polynomiale partielle. Calculs avec  $No = 2$  et différents  $Nr$ .

ale partielle que pour la base polynomiale complète. La base polynomiale partielle peut présenter un intérêt dans le cas d’une discrétisation stochastique adaptative, cf chapitre 6, puisqu’elle permet de diminuer considérablement le nombre de degrés de liberté dans les zones où la solution est régulière.

### 3.1.2.4 Convergence de l’erreur stochastique

La simplicité de ce cas test permet d’étudier la convergence de la solution de Galerkin vers la solution stochastique. En effet, pour le problème de Riemann, on peut facilement construire la solution stochastique  $U(x, t, \xi)$  pour tout  $\xi$ , tant que le choc n’a pas atteint une des frontières du domaine. On utilise une stratégie d’échantillonnage de Monte Carlo pour estimer les deux premiers moments de  $U$ . On procède comme suit. Tout d’abord, un ensemble de  $M$  réalisations de  $\xi$  est généré par échantillonnage uniforme de  $[0, 1]^2$ . Puis, pour chaque élément  $\xi^{(j)}$  de l’échantillon, on définit  $U^{(j)}(x, t) := U(x, t, \xi^{(j)})$  pour  $j \in \{1, \dots, M\}$ . L’estimateur Monte Carlo de l’espérance est

$$\langle U \rangle(x, t) \approx \frac{1}{M} \sum_{j=1}^M U^{(j)}(x, t) =: E_s(U)(x, t), \quad (3.8)$$

et l’estimateur Monte Carlo de l’écart type est

$$\sigma^2(U)(x, t) \approx \frac{1}{M} \sum_{j=1}^M \left( U^{(j)} - E_s(U) \right)^2(x, t) =: \sigma_s^2(U)(x, t). \quad (3.9)$$

Pour s’affranchir de l’erreur d’échantillonnage dans l’estimateur empirique, on utilise  $M = 100000$ .

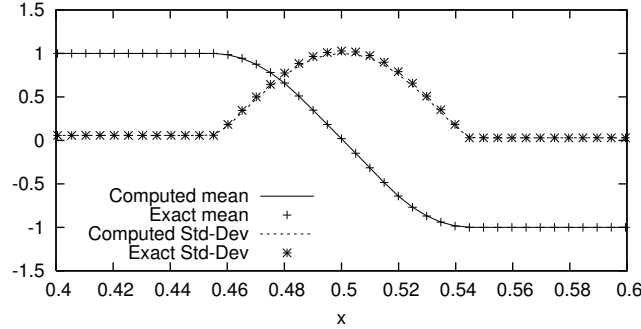


FIGURE 3.15 – Cas test 2 (Burgers) : Comparaison de l’espérance et de l’écart type de la solution de Galerkin à l’instant  $t = 0.6$ , calculée avec  $N_o = 2$ ,  $N_r = 4$  et  $N_c = 201$ , avec les estimations MC correspondantes de l’espérance et de l’écart type de la solution stochastique. Seule la portion du domaine de calcul la plus intéressante est montrée.

Dans la figure 3.15, on compare l’espérance et l’écart type de la solution stochastique et de la solution de Galerkin pour  $N_o = 2$  et  $N_r = 4$  à l’instant  $t = 0.6$  sur un maillage de  $N_c = 201$  cellules dans l’espace physique. On observe que les espérances de la solution de Galerkin et de la solution stochastique sont en excellent accord. Pour les écarts type, l’accord est bon, bien que la solution de Galerkin sous-estime légèrement l’écart type avec toutefois moins de 5% d’erreur relative. Le graphe en haut de la figure 3.16 présente la distribution spatiale de l’erreur sur l’écart type pour différents niveaux de résolution de la grille spatiale et pour des paramètres de discrétisation stochastique fixés à  $N_o = 2$  et  $N_r = 4$ . On observe que raffiner la grille spatiale améliore la précision. Le graphe du bas de la figure 3.16 présente, pour  $N_o = 2$  et différents  $N_r$ , l’erreur spatiale intégrée définie par

$$S_h^2 := \sum_{i=1}^{N_c} \Delta x (\sigma(U_i) - \sigma(U_i^P))^2, \quad (3.10)$$

où l’indice  $i$  fait référence à la cellule de discrétisation spatiale. Les résultats montrent qu’excepté pour le niveau de résolution stochastique le plus bas et la grille spatiale la plus fine, l’erreur est dominée par l’erreur de discrétisation spatiale.

Pour analyser davantage la convergence stochastique de la méthode, on étudie la convergence sur un maillage spatial fixé avec  $N_c = 201$ . On considère la mesure d’erreur

$$\epsilon_h^2(t) := \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^{N_c} \Delta x \left( U_j^P(t, \xi^{(j)}) - U_i^h(t, \xi^{(j)}) \right)^2, \quad (3.11)$$

où, pour chaque élément  $\xi^{(j)}$  d’un échantillon,  $U_i^P(t, \xi^{(j)})$  et  $U_i^h(t, \xi^{(j)})$  sont respectivement évaluées à partir du développement stochastique de la solution de Galerkin à la cellule  $i$  de l’espace physique et à partir de la solution du problème déterministe de Burgers correspondant sur la cellule  $i$  de l’espace physique calculée avec le

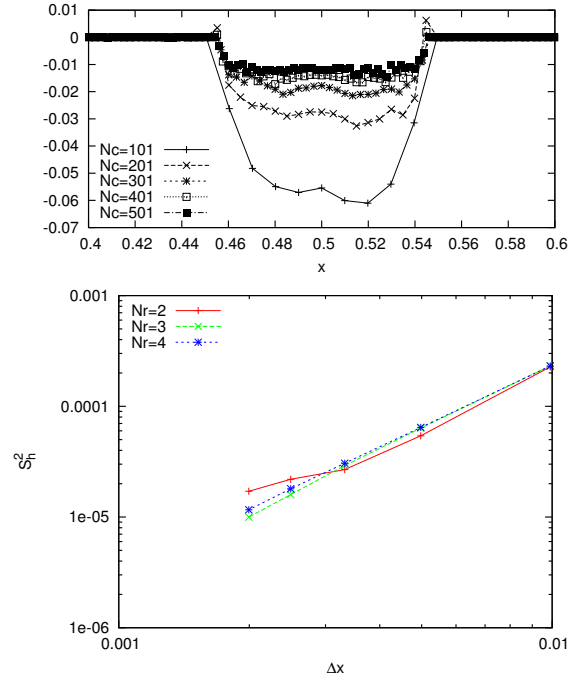


FIGURE 3.16 – Cas test 2 (Burgers) : Erreurs sur l'écart type à l'instant  $t = 0.6$  pour différents maillages spatiaux. En haut : Distribution de l'erreur sur le domaine spatial pour  $No = 2$  et  $Nr = 4$ . En bas : Erreur intégrée en espace pour  $No = 2$  et différents  $Nr$ .

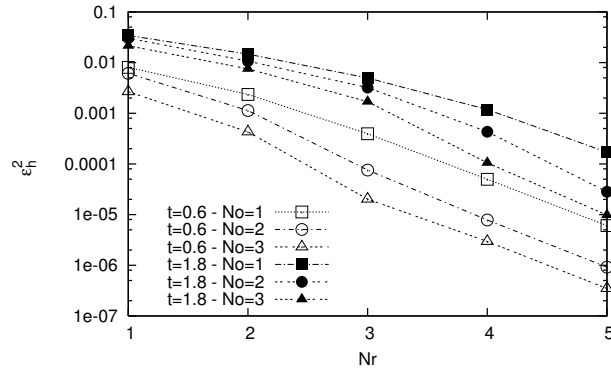


FIGURE 3.17 – Cas test 2 (Burgers) : Erreur stochastique  $\epsilon_h^2$  aux instants  $t = 0.6$  et  $t = 1.8$  en fonction du niveau de résolution  $Nr$  et pour différents ordres polynomiaux  $No$ . Calculs avec  $Nc = 201$ .

solveur déterministe. On utilise un échantillon de taille  $M = 10000$ . La Figure 3.17 présente l'erreur stochastique  $\epsilon_h^2$  aux instants  $t = 0.6$  et  $t = 1.8$ , en fonction du niveau de résolution  $Nr$  et pour les ordres polynomiaux  $No = 1, 2$  et  $3$ . Pour ces simulations, l'approximation de la matrice de décentrement utilise un degré polynomial défini par  $d_{\text{fit}} = \min(8, (No + 1)^2)$ . Pour les deux instants, on observe un

ordre de décroissance similaire de l'erreur stochastique en fonction du niveau de résolution. Les erreurs sont plus importantes pour des temps plus longs puisque les chocs ont lieu sur une plus large portion du domaine spatial.

### 3.1.3 Cas test 3 : point sonique

Le but de ce cas test est de valider le correcteur entropique construit dans la section 2.4 sur la loi de conservation scalaire de Burgers dans le cas où un point sonique est présent presque sûrement.

#### 3.1.3.1 Définition du problème

On considère à nouveau l'équation de Burgers avec une condition initiale incertaine  $U^0(x, \xi)$  définie en utilisant deux états incertains,  $U^+(\xi_1)$  et  $U^-(\xi_2)$ , le premier presque sûrement négatif et le second presque sûrement positif. On prend pour  $x \in [0, 1]$ ,

$$U^0(x, \xi) = \begin{cases} U^-(\xi_1) & x < 1/2, \\ U^+(\xi_2) & x \geq 1/2, \end{cases} \quad (3.12)$$

avec

$$\begin{aligned} U^-(\xi_1) &= -1 + 0.05(2\xi_1 - 1), & \xi_1 &\sim \mathcal{U}[0, 1] \rightarrow U^- \sim \mathcal{U}[-1.05, -0.95], \\ U^+(\xi_2) &= 1 + 0.1(2\xi_2 - 1), & \xi_2 &\sim \mathcal{U}[0, 1] \rightarrow U^+ \sim \mathcal{U}[0.9, 1.1], \end{aligned} \quad (3.13)$$

et on résout l'équation de Burgers stochastique ( $N = 2$ ) avec des conditions aux bords de sortie libre en  $x = 0$  et en  $x = 1$ . La condition initiale est illustrée sur la figure 3.18. L'équation de Burgers stochastique est intégrée en temps en utilisant le solveur de Roe avec correcteur entropique. On utilise  $N_c = 250$  cellules pour la discrétisation spatiale et les paramètres stochastiques sont  $N_o = 3$  pour l'ordre polynomial et  $N_r = 3$  pour le niveau de résolution, si bien que la dimension de l'espace d'approximation stochastique est  $P = (N_o + 1)^{N_r} = 1024$ .

Pour les conditions initiales spécifiées par (3.13) et en posant  $t_s := \frac{x-1/2}{t}$ , la solution entropique stochastique de l'équation de Burgers consiste en une onde de détente dont l'expression peut être facilement calculée pour  $t > 0$  et  $\xi = (\xi_1, \xi_2) \in [0, 1]^2$  :

$$U(x, t, (\xi_1, \xi_2)) = \begin{cases} U^-(\xi_1) & t_s \leq U^-(\xi_1), \\ t_s & U^-(\xi_1) < t_s < U^+(\xi_2), \\ U^+(\xi_2) & t_s \geq U^+(\xi_2). \end{cases} \quad (3.14)$$

Une remarque importante est que la solution est indépendante de l'incertitude pour  $(x, t)$  tel que  $t_s$  est dans l'intervalle

$$\sup_{\xi_1 \in [0, 1]} U^-(\xi_1) < t_s < \inf_{\xi_2 \in [0, 1]} U^+(\xi_2), \quad (3.15)$$

de sorte que l'écart type vaut zéro pour  $x \in I(t)$  avec  $I(t) := [-0.95t + 0.5, 0.9t + 0.5]$ . De plus, on remarque que l'espérance et l'écart type de la solution stochastique

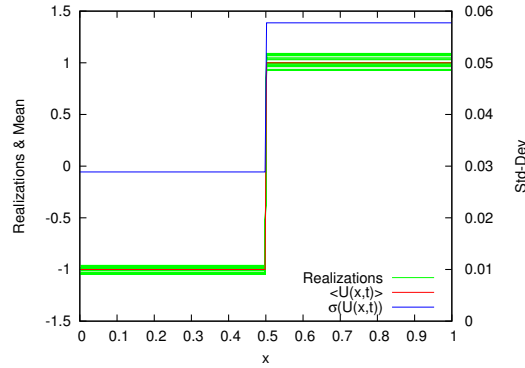


FIGURE 3.18 – Cas test 3 (Burgers) : Échantillon de 20 réalisations de la condition initiale, espérance et écart type. Échelle à droite pour l'écart type.

sont tous deux des fonctions affines par morceaux en la variable spatiale  $x$  à chaque instant.

### 3.1.3.2 Résultats

On compare dans la figure 3.19 la solution de Galerkin à l'instant  $t = 0.3$  obtenue sans correcteur entropique (à gauche) et avec le correcteur entropique proposé dans la section 2.4 (à droite). L'espérance de la solution de Galerkin et l'écart type sont tracés ainsi qu'un échantillon de réalisations. Les réalisations sont, comme précédemment, reconstruites à partir du développement stochastique de la solution en utilisant un unique échantillon de réalisations de  $\xi \in [0, 1]^2$ . Comme dans le cas déterministe, on observe dans le graphe de gauche de la figure 3.19 que le solveur de Roe stochastique sans correcteur entropique ne capture pas la solution entropique du problème. En fait, il capture un choc non entropique ainsi que des petites perturbations dues au fait que les états initiaux gauche et droit sont presque sûrement non symétriques. Avec la correction entropique proposée dans la section 2.4, la solution entropique est bien capturée, comme le montre le graphe de droite de la figure 3.19. De plus, on a vérifié que les points soniques du système de Galerkin sont, comme dans le cas déterministe, détectés seulement aux interfaces localisées en  $x = 0.5$ . En fait, l'incertitude n'est pas vraiment importante pour ce cas test de sorte que les ondes soniques ont presque sûrement la même localisation physique pour les instants que l'on considère. En outre, il s'avère que l'ensemble des indices soniques dans un élément stochastique correspondant à  $x = 0.5$  a pour cardinal  $(N_0 + 1)^2$ , indiquant que toutes les valeurs propres de la matrice Jacobienne de Galerkin changent de signe à l'interface et par conséquent toutes les composantes du flux numérique de Galerkin ont besoin d'être complétées par une correction entropique.

Sur la figure 3.20, les deux premiers moments de la solution de Galerkin sont comparés aux deux premiers moments de la solution stochastique pour deux choix différents du paramètre  $C_{CFL}$ . On voit que l'espérance de la solution de Galerkin et celle de la solution stochastique sont en accord dans les deux cas, même si, comme



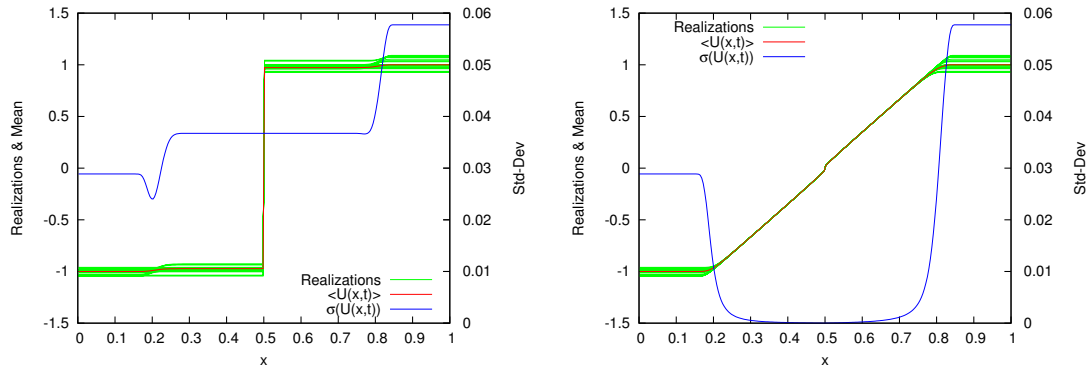


FIGURE 3.19 – Cas test 3 (Burgers) : Solution de Galerkin à l’instant  $t = 0.3$  obtenue sans (à gauche) et avec (à droite) correcteur entropique. L’espérance et l’écart type sont tracés en fonction de  $x$ , ainsi qu’une reconstruction de 20 réalisations. Échelle à droite pour l’écart type.

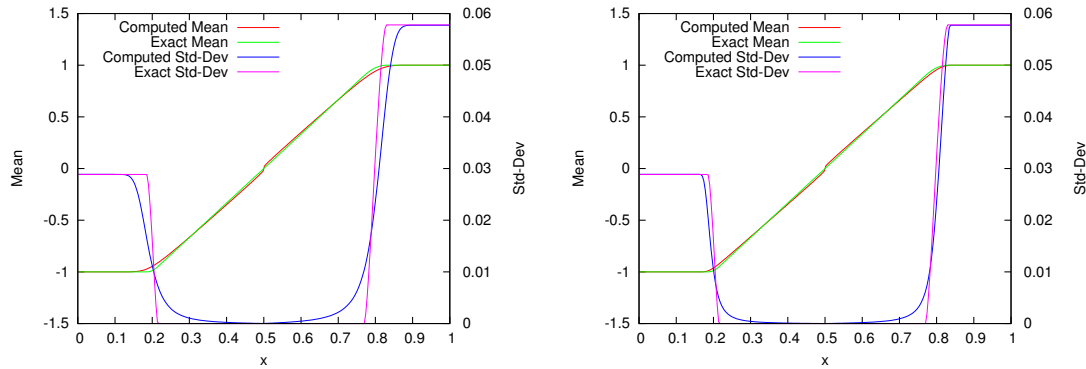


FIGURE 3.20 – Cas test 3 (Burgers) : Comparaison de l’espérance et de l’écart type de la solution de Galerkin à l’instant  $t = 0.3$  avec l’espérance et l’écart type de la solution stochastique pour  $C_{CFL} = 0.7$  (à gauche) et  $C_{CFL} = 0.99$  (à droite). Échelle à droite pour l’écart type.

dans le cas déterministe, un léger saut subsiste en  $x = 0.5$  pour l’espérance de la solution de Galerkin. L’écart type est correctement capturé par le solveur de Roe stochastique : il vaut quasiment zéro sur un intervalle inclus dans  $I(t = 0.3) = [0.215, 0.77]$  (où l’écart type exact vaut zéro). Cependant, des erreurs significatives subsistent au voisinage de  $x = 0.2$  et  $x = 0.8$  en particulier pour le cas  $C_{CFL} = 0.7$  car le solveur de Roe produit un étalement de la solution de Galerkin sur l’espace physique. Ce problème peut être en partie résolu en augmentant le paramètre CFL, comme illustré dans le graphe de droite de la figure 3.20.

Une autre observation intéressante est que la solution de Galerkin dépend presque sûrement uniquement de  $\xi_1$  pour  $x < 0.5$  et uniquement de  $\xi_2$  pour  $x > 0.5$ , comme c’est le cas pour la solution exacte. Cela est montré sur la figure 3.21 où la solution de Galerkin est tracée en fonction de  $\xi = (\xi_1, \xi_2)$  pour  $x = 0.2$  (à gauche) et

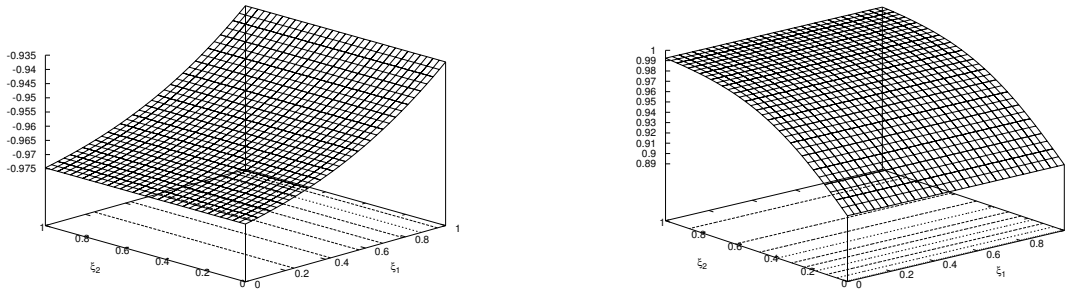


FIGURE 3.21 – Cas test 3 (Burgers) : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  à l'instant  $t = 0.3$  en  $x = 0.2$  (à gauche) et en  $x = 0.8$  (à droite).

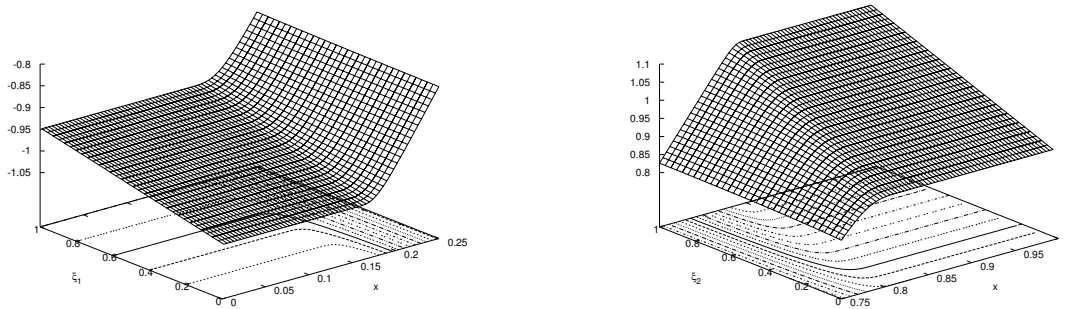


FIGURE 3.22 – Cas test 3 (Burgers) : Solution de Galerkin à l'instant  $t = 0.3$  en fonction de  $(x, \xi_1)$  pour  $0 \leq x \leq 0.25$  (à gauche) et  $(x, \xi_2)$  pour  $0.75 \leq x \leq 1$  (à droite).

$x = 0.8$  (à droite). Sur la figure 3.22, on présente la solution de Galerkin à l'instant  $t = 0.3$  en fonction de  $(x, \xi_1)$  pour  $0 \leq x \leq 0.25$  (à gauche) et en fonction de  $(x, \xi_2)$  pour  $0.75 \leq x \leq 1$  (à droite). Cette figure confirme la pertinence du développement stochastique puisque la solution de Galerkin est continue comme on s'y attend.

Le léger saut qui subsiste au voisinage du point sonique en  $x = 0.5$  n'est pas surprenant. Comme dans le cas déterministe, il résulte de l'utilisation d'un solveur numérique du premier ordre. Puisque l'écart type de la solution est presque sûrement nul au centre du domaine, le saut coïncide presque sûrement avec la valeur obtenue dans le cas déterministe. Le saut peut par conséquent être représenté dans le graphe de gauche de la figure 3.23 en fonction du pas d'espace en prenant la différence entre les valeurs des espérances de la solution calculée sur la cellule de droite et la cellule de gauche sur l'interface localisée en  $x = 0.5$ . Il décroît linéairement avec le pas d'espace ; il s'agit du même comportement que dans le cas déterministe.

Pour analyser davantage la convergence spatiale de la méthode, on considère la

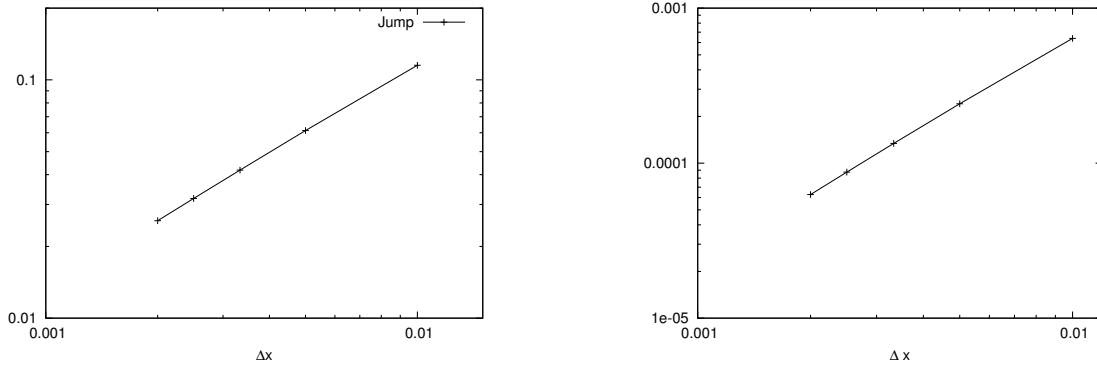


FIGURE 3.23 – Cas test 3 (Burgers) : A gauche : Amplitude du saut de la solution de Galerkin en fonction du pas d'espace  $\Delta x$ . A droite : Erreur stochastique  $\epsilon_{\text{ex}}^2$  en fonction du pas d'espace  $\Delta x$ . Calculs à l'instant  $t = 0.3$ .

mesure d'erreur

$$\epsilon_{\text{ex}}^2(t) := \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^{N_c} \Delta x \left( U_i^P(t, \xi^{(j)}) - U^{\text{ex}}(x_i, t, \xi^{(j)}) \right)^2, \quad (3.16)$$

où, pour chaque élément  $\xi^{(j)}$  d'un échantillon,  $U_i^P(t, \xi^{(j)})$  et  $U^{\text{ex}}(x_i, t, \xi^{(j)})$  sont respectivement évaluées à partir de la solution de Galerkin à la cellule  $i$  de l'espace physique et à partir de la solution exacte du problème déterministe correspondant au centre de la cellule  $i$  de l'espace physique. On utilise un échantillon de taille  $M = 100000$ . Le graphe de droite de la figure 3.23 présente l'erreur  $\epsilon_{\text{ex}}^2$  à l'instant  $t = 0.3$  en fonction du pas d'espace. On observe un taux de décroissance linéaire de l'erreur car l'erreur spatiale domine largement dans ce cas test où la solution est régulière aussi bien en espace physique que stochastique.

### 3.1.3.3 Comparaison de la méthode de Galerkin avec une méthode de Monte Carlo

Finalement, on compare la solution de Galerkin avec une solution Monte Carlo (MC) obtenue à partir de la résolution de 100000 problèmes déterministes, chacun de ces problèmes correspondant à la réalisation de  $\xi$  dans un échantillon. Les deux méthodologies donnent des résultats différents, comme on le voit sur la figure 3.24. Dans le graphe de gauche, la valeur absolue de la différence entre l'écart type de la solution stochastique et celui de la solution de Galerkin est comparée avec la valeur absolue de la différence entre l'écart type de la solution stochastique et celui de la solution MC. On remarque que la méthode de Galerkin donne de meilleurs résultats que la méthode MC au voisinage du point sonique en  $x = 0.5$ . En effet, l'écart type ne vaut pas zéro dans le cas de la méthode MC car celle-ci considère un pas de temps  $\Delta t$  spécifique pour chaque réalisation de  $\xi$ , ce qui crée une variabilité du saut à  $x = 0.5$ . La méthode de Galerkin considère le même pas de temps  $\Delta t$  pour tous les éléments stochastiques (*cf* discussion ci-dessous) et par conséquent surmonte

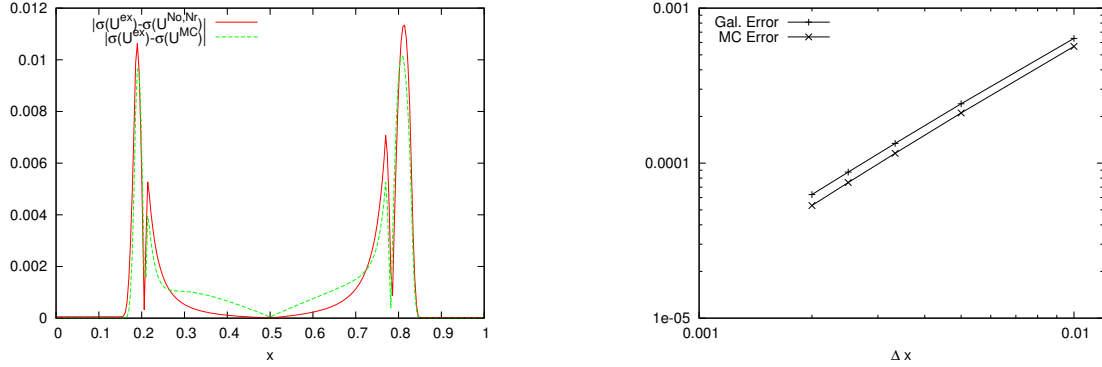


FIGURE 3.24 – Cas test 3 (Burgers) : A gauche : Différence entre l'écart type de la solution stochastique et celui de la solution de Galerkin en  $x = 0.5$  comparée avec la différence entre l'écart type de la solution stochastique et celui de la solution MC. A droite : Erreur stochastique  $\epsilon_{\text{ex}}^2$  en fonction du pas d'espace  $\Delta x$  pour la solution de Galerkin et la solution MC. Calculs à l'instant  $t = 0.3$ .

ces difficultés en introduisant une certaine diffusion numérique. Cependant, cette diffusion crée des erreurs numériques au voisinage de  $x = 0.2$  et  $x = 0.8$  qui sont plus importantes que pour la méthode MC. Pour une comparaison plus précise, on considère la mesure d'erreur

$$\epsilon_{\text{ex,MC}}^2(t) := \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^{N_c} \Delta x \left( U_i^{\text{MC}}(t, \xi^{(j)}) - U^{\text{ex}}(x_i, t, \xi^{(j)}) \right)^2, \quad (3.17)$$

où, pour chaque élément  $\xi^{(j)}$  d'un échantillon,  $U_i^{\text{MC}}(t, \xi^{(j)})$  et  $U^{\text{ex}}(x_i, t, \xi^{(j)})$  sont respectivement évaluées à partir de la solution du problème déterministe correspondant sur la cellule  $i$  de l'espace physique calculée avec le solveur déterministe et à partir de la solution exacte du problème déterministe correspondant au centre de la cellule  $i$  de l'espace physique. À nouveau, un échantillon de taille  $M = 100000$  est utilisé. Dans le graphe de droite de la figure 3.24, on compare les quantités  $\epsilon_{\text{ex,MC}}^2$  et  $\epsilon_{\text{ex}}^2$  en fonction du pas d'espace. L'ordre de convergence observé est un, comme attendu pour le schéma volumes finis de type Roe utilisé pour la discrétisation en espace. L'erreur MC (donnée par  $\epsilon_{\text{ex,MC}}^2(t)$ ) est légèrement plus petite que l'erreur de Galerkin (donnée par  $\epsilon_{\text{ex}}^2(t)$ ). Finalement, on remarque que si l'on prend un  $\Delta t$  constant pour la méthode MC correspondant au  $\Delta t$  utilisé dans la méthode de Galerkin, on obtient les mêmes résultats avec la méthode MC et avec la méthode de Galerkin.

## 3.2 Euler

Cette section présente les résultats numériques pour les équations d'Euler.

### 3.2.1 Cas test 4 : tube à choc sans point sonique

Dans cette section, la méthode est testée sur les équations d'Euler sans point sonique avec un paramètre aléatoire. Le but de ce cas test est de valider dans un premier temps le solveur de Roe stochastique sans correcteur entropique de la section 2.3 sur un système hyperbolique non linéaire de lois de conservation, de sorte que l'hyperbolicité du système de Galerkin n'est pas garantie. On considère le problème du tube à choc de Sod unidimensionnel, où l'écoulement d'un gaz parfait est régi par les équations d'Euler. Pour les équations d'Euler, on adopte les conventions de notation thermodynamique à la place de la convention majuscule/minuscule utilisée jusqu'à présent. Les quantités conservées sont la densité du fluide  $\rho$ , la quantité de mouvement  $q = \rho v$  (où  $v$  est la vitesse) et l'énergie totale  $E = 1/2 \rho v^2 + \rho e$ , où le premier terme est l'énergie cinétique et le second l'énergie interne (par unité de volume). Le tube s'étend sur une unité de longueur et est fermé par deux extrémités rigides en  $x = 0$  et  $x = 1$ . Les conditions aux limites sont  $q = 0$  et  $\frac{\partial \rho}{\partial x} = \frac{\partial E}{\partial x} = 0$  aux extrémités du tube. On utilise  $N_c = 250$  cellules pour la discrétisation du domaine spatial.

#### 3.2.1.1 Définition du problème

On considère une incertitude sur le coefficient adiabatique  $\gamma$  paramétrée en utilisant une unique variable aléatoire  $\xi$  de distribution uniforme sur  $[0, 1]$ . On considère une probabilité uniforme de  $\gamma$  dans l'intervalle  $[1.4, 1.6]$ , telle que la paramétrisation est

$$\gamma(\xi) = 1.4 + 0.2 \xi, \quad \xi \sim \mathcal{U}[0, 1]. \quad (3.18)$$

Avec les notations introduites ci-dessus, on pose

$$U(x, t, \xi) := (\rho(x, t, \xi), q(x, t, \xi), E(x, t, \xi)) \in \mathcal{A}_U, \quad (3.19)$$

où  $\mathcal{A}_U \subset \mathbb{R}^3$  est l'ensemble d'états admissibles tels que la densité et la pression sont positives, et

$$\begin{aligned} F(U; \xi) &:= (F_\rho(U; \xi), F_q(U; \xi), F_E(U; \xi)) \\ &:= (q(\xi), \left(\frac{q^2}{\rho} + p\right)(\xi), \left(\frac{q}{\rho}(E + p)\right)(\xi)) \in \mathbb{R}^3, \end{aligned} \quad (3.20)$$

avec la pression  $p$  donnée par la loi des gaz parfaits

$$p(\rho, q, E) = (\gamma - 1) \left( E - \frac{1}{2} \frac{q^2}{\rho} \right). \quad (3.21)$$

Les conditions initiales sont

$$\rho^0(x) = \begin{cases} 1 & x \in [0, 1/2], \\ 0.125 & x \in ]1/2, 1], \end{cases} \quad v^0(x) = 0, \quad p^0(x) = \begin{cases} 1 & x \in [0, 1/2], \\ 0.125 & x \in ]1/2, 1]. \end{cases} \quad (3.22)$$

Le problème ainsi défini est adimensionné et toutes les composantes de la solution prennent des valeurs au plus d'ordre un.

### 3.2.1.2 Solveur numérique

Les calculs ci-dessous sont en fait réalisés indépendamment sur chaque élément stochastique  $\alpha_\sigma \in \mathcal{P}_\sigma$ .

**Calcul du flux de Galerkin**  $f(u) \in \mathbb{R}^{3\mathcal{P}}$ . On considère le vecteur des modes stochastiques

$$u = (\rho_\alpha, q_\alpha, E_\alpha)_{\alpha \in \mathcal{P}} \in \mathbb{R}^{3\mathcal{P}} \quad (3.23)$$

fournissant les développements

$$\rho^{\mathcal{P}}(\xi) = \sum_{\alpha \in \mathcal{P}} \rho_\alpha \Phi_\alpha(\xi), \quad q^{\mathcal{P}}(\xi) = \sum_{\alpha \in \mathcal{P}} q_\alpha \Phi_\alpha(\xi) \quad \text{et} \quad E^{\mathcal{P}}(\xi) = \sum_{\alpha \in \mathcal{P}} E_\alpha \Phi_\alpha(\xi). \quad (3.24)$$

Alors, le flux de Galerkin a pour composantes  $(f_{\rho_\alpha}, f_{q_\alpha}, f_{E_\alpha})_{\alpha \in \mathcal{P}}$  telles que  $(f_{\rho_\alpha})_{\alpha \in \mathcal{P}}$ ,  $(f_{q_\alpha})_{\alpha \in \mathcal{P}}$  et  $(f_{E_\alpha})_{\alpha \in \mathcal{P}}$  sont les modes stochastiques des composantes  $F_\rho(U^{\mathcal{P}}; \xi)$ ,  $F_q(U^{\mathcal{P}}; \xi)$  et  $F_E(U^{\mathcal{P}}; \xi)$  du flux  $F(U^{\mathcal{P}}; \xi)$  avec  $U^{\mathcal{P}} = (\rho^{\mathcal{P}}, q^{\mathcal{P}}, E^{\mathcal{P}})$ .

Comme motivé dans la section 2.1.4, où on a introduit le produit de Galerkin  $(a * b)$  et l'inverse de Galerkin  $(a^{-*})$ , on utilise une approche pseudo-spectrale [18] pour calculer l'approximation du flux de Galerkin. En procédant étape par étape, on obtient des quantités intermédiaires, telles que l'énergie cinétique et la vitesse du son, qui peuvent être réutilisées à des étapes ultérieures du schéma numérique, en particulier lors du calcul de l'état de Roe. On définit les composantes  $F_\rho^*(U^{\mathcal{P}}; \xi)$ ,  $F_q^*(U^{\mathcal{P}}; \xi)$  et  $F_E^*(U^{\mathcal{P}}; \xi)$  du flux stochastique approché  $F^*(U^{\mathcal{P}}; \xi)$  comme suit :

$$\begin{aligned} F_\rho^*(U^{\mathcal{P}}; \cdot) &= q^{\mathcal{P}}, & F_q^*(U^{\mathcal{P}}; \cdot) &= (q^{\mathcal{P}} * q^{\mathcal{P}}) * (\rho^{\mathcal{P}})^{-*} + p^*, \\ F_E^*(U^{\mathcal{P}}; \cdot) &= v^* * (E^{\mathcal{P}} + p^*), \end{aligned} \quad (3.25)$$

où  $v^* := q^{\mathcal{P}} * (\rho^{\mathcal{P}})^{-*}$  et  $p^* := (\gamma - 1) * (E^{\mathcal{P}} - (q^{\mathcal{P}} * q^{\mathcal{P}}) * (\rho^{\mathcal{P}})^{-*} / 2)$ . On observe que  $F^*(U^{\mathcal{P}}; \xi)$  est seulement une approximation de la projection de Galerkin de  $F(U^{\mathcal{P}}; \xi)$  puisque la composition des opérations de Galerkin élémentaires (produit, inversion) introduit une erreur d'approximation pseudo-spectrale. Le flux de Galerkin pseudo-spectral est alors noté

$$f^*(u) = (f_\alpha^*(u))_{\alpha \in \mathcal{P}} = (f_{\rho_\alpha}^*, f_{q_\alpha}^*, f_{E_\alpha}^*)_\alpha, \quad (3.26)$$

où  $(f_{\rho_\alpha}^*)_{\alpha \in \mathcal{P}}$ ,  $(f_{q_\alpha}^*)_{\alpha \in \mathcal{P}}$  et  $(f_{E_\alpha}^*)_{\alpha \in \mathcal{P}}$  sont les modes stochastiques de  $F^*(U^{\mathcal{P}}; \xi)$ . Au total, le calcul du flux de Galerkin pseudo-spectral requiert quatre produits de Galerkin et une inversion de Galerkin.

**Rappels dans le cas déterministe.** En introduisant l'enthalpie  $H = (E + p)/\rho$ , la matrice Jacobienne de la fonction flux est donnée par

$$\nabla_u f(u) = \begin{pmatrix} 0 & 1 & 0 \\ 1/2(\gamma - 3)v^2 & -(\gamma - 3)v & \gamma - 1 \\ 1/2(\gamma - 1)v^3 - vH & H - (\gamma - 1)v^2 & \gamma v \end{pmatrix}. \quad (3.27)$$

De plus, l'état de Roe à une interface  $LR$  est donné par

$$\rho_{LR}^{\text{Roe}} := \sqrt{\rho_L \rho_R}, \quad (3.28)$$

$$v_{LR}^{\text{Roe}} := (\sqrt{\rho_L} v_L + \sqrt{\rho_R} v_R) / (\sqrt{\rho_L} + \sqrt{\rho_R}), \quad (3.29)$$

$$H_{LR}^{\text{Roe}} := (\sqrt{\rho_L} H_L + \sqrt{\rho_R} H_R) / (\sqrt{\rho_L} + \sqrt{\rho_R}). \quad (3.30)$$

Enfin, la vitesse du son pour l'état de Roe est

$$(c_{LR}^{\text{Roe}})^2 := (\gamma - 1)(H_{LR}^{\text{Roe}} - (v_{LR}^{\text{Roe}})^2/2). \quad (3.31)$$

**Calcul de la matrice Jacobienne de Galerkin.** La matrice Jacobienne de Galerkin pseudo-spectrale  $\nabla_u f^*(u) \in \mathbb{R}^{3P, 3P}$  est donnée par

$$\nabla_u f^*(u) = \left( \sum_{\delta \in \mathcal{P}} (\nabla_U F^*(U^{\text{P}}; \cdot))_{\delta} \mathcal{M}_{\alpha\beta\delta} \right)_{\alpha, \beta \in \mathcal{P}}. \quad (3.32)$$

En introduisant l'enthalpie  $H^* := (E^{\text{P}} + p^*) * \rho^{-*}$ ,  $\nabla_U F^*(U^{\text{P}}; \cdot)$  est définie par

$$\nabla_U F^*(U^{\text{P}}; \cdot) = \begin{pmatrix} 0 & 1 & 0 \\ 1/2(\gamma - 3) * v^{**} & -(\gamma - 3) * v^* & \gamma - 1 \\ 1/2(\gamma - 1) * (v^* * v^{**}) - v^* * H^* & H^* - (\gamma - 1) * v^{**} & \gamma * v^* \end{pmatrix} \quad (3.33)$$

En terme de coûts de calcul, cette approche requiert neuf produits de Galerkin et une inversion de Galerkin, en particulier grâce à la réutilisation du produit de Galerkin  $v^{**} = (v^* * v^*)$ . En outre, on a vérifié numériquement que  $\nabla_u f^*(u) \in \mathbb{R}^{3P, 3P}$  est diagonalisable dans  $\mathbb{R}$ .

**Calcul de  $a_{LR}^{\text{Roe}}$  et de sa valeur absolue.** La densité de Roe, la vitesse, l'enthalpie et la vitesse du son sont approchées sur l'espace d'approximation stochastique  $S^{\text{P}}$  pour une interface donnée  $LR$  sous la forme

$$\rho_{LR}^{\text{Roe},*} := (\rho_L^{\text{P}})^{*/2} * (\rho_R^{\text{P}})^{*/2}, \quad (3.34)$$

$$v_{LR}^{\text{Roe},*} := ((\rho_L^{\text{P}})^{*/2} * v_L^* + (\rho_R^{\text{P}})^{*/2} * v_R^*) * ((\rho_L^{\text{P}})^{*/2} + (\rho_R^{\text{P}})^{*/2})^{-*}, \quad (3.35)$$

$$H_{LR}^{\text{Roe},*} := ((\rho_L^{\text{P}})^{*/2} * H_L^* + (\rho_R^{\text{P}})^{*/2} * H_R^*) * ((\rho_L^{\text{P}})^{*/2} + (\rho_R^{\text{P}})^{*/2})^{-*}, \quad (3.36)$$

$$(c_{LR}^{\text{Roe},*})^2 := (\gamma - 1) * (H_{LR}^{\text{Roe},*} - (v_{LR}^{\text{Roe},*} * v_{LR}^{\text{Roe},*})/2). \quad (3.37)$$

On obtient ainsi l'état de Roe  $U_{LR}^{\text{Roe},*}$  et la matrice de Roe pseudo-spectrale  $a_{LR}^{\text{Roe},*}$  est donnée par

$$a_{LR}^{\text{Roe},*} := \left( \sum_{\delta \in \mathcal{P}} (\nabla_U F^*(U_{LR}^{\text{Roe},*}; \cdot))_{\delta} \mathcal{M}_{\alpha\beta\delta} \right)_{\alpha, \beta \in \mathcal{P}}. \quad (3.38)$$

Finalemnt, la valeur absolue de  $a_{LR}^{\text{Roe},*}$  est calculée comme décrit dans les sections 2.3.2 et 2.3.3, en utilisant les valeurs propres de  $\nabla_U F^*(U_{LR}^{\text{Roe},*}(\xi); \xi)$  aux points de Gauss dans chaque élément stochastique. On rappelle que ces valeurs propres sont données par

$$(v_{LR}^{\text{Roe},*} \pm c_{LR}^{\text{Roe},*})(\xi_{\eta})_{\eta \in \mathcal{P}_{\pi}} \quad \text{et} \quad v_{LR}^{\text{Roe},*}(\xi_{\eta})_{\eta \in \mathcal{P}_{\pi}}. \quad (3.39)$$

### 3.2.1.3 Résultats

Dans cette section, on présente les résultats pour le problème du tube à choc avec une incertitude sur le coefficient adiabatique. On commence par une analyse générale des résultats, en prenant  $\text{No} = 2$  et  $\text{Nr} = 3$  comme paramètres de discrétisation stochastique, de sorte que la dimension de l'espace d'approximation stochastique est  $P = 24$ .

Dans le cas déterministe et pour la condition initiale (3.22) avec une certaine réalisation de  $\gamma(\xi)$ , une onde de choc issue de la discontinuité initiale se propage vers la droite, tandis qu'une onde de détente plus lente se propage vers la gauche du domaine et une discontinuité de contact (également plus lente) se propage vers la droite. Lorsque les ondes atteignent les parois du tube, elles sont réfléchies vers l'intérieur et se propagent l'une vers l'autre, se rencontrent et interagissent.

Ici, la vitesse du son incertaine va affecter la vitesse du choc, de la discontinuité de contact et de l'onde de détente. Des solutions pour différentes réalisations de  $\gamma(\xi)$  exhibent des comportements similaires à ceux du cas déterministe, mais avec des pentes différentes pour le choc, la discontinuité de contact et l'onde de détente dans le diagramme espace-temps. On peut observer ces comportements sur la figure 3.25 où sont tracées la densité dans le cas déterministe (avec un coefficient adiabatique fixé à  $\langle \gamma \rangle$ ) et l'espérance de la solution de Galerkin dans le cas stochastique. L'étalement de la localisation du choc et de la discontinuité de contact avec le temps est visible, alors que pour l'onde de détente, qui est déjà continue dans le cas déterministe, l'impact de la vitesse du son incertaine est moins prononcé.

L'impact de l'incertitude peut également être apprécié sur les écarts type de la densité, présentés sur la figure 3.26 pour des temps courts et plus longs. Les valeurs les plus importantes de l'écart type sont observées le long du chemin de l'onde de choc, les valeurs maximales correspondant aux instants où l'onde de choc est réfléchi sur les parois du tube. Pour des temps courts ( $t \leq 0.25$ ), l'incertitude est présente uniquement dans les zones qui peuvent être atteintes par l'onde de choc ou de détente dans le temps imparti pour au moins une réalisation. Le premier élément à l'origine de niveaux d'incertitude plus élevés est l'interaction du choc avec



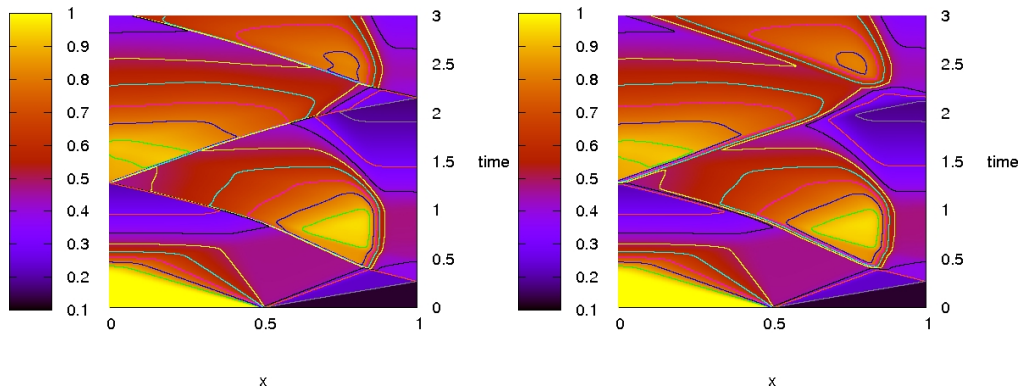


FIGURE 3.25 – Cas test 4 (Euler) : Diagramme espace-temps de la densité déterministe (à gauche) et de l'espérance de la densité de Galerkin calculée avec les paramètres  $No = 2$  et  $Nr = 3$  (à droite).

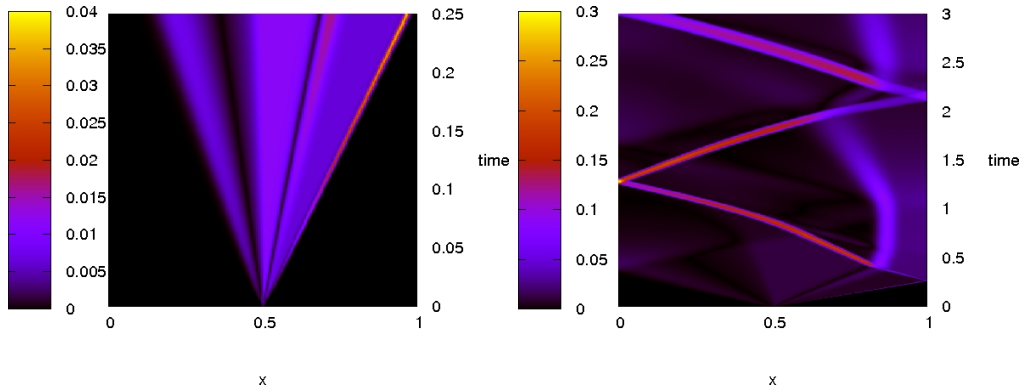


FIGURE 3.26 – Cas test 4 (Euler) : Diagramme espace-temps de l'écart type de la densité de Galerkin pour des temps courts (à gauche) et plus longs (à droite) calculés avec les paramètres  $No = 2$  et  $Nr = 3$ . Différentes échelles de couleur sont utilisées.

la paroi puisque l'arrivée du choc sur la paroi cause une augmentation abrupte de la densité sur une courte période de temps. L'incertitude dans l'instant d'arrivée du choc induit par conséquent une large variabilité dans la solution. Le second événement à l'origine de niveaux d'incertitude plus élevés est l'interaction entre le choc incertain, la discontinuité de contact et l'onde de détente.

Pour vérifier la validité du développement stochastique, on montre dans la figure 3.27 une reconstruction de la densité de Galerkin  $\rho^P$  à divers instants. La densité est initialement discontinue dans la direction  $x$ . Lorsque le temps augmente, la densité devient discontinue aussi bien dans la direction  $x$  que dans la direction  $\xi$  puisque l'onde de choc se propage avec une vitesse incertaine. Dans le plan  $(x, \xi)$ , la discontinuité devient de plus en plus oblique, conséquence d'une dépendance monotone de la vitesse du choc en  $\xi$ . Pour les points  $(x, \xi)$  qui ne sont pas situés trop près de la discontinuité, la solution est régulière et apparaît approchée de façon précise

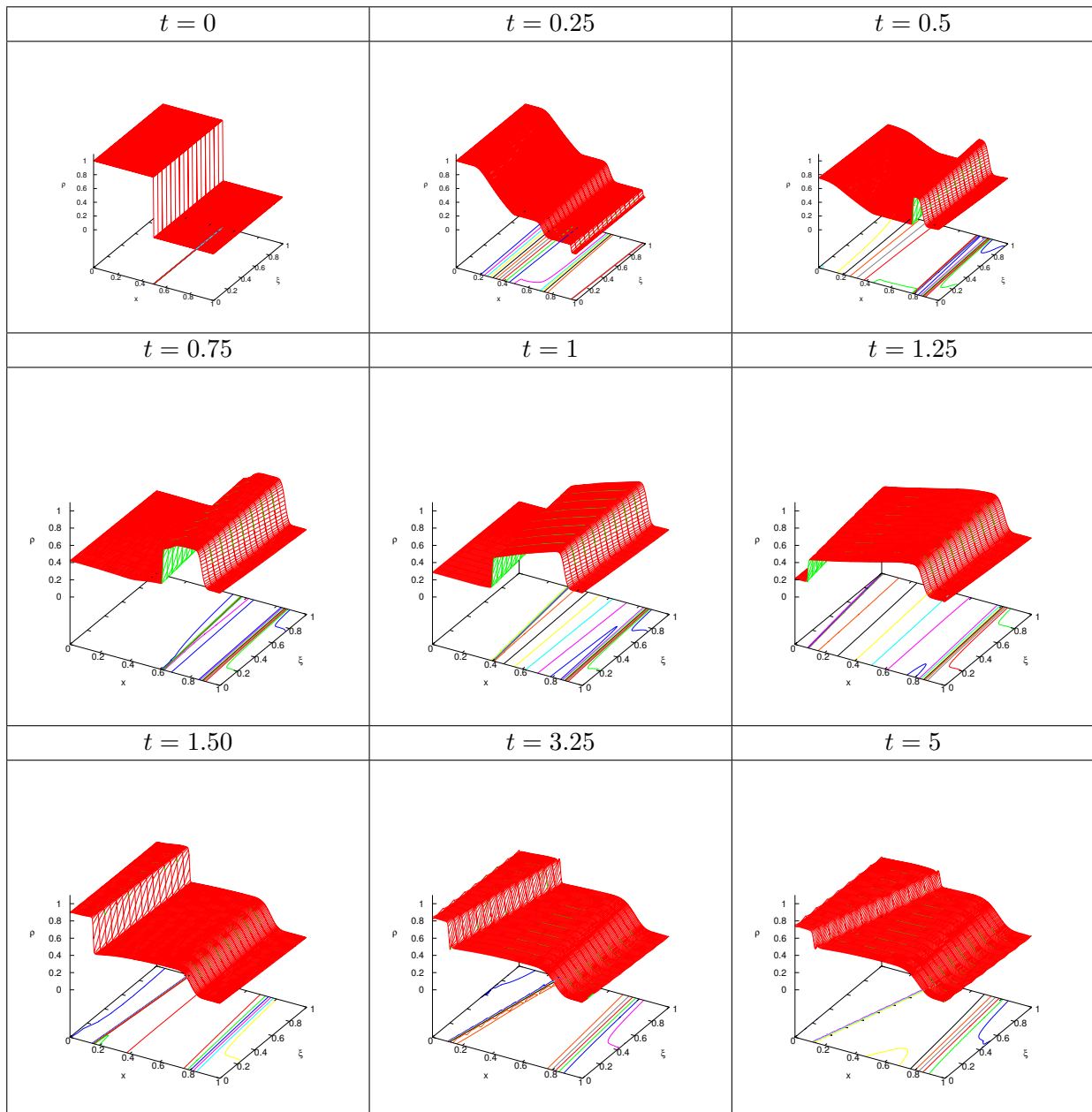


FIGURE 3.27 – Cas test 4 (Euler) : Reconstruction de la densité de Galerkin  $\rho^P$  à des instants différents.

dans l'espace d'approximation stochastique. Dans le voisinage de la discontinuité, la solution exhibe des petites oscillations non physiques qui sont déclenchées par le phénomène de Gibbs, de sorte que la densité prend des valeurs légèrement en dehors des bornes fournies par la condition initiale. De telles oscillations sont en fait causées par une résolution stochastique insuffisante et peuvent être réduites en augmentant le niveau de résolution et/ou l'ordre polynomial du développement

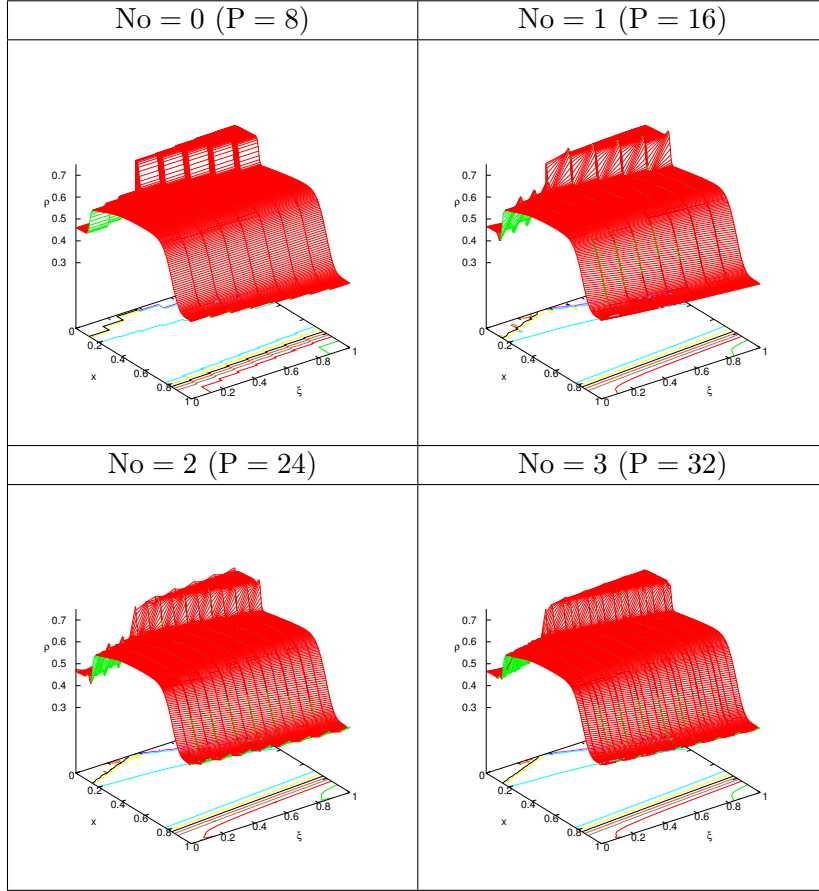


FIGURE 3.28 – Cas test 4 (Euler) : Convergence de la densité de Galerkin  $\rho^P$  avec No. Nr = 3,  $t = 6.5$ .

stochastique. Cela est illustré dans les Figures 3.28 et 3.29, qui montrent la convergence de la densité de Galerkin lorsque les valeurs de No et Nr augmentent. Les oscillations deviennent plus petites lorsque la résolution stochastique augmente.

On définit la mesure d'erreur sur la densité

$$\epsilon_h(x_i, t) := \left( \frac{1}{M} \sum_{j=1}^M \left( \rho_i^P(t, \xi^{(j)}) - \rho_i^h(t, \xi^{(j)}) \right)^2 \right)^{1/2}, \quad (3.40)$$

où, pour chaque élément  $\xi^{(j)}$  d'un échantillon,  $\rho_i^P(t, \xi^{(j)})$  et  $\rho_i^h(t, \xi^{(j)})$  sont respectivement évaluées à partir du développement stochastique de la solution de Galerkin de la cellule  $i$  de l'espace physique et à partir de la solution du problème déterministe d'Euler correspondant sur la cellule  $i$  de l'espace physique calculée avec le solveur déterministe. On utilise un échantillon de taille  $M = 10000$ . La figure 3.30 présente l'erreur  $\epsilon_h(x_i, t)$  pour des instants courts et plus longs en utilisant les paramètres  $N_c = 250$ ,  $N_r = 3$  et  $N_o = 2$ . Pour des temps courts, l'incertitude ne s'est pas encore propagée sur tout le domaine spatial. On peut distinguer trois zones d'erreur correspondant aux voisinages des trois ondes. Comme attendu, l'erreur atteint

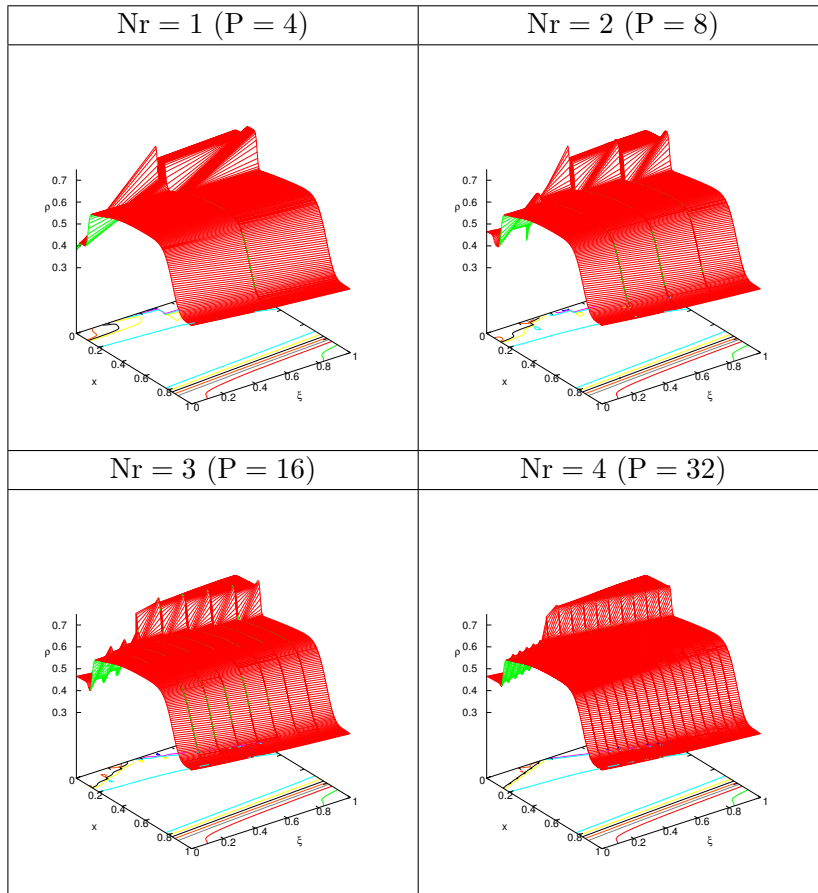


FIGURE 3.29 – Cas test 4 (Euler) : Convergence de la densité de Galerkin  $\rho^P$  avec Nr. No = 1,  $t = 6.5$ .

son maximum au voisinage du choc. Lorsque le temps augmente, les incertitudes se propagent dans tout le domaine. La zone d'erreur correspondant au choc s'étale et la discontinuité dans le plan  $(x, \xi)$  devient plus oblique. De plus, après plusieurs réflexions des ondes sur les parois ( $t = 6.5$ ), l'erreur reste faible indiquant qu'aucune instabilité ne se développe pour des temps plus longs. Dans la figure 3.31, on examine la convergence de l'erreur  $\epsilon_h(x, t = 6.5)$  lorsque la valeur de No ou de Nr augmente, confirmant que ces deux paramètres peuvent être utilisés pour améliorer la résolution stochastique.

Pour compléter la discussion, on fournit une brève estimation de l'efficacité de la méthode numérique pour ce cas test. On montre dans le tableau 3.1 l'évolution des temps de calcul pour différentes discrétisations stochastiques. Les temps CPU ( $T_{\text{CPU}}$ ) sont évalués pour une intégration des équations d'Euler jusqu'à l'instant  $t = 3$  sur une grille spatiale comprenant  $N_c = 250$  cellules. Ces temps sont par ailleurs normalisés par le temps de calcul utilisant No = Nr = 0, qui représente le temps de calcul pour le problème déterministe. Puisque le pas de temps pour l'intégration est basé sur une CFL fixée, celui-ci dépend également de la discrétisation

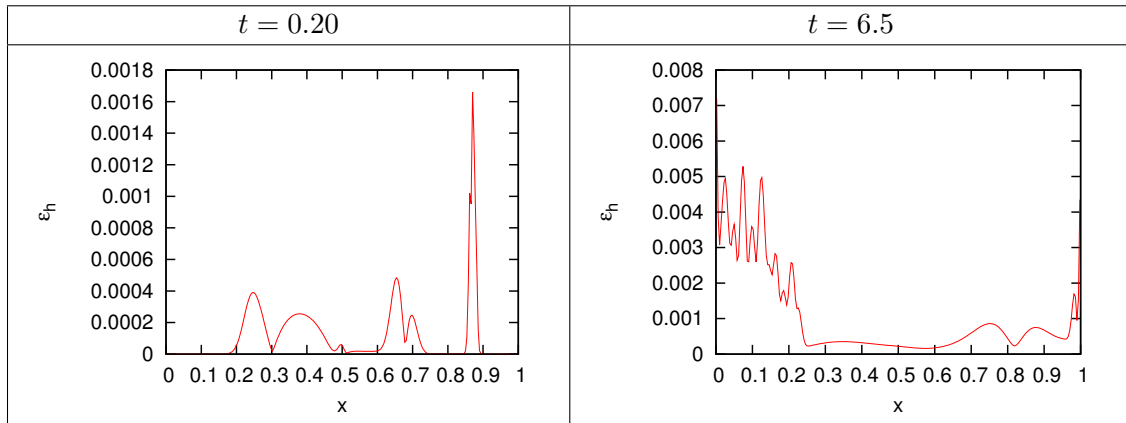


FIGURE 3.30 – Cas test 4 (Euler) : Erreur stochastique  $\epsilon_h(x_i, t)$  à des instants courts (à gauche) et plus longs (à droite). Calculs avec  $N_c = 250$ ,  $N_r = 3$  et  $N_o = 2$ .

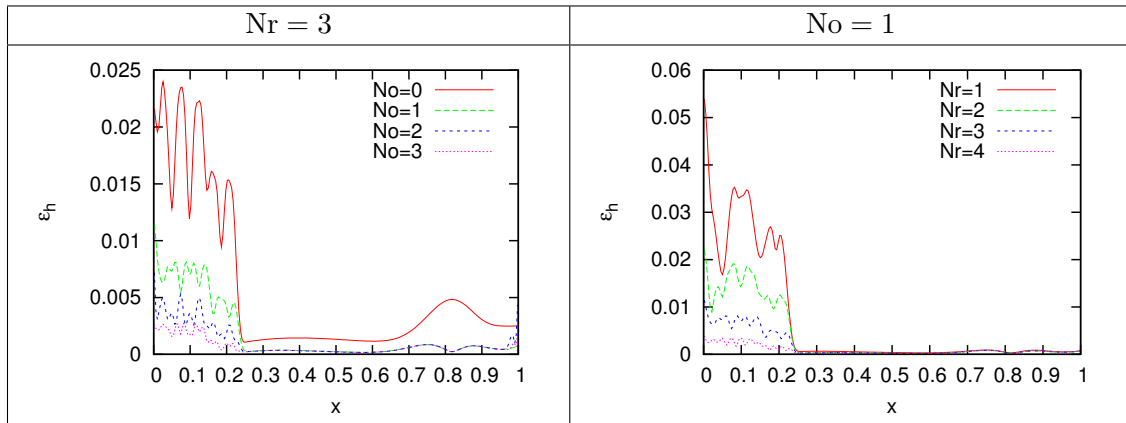


FIGURE 3.31 – Cas test 4 (Euler) : Erreur stochastique  $\epsilon_h(x_i, t = 6.5)$  pour différents  $N_o$  et  $N_r$ . Calculs avec  $N_c = 250$ .

sation stochastique : les temps mesurés correspondent à l'exécution d'un nombre d'itérations différents. Toutefois, on observe uniquement 0.5% de variabilité environ dans le nombre d'itérations en temps entre les simulations utilisant une discrétisation stochastique plus ou moins fine. Les temps CPU peuvent donc être interprétés comme les temps d'exécution d'un nombre fixé d'itérations en temps.

Le tableau 3.1 montre un scaling linéaire avec le nombre  $2^{N_r}$  d'éléments stochastiques pour un ordre polynomial  $N_o$  fixé. Ce scaling était attendu et est obtenu grâce au découplage du problème de Galerkin sur les éléments stochastiques grâce à l'introduction de la base SE. Ce scaling linéaire par rapport au nombre d'éléments stochastiques est également attendu pour des problèmes de dimension stochastique plus grande ( $N > 1$ ); cependant, le temps CPU global augmente significativement avec  $N$  à cause de la croissance exponentielle de la base stochastique locale (sauf pour  $N_o = 0$ ). Pour un niveau de résolution  $N_r$  fixé, le scaling de  $T_{CPU}$  avec la dimension de  $S^P$  est approximativement linéaire au moins pour  $N_o \leq 4$ . Il y a deux effets. Pre-

**Tableau 3.1** Cas test 4 (Euler) : Temps de calcul normalisés  $T_{\text{CPU}}$  pour différents paramètres de discrétisation stochastique  $N_r$  et  $N_o$ .

	$N_r = 2$		$N_r = 3$		$N_r = 4$	
	$T_{\text{CPU}}$	P	$T_{\text{CPU}}$	P	$T_{\text{CPU}}$	P
$N_o = 0$	4.0	(4)	8.1	(8)	16.1	(16)
$N_o = 1$	6.9	(8)	13.9	(16)	27.8	(32)
$N_o = 2$	11.8	(12)	23.2	(24)	46.5	(48)
$N_o = 3$	17.1	(16)	34.1	(32)	68.1	(64)
$N_o = 4$	24.8	(20)	49.3	(40)	98.0	(80)

mièrement, l'évaluation spectrale des non-linéarités dans le flux et dans les états de Roe a une complexité (nombre d'opérations) qui est essentiellement proportionnelle au nombre de termes non nuls dans le tenseur de multiplication du troisième ordre  $\mathcal{M}_{\alpha\beta\delta}$ , qui lui-même augmente exponentiellement avec  $N_o$ . Deuxièmement, lorsque  $N_o$  augmente, on doit utiliser un degré plus élevé  $d_{\text{fit}}$  pour l'approximation polynomiale des matrices de décentrement, entraînant des coûts de calcul plus élevés. Le deuxième effet peut être tempéré, suite aux expériences numériques sur l'équation de Burgers, en limitant  $d_{\text{fit}}$  à une petite valeur ; les simulations présentées dans cette section utilisent  $d_{\text{fit}} + 1 = \min(9, 3(N_o + 1))$ . Aucun effet significatif sur la solution n'a été observé en prenant des degrés  $d_{\text{fit}}$  plus élevés pour  $N_o \geq 3$ . Dans l'ensemble, la complexité des calculs des non-linéarités apparaît comme le facteur le plus limitant de la méthode, et on s'attend à ce que la situation soit encore plus défavorable pour des problèmes à grande dimension stochastique  $N$  (*cf* par exemple [51]) : cette tendance plaide pour l'utilisation d'espaces d'approximation stochastique de degré modéré pour les problèmes stochastiques non réguliers.

### 3.2.2 Cas test 5 : tube à choc avec point sonique

Dans cette section, la méthode est testée sur les équations stochastiques d'Euler avec un paramètre aléatoire en présence de points soniques. Le but de ce cas test est de valider le solveur de Roe avec correcteur entropique de la section 2.4 sur un système hyperbolique non linéaire de lois de conservation, tel que l'ensemble des vecteurs propres du système de Galerkin est plus complexe que dans le cas scalaire. De plus, on propose un cas test où des points soniques apparaissent juste dans une portion du domaine stochastique. La configuration du problème considéré est la même que précédemment, à savoir qu'on considère le problème du tube à choc de Sod avec les mêmes conditions aux limites et on utilise les mêmes conventions de notation. On utilise à nouveau  $N_c = 250$  cellules pour la discrétisation du domaine spatial.

### 3.2.2.1 Définition du problème

On considère à présent une incertitude sur le nombre de Mach initial paramétré par une variable aléatoire unique  $\xi$  de distribution uniforme dans  $[0, 1]$  :

$$\text{Ma}^0(\xi) = \begin{cases} 0.7 + 0.5\xi & x \in [0, 1/4], \\ 2.46 \times (0.7 + 0.5\xi) & x \in ]1/4, 1]. \end{cases} \quad (3.41)$$

On rappelle les vecteurs des variables conservatives

$$U(x, t, \xi) = (\rho(x, t, \xi), q(x, t, \xi), E(x, t, \xi)) \in \mathcal{A}_U, \quad (3.42)$$

et les fonctions flux

$$\begin{aligned} F(U; \xi) &= (F_\rho(U; \xi), F_q(U; \xi), F_E(U; \xi)) \\ &= (q(\xi), (q^2/\rho + p)(\xi), (v(E + p))(\xi)) \in \mathbb{R}^3, \end{aligned} \quad (3.43)$$

où  $p$  est la pression donnée par la loi des gaz parfaits (3.21) avec un coefficient adiabatique  $\gamma = 1.4$ . Les autres conditions initiales sont

$$p^0(x) = \begin{cases} 0.05 & x \in [0, 1/4], \\ 0.0004 & x \in ]1/4, 1], \end{cases} \quad \rho^0(x) = \begin{cases} 1.4 & x \in [0, 1/4], \\ 0.042 & x \in ]1/4, 1]. \end{cases} \quad (3.44)$$

### 3.2.2.2 Solveur numérique

À chaque interface LR, le flux numérique  $\varphi^{\text{DM}}(u_L, u_R)$  défini par (2.55) doit être évalué. On approche les quantités stochastiques présentes dans le flux numérique par leur projection dans  $S^{\text{P}}$  que l'on calcule de manière pseudo-spectrale comme décrit ci-dessus. La première composante de  $\varphi^{\text{DM}}(u_L, u_R)$  est le flux numérique de type Roe  $\varphi^{\text{Roe}}(u_L, u_R)$  dont la construction a été détaillée en section 3.2.1.2. On décrit ici l'évaluation de la partie de correction de  $\varphi^{\text{DM}}(u_L, u_R)$  pour le cas des équations d'Euler. À cette fin, on a besoin des valeurs propres et des vecteurs propres approchés de  $a_{LR}^{\text{Roe},*}$ . On rappelle que la matrice Jacobienne de Galerkin  $\nabla_u f(u) \in \mathbb{R}^{3\text{P}, 3\text{P}}$  est approchée par (3.32). Par conséquent, les valeurs propres de  $a_{LR}^{\text{Roe},*}$  peuvent être approchées par l'approximation dans  $S^{\text{P}}$  des valeurs propres stochastiques de  $\nabla_U F^*(U_{LR}^{\text{Roe},*}(\xi); \xi)$  aux points de Gauss dans chaque élément stochastique, à savoir

$$(v_{LR}^{\text{Roe},*} \pm c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta \in \mathcal{P}_\pi}, \text{ et } v_{LR}^{\text{Roe},*}(\xi_\eta)_{\eta \in \mathcal{P}_\pi}, \quad (3.45)$$

où  $v_{LR}^{\text{Roe}}(\xi)$  et  $c_{LR}^{\text{Roe}}(\xi)$  sont respectivement le vitesse et la vitesse du son pour l'état de Roe  $U_{LR}^{\text{Roe},*}$ . De plus, les vecteurs propres de  $a_{LR}^{\text{Roe},*}$  peuvent être approchés par les vecteurs  $(\omega_\eta R^{k,*}(\xi_\eta) \Phi_\beta)_{\beta, \eta \in \mathcal{P}_\pi, k \in \{1, 2, 3\}}$ , où  $(R^{k,*}(\xi))_{k \in \{1, 2, 3\}}$  sont les approximations dans  $S^{\text{P}}$  des vecteurs propres stochastiques de  $\nabla_U F^*(U_{LR}^{\text{Roe},*}(\xi); \xi)$ , données par

$$R^{1,*} = (1, v_{LR}^{\text{Roe},*} + c_{LR}^{\text{Roe},*}, H_{LR}^{\text{Roe},*} + v_{LR}^{\text{Roe},*} * c_{LR}^{\text{Roe},*})^T, \quad (3.46)$$

$$R^{2,*} = (1, v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*}, H_{LR}^{\text{Roe},*} - v_{LR}^{\text{Roe},*} * c_{LR}^{\text{Roe},*})^T, \quad (3.47)$$

$$R^{3,*} = (1, v_{LR}^{\text{Roe},*}, H_{LR}^{\text{Roe},*} - (c_{LR}^{\text{Roe},*} * c_{LR}^{\text{Roe},*})/(\gamma - 1))^T, \quad (3.48)$$

où  $H^{\text{Roe},*}$  est définie par (3.36). Ainsi, on a tous les outils nécessaires pour calculer les états intermédiaires approchés  $\{u'_\gamma\}_{\gamma \in \{1, \dots, mP_\pi\}}$  pour toutes les interfaces  $LR$  et pour déterminer l'ensemble des indices soniques  $S'$ .

### 3.2.2.3 Résultats

Dans cette section, on présente les résultats pour le problème du tube à choc de Sod avec une incertitude sur le nombre de Mach initial. On commence par une analyse générale des résultats, en prenant pour paramètres de discrétisation stochastique  $No = 2$  et  $Nr = 3$ , si bien que la dimension de l'espace d'approximation stochastique est  $P = 24$ .

Dans le cas déterministe et pour la condition initiale (3.44), on obtient le même comportement que pour le cas test précédent avec la génération d'une onde de choc, d'une onde de détente et d'une discontinuité de contact. Ici, c'est le nombre de Mach initial incertain qui va affecter la vitesse des ondes. De plus, l'incertitude dans le nombre de Mach initial est telle que des points soniques sont générés uniquement pour  $\xi \in [0, 0.6]$ . Pour vérifier la validité du développement stochastique et l'efficacité du correcteur entropique, on montre sur la figure 3.32 une reconstruction de la densité de Galerkin  $\rho^P$  à l'instant  $t = 1$  obtenue sans (à gauche) et avec (à droite) correcteur entropique. Sur le graphe de gauche, on observe qu'un choc non entropique est généré au voisinage de  $x = 0.2$  pour  $\xi \in [0, 0.6]$ , alors que pour  $\xi \in [0.6, 1]$  l'onde de détente est correctement capturée. Lorsque l'on utilise le correcteur entropique, comme dans le graphe de droite de la figure 3.32, l'onde de détente sonique est correctement capturée pour  $\xi \in [0, 1]$ , même si un léger saut subsiste à cause du solveur numérique qui est du premier ordre. Comme pour l'équation de Burgers (*cf* le graphe de gauche de la figure 3.23), ce saut converge vers zéro au premier ordre avec le pas d'espace  $\Delta x$ . On observe également que le solveur capture correctement la discontinuité aussi bien dans l'espace physique que

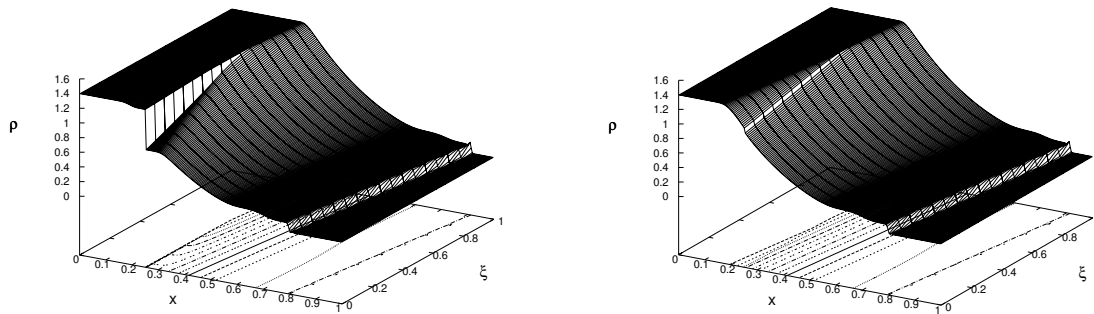


FIGURE 3.32 – Cas test 5 (Euler) : Reconstruction de la densité de Galerkin  $\rho^P$  à l'instant  $t = 1$  obtenue sans (à gauche) et avec (à droite) correcteur entropique. Calculs avec  $Nr = 3$  et  $No = 2$ .



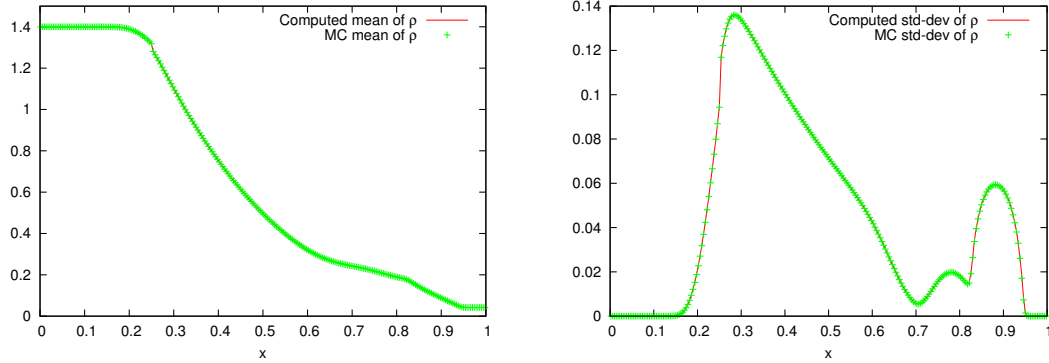


FIGURE 3.33 – Cas test 5 (Euler) : Comparaison à l’instant  $t = 1$  de l’espérance et de l’écart type de la densité de Galerkin (utilisant  $N_r = 3$  et  $N_o = 2$ ) et ceux par une méthode de MC (utilisant un échantillon de taille 10000).

dans l’espace stochastique au voisinage de l’onde de choc ; ce point a déjà été étudié dans le cas test précédent.

On compare dans la figure 3.33 à l’instant  $t = 1$  l’espérance et l’écart type de la densité de Galerkin avec ceux calculés par la méthode MC à partir de la résolution de problèmes d’Euler déterministes (discrets) correspondant aux réalisations d’un échantillon de taille 10000. Suivant les remarques de la section 3.1.3.3, on calcule la solution pour les deux méthodes avec un pas de temps fixé  $\Delta t$  ; pour simplifier, ce pas de temps est pris égal au minimum de tous les pas de temps obtenus avec le solveur de Galerkin. On remarque que les espérances et écarts type sont en excellent accord.

### 3.2.2.4 Amélioration des performances CPU

On décrit dans cette section une technique permettant de réduire les temps de calcul lors de l’étape de détection des points soniques. Pour illustrer l’idée, on présente à la figure 3.34 à l’instant  $t = 1$  et pour des interfaces sélectionnées à différentes positions dans l’espace physique, les valeurs propres approchées  $(v_{LR}^{\text{Roe},*} \pm c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta \in \mathcal{P}_\pi}$  et  $v_{LR}^{\text{Roe},*}(\xi_\eta)_{\eta \in \mathcal{P}_\pi}$ . On évalue ces valeurs propres sur chaque élément stochastique et on présente les fonctions de densité associées à ces valeurs propres. Cette figure mérite différents commentaires. Tout d’abord, on remarque que les fonctions de densité ont des comportements complètement différents pour les différentes interfaces considérées. Pour  $x = 0.15$ ,  $x = 0.2$  et  $x = 0.25$ , les ensembles de valeurs propres correspondant à  $v - c$ ,  $v$  et  $v + c$  ont des supports distincts, alors que les trois ensembles ont des supports qui se chevauchent pour  $x = 0.9$  (et sont par conséquent représentés sur trois graphes différents au bas de la figure 3.34). Pour l’interface située en  $x = 0.9$ , qui est affectée par l’onde de choc, il est intéressant de remarquer que chaque ensemble de valeurs propres peut être séparé en deux sous-ensembles différents, excepté pour quelques valeurs propres qui correspondent en fait à la discontinuité dans l’espace stochastique. A l’inverse, pour

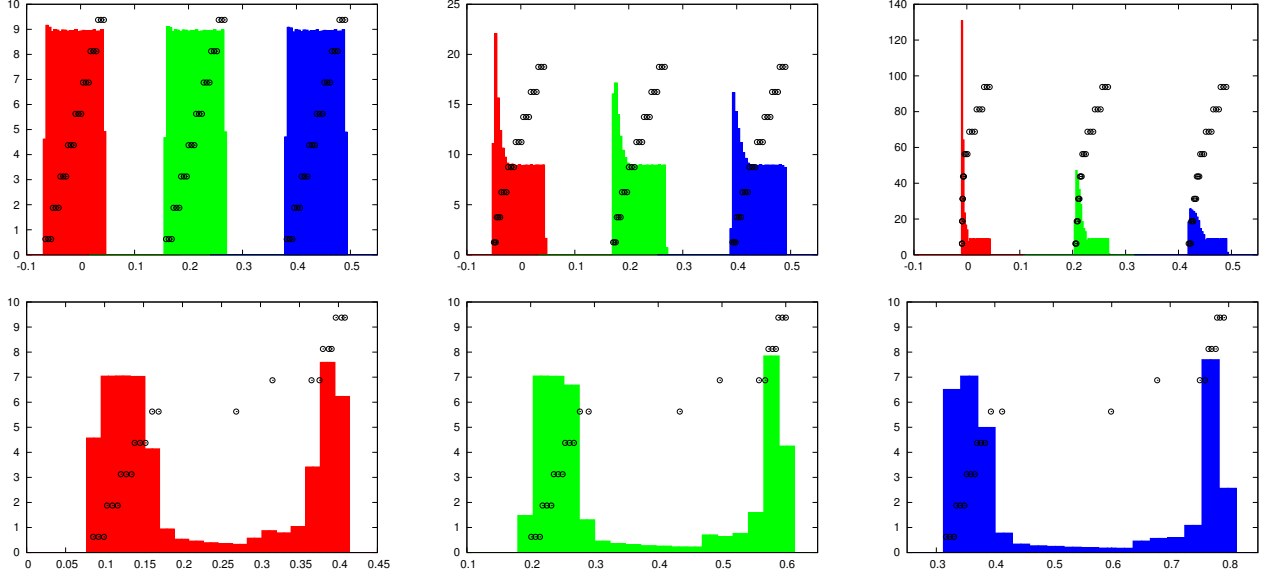


FIGURE 3.34 – Cas test 5 (Euler) : Valeurs propres approchées  $(v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta \in \mathcal{P}_\pi}$  (rouge/gris),  $v_{LR}^{\text{Roe},*}(\xi_\eta)_{\eta \in \mathcal{P}_\pi}$  (vert/gris clair) et  $(v_{LR}^{\text{Roe},*} + c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta \in \mathcal{P}_\pi}$  (bleu/gris foncé) correspondant à chaque élément stochastique avec leurs fonctions de densité pour  $x = 0.15$ ,  $x = 0.2$ ,  $x = 0.25$  (en haut) et  $x = 0.9$  (en bas). Calculs à l’instant  $t = 1$  avec  $\text{Nr} = 3$  et  $\text{No} = 2$ .

les trois graphes en haut de la figure 3.34, les valeurs propres de chaque élément stochastique sont très proches et les fonctions de densité atteignent leur valeur maximale lorsque les valeurs propres sont pratiquement égales, *i.e.*, pour les éléments stochastiques supportant une faible variabilité. Ce phénomène est particulièrement visible en  $x = 0.2$  et  $x = 0.25$ . En outre, les plateaux dans la fonction de densité correspondent aux éléments stochastiques où les valeurs propres ont localement des distributions uniformes. Une autre remarque importante est que les trois premiers graphes correspondent aux positions localisées au voisinage des points soniques, où seules les valeurs propres  $(v_{LR}^{\text{Roe},*} - c_{LR}^{\text{Roe},*})(\xi_\eta)_{\eta \in \mathcal{P}_\pi}$  prennent des valeurs négatives.

Sur la base de ces observations, l’idée clé pour réduire le temps CPU est de tester la présence de points soniques uniquement sur les éléments stochastiques  $\alpha_\sigma \in \mathcal{P}_\sigma$  où  $E^{\alpha_\sigma}[v - c]$  change de signe à l’interface :

$$E^{\alpha_\sigma}[(v_L^* - c_L^*)] - \text{ctol} * v_{ref} < 0 < E^{\alpha_\sigma}[(v_R^* - c_R^*)] + \text{ctol} * v_{ref}, \quad (3.49)$$

où  $\text{ctol}$  est un paramètre fixé et  $v_{ref}$  est une vitesse de référence ici prise égale à un conformément à la taille des valeurs propres présentées dans la figure 3.34.

On montre dans la figure 3.35 la portion du domaine  $(x, \xi)$  qui est sélectionnée pour la correction entropique et le nombre de points soniques détectés ( $\text{card } S'$ ) pour deux valeurs différentes du paramètre  $\text{ctol}$ . Pour  $\text{ctol} = 10^{-2}$  (à gauche), on observe une portion rectangulaire du domaine sur la gauche qui est positive au test (3.49) mais qui ne correspond pas en réalité à des points soniques. Les

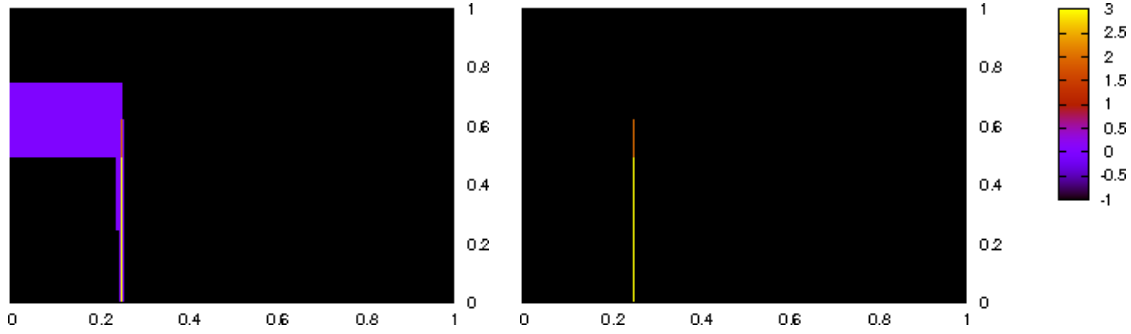


FIGURE 3.35 – Cas test 5 (Euler) : Portion du domaine  $(x, \xi)$  sélectionnée pour la correction entropique et card  $S'$ . Valeur=-1 si pas de correction. Valeur=0 si le test (3.49) est positif sans détection de point sonique. Card  $S'$  sinon. Calculs avec  $ctol = 10^{-2}$  (à gauche) et  $ctol = 10^{-5}$  (à droite).

interfaces marquées correspondent en fait à la portion du domaine où la condition initiale passe du cas sonique au cas subsonique. Pour  $ctol = 10^{-5}$  (à droite), la portion du domaine positive au test (3.49) correspond bien aux points soniques. Dans les deux cas, trois points soniques sont détectés dans les quatre éléments stochastiques correspondant à  $\xi \in [0, 0.5]$ , et deux points soniques sont détectés dans le cinquième élément stochastique.

Pour compléter la discussion, on fournit une estimation des temps CPU dans le tableau 3.2. Les temps CPU ( $T_{CPU}$ ) correspondent à une intégration des équations d'Euler jusqu'à l'instant  $t = 1$  sur un maillage fixé de  $N_c = 250$  cellules, et sont normalisés par le temps de calcul utilisant  $No = Nr = 0$  et  $ctol = +\infty$ , *i.e.* pour le problème déterministe pour lequel la correction entropique est testée partout. On observe que les idées introduites ci-dessus permettent d'économiser un temps de calcul considérable puisque les temps CPU peuvent être divisés par un facteur deux pour  $No$  et  $Nr$  fixés. En outre, on indique dans le tableau 3.2 la fraction d'éléments stochastiques qui sont réellement testés pour la correction. Cette quantité ne change pas significativement lorsque l'on augmente  $ctol$  au delà de  $10^{-3}$ . Enfin, on définit la mesure d'erreur sur la densité par

$$\epsilon_h^2(t) \equiv \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^{N_c} \Delta x \left( \rho_i^P(t, \xi^{(j)}) - \rho_i^h(t, \xi^{(j)}) \right)^2, \quad (3.50)$$

où, pour chaque élément  $\xi^{(j)}$  dans un échantillon,  $\rho_i^P(t, \xi^{(j)})$  et  $\rho_i^h(t, \xi^{(j)})$  sont respectivement évaluées à partir de la solution de Galerkin sur la cellule  $i$  de l'espace physique et à partir de la solution du problème déterministe d'Euler correspondant sur la cellule  $i$  de l'espace physique calculée avec le solveur déterministe. On utilise un échantillon de taille  $M = 100000$ . On a observé que la valeur de  $ctol$  n'affectait pas l'erreur  $\epsilon_h$ , que l'on indique dans la dernière ligne du tableau 3.2. Cela confirme l'efficacité du critère (3.49) puisqu'il permet un gain en temps de calcul considérable sans dégrader la précision.

**Tableau 3.2** Cas test 5 (Euler) : Temps de calculs normalisés  $T_{\text{CPU}}$ , fraction des cellules stochastiques réellement testées et mesure d'erreur  $\epsilon_h$  pour différents paramètres de discrétisation stochastique Nr et No.

P	No = 1, Nr = 3		No = 2, Nr = 3		No = 3, Nr = 3	
	16		24		32	
	$T_{\text{CPU}}$	fraction	$T_{\text{CPU}}$	fraction	$T_{\text{CPU}}$	fraction
$ctol = +\infty$	11.7	1.0e-0	16.1	1.0e-0	21.6	1.0e-0
$ctol = 10^{-1}$	8.2	3.7e-1	11.8	3.7e-1	16.4	3.7e-1
$ctol = 10^{-2}$	6.5	7.1e-2	9.8	7.1e-2	13.9	7.1e-2
$ctol = 10^{-3}$	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	

**Tableau 3.3** Cas test 5 (Euler) : Fraction du temps de calcul utilisé pour la correction entropique,  $R_{\text{cor}}(ctol)$ , pour différentes valeurs de l'ordre polynomial No et du paramètre  $ctol$ . Calculs avec Nr = 3.

P	No = 1	No = 2	No = 3
	16	24	32
$ctol = +\infty$	0.49	0.42	0.38
$ctol = 10^{-1}$	0.19	0.16	0.14
$ctol = 10^{-2}$	0.04	0.03	0.02

On termine la discussion sur les temps CPU avec une brève estimation du coût de calcul général résultant de la correction entropique. À cette fin, on évalue les temps CPU dédiés à la correction entropique en calculant le rapport

$$R_{\text{cor}}(ctol) = \frac{T_{\text{CPU}}(ctol) - T_{\text{CPU}}(ctol = 0)}{T_{\text{CPU}}(ctol = +\infty)}. \quad (3.51)$$

Puisque pour  $ctol = 0$  l'ensemble des interfaces où la correction est testée est une fraction négligeable de l'ensemble total des interfaces,  $T_{\text{CPU}}(ctol = 0)$  est essentiellement égal au temps CPU pour le solveur de Roe sans correction entropique. A l'inverse,  $T_{\text{CPU}}(ctol = +\infty)$  est le temps de calcul où le correcteur entropique est testé sur l'ensemble des points soniques. Par conséquent  $R_{\text{cor}}(ctol)$  est la fraction de temps CPU dédié à la correction entropique dans le calcul du flux. Les valeurs de  $R_{\text{cor}}(ctol)$  sont fournies dans le tableau 3.3 pour différentes valeurs de  $ctol$ , d'ordre polynomial No et pour un niveau de résolution fixé Nr = 3. La première ligne, correspondant à  $ctol = +\infty$ , montre que la correction entropique génère un surplus de calcul significatif lorsque celle-ci est testée partout. Chose intéressante, lorsque l'ordre polynomial No augmente,  $R_{\text{cor}}(ctol)$  diminue, indiquant que pour une interface donnée, le calcul de la partie de Roe du flux devient relativement plus coûteuse que sa partie corrective. Cette tendance peut être attribuée aux calculs pseudo-spectraux, si bien que celle-ci dépend du système conservatif considéré, et plus particulièrement des non-linéarités présentes dans les fonctions de flux. En outre, le

tableau 3.3 indique que diminuer  $ctol$  permet de réduire considérablement la part de temps de calcul pour la correction entropique : pour  $ctol = 10^{-2}$ , seulement une quantité marginale du temps CPU total est dédiée à la correction entropique.

### 3.3 Conclusion

Dans ce chapitre, le solveur de Roe stochastique avec correcteur entropique a été testé avec succès sur les équations de Burgers et d'Euler, respectivement en deux et une dimensions stochastiques. Le solveur s'est avéré apte à capturer de manière précise et robuste la dynamique de ces équations, en particulier les discontinuités pouvant émerger au cours du temps dans le domaine stochastique. On a étudié la stabilité temporelle du schéma ainsi que la précision du développement stochastique de la solution de Galerkin obtenue en utilisant des comparaisons avec des méthodes NISP et Monte Carlo. De plus, les calculs de la matrice de décentrement par fit polynomial et du correcteur entropique se sont avérés robustes et efficaces. Finalement, en faisant varier les paramètres  $N_r$  et  $N_o$ , on a montré qu'on pouvait contrôler l'erreur stochastique et les overshoots de la solution de Galerkin au voisinage des discontinuités dans le domaine stochastique.

Néanmoins, cette approche reste coûteuse si on veut une solution précise au voisinage des chocs, puisqu'une discrétisation stochastique suffisamment fine est nécessaire. On propose donc dans la partie suivante de construire un solveur de Roe stochastique adaptatif, basé sur le concept d'analyse multi-résolution, où le niveau de résolution  $N_r$  de l'espace d'approximation stochastique est adapté à la régularité stochastique locale de la solution. Compte tenu de la localisation des discontinuités dans le domaine spatial, l'adaptation stochastique dépendra également de l'espace et du temps.

Le concept d'analyse multi-résolution dans  $L^2(\Xi, p_\xi)$  est présenté dans le chapitre 4 dans le cas d'une dimension stochastique ( $N = 1$ ) pour des variables aléatoires  $U(\xi)$  à valeurs dans  $\mathbb{R}$  ( $m = 1$ ). La construction du solveur de Roe stochastique adaptatif, basé sur ce concept, est discuté dans le chapitre 5 ; on présente ensuite l'application de ce solveur au cas test 4 des équations d'Euler. Enfin, une extension du solveur de Roe stochastique adaptatif dans le cas multidimensionnel est proposée au chapitre 6.

Deuxième partie

**Adaptation stochastique**



# Outils d'Analyse Multi-Résolution 1D

---

## Sommaire

<b>4.1</b>	<b>Analyse Multi-Résolution Multi-Ondelettes</b>	<b>98</b>
4.1.1	Ingrédients de base	98
4.1.2	Système orthogonal MW d'Alpert	100
4.1.3	Seuillage	108
4.1.4	Complétion à la Harten	109
<b>4.2</b>	<b>Arbres binaires</b>	<b>111</b>
4.2.1	Principales définitions	111
4.2.2	Grille associée à un arbre binaire	112
4.2.3	Espace d'approximation stochastique associé à un arbre binaire	113
4.2.4	Inclusion d'arbres	114
<b>4.3</b>	<b>Adaptation à une fonction connue</b>	<b>114</b>
4.3.1	Principe	114
4.3.2	Aspects algorithmiques	117
4.3.3	Résultats	119
<b>4.4</b>	<b>Synthèse</b>	<b>124</b>

---

Dans les chapitres précédents, on a considéré des solutions stochastiques  $U$  de variance finie ( $U \in L^2(\Xi, p_\xi)$ ), admettant donc un développement stochastique sur la base SE. Typiquement, les degrés de liberté de la discrétisation stochastique étaient les  $P$  modes stochastiques de la solution de Galerkin  $U^P$ . On souhaiterait maintenant conserver un certain niveau de précision tout en diminuant le nombre de degrés de liberté  $P$ , en utilisant des éléments stochastiques de grande taille dans les parties de  $\Xi$  où  $U$  est régulière et des éléments stochastiques plus petits au voisinage des discontinuités et des fortes variations de  $U$ ; on souhaiterait ainsi mettre en œuvre une “h-adaptation” de la discrétisation stochastique. On rappelle que la fonction  $U$  peut être développée de façon équivalente sur la base MW (Multi-Ondelettes de Alpert) [48]. On propose donc d'utiliser une analyse du comportement régulier local de  $U$  grâce au concept d'Analyse Multi-Résolution Multi-Ondelettes, cf Keiner [42], qu'on notera AMR-MW. Dans le contexte déterministe, Hovhannisyan et. al ont récemment appliqué ce concept à des schémas de Galerkin Discontinu [40]. Dans ce chapitre, on se restreint pour simplifier à des variables aléatoires scalaires  $U : \Xi \mapsto \mathbb{R}$ . L'extension aux vecteurs aléatoires et aux processus stochastiques est



directe en travaillant composante par composante ou en se plaçant à  $(x, t)$  fixés. De plus, on se place dans le cas unidimensionnel ( $N = 1$ ), de sorte que  $\Xi = [0, 1]$ . L'extension au cas multidimensionnel est discutée dans le chapitre 6. Par la suite, l'espace des variables aléatoires du second ordre définies sur  $\mathcal{P}_\xi$  est noté  $L^2(\Xi)$  pour simplifier.

## 4.1 Analyse Multi-Résolution Multi-Ondelettes

L'AMR-MW est une généralisation à des fonctions de base de degré  $N_0 > 0$  de l'AMR proposée par Harten pour des moyennes de cellules [38], utilisée en particulier pour la construction de schémas Volumes Finis adaptatifs pour les systèmes hyperboliques déterministes, *cf.* entre autres, Müller [63] et Cohen et. al [13].

### 4.1.1 Ingrédients de base

Les deux ingrédients essentiels d'une AMR-MW sont les suivants :

- une hiérarchie d'espaces d'approximation construite à partir d'une hiérarchie de grilles emboîtées,
- une transformation multi-échelles, basée sur la notion de détails et utilisant deux bases hilbertiennes de  $L^2(\Xi)$  dont les éléments sont appelés fonctions d'échelle et multi-ondelettes.

#### 4.1.1.1 Hiérarchie d'espaces d'approximation

Le point de départ de la construction d'une AMR-MW est une séquence de grilles emboîtées. On considère un ensemble hiérarchique de partitions dyadiques uniformes de  $\Xi = [0, 1]$ , illustré à la figure 4.1. Soit  $G_j := \bigcup_{l \in \mathcal{I}_j} \{K_{j,l}\}$ ,  $\mathcal{I}_j \in \{1, \dots, N_j\}$ ,  $j \in \mathbb{N}$ ,  $N_j \in \mathbb{N}_0 := \mathbb{N} \setminus \{0\}$ , une séquence de grilles de résolution croissante. Ces grilles sont composées d'éléments stochastiques (intervalles en 1D)  $K_{j,l} = [\xi_{j,l-1}, \xi_{j,l}]$  déterminés par les points de discrétisation  $\xi_{j,l} := l/N_j$  avec  $l \in \{0, \dots, N_j\}$ . Le centre des éléments stochastiques est donné par  $\hat{\xi}_{j,l} = (\xi_{j,l-1} + \xi_{j,l})/2$ . On suppose que  $N_j = O(2^j)$  et que la hiérarchie de grilles  $\{G_j\}_{j \in \mathbb{N}}$  est emboîtée si bien que

$$K_{j,l} = K_{j+1,2l-1} \cup K_{j+1,2l}, \quad \forall j \in \mathbb{N}, l \in \mathcal{I}_j. \quad (4.1)$$

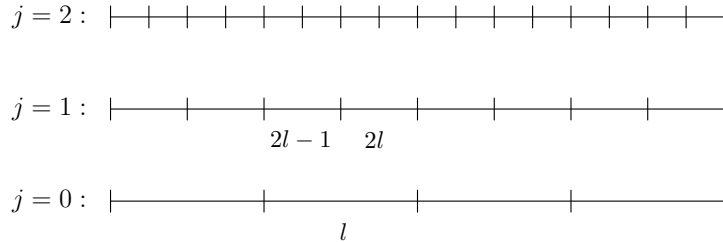


FIGURE 4.1 – Séquence de grilles dyadiques emboîtées.

**Remarque 4.1.** La grille grossière  $G_0$  se réduit à l'intervalle  $\Xi = [0, 1]$  lorsque l'on définit  $N_j := 2^j$ . On adoptera ce choix pour le système MW d'Alpert à la section 4.1.2.

Sur la hiérarchie de grilles emboîtées  $\{G_j\}_{j \in \mathbb{N}}$ , on introduit les espaces de fonctions polynomiales par morceaux de degré plus petit ou égal à  $N_0$  :

$$S_j^{N_0} = \{f \in L^2(\Xi), \forall l \in \mathcal{I}_j, f|_{K_{j,l}} \in \Pi_{N_0}[\xi]\}, \quad \forall j \in \mathbb{N}, \quad (4.2)$$

qui correspondent en fait aux espaces  $S^{N_0, N_r}$  définis dans les chapitres précédents avec un niveau de résolution  $N_r$  qui varie (on a simplement  $N_r = j$  lorsque  $N_j = 2^j$ ). Ces espaces forment une famille hiérarchique d'espaces d'approximation stochastique emboîtés,

$$S_0^{N_0} \subset S_1^{N_0} \subset \dots \subset S_j^{N_0} \subset \dots \subset L^2(\Xi), \quad (4.3)$$

et on a

$$L^2(\Xi) = \overline{\bigcup_{j \in \mathbb{N}} S_j^{N_0}}^{L^2}. \quad (4.4)$$

#### 4.1.1.2 Notion de détails

Une conséquence immédiate de (4.3) est qu'il existe des supplémentaires à  $S_j^{N_0}$  dans  $S_{j+1}^{N_0}$ ,  $\forall j \in \mathbb{N}$ , appelés espaces de détail et notés  $W_j^{N_0}$ , tels que

$$S_{j+1}^{N_0} = S_j^{N_0} \oplus W_j^{N_0}. \quad (4.5)$$

Cette somme n'est pas nécessairement orthogonale. On verra à la section 4.1.4 un exemple de somme non orthogonale lorsque  $N_0 = 0$ . Lorsque celle-ci est  $L^2(\Xi)$ -orthogonale, on parle d'AMR-MW orthogonale.

Par récurrence, on déduit de (4.5) que pour tout  $j \in \mathbb{N}_0$ ,

$$S_j^{N_0} = S_0^{N_0} \bigoplus_{j'=0}^{j-1} W_{j'}^{N_0}. \quad (4.6)$$

Les espaces  $S_j^{N_0}$  et  $W_j^{N_0}$  sont engendrés par des fonctions de base appelées respectivement fonctions d'échelle et multi-ondelettes d'ordre  $N_0$  (ou plus simplement ondelettes pour  $N_0 = 0$ ), notées

$$\{\Phi_{j,l,\alpha}\}_{l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi} \quad \text{et} \quad \{\Psi_{j,l,\alpha}\}_{l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi}, \quad (4.7)$$

i.e.  $S_j^{N_0} = \text{span}_{l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi} \{\Phi_{j,l,\alpha}\}$  et  $W_j^{N_0} = \text{span}_{l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi} \{\Psi_{j,l,\alpha}\}$ . On rappelle que l'on a posé  $\mathcal{P}_\pi = \{1, \dots, P_\pi\}$  avec  $P_\pi = (N_0 + 1)$  pour  $N = 1$ . On notera la double indexation  $(l, \alpha)$  des fonctions de base, qui provient du fait que sur chaque élément stochastique, on construit  $P_\pi$  fonctions de base.

Soit  $J \in \mathbb{N}$  un niveau de résolution fixé et soit  $U^J \in S_J^{N_0}$ . La fonction  $U^J$  peut être développée sur la base des fonctions d'échelle  $J$  sous la forme

$$U^J(\xi) = \sum_{l \in \mathcal{I}_J} \sum_{\alpha \in \mathcal{P}_\pi} u_{J,l,\alpha} \Phi_{J,l,\alpha}(\xi), \quad (4.8)$$

ou, de façon équivalente, sur la base des fonctions d'échelle grossière et des multi-ondelettes jusqu'au niveau de résolution  $(J - 1)$  :

$$U^J(\xi) = \sum_{l \in \mathcal{I}_0} \sum_{\alpha \in \mathcal{P}_\pi} u_{0,l,\alpha} \Phi_{0,l,\alpha}(\xi) + \sum_{j=0}^{J-1} \sum_{l \in \mathcal{I}_j} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi). \quad (4.9)$$

Pour simplifier, on réécrit (4.9) sous la forme

$$U^J(\xi) = \sum_{j=-1}^{J-1} \sum_{l \in \mathcal{I}_j} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi), \quad (4.10)$$

où on définit  $\mathcal{I}_{-1} := \mathcal{I}_0$ ,  $d_{-1,l,\alpha} := u_{0,l,\alpha}$  et  $\Psi_{-1,l,\alpha} := \Phi_{0,l,\alpha}$ , pour tout  $l \in \mathcal{I}_0$ . Les coefficients  $\{u_{j,l,\alpha}\}$  sont appelés les **coefficients d'échelle** et les coefficients  $\{d_{j,l,\alpha}\}$  les coefficients multi-ondelettes ou **détails**.

## 4.1.2 Système orthogonal MW d'Alpert

### 4.1.2.1 Construction des fonctions de base

On pose  $N_j := 2^j$ , *i.e.*, on choisit comme grille grossière  $G_0 = \Xi$ . Soit  $\{\Phi_\alpha^{N_0}\}_{\alpha \in \mathcal{P}_\pi}$  l'ensemble des polynômes de Legendre normalisés sur  $\Xi = [0, 1]$ . Ces polynômes forment une base de  $\Pi_{N_0}[\xi] \equiv S_0^{N_0}$  et sont représentés sur la figure 4.2 pour  $N_0 = 5$ . Puis, pour tout  $j \in \mathbb{N}_0$ , l'espace  $S_j^{N_0}$  est engendré par les  $N_j \mathcal{P}_\pi$  fonctions d'échelle obtenues par dilatation et translation des  $\Phi_\alpha^{N_0}$  :

$$\Phi_{j,l,\alpha} = \begin{cases} 2^{j/2} \Phi_\alpha^{N_0}(2^j \xi - l + 1), & \xi \in K_{j,l}, \\ 0, & \text{sinon,} \end{cases} \quad \forall l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi. \quad (4.11)$$

On rappelle que ces fonctions correspondent aux fonctions de base SE monodimensionnelles pour un niveau de résolution  $N_r = j$  fixé et qu'elles forment un système orthonormal, *i.e.*

$$\langle \Phi_{j,l,\alpha}, \Phi_{j,l',\beta} \rangle = \delta_{\alpha\beta} \delta_{ll'}. \quad (4.12)$$

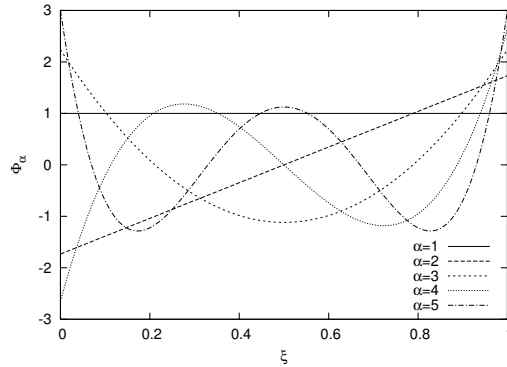
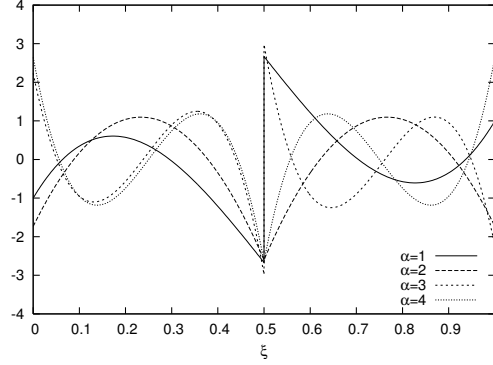


FIGURE 4.2 – Polynômes de Legendre pour  $\alpha \in \{1, \dots, 5\}$ .

FIGURE 4.3 – Multi-ondelettes mère de Le Maître  $\{\Psi_\alpha^{\text{No}}\}$  pour  $\text{No} = 3$ .

Le choix d'Alpert est de considérer une complétion orthogonale pour (4.5). Les fonctions  $\{\Psi_\alpha^{\text{No}}\}_{\alpha \in \mathcal{P}_\pi}$  qui engendrent  $W_0^{\text{No}}$  sont des polynômes de degré  $\leq \text{No}$  sur  $[0, 1/2]$  et  $[1/2, 1]$ . Ces fonctions sont choisies orthonormales :

$$\langle \Psi_\alpha^{\text{No}}, \Psi_\beta^{\text{No}} \rangle = \delta_{\alpha\beta}. \quad (4.13)$$

De plus, comme  $W_0^{\text{No}} \perp S_0^{\text{No}}$ , les  $(\text{No} + 1)$  premiers moments des  $\Psi_\alpha^{\text{No}}$  s'annulent :

$$\langle \Psi_\alpha^{\text{No}}, \xi^{\beta-1} \rangle = 0, \quad \forall \alpha, \beta \in \{1, \dots, \text{No} + 1\}. \quad (4.14)$$

Les  $2(\text{No} + 1)$  conditions (4.13) et (4.14) sont suffisantes pour construire une base orthonormale de la complétion  $W_0^{\text{No}}$ . Les fonctions ainsi construites sont les multi-ondelettes mère utilisées par Le Maître *et al.* [48]. Elles sont représentées sur la figure 4.3 pour  $\text{No} = 3$ . Enfin, pour tout  $j \in \mathbb{N}_0$ , l'espace  $W_j^{\text{No}}$  est engendré par les  $N_j \mathcal{P}_\pi$  fonctions obtenues par dilatation et translation des  $\Psi_\alpha^{\text{No}}$  :

$$\Psi_{j,l,\alpha} = \begin{cases} 2^{j/2} \Psi_\alpha^{\text{No}}(2^j \xi - l + 1), & \xi \in K_{j,l}, \\ 0, & \text{sinon,} \end{cases} \quad \forall l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi. \quad (4.15)$$

Ces fonctions forment un système orthonormal, *i.e.*

$$\langle \Psi_{j,l,\alpha}, \Psi_{j',l',\beta} \rangle = \delta_{\alpha\beta} \delta_{ll'}. \quad (4.16)$$

Lorsque  $\text{No} = 0$ , on obtient les ondelettes de Haar.

**Remarque 4.2.** Alpert a proposé une construction intéressante qui permet l'annulation de moments additionnels pour certaines des fonctions de base [4]. La méthodologie permet de construire  $(\text{No}+1)$  fonctions de base  $\Upsilon_\alpha^{\text{No}}$  de  $W_0^{\text{No}}$ , à support dans l'intervalle  $[0, 1]$ , qui ont les propriétés suivantes :

1. La restriction de  $\Upsilon_\alpha^{\text{No}}$  à l'intervalle  $[0, 1/2]$  est un polynôme de degré  $\text{No}$ .
2. La fonction  $\Upsilon_\alpha^{\text{No}}$  est étendue à l'intervalle  $[1/2, 1]$  comme fonction paire ou impaire par rapport à  $\xi = 1/2$  selon la parité de  $(\alpha + \text{No})$ .

3. Les fonctions  $\Upsilon_\alpha^{\text{No}}$  sont orthonormales, i.e., elles satisfont (4.13).

4. La fonction  $\Upsilon_\alpha^{\text{No}}$  a  $(\alpha + \text{No})$  moments qui s'annulent, i.e.

$$\langle \Upsilon_\alpha^{\text{No}}, \xi^{\beta-1} \rangle = 0, \quad \forall \beta \in \{1, \dots, \alpha + \text{No}\}.$$

La construction est la suivante. On définit  $\{f_\alpha^0\}_{\alpha \in \{1, \dots, \text{No}+1\}}$  sur  $\Xi = [0, 1]$  par

$$f_\alpha^0(\xi) = \begin{cases} \xi^\alpha, & \text{si } \xi \geq 1/2, \\ -\xi^\alpha, & \text{si } \xi < 1/2, \end{cases} \quad \forall \alpha \in \{1, \dots, \text{No} + 1\}. \quad (4.17)$$

On remarque que les  $2(\text{No} + 1)$  fonctions  $\{1, \xi, \dots, \xi^{\text{No}}, f_1^0, f_2^0, \dots, f_{\text{No}+1}^0\}$  sont linéairement indépendantes et qu'elles engendrent donc l'espace  $S_1^{\text{No}}$ .

1. Par le processus de Gram–Schmidt, on orthogonalise les  $\{f_\alpha^0\}_{\alpha \in \{1, \dots, \text{No}+1\}}$  par rapport aux fonctions  $1, \xi, \dots, \xi^{\text{No}}$ , pour obtenir des nouvelles fonctions  $\{f_\alpha^1\}_{\alpha \in \{1, \dots, \text{No}+1\}}$ . Cette orthogonalisation est préservée lors des étapes suivantes, qui sont basées sur des combinaisons linéaires des  $\{f_\alpha^1\}_{\alpha \in \{1, \dots, \text{No}+1\}}$ .

2. La séquence d'étapes suivantes fournit  $\text{No}$  fonctions orthogonales à  $\xi^{\text{No}+1}$ , parmi lesquelles  $(\text{No} - 1)$  fonctions sont orthogonales à  $\xi^{\text{No}+2}$ , et ainsi de suite, jusqu'à ce qu'une fonction soit orthogonale à  $\xi^{2\text{No}}$ . En général, il existe au moins une fonction parmi les  $f_\alpha^1$  qui n'est pas orthogonale à  $\xi^{\text{No}+1}$ . On choisit (par exemple) la première qui se présente. (Si elles sont toutes orthogonales, on garde l'énumération telle quelle et on pourrait aller plus loin dans l'annulation des moments.) On utilise une permutation entre cette fonction et  $f_1^1$  de sorte qu'elle apparaisse en premier. On définit ensuite  $f_\alpha^2 = f_\alpha^1 - a_\alpha f_1^1$  où  $a_\alpha$  est choisi de sorte que  $\langle f_\alpha^2, \xi^{\text{No}+1} \rangle = 0$  pour tout  $\alpha \in \{2, \dots, \text{No} + 1\}$  réalisant l'orthogonalité désirée par rapport à  $\xi^{\text{No}+1}$ . Cela conduit à une famille de fonctions  $\{f_\alpha^2\}_{\alpha \in \{2, \dots, \text{No}+1\}}$ . De façon similaire, on orthogonalise par rapport à  $\xi^{\text{No}+2}, \dots, \xi^{2\text{No}}$ , à tour de rôle, pour obtenir les fonctions  $f_1^1, f_2^2, f_3^3, \dots, f_{\text{No}+1}^{\text{No}}$  telles que  $\langle f_\alpha^2, \xi^{\beta-1} \rangle = 0, \quad \forall \beta \in \{1, \dots, \alpha + \text{No}\}.$

3. Finalement, on procède à l'orthogonalisation de Gram–Schmidt de  $f_{\text{No}+1}^{\text{No}+1}, f_{\text{No}}^{\text{No}}, \dots, f_1^1$  dans cet ordre, et on normalise pour obtenir les fonctions  $\Upsilon_{\text{No}+1}^{\text{No}}, \Upsilon_{\text{No}}^{\text{No}}, \dots, \Upsilon_1^{\text{No}}$ . On vérifie également que ces fonctions sont paire ou impaire par rapport à  $\xi = 1/2$  selon la parité de  $(\alpha + \text{No})$ .

Les fonctions  $\{\Upsilon_\alpha^{\text{No}}\}_{\alpha \in \mathcal{P}_\pi}$  construites ci-dessus sont appelées multi-ondelettes mère d'Alpert et engendrent  $W_0^{\text{No}}$ . Elles sont représentées sur la figure 4.4 pour  $\text{No} = 3$ . Ces multi-ondelettes sont intéressantes si l'on souhaite compresser des données, par exemple pour représenter des opérateurs de façon creuse (cf Alpert [4]). Cependant, on souhaite ici adapter le maillage stochastique en fonction de développements locaux sur les éléments stochastiques; on choisira donc par la suite de travailler avec les multi-ondelettes utilisées par Le Maître et al. [48].

#### 4.1.2.2 Principales propriétés

Les fonctions d'échelle (SE) et les multi-ondelettes de Le Maître (MW) vérifient les propriétés suivantes :

– elles ont un support local, *i.e.*,

$$\text{supp}(\Phi_{j,l,\alpha}) = \text{supp}(\Psi_{j,l,\alpha}) = K_{j,l}, \quad (4.18)$$

– elles satisfont les propriétés d'orthogonalité suivantes

$$\langle \Phi_{j,l,\alpha}, \Phi_{j,l',\beta} \rangle = \delta_{\alpha\beta} \delta_{ll'} \quad (\Phi_j^{\text{No}} \text{ base orthonormale de } S_j^{\text{No}}), \quad (4.19)$$

$$\langle \Psi_{j,l,\alpha}, \Phi_{j,l',\beta} \rangle = 0 \quad (W_j^{\text{No}} \perp S_j^{\text{No}}), \quad (4.20)$$

$$\langle \Psi_{j,l,\alpha}, \Psi_{j',l',\beta} \rangle = 0 \text{ pour } j \neq j' \quad (W_j^{\text{No}} \perp W_{j'}^{\text{No}} \text{ pour } j \neq j'). \quad (4.21)$$

– les  $(\text{No} + 1)$  premiers moments des fonctions multi-ondelettes s'annulent, *i.e.*

$$\langle \Psi_{j,l,\alpha}, \xi^{\beta-1} \rangle = 0, \quad \forall \beta \in \{1, \dots, \text{No} + 1\}. \quad (4.22)$$

Soit  $J \in \mathbb{N}$  un niveau de résolution fixé et  $U^J \in S_J^{\text{No}}$ . La fonction  $U^J$  peut être développée sur la base des fonctions d'échelle  $J$ , correspondant ici à la base SE de  $S_J^{\text{No}}$ ,

$$U^J(\xi) = \sum_{l \in \mathcal{I}_J} \sum_{\alpha \in \mathcal{P}_\pi} u_{J,l,\alpha} \Phi_{J,l,\alpha}(\xi), \quad (4.23)$$

ou, de façon équivalente, sur la base des fonctions d'échelle grossière, correspondant ici aux  $P_\pi$  polynômes de Legendre sur  $[0, 1]$ , et des MW (multi-ondelettes de Le Maître) jusqu'au niveau de résolution  $(J - 1)$ ,

$$U^J(\xi) = \sum_{\alpha \in \mathcal{P}_\pi} u_{0,\alpha} \Phi_\alpha^{\text{No}}(\xi) + \sum_{j=0}^{J-1} \sum_{l \in \mathcal{I}_j} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi). \quad (4.24)$$

Les coefficients d'échelle  $J$  sont déterminés par

$$u_{J,l,\alpha} := \langle U^J, \Phi_{J,l,\alpha} \rangle, \quad (4.25)$$

tandis que les coefficients d'échelle grossière et les détails sont déterminés par

$$u_{0,\alpha} := \langle U^J, \Phi_\alpha^{\text{No}} \rangle \quad \text{et} \quad d_{j,l,\alpha} := \langle U^J, \Psi_{j,l,\alpha} \rangle. \quad (4.26)$$

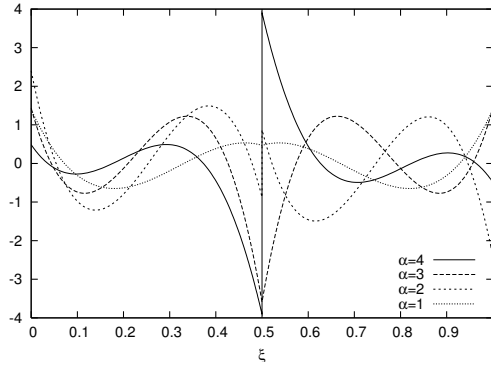


FIGURE 4.4 – Multi-ondelettes mère d'Alpert  $\{\Psi_\alpha^{\text{No}}\}$  pour  $\text{No} = 3$ .

Comme précédemment, on réécrit (4.24) sous la forme

$$U^J(\xi) = \sum_{j=-1}^{J-1} \sum_{l \in \mathcal{I}_j} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi), \quad (4.27)$$

où on définit  $\mathcal{I}_{-1} = \mathcal{I}_0 = \{1\}$ ,  $d_{-1,1,\alpha} := u_{0,\alpha}$  et  $\Psi_{-1,1,\alpha} := \Phi_\alpha^{\text{No}}$ .

On considère à présent  $U \in L^2(\Xi)$ . Pour tout  $j \in \mathbb{N}$ , la projection orthogonale  $P_{S_j^{\text{No}}}U$  de  $U$  sur  $S_j^{\text{No}}$  est donnée par

$$P_{S_j^{\text{No}}}U(\xi) := U^j(\xi) = \sum_{l \in \mathcal{I}_j} \sum_{\alpha \in \mathcal{P}_\pi} u_{j,l,\alpha} \Phi_{j,l,\alpha}(\xi), \quad (4.28)$$

$$= \sum_{j'=-1}^{j-1} \sum_{l \in \mathcal{I}_{j'}} \sum_{\alpha \in \mathcal{P}_\pi} d_{j',l,\alpha} \Psi_{j',l,\alpha}(\xi), \quad (4.29)$$

où les coefficients d'échelle et les détails sont déterminés par

$$u_{j,l,\alpha} := \langle U, \Phi_{j,l,\alpha} \rangle \quad \text{et} \quad \forall j' \in \{-1, \dots, j-1\}, \quad d_{j',l,\alpha} := \langle U, \Psi_{j',l,\alpha} \rangle. \quad (4.30)$$

De plus, par densité, on a

$$U(\xi) = \sum_{j=-1}^{+\infty} \sum_{l \in \mathcal{I}_j} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi), \quad (4.31)$$

où les détails sont déterminés par

$$d_{j,l,\alpha} := \langle U, \Psi_{j,l,\alpha} \rangle. \quad (4.32)$$

La propriété d'annulation des moments (4.22) et la normalisation (4.16) impliquent que les détails sont d'autant plus petits que le niveau de résolution est grand, pourvu que la fonction considérée soit localement régulière, *i.e.*,

$$\begin{aligned} |d_{j,l,\alpha}| &\leq \inf_{P \in \Pi_{\text{No}}[\xi]} |\langle U - P, \Psi_{j,l,\alpha} \rangle| \\ &\leq C 2^{-j(\text{No}+1)} \|U\|_{H^{\text{No}+1}(K_{j,l})}. \end{aligned} \quad (4.33)$$

Plus précisément, les détails décroissent avec un facteur d'au moins  $2^{-j(\text{No}+1)}$ , pourvu que  $U$  ait une régularité suffisante sur le support de la multi-ondelette, *i.e.*  $U$  appartient à l'espace de Sobolev  $H^{\text{No}+1}(K_{j,l})$ , équipé de sa norme canonique  $\|\cdot\|_{H^{\text{No}+1}(K_{j,l})}$ . En fait, plus  $\text{No}$  est grand, plus les détails sont *a priori* petits dans les régions régulières.

#### 4.1.2.3 Transformation multi-échelles

Dans le but d'exploiter la propriété de décroissance (4.33), on souhaite transformer le vecteur des coefficients d'échelle à niveau de résolution  $J$  fixé

$$u_J := (u_{J,l})_{l \in \mathcal{I}_J} \in \mathbb{R}^{N_J P_\pi}, \quad u_{J,l} := (u_{J,l,\alpha})_{\alpha \in \mathcal{P}_\pi} \in \mathbb{R}^{P_\pi}, \quad (4.34)$$

en un vecteur de coefficients d'échelle grossière et de détails

$$d_{(J)} := (d_j)_{j \in \{-1, \dots, J-1\}} \in \mathbb{R}^{N_{JP_\pi}}, \quad (4.35)$$

$$d_j := (d_{j,l})_{l \in \mathcal{I}_j} \in \mathbb{R}^{N_{jP_\pi}}, \quad d_{j,l} := (d_{j,l,\alpha})_{\alpha \in \mathcal{P}_\pi} \in \mathbb{R}^{P_\pi}, \quad (4.36)$$

qui représente les mises à jour successives de la solution d'une résolution grossière à une résolution plus fine. Pour déterminer facilement (4.35)-(4.36) à partir de (4.34), on remarque tout d'abord que l'emboîtement des espaces (4.3) et les conditions d'orthogonalité (4.19), (4.20) et (4.21) assurent l'existence d'une décomposition à deux échelles : pour tout  $j \in \mathbb{N}$ ,

$$\Phi_{j,l,\alpha} = \sum_{\beta \in \mathcal{P}_\pi} \sum_{s \in \{-1,0\}} m_{\alpha\beta}^{\Phi,s} \Phi_{j+1,2l+s,\beta}, \quad \forall l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi, \quad (4.37)$$

$$\Psi_{j,l,\alpha} = \sum_{\beta \in \mathcal{P}_\pi} \sum_{s \in \{-1,0\}} m_{\alpha\beta}^{\Psi,s} \Phi_{j+1,2l+s,\beta}, \quad \forall l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi, \quad (4.38)$$

où les réels  $m_{\alpha\beta}^{\Phi,s}$  et  $m_{\alpha\beta}^{\Psi,s}$  sont appelés coefficients de filtre et sont déterminés pour  $s \in \{-1,0\}$  par

$$m_{\alpha\beta}^{\Phi,s} = \langle \Phi_{j,l,\alpha}, \Phi_{j+1,2l+s,\beta} \rangle, \quad m_{\alpha\beta}^{\Psi,s} = \langle \Psi_{j,l,\alpha}, \Phi_{j+1,2l+s,\beta} \rangle. \quad (4.39)$$

Par construction, ces coefficients sont invariants par rapport au niveau de résolution  $j$  et à l'élément  $K_{j,l}$ . En effet, en posant  $\Xi_{-1} := [0, 1/2]$  et  $\Xi_0 := [1/2, 1]$ , on obtient pour  $s \in \{-1,0\}$ ,

$$m_{\alpha\beta}^{\Phi,s} = \int_{\Xi_s} \Phi_\alpha^{\text{No}}(\xi) \Phi_\beta^{\text{No}}(2\xi - s - 1) d\xi, \quad (4.40)$$

$$m_{\alpha\beta}^{\Psi,s} = \int_{\Xi_s} \Psi_\alpha^{\text{No}}(\xi) \Phi_\beta^{\text{No}}(2\xi - s - 1) d\xi, \quad (4.41)$$

où l'on a utilisé les supports des fonctions d'échelles et des multi-ondelettes ainsi que leur dilatation et translation. On introduit les matrices des coefficients de filtre

$$M^{\Phi,s} := (m_{\alpha\beta}^{\Phi,s})_{\alpha,\beta \in \mathcal{P}_\pi} \in \mathbb{R}^{P_\pi, P_\pi}, \quad (4.42)$$

$$M^{\Psi,s} := (m_{\alpha\beta}^{\Psi,s})_{\alpha,\beta \in \mathcal{P}_\pi} \in \mathbb{R}^{P_\pi, P_\pi}. \quad (4.43)$$

Ces matrices sont triangulaires inférieures. En effet, si on considère par exemple la ligne de coefficients  $\{m_{\alpha,\beta}^{\Phi,-1}\}_{\beta \in \mathcal{P}_\pi}$  pour  $\alpha \in \mathcal{P}_\pi$  fixé, on a

$$\int_0^{1/2} \Phi_\alpha^{\text{No}}(\xi) \Phi_\beta^{\text{No}}(2\xi) d\xi = \int_0^1 \Phi_\alpha^{\text{No}}\left(\frac{\xi}{2}\right) \Phi_\beta^{\text{No}}(\xi) d\xi = 0, \quad \text{si } \beta > \alpha, \quad (4.44)$$

puisque  $\Phi_\alpha^{\text{No}}(\frac{\xi}{2}) \in \Pi_\alpha[\xi]$  et que les polynômes de Legendre sont tels que  $\Phi_\beta^{\text{No}} \perp \Pi_\alpha[\xi]$  pour  $\beta > \alpha$ . On montre de façon similaire le résultat pour les autres coefficients. De plus, en utilisant les propriétés de symétrie des fonctions d'échelle de Legendre et des multi-ondelettes, on obtient

$$m_{\alpha\beta}^{\Phi,0} = (-1)^{\alpha+\beta} m_{\alpha\beta}^{\Phi,-1}, \quad m_{\alpha\beta}^{\Psi,0} = (-1)^{\alpha+\beta+P_\pi} m_{\alpha\beta}^{\Psi,-1}. \quad (4.45)$$



Par la suite, on utilisera uniquement les matrices de coefficients de filtre données par (4.42) et (4.43). Néanmoins, pour compléter la présentation, on décrit brièvement la transformation multi-échelles permettant de passer de  $u_j$  à  $d_{(j)}$ .

Pour réécrire les relations d'échelles (4.37) et (4.38) sous forme matricielle, on définit les vecteurs des fonctions d'échelle et des multi-ondelettes suivants :

$$\Phi_j^{\text{No}} := (\Phi_{j,l})_{l \in I_j} \in (L^2(\Xi))^{N_j P_\pi}, \quad \Phi_{j,l} := (\Phi_{j,l,\alpha})_{\alpha \in \mathcal{P}_\pi} \in (L^2(\Xi))^{P_\pi}, \quad (4.46)$$

$$\Psi_j^{\text{No}} := (\Psi_{j,l})_{l \in I_j} \in (L^2(\Xi))^{N_j P_\pi}, \quad \Psi_{j,l} := (\Psi_{j,l,\alpha})_{\alpha \in \mathcal{P}_\pi} \in (L^2(\Xi))^{P_\pi}, \quad (4.47)$$

ainsi que les matrices rectangulaires suivantes  $M^{\Phi,j}, M^{\Psi,j} \in \mathbb{R}^{N_j P_\pi, N_{j+1} P_\pi}$  :

$$M^{\Phi,j} = \left( \begin{array}{cc|cc|ccc} M^{\Phi,-1} & M^{\Phi,0} & 0_{P_\pi} & 0_{P_\pi} & \dots & & & \\ 0_{P_\pi} & 0_{P_\pi} & M^{\Phi,-1} & M^{\Phi,0} & 0_{2P_\pi} & \dots & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & \dots & 0_{2P_\pi} & M^{\Phi,-1} & M^{\Phi,0} & 0_{P_\pi} & 0_{P_\pi} \\ & & & & \dots & 0_{P_\pi} & 0_{P_\pi} & M^{\Phi,-1} & M^{\Phi,0} \end{array} \right), \quad (4.48)$$

$$M^{\Psi,j} = \left( \begin{array}{cc|cc|ccc} M^{\Psi,-1} & M^{\Psi,0} & 0_{P_\pi} & 0_{P_\pi} & \dots & & & \\ 0_{P_\pi} & 0_{P_\pi} & M^{\Psi,-1} & M^{\Psi,0} & 0_{2P_\pi} & \dots & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & \dots & 0_{2P_\pi} & M^{\Psi,-1} & M^{\Psi,0} & 0_{P_\pi} & 0_{P_\pi} \\ & & & & \dots & 0_{P_\pi} & 0_{P_\pi} & M^{\Psi,-1} & M^{\Psi,0} \end{array} \right), \quad (4.49)$$

où  $0_{P_\pi}$  désigne la matrice nulle de  $\mathbb{R}^{P_\pi}$ . On peut alors réécrire (4.37)-(4.38) sous la forme

$$\Phi_j^{\text{No}} = M^{\Phi,j} \Phi_{j+1}^{\text{No}} \quad \text{et} \quad \Psi_j^{\text{No}} = M^{\Psi,j} \Phi_{j+1}^{\text{No}}. \quad (4.50)$$

**Proposition 4.1.** *La matrice  $M^j$  définie par*

$$M^j := \begin{pmatrix} M^{\Phi,j} \\ M^{\Psi,j} \end{pmatrix} \in \mathbb{R}^{N_{j+1} P_\pi, N_{j+1} P_\pi} \quad (4.51)$$

*est orthogonale.*

*Démonstration.* Pour tout  $j \in \mathbb{N}$  et  $U \in (L^2(\Xi))^{N_j P_\pi}$ ,  $UU^T$  désigne la matrice de  $\mathbb{R}^{N_j P_\pi, N_j P_\pi}$  de coefficients

$$(UU^T)_{l\alpha, l'\alpha'} = \langle U_{l,\alpha}, U_{l',\alpha'} \rangle.$$

En omettant provisoirement l'indice No, il vient

$$\begin{aligned} I_{N_j P_\pi} &= \Phi_j \Phi_j^T = M^{\Phi,j} \Phi_{j+1} \Phi_{j+1}^T (M^{\Phi,j})^T, \\ I_{N_j P_\pi} &= \Psi_j \Psi_j^T = M^{\Psi,j} \Phi_{j+1} \Phi_{j+1}^T (M^{\Psi,j})^T, \\ 0_{N_j P_\pi} &= \Phi_j \Psi_j^T = M^{\Phi,j} \Phi_{j+1} \Phi_{j+1}^T (M^{\Psi,j})^T, \end{aligned}$$

où  $I_{N_j P_\pi}$  et  $0_{N_j P_\pi}$  désignent respectivement la matrice identité et la matrice nulle de  $\mathbb{R}^{N_j P_\pi}$ . En utilisant  $\Phi_{j+1} \Phi_{j+1}^T = I_{N_{j+1} P_\pi}$ , on obtient

$$\begin{aligned} M^{\Phi,j} (M^{\Phi,j})^T &= I_{N_j P_\pi}, \\ M^{\Psi,j} (M^{\Psi,j})^T &= I_{N_j P_\pi}, \\ M^{\Phi,j} (M^{\Psi,j})^T &= 0_{N_j P_\pi}. \end{aligned}$$

Par conséquent,

$$M^j (M^j)^T = \begin{pmatrix} M^{\Phi,j} (M^{\Phi,j})^T & M^{\Phi,j} (M^{\Psi,j})^T \\ M^{\Psi,j} (M^{\Phi,j})^T & M^{\Psi,j} (M^{\Psi,j})^T \end{pmatrix} = \begin{pmatrix} I_{N_j P_\pi} & 0_{N_j P_\pi} \\ 0_{N_j P_\pi} & I_{N_j P_\pi} \end{pmatrix},$$

ce qui conclut la preuve.  $\square$

La proposition 4.1 signifie que les relations (4.50) réalisent un changement de base d'inverse

$$\Phi_{j+1}^{\text{No}} = (M^{\Phi,j})^T \Phi_j^{\text{No}} + (M^{\Psi,j})^T \Psi_j^{\text{No}}. \quad (4.52)$$

Le changement de base fournit des relations à deux échelles pour les coefficients d'échelle et les détails qui reflètent le format de cascade typique d'une transformation en ondelettes

$$\begin{pmatrix} u_j \\ \dots \\ d_j \end{pmatrix} = M^j u_{j+1}, \quad u_{j+1} = (M^j)^T \begin{pmatrix} u_j \\ \dots \\ d_j \end{pmatrix}. \quad (4.53)$$

Une application successive des relations (4.53) décompose le vecteur  $u_J$  en coefficients d'échelle grossière et en détails de résolution plus élevée  $d_{(J)}$  définies par (4.35). Cette transformation est communément appelée transformation multi-échelles et est déterminée par l'opérateur multi-échelles linéaire bijectif  $\mathcal{M}^J : \mathbb{R}^{N_J P_\pi} \mapsto \mathbb{R}^{N_J P_\pi}$  défini par

$$d_{(J)} = \mathcal{M}^J u_J, \quad \mathcal{M}^J := \overline{M}^0 \dots \overline{M}^{J-1}, \quad (4.54)$$

$$\overline{M}^j = \begin{pmatrix} M^j & 0 \\ 0 & I_{(N_J - N_{j+1}) P_\pi} \end{pmatrix}, \quad j \in \{0, \dots, J-1\}, \quad (4.55)$$

que l'on peut représenter par le schéma suivant :

$$\mathcal{M}^J : \begin{array}{ccccccc} u_J & \rightarrow & u_{J-1} & \rightarrow & \dots & \rightarrow & u_1 & \rightarrow & u_0 \\ & & \searrow & & \searrow & & \searrow & & \searrow \\ & & d_{J-1} & & \dots & & d_1 & & d_0 \end{array}$$

et d'inverse

$$u_J = (\mathcal{M}^J)^{-1} d_{(J)}, \quad (\mathcal{M}^J)^{-1} = \check{M}^{J-1} \dots \check{M}^0, \quad (4.56)$$

$$\check{M}^j = \begin{pmatrix} I_{(N_J - N_{j+1}) P_\pi} & 0 \\ 0 & (M^j)^T \end{pmatrix}, \quad j \in \{0, \dots, J-1\}. \quad (4.57)$$

### 4.1.3 Seuillage

On se donne un niveau de résolution maximale  $J$  et une fonction  $U \in L^2(\Xi)$ . On pose

$$\mathcal{D}_J = \{(j, l) : j = \{-1, 0, \dots, J-1\}, l \in \mathcal{I}_j\}. \quad (4.58)$$

La projection  $L^2(\Xi)$ -orthogonale  $U_J := P_{S_J^{\text{No}}}U$  définie par (4.29) avec  $j = J$  se réécrit sous la forme compacte

$$U^J(\xi) = \sum_{(j,l) \in \mathcal{D}_J} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi). \quad (4.59)$$

Grâce à la propriété de décroissance (4.33), les détails sont susceptibles de devenir suffisamment petits pour être négligés lorsque la fonction  $U$  est localement régulière. Cela donne lieu à un seuillage caractérisé par l'ensemble d'indices

$$\mathcal{D}_{J,\eta} := \left\{ (j, l) \in \mathcal{D}_J : \|d_{j,l}\|_{l^2} = \left( \sum_{\alpha \in \mathcal{P}_\pi} |d_{j,l,\alpha}|^2 \right)^{1/2} > \epsilon_J(\eta) \right\}, \quad (4.60)$$

où  $\eta > 0$  est le paramètre de seuillage et  $\epsilon_J(\eta)$  est la valeur de seuillage déterminée par un scaling de  $\eta$ . On dit que  $\mathcal{D}_{J,\eta}$  est l'ensemble des détails significatifs. Par convention, on suppose que les indices  $(-1, 1)$  et  $(0, 1)$  appartiennent à  $\mathcal{D}_{J,\eta}$ . On définit alors la projection seuillée  $U^{J,\eta}$  par le développement MW suivant

$$U^{J,\eta}(\xi) = \sum_{(j,l) \in \mathcal{D}_{J,\eta}} \sum_{\alpha \in \mathcal{P}_\pi} d_{j,l,\alpha} \Psi_{j,l,\alpha}(\xi). \quad (4.61)$$

**Erreur d'approximation.** On souhaite évaluer l'erreur d'approximation introduite par le seuillage lorsque l'on approche  $U$  par sa projection seuillée  $U^{J,\eta}$  plutôt que par sa projection orthogonale  $U^J$ . On s'intéresse donc à la quantité  $U^J - U^{J,\eta}$ .

**Proposition 4.2.** *Soit  $\eta > 0$  et  $\epsilon_J(\eta) = 2^{-J/2}\eta$ . Alors, pour tout  $U \in L^2(\Xi)$ , on a*

$$\|U^J - U^{J,\eta}\|_{L^2(\Xi)} \leq \eta. \quad (4.62)$$

*Démonstration.* On rappelle que  $\|\Psi_{j,l,\alpha}\|_{L^2(\Xi)} = 1$ . Les coefficients multi-échelles non significatifs sont bornés par  $\|d_{j,l}\|_{l^2} \leq \epsilon_J(\eta)$  pour tout  $(j, l) \notin \mathcal{D}_{J,\eta}$  et il existe au plus  $N_j = 2^j$  éléments stochastiques au niveau de résolution  $j$  pour lesquelles  $(j, l) \notin \mathcal{D}_{J,\eta}$ . Par conséquent, l'erreur d'approximation est estimée en norme  $L^2$  par

$$\begin{aligned} \|U^J - U^{J,\eta}\|_{L^2(\Xi)}^2 &= \sum_{(j,l) \notin \mathcal{D}_{J,\eta}} \sum_{\alpha \in \mathcal{P}_\pi} |d_{j,l,\alpha}|^2 \|\Psi_{j,l,\alpha}\|_{L^2(\Xi)}^2 \\ &\leq \sum_{(j,l) \notin \mathcal{D}_{J,\eta}} \epsilon_J(\eta)^2 \leq \left( 1 + \sum_{j=0}^{J-1} 2^j \right) \epsilon_J(\eta)^2. \end{aligned}$$

Puisque  $\epsilon_J = 2^{-J/2}\eta$  et  $1 + \sum_{j=0}^{J-1} 2^j = 2^J$ , on obtient  $\|U^J - U^{J,\eta}\|_{L^2(\Xi)}^2 \leq \eta^2$ , ce qui conclut la preuve.  $\square$

**Remarque 4.3.** Pour simplifier la présentation, on a considéré que les bases  $\{\Phi_{j,l,\alpha}\}_{l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi}$  et  $\{\Psi_{j,l,\alpha}\}_{l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi}$  sont orthonormales pour tout  $j \in \mathbb{N}$ . Dans la pratique, ainsi que dans le cas des AMR et AMR-MW présentées dans la littérature, les bases sont en fait orthogonales et les normes  $L^2$  des fonctions de base sont dépendantes du niveau de résolution ( $\|\Phi_{j,l,\alpha}\|_{L^2(\Xi)} = \|\Psi_{j,l,\alpha}\|_{L^2(\Xi)} = 2^{-j/2}$  pour tout  $l \in \mathcal{I}_j, \alpha \in \mathcal{P}_\pi$ ). On choisit alors des valeurs de seuillage dépendantes du niveau de résolution, à savoir  $\epsilon_j(\eta) = 2^{(j-J)/2}\eta$ .

#### 4.1.4 Complétion à la Harten

Si l'on considère l'AMR(-MW) d'Alpert pour  $N_0 = 0$  (ondelettes de Haar), la propriété d'annulation des moments (4.22) devient  $\langle 1, \Psi_{j,l,\alpha} \rangle = 0$ , i.e les détails s'annulent uniquement contre les constantes, ce qui n'est pas satisfaisant. L'idée de Harten consiste à considérer une complétion différente (non orthogonale) pour (4.5) dans le cas  $N_0 = 0$ , cf, entre autres, le livre de Cohen [12]. La grille grossière  $G_0$  est choisie telle que  $N_0 \geq 4$ . Pour éviter toute confusion, on note  $\tilde{W}_j^0$  les nouveaux supplémentaires à  $S_j^0$  dans  $S_{j+1}^0$ ,  $\forall j \in \mathbb{N}$ , tels que

$$S_{j+1}^0 = S_j^0 \oplus \tilde{W}_j^0. \quad (4.63)$$

Ces espaces sont toujours appelés espaces de détail. Par récurrence, on déduit de (4.63) que pour tout  $j \in \mathbb{N}_0$ ,

$$S_j^0 = S_0^0 \bigoplus_{j'=0}^{j-1} \tilde{W}_{j'}^0. \quad (4.64)$$

L'espace  $S_j^0$  est toujours engendré par les fonctions  $\{\Phi_{j,l}\}_{l \in \mathcal{I}_j}$  (on omet l'indice polynomial pour  $N_0 = 0$ ), appelées ici fonctions d'échelle de synthèse, qui correspondent en fait aux fonctions indicatrices des intervalles  $K_{j,l}$ , et on note  $\{\tilde{\Psi}_{j,l}\}_{l \in \mathcal{I}_j}$  les fonctions qui engendrent  $\tilde{W}_j^0$ , appelées ici ondelettes de synthèse. Pour  $N_0 = 0$ , les relations de décomposition à deux échelles (4.37)-(4.38) deviennent pour tout  $j \in \mathbb{N}$  et  $l \in \mathcal{I}_j$ ,

$$\Phi_{j,l} = \sum_{s \in \{-1,0\}} M_{11}^{\Phi,s} \Phi_{j+1,2l+s}, \quad \forall l \in \mathcal{I}_j, \quad (4.65)$$

$$\Psi_{j,l} = \sum_{s \in \{-1,0\}} M_{11}^{\Psi,s} \Phi_{j+1,2l+s}, \quad \forall l \in \mathcal{I}_j, \quad (4.66)$$

où  $M_{11}^{\Phi,s}$  et  $M_{11}^{\Psi,s}$  sont des scalaires.

Pour obtenir une propriété d'annulation des moments de la forme  $\langle \tilde{\Psi}_{j,l}, \xi^{\beta-1} \rangle = 0, \forall \beta \in \mathcal{P}_{M_0} := \{1, \dots, M_0\}$ , avec  $M_0 > 1$  fixé, on modifie linéairement la relation de décomposition à deux échelles (4.66) en ajoutant des combinaisons des fonctions  $\Phi_{j,l}$ . On considère pour tout  $j \in \mathbb{N}$  et  $l \in \mathcal{I}_j$  un stencil  $\mathcal{L}_{j,l} \subset \mathcal{I}_j$  tel que  $\text{card}(\mathcal{L}_{j,l}) = M_0$  et un vecteur  $L^{j,l} \in \mathbb{R}^{M_0}$ . On cherche l'ondelette de synthèse  $\tilde{\Psi}_{j,l}$  sous la forme

$$\tilde{\Psi}_{j,l} = \Psi_{j,l} + \sum_{l' \in \mathcal{L}_{j,l}} L_{l'}^{j,l} \Phi_{j,l'}, \quad (4.67)$$

et on détermine  $L^{j,l}$  de sorte que

$$\sum_{l' \in \mathcal{L}_{j,l}} \int_{K_{j,l'}} \xi^{\beta-1} \tilde{\Psi}_{j,l'}(\xi) d\xi = 0, \quad \forall \beta \in \mathcal{P}_{\text{Mo}}. \quad (4.68)$$

On obtient donc le vecteur  $L^{j,l}$  en résolvant le système linéaire  $A_l L^{j,l} = b_l$ , où  $A_l \in \mathbb{R}^{\text{Mo}, \text{Mo}}$  et  $b_l \in \mathbb{R}^{\text{Mo}}$  sont définis par

$$A_l := \left( \int_{K_{j,l'}} \xi^{\beta-1} \Phi_{j,l'}(\xi) d\xi \right)_{\beta \in \mathcal{P}_{\text{Mo}}, l' \in \mathcal{L}_{j,l}}, \quad (4.69)$$

$$b_l := \left( - \int_{K_{j,l}} \xi^{\beta-1} \Psi_{j,l}(\xi) d\xi \right)_{\beta \in \mathcal{P}_{\text{Mo}}}. \quad (4.70)$$

On améliore ainsi la propriété d'annulation des moments en élargissant les supports des fonctions de détails, plutôt qu'en augmentant le degré des fonctions d'échelle. Typiquement, dans le cas de l'AMR de Harten utilisée pour construire des schémas VF adaptatifs (cf Müller [63] et Cohen et. al [13]),  $\text{Mo} = 2t + 1$ , avec  $t \in \mathbb{N}_0$  et

$$\mathcal{L}_{j,l} := \{l - t + \mu(l), \dots, l + t + \mu(l)\}, \quad (4.71)$$

où  $\mu(l)$  permet d'effectuer un décentrement au bord. Par exemple, pour  $t = 1$ , qui est le cas le plus utilisé en pratique,  $\mu(l) = 0$  si  $l \in \{2, \dots, N_j - 1\}$  (stencil symétrique), tandis que  $\mu(1) = 1$  et  $\mu(N_j) = -1$ . On considère par la suite pour simplifier que  $t = 1$  ( $\text{Mo} = 3$ ) et que  $\mu(l)$  prend ses valeurs dans l'ensemble  $\{-1, 0, 1\}$ .

On réécrit (4.67) sous la forme

$$\tilde{\Psi}_{j,l} = \sum_{l' \in \tilde{\mathcal{L}}_{j,l}} \tilde{L}_{l'}^{j,l} \Phi_{j+1,l'}, \quad (4.72)$$

où  $\tilde{\mathcal{L}}_{j,l} \subset \mathcal{I}_{j+1}$  est le stencil de niveau  $(j + 1)$  de cardinal  $2\text{Mo} = 6$  tel que  $\bigcup_{l' \in \tilde{\mathcal{L}}_{j,l}} K_{j+1,l'} = \bigcup_{l' \in \mathcal{L}_{j,l}} K_{j,l'}$ . En posant  $n_{\min} = 2(-1 + \mu(l))$  et  $n_{\max} = 2(1 + \mu(l)) + 1$ , on a

$$\tilde{\Psi}_{j,l} = \sum_{n=n_{\min}}^{n_{\max}} \tilde{L}_{n+(2l-1)}^{j,l} \Phi_{j+1,n+(2l-1)}. \quad (4.73)$$

On observe que  $n_{\max} - n_{\min} = 5$ , ce qui donne 6 valeurs pour  $n$ . On introduit trois vecteurs de coefficients  $\left( \tilde{L}_{n+(2l-1)}^{j,l} \right)_{n \in \{n_{\min}, \dots, n_{\max}\}} \in \mathbb{R}^6$  qui dépendent donc de la valeur prise par  $\mu(l) \in \{-1, 0, 1\}$ . On les note  $\tilde{L}^\mu$  pour  $\mu \in \{-1, 0, 1\}$ . On considère ensuite les fonctions  $\bar{\Phi}^\mu$  définies sur  $\mathbb{R}$  et à support égal à  $[-1 + \mu, 2 + \mu]$ , solutions de l'équation fonctionnelle

$$\bar{\Phi}^\mu(\xi) = \sum_{n=2(-1+\mu)}^{2(1+\mu)+1} (-1)^{1-n} \tilde{L}_{1-n}^\mu \bar{\Phi}^\mu(2\xi - n), \quad (4.74)$$

et les fonctions  $\bar{\Psi}^\mu$  définies sur  $\mathbb{R}$  et à support égal à  $[\frac{-1+\mu}{2}, \frac{3+\mu}{2}]$ , vérifiant la relation

$$\bar{\Psi}^\mu(\xi) = M_{11}^{\Phi,0} \bar{\Phi}^\mu(2\xi) - M_{11}^{\Phi,-1} \bar{\Phi}^\mu(2\xi - 1). \quad (4.75)$$

Les fonctions  $\bar{\Phi}^\mu$  et  $\bar{\Psi}^\mu$  sont appelées respectivement fonction d'échelle mère d'analyse et ondelette mère d'analyse. Ces fonctions ne sont pas connues explicitement mais peuvent être approchées par des algorithmes itératifs [12]. De plus, elles fournissent par dilatation et translation des fonctions d'échelle d'analyse  $\bar{\Phi}_{j,l}^\mu$  et des ondelettes d'analyse  $\bar{\Psi}_{j,l}^\mu$ , qui vérifient les relations de biorthogonalité suivantes :

$$\langle \bar{\Phi}_{j,l}, \bar{\Phi}_{j,l'} \rangle = \delta_{ll'}, \quad \langle \bar{\Phi}_{j,l}, \bar{\Psi}_{j,l'} \rangle = 0, \quad (4.76)$$

$$\langle \tilde{\Psi}_{j,l}, \bar{\Phi}_{j,l'} \rangle = 0, \quad \langle \tilde{\Psi}_{j,l}, \bar{\Psi}_{j,l'} \rangle = \delta_{ll'}. \quad (4.77)$$

L'intérêt d'introduire les fonctions d'analyse et de synthèse ci-dessus est le suivant. Soit  $J$  un niveau de résolution fixé et une fonction  $U \in L^2(\Xi)$ . De façon similaire à (4.8) pour  $\text{No} = 0$ , on pose

$$\tilde{U}^J(\xi) = \sum_{l \in \mathcal{I}_J} u_{J,l} \bar{\Phi}_{J,l}(\xi), \quad (4.78)$$

où, pour tout  $j \in \mathbb{N}$ , les coefficients d'échelle sont donnés par  $u_{J,l} := \langle U, \bar{\Phi}_{J,l} \rangle$ . On vérifie que, de façon similaire à (4.10) pour  $\text{No} = 0$ , on a

$$\tilde{U}^J(\xi) = \sum_{j=-1}^{J-1} \sum_{l \in \mathcal{I}_j} d_{j,l} \bar{\Psi}_{j,l}(\xi), \quad (4.79)$$

où les coefficients d'ondelettes (détails) sont donnés par  $d_{j,l} := \langle U, \tilde{\Psi}_{j,l} \rangle$ . Il est important de noter que  $\tilde{U}^J$  n'est plus la projection orthogonale de  $U$  sur  $S_j^0$ . En revanche, l'expression (4.79) permet d'exploiter la propriété de décroissance des détails dans le contexte de schémas VF adaptatifs. Par la suite, on préférera utiliser l'AMR-MW d'Alpert afin d'exploiter le moindre recouvrement du support des multi-ondelettes à un niveau de résolution donné.

## 4.2 Arbres binaires

Les outils d'AMR présentés à la section 4.1 sont basés sur une grille uniforme de  $\Xi = [0, 1]$ . L'objectif est d'utiliser des grilles non-uniformes, celles-ci étant construites de manière adaptative. Une façon commode de représenter des grilles non-uniformes de  $\Xi$  est de considérer des arbres binaires. Ce choix est guidé par la structure dyadique de la hiérarchie de grilles et de détails. Il sera maintenu dans le cas multidimensionnel au chapitre 6.

### 4.2.1 Principales définitions

**Définition 4.1.** *Un arbre binaire  $\mathbf{T}$  est un ensemble fini de nœuds  $\mathcal{N}(\mathbf{T})$  munis d'une relation de filiation telle que*

- tout nœud a zéro ou deux enfants,
- tout nœud, sauf un seul localisé au sommet de l'arbre, appelé racine et noté  $\mathbf{n}_0$ , est fils d'un autre nœud.

Les nœuds qui n'ont pas d'enfants sont appelés les feuilles et sont rassemblés dans l'ensemble  $\mathcal{L}(\mathbf{T}) \subset \mathcal{N}(\mathbf{T})$ . On définit

- le parent d'un nœud  $\mathbf{n} \in \mathcal{N}(\mathbf{T}) \setminus \{\mathbf{n}_0\}$  :  $\mathbf{p}(\mathbf{n})$ ,
- les enfants “gauche” et “droit” d'un nœud  $\mathbf{n} \in \mathcal{N}(\mathbf{T}) \setminus \mathcal{L}(\mathbf{T})$  :  $\mathbf{c}^-(\mathbf{n})$  et  $\mathbf{c}^+(\mathbf{n})$ .

Par la suite, on considère toujours des arbres binaires qui ne sont pas réduits au nœud racine.

Pour un nœud  $\mathbf{n} \in \mathcal{N}(\mathbf{T})$ , sa profondeur  $|\mathbf{n}|$  est définie relativement au nœud racine ; le nœud racine  $\mathbf{n}_0$  a une profondeur  $|\mathbf{n}_0| = 0$  et ses deux enfants ont une profondeur  $|\mathbf{c}^-(\mathbf{n}_0)| = |\mathbf{c}^+(\mathbf{n}_0)| = 1$ . De manière générale, on a la relation

$$|\mathbf{n}| = |\mathbf{p}(\mathbf{n})| + 1. \quad (4.80)$$

La profondeur d'un arbre est la profondeur maximale sur l'ensemble de ses nœuds. On la note  $|\mathbf{T}|$  et on a

$$|\mathbf{T}| = \max_{\mathbf{n} \in \mathcal{N}(\mathbf{T})} |\mathbf{n}| = \max_{\mathbf{l} \in \mathcal{L}(\mathbf{T})} |\mathbf{l}|. \quad (4.81)$$

**Définition 4.2.** *Un arbre binaire est dit complet si tout nœud  $\mathbf{n}$  de profondeur  $|\mathbf{n}| < |\mathbf{T}|$  a deux enfants.*

On montre facilement que dans un arbre complet, il y a  $2^j$  nœuds ayant une profondeur  $j \leq |\mathbf{T}|$ . Les  $2^{|\mathbf{T}|}$  nœuds de profondeur  $|\mathbf{T}|$  n'ont pas d'enfants et sont donc les feuilles de  $\mathbf{T}$ . On notera par la suite  $\mathbf{T}_J$  l'arbre binaire complet de profondeur  $J$ .

#### 4.2.2 Grille associée à un arbre binaire

Soit  $\mathbf{T}$  un arbre binaire. À chaque nœud  $\mathbf{n} \in \mathcal{N}(\mathbf{T})$ , on peut associer un support  $\mathbf{S}(\mathbf{n}) \subset \Xi$  défini récursivement. Tout d'abord, le support du nœud racine est  $\mathbf{S}(\mathbf{n}_0) = \Xi$ . Soit maintenant  $\mathbf{S}(\mathbf{n}) = [\xi_n^-, \xi_n^+]$  le support d'un nœud  $\mathbf{n}$ , avec  $\mathbf{n} \notin \mathcal{L}(\mathbf{T})$ . Alors, les supports de ses enfants gauche et droit sont respectivement

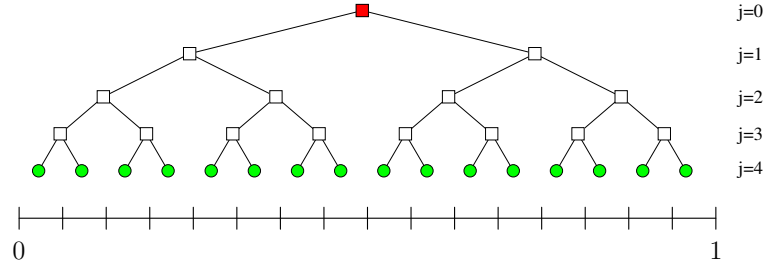
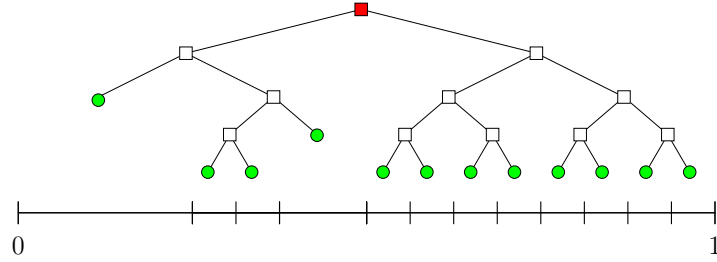
$$\mathbf{S}(\mathbf{c}^-(\mathbf{n})) = [\xi_n^-, (\xi_n^- + \xi_n^+)/2] \quad \text{et} \quad \mathbf{S}(\mathbf{c}^+(\mathbf{n})) = [(\xi_n^- + \xi_n^+)/2, \xi_n^+]. \quad (4.82)$$

On définit ainsi la grille associée à l'arbre binaire  $\mathbf{T}$  sous la forme

$$G(\mathbf{T}) = \bigcup_{\mathbf{l} \in \mathcal{L}(\mathbf{T})} \{\mathbf{S}(\mathbf{l})\}. \quad (4.83)$$

On vérifie facilement que  $\bigcup_{\mathbf{l} \in \mathcal{L}(\mathbf{T})} \mathbf{S}(\mathbf{l}) = \Xi$  et que  $|\mathbf{S}(\mathbf{l}) \cap \mathbf{S}(\mathbf{l}')| = 0$  si  $\mathbf{l} \neq \mathbf{l}'$ .

L'arbre binaire complet  $\mathbf{T}_J$  et sa grille uniforme  $G(\mathbf{T}_J)$  sont représentés sur la figure 4.5 pour  $J = 4$ . Un exemple d'arbre binaire (incomplet)  $\mathbf{T}$  de profondeur 4 et sa grille non uniforme  $G(\mathbf{T})$  sont représentés sur la figure 4.6.

FIGURE 4.5 – Arbre binaire complet  $T_J$  et sa grille uniforme pour  $J = 4$ .FIGURE 4.6 – Exemple d'un arbre binaire (incomplet)  $T$  de profondeur 4 et sa grille non uniforme  $G(T)$ .

### 4.2.3 Espace d'approximation stochastique associé à un arbre binaire

L'espace d'approximation stochastique associé à l'arbre binaire  $T$  est défini par

$$S^{\text{No}}(T) := \{f \in L^2(\Xi), \forall 1 \in \mathcal{L}(T), f|_{\mathcal{S}(1)} \in \Pi_{\text{No}}[\xi]\}. \quad (4.84)$$

Sa dimension vaut

$$\dim(S^{\text{No}}(T)) = P_\pi \times \text{card}(\mathcal{L}(T)) =: P(T), \quad (4.85)$$

où  $P_\pi = (\text{No} + 1) = \text{card}(\mathcal{P}_\pi)$ . Pour un arbre binaire complet  $T_J$ , on observe que  $S^{\text{No}}(T) = S_J^{\text{No}}$ , où les espaces  $S_J^{\text{No}}$  sont ceux définis par (4.2).

Pour un arbre binaire  $T$ , on réécrit les développements SE et MW de la projection orthogonale d'une fonction  $U \in L^2(\Xi)$  sur  $S^{\text{No}}(T)$  sous la forme

$$P_{S^{\text{No}}(T)}U(\xi) := U^T(\xi) = \sum_{1 \in \mathcal{L}(T)} \sum_{\alpha \in \mathcal{P}_\pi} u_{1,\alpha} \Phi_{1,\alpha}(\xi) \quad (4.86)$$

$$= \sum_{\alpha \in \mathcal{P}_\pi} u_{n_0,\alpha} \Phi_{n_0,\alpha}(\xi) + \sum_{\substack{\mathbf{n} \in \mathcal{N}(T) \\ \mathbf{n} \notin \mathcal{L}(T)}} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}(\xi), \quad (4.87)$$

où pour toute feuille  $1 \in \mathcal{L}(T)$ ,  $\{\Phi_{1,\alpha}\}_{\alpha \in \mathcal{P}_\pi}$  sont les fonctions de base SE sur  $\mathcal{S}(1)$ ,  $\{\Phi_{n_0,\alpha}\}_{\alpha \in \mathcal{P}_\pi}$  est la base de Legendre sur  $[0, 1]$  et pour tout nœud  $\mathbf{n} \in \mathcal{N}(T) \setminus \mathcal{L}(T)$ ,  $\{\Psi_{\mathbf{n},\alpha}\}_{\alpha \in \mathcal{P}_\pi}$  sont les fonctions de base MW sur  $\mathcal{S}(\mathbf{n})$ . Afin d'écrire (4.87) sous forme



plus compacte, on note  $\mathbf{n}_0^*$  un dédoublement de la racine de l'arbre et on pose  $d_{\mathbf{n}_0^*,\alpha} = u_{\mathbf{n}_0,\alpha}$  et  $\Psi_{\mathbf{n}_0^*,\alpha} = \Phi_{\mathbf{n}_0,\alpha}$ , si bien que (4.87) devient

$$U^T(\xi) = \sum_{\mathbf{n} \in \mathcal{N}^*(\mathbb{T})} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}(\xi), \quad (4.88)$$

où on a posé

$$\mathcal{N}^*(\mathbb{T}) := \{\mathbf{n}_0^*\} \cup \mathcal{N}(\mathbb{T}) \setminus \mathcal{L}(\mathbb{T}). \quad (4.89)$$

Les coefficients d'échelle et les détails sont déterminés par

$$u_{\mathbf{1},\alpha} := \langle U, \Phi_{\mathbf{1},\alpha} \rangle, \quad d_{\mathbf{n},\alpha} := \langle U, \Psi_{\mathbf{n},\alpha} \rangle. \quad (4.90)$$

En utilisant un multi-indice  $\mathbf{k}$ , on reformule simplement ces développements par

$$U^T(\xi) = \sum_{\mathbf{k} \in \text{SE}(\mathbb{T})} u_{\mathbf{k}} \Phi_{\mathbf{k}}(\xi) = \sum_{\mathbf{k} \in \text{MW}(\mathbb{T})} d_{\mathbf{k}} \Psi_{\mathbf{k}}(\xi), \quad (4.91)$$

où  $\text{SE}(\mathbb{T})$  et  $\text{MW}(\mathbb{T})$  sont les ensembles de multi-indices  $(\mathbf{1}, \alpha) \in \mathcal{L}(\mathbb{T}) \times \mathcal{P}_\pi$  et  $(\mathbf{n}, \alpha) \in \mathcal{N}^*(\mathbb{T}) \times \mathcal{P}_\pi$  des développements SE et MW respectivement, tous deux de cardinal  $P(\mathbb{T})$ , car  $\text{card}(\mathcal{N}^*(\mathbb{T})) = \text{card}(\mathcal{N}(\mathbb{T}) \setminus \mathcal{L}(\mathbb{T})) + 1 = \text{card}(\mathcal{L}(\mathbb{T}))$ , ce qu'on peut vérifier par récurrence en ajoutant à chaque fois deux enfants à une feuille d'un arbre.

#### 4.2.4 Inclusion d'arbres

On définit la relation d'inclusion entre deux arbres  $\mathbb{T}$  et  $\mathbb{T}'$  suivante :

$$\mathbb{T} \subset \mathbb{T}' \iff \mathcal{N}(\mathbb{T}) \subset \mathcal{N}(\mathbb{T}'). \quad (4.92)$$

Il s'agit d'une relation d'ordre partiel sur les arbres binaires. On vérifie facilement que

$$\mathbb{T} \subset \mathbb{T}' \Rightarrow S(\mathbb{T}) \subset S(\mathbb{T}'). \quad (4.93)$$

### 4.3 Adaptation à une fonction connue

Soit  $J$  un niveau de résolution maximale et soit une fonction  $U \in L^2(\Xi)$ . On note  $\mathbb{T}_J$  l'arbre binaire complet de profondeur  $J$  et  $U^{\mathbb{T}_J}$  la projection  $L^2(\Xi)$ -orthogonale de  $U$  sur l'espace d'approximation stochastique  $S(\mathbb{T}_J)$ . On souhaite construire un arbre  $\mathbb{T}_{J,\eta}$ , de profondeur  $|\mathbb{T}_{J,\eta}| \leq J$ , qui contienne uniquement les nœuds de  $\mathcal{N}^*(\mathbb{T}_J)$  dont les détails  $d_{\mathbf{n},\alpha}$  sont significatifs.

#### 4.3.1 Principe

On redéfinit l'ensemble des double-indices  $\mathcal{D}_{J,\eta}$ , auparavant défini par (4.60), comme un ensemble de nœuds

$$\mathcal{D}_{J,\eta} := \left\{ \mathbf{n} \in \mathcal{N}^*(\mathbb{T}_J) : \|d_{\mathbf{n}}\|_{l^2} = \left( \sum_{\alpha \in \mathcal{P}_\pi} |d_{\mathbf{n},\alpha}|^2 \right)^{1/2} > \epsilon_J(\eta) \right\}, \quad (4.94)$$

où  $\epsilon_J(\eta) = 2^{-J/2}\eta$ . Par convention,  $\mathbf{n}_0$  et  $\mathbf{n}_0^*$  sont dans  $\mathcal{D}_{J,\eta}$ . En général, l'ensemble  $\mathcal{D}_{J,\eta}$  n'engendre pas nécessairement une structure d'arbre binaire. Voici un exemple. On représente au centre de la figure 4.7 un arbre binaire complet de profondeur 3. Les carrés représentent les nœuds dans  $\mathcal{N}^*(\mathbb{T}_J)$  (le dédoublement de la racine est omis) et les ronds (verts) les feuilles dans  $\mathcal{L}(\mathbb{T}_J)$ . Un nœud dans  $\mathcal{N}^*(\mathbb{T}_J)$  est coloré en bleu si il appartient à  $\mathcal{D}_{J,\eta}$ ; sinon, il est laissé blanc. L'exemple d'arbre donné ici possède un nœud au niveau de résolution  $j = 1$  dont les détails ne sont pas significatifs alors que les détails de ses enfants le sont. L'ensemble  $\mathcal{D}_{J,\eta}$  n'a donc pas une structure d'arbre binaire puisque l'élimination du nœud représenté par le carré blanc laisse des orphelins.

On introduit par conséquent deux nouveaux ensembles de nœuds  $\mathcal{D}_{J,\eta}^b$  et  $\mathcal{D}_{J,\eta}^\sharp$  définis par

$$\mathcal{D}_{J,\eta}^b := \left\{ \mathbf{n} \in \mathcal{N}^*(\mathbb{T}_J), \forall \mathbf{n}' \in \mathbf{A}(\mathbf{n}) : \|d_{\mathbf{n}'}\|_{l^2} = \left( \sum_{\alpha \in \mathcal{P}_\pi} |d_{\mathbf{n}',\alpha}|^2 \right)^{1/2} > \epsilon_J(\eta) \right\}, \quad (4.95)$$

$$\mathcal{D}_{J,\eta}^\sharp := \left\{ \mathbf{n} \in \mathcal{N}^*(\mathbb{T}_J), \exists \mathbf{n}' \in \mathbf{D}(\mathbf{n}) : \|d_{\mathbf{n}'}\|_{l^2} = \left( \sum_{\alpha \in \mathcal{P}_\pi} |d_{\mathbf{n}',\alpha}|^2 \right)^{1/2} > \epsilon_J(\eta) \right\}, \quad (4.96)$$

où  $\mathbf{A}(\mathbf{n})$  et  $\mathbf{D}(\mathbf{n})$  désignent respectivement les ancêtres et les descendants d'un nœud :

$$\mathbf{A}(\mathbf{n}) := \{ \mathbf{n}' \in \mathcal{N}^*(\mathbb{T}_J) : |\mathbf{n}'| \leq |\mathbf{n}|, \mathbf{S}(\mathbf{n}') \supseteq \mathbf{S}(\mathbf{n}) \}, \quad (4.97)$$

$$\mathbf{D}(\mathbf{n}) := \{ \mathbf{n}' \in \mathcal{N}^*(\mathbb{T}_J) : |\mathbf{n}'| \geq |\mathbf{n}|, \mathbf{S}(\mathbf{n}') \subseteq \mathbf{S}(\mathbf{n}) \}. \quad (4.98)$$

On note que  $\mathbf{n} \in \mathbf{A}(\mathbf{n})$  et  $\mathbf{n} \in \mathbf{D}(\mathbf{n})$  par définition. On vérifie facilement que  $\mathcal{D}_{J,\eta}^b$  et  $\mathcal{D}_{J,\eta}^\sharp$  engendrent des arbres binaires que l'on note respectivement  $\mathbb{T}_{J,\eta}^b$  et  $\mathbb{T}_{J,\eta}^\sharp$ . De plus, on a pour l'AMR-MW d'Alpert

$$\mathcal{D}_{J,\eta}^b \subseteq \mathcal{D}_{J,\eta} \subseteq \mathcal{D}_{J,\eta}^\sharp. \quad (4.99)$$

Les arbres  $\mathbb{T}_{J,\eta}^b$  et  $\mathbb{T}_{J,\eta}^\sharp$  sont illustrés respectivement à gauche et à droite sur la figure 4.7. Les projections seuillées  $U^{\mathbb{T}_{J,\eta}^b}$  et  $U^{\mathbb{T}_{J,\eta}^\sharp}$  associées respectivement aux ensembles  $\mathcal{D}_{J,\eta}^b$  et  $\mathcal{D}_{J,\eta}^\sharp$  sont définies par les développements MW suivants :

$$U^{\mathbb{T}_{J,\eta}^b}(\xi) = \sum_{\mathbf{n} \in \mathcal{D}_{J,\eta}^b} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}(\xi), \quad (4.100)$$

$$U^{\mathbb{T}_{J,\eta}^\sharp}(\xi) = \sum_{\mathbf{n} \in \mathcal{D}_{J,\eta}^\sharp} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}(\xi). \quad (4.101)$$

Outre les développements MW (4.100) et (4.101), les projections seuillées  $U^{\mathbb{T}_{J,\eta}^b}$  et  $U^{\mathbb{T}_{J,\eta}^\sharp}$  admettent des développements sur les bases SE locales associées aux feuilles

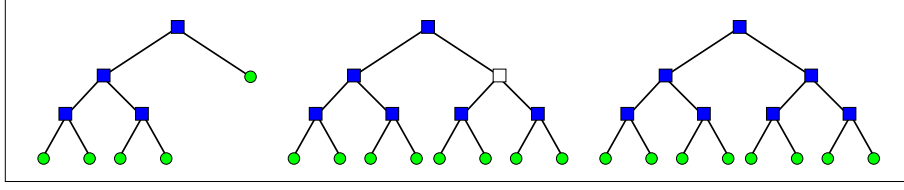


FIGURE 4.7 – Exemple d'arbre binaire complet et son ensemble  $\mathcal{D}_{J,\eta}$  (au centre). Les carrés représentent les nœuds dans  $\mathcal{N}^*(\mathbb{T}_J)$  (le dédoublement de la racine est omis) et les ronds (verts) les feuilles dans  $\mathcal{L}(\mathbb{T}_J)$ . Un nœud dans  $\mathcal{N}^*(\mathbb{T}_J)$  est coloré en bleu si il appartient à  $\mathcal{D}_{J,\eta}$  ; sinon, il est laissé blanc. Ensembles  $\mathcal{D}_{J,\eta}^b$  (à gauche) et  $\mathcal{D}_{J,\eta}^\sharp$  (à droite) correspondants.

des arbres  $\mathbb{T}_{J,\eta}^b$  et  $\mathbb{T}_{J,\eta}^\sharp$  sous la forme

$$U^{\mathbb{T}_{J,\eta}^b}(\xi) = \sum_{1 \in \mathcal{L}(\mathbb{T}_{J,\eta}^b)} \sum_{\alpha \in \mathcal{P}_\pi} u_{1,\alpha} \Phi_{1,\alpha}(\xi), \quad (4.102)$$

$$U^{\mathbb{T}_{J,\eta}^\sharp}(\xi) = \sum_{1 \in \mathcal{L}(\mathbb{T}_{J,\eta}^\sharp)} \sum_{\alpha \in \mathcal{P}_\pi} u_{1,\alpha} \Phi_{1,\alpha}(\xi). \quad (4.103)$$

On souhaite évaluer l'erreur d'approximation introduite par le seuillage lorsque l'on approche  $U^{\mathbb{T}_J}$  par l'une ou l'autre de ses projections seuillées  $U^{\mathbb{T}_{J,\eta}^b}$  ou  $U^{\mathbb{T}_{J,\eta}^\sharp}$ . Typiquement, on aimerait obtenir des propriétés d'approximation semblables à la proposition 4.2. La projection seuillée  $U^{J,\eta}$  définie par (4.61) se réécrit sous la forme

$$U^{J,\eta}(\xi) = \sum_{\mathbf{n} \in \mathcal{D}_{J,\eta}} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}(\xi). \quad (4.104)$$

**Proposition 4.3.** *Soit  $\eta > 0$  et  $\epsilon_J(\eta) = 2^{-J/2}\eta$ . Alors, pour tout  $U \in L^2(\Xi)$ , on a*

$$\|U^{\mathbb{T}_J} - U^{\mathbb{T}_{J,\eta}^\sharp}\|_{L^2(\Xi)} \leq \eta. \quad (4.105)$$

*Démonstration.* Comme  $\mathcal{D}_{J,\eta} \subseteq \mathcal{D}_{J,\eta}^\sharp$ , on a

$$\|U^{\mathbb{T}_J} - U^{\mathbb{T}_{J,\eta}^\sharp}\|_{L^2(\Xi)} \leq \|U^{\mathbb{T}_J} - U^{J,\eta}\|_{L^2(\Xi)}.$$

On conclut grâce à la proposition 4.2.  $\square$

Le cas de  $U^{\mathbb{T}_{J,\eta}^b}$  est plus délicat puisque l'ensemble  $\mathcal{D}_{J,\eta}^b$  peut être plus petit que  $\mathcal{D}_{J,\eta}$ . On prouve ici un résultat d'approximation par une classe raisonnable de fonctions dont les détails ont une propriété de “non-croissance”.

**Proposition 4.4.** *Soit  $\rho$  un réel strictement positif. Soit  $\mathcal{C}_\rho^b$  la classe de fonctions de  $L^2(\Xi)$  telles que*

$$\|d_{\mathbf{n}}\|_{l^2} \leq \rho \|d_{\mathbf{n}'}\|_{l^2}, \quad \forall \mathbf{n} \in \mathcal{N}^*(\mathbb{T}_J), \mathbf{n}' \in \mathbf{A}(\mathbf{n}) \text{ t.q. } |\mathbf{n}'| = |\mathbf{n}| - 1. \quad (4.106)$$

Soit  $\eta > 0$  et  $\epsilon_{J,\rho}(\eta) = (2 \max(1, \rho)^2)^{-J/2} \eta$ . Alors, pour toute fonction  $U \in \mathcal{C}_\rho^b$ , on a

$$\|U^{\mathbf{T}^J} - U^{\mathbf{T}_{J,\eta}^b}\|_{L^2(\Xi)} \leq \eta. \quad (4.107)$$

*Démonstration.* Soit  $U \in \mathcal{C}_\rho^b$ . Pour tout  $\mathbf{n} \notin \mathcal{D}_{J,\eta}^b$ , soit

$$\mathbf{n}' = \arg \min_{\substack{\mathbf{n}'' \in \mathbf{A}(\mathbf{n}) \\ \mathbf{n}'' \notin \mathcal{D}_{J,\eta}^b}} |\mathbf{n}''|.$$

On a  $\|d_{\mathbf{n}'}\|_{l^2} \leq \epsilon_{J,\rho}(\eta)$  et puisque  $U \in \mathcal{C}_\rho^b$ , il vient

$$\|d_{\mathbf{n}}\|_{l^2} \leq \rho^{(|\mathbf{n}|-|\mathbf{n}'|)} \epsilon_{J,\rho}(\eta) \leq \max(1, \rho)^{|\mathbf{n}|} \epsilon_{J,\rho}(\eta)$$

Par conséquent, l'erreur d'approximation est estimée en norme  $L^2$  par

$$\begin{aligned} \|U^{\mathbf{T}^J} - U^{\mathbf{T}_{J,\eta}^b}\|_{L^2(\Xi)}^2 &= \sum_{\mathbf{n} \notin \mathcal{D}_{J,\eta}^b} (\max(1, \rho)^{|\mathbf{n}|} \epsilon_{J,\rho}(\eta))^2 \\ &\leq \left(1 + \sum_{j=0}^{J-1} (2 \max(1, \rho)^2)^j\right) (\epsilon_{J,\rho}(\eta))^2, \end{aligned}$$

et on conclut comme précédemment.  $\square$

### 4.3.2 Aspects algorithmiques

Les deux ensembles de nœuds  $\mathcal{D}_{J,\eta}^b$  et  $\mathcal{D}_{J,\eta}^\sharp$  fournissent deux stratégies différentes pour construire un arbre des détails significatifs  $\mathbf{T}_{J,\eta}^b$  ou  $\mathbf{T}_{J,\eta}^\sharp$  pour une fonction  $U \in L^2(\Xi)$ . Dans cette section, on décrit les deux algorithmes permettant simultanément la construction de  $\mathbf{T}_{J,\eta}^b$  ou  $\mathbf{T}_{J,\eta}^\sharp$  et la détermination des coefficients d'échelle  $(u_{\mathbf{l},\alpha})_{\alpha \in \mathcal{P}_\pi}$  pour  $\mathbf{l} \in \mathcal{L}(\mathbf{T}_{J,\eta}^b)$  ou  $\mathbf{l} \in \mathcal{L}(\mathbf{T}_{J,\eta}^\sharp)$ .

#### 4.3.2.1 Algorithme correspondant à $\mathcal{D}_{J,\eta}^b$

On construit itérativement  $\mathbf{T}_{J,\eta}^b$  et les coefficients d'échelle de  $U^{\mathbf{T}_{J,\eta}^b}$  en partant de l'arbre  $\mathbf{T}_1 := \{\mathbf{n}_0, \mathbf{c}^-(\mathbf{n}_0), \mathbf{c}^+(\mathbf{n}_0)\}$  et en recalculant les projections orthogonales locales de  $U$  sur des espaces d'approximation stochastiques associés à de nouvelles feuilles tant que les détails associés aux nœuds parents sont significatifs.

L'algorithme de construction est le suivant :

#### Algorithme 4.1.

*%Initialisation :*

$$\mathbf{T}_{J,\eta}^b = \mathbf{T}_1 := \{\mathbf{n}_0, \mathbf{c}^-(\mathbf{n}_0), \mathbf{c}^+(\mathbf{n}_0)\}$$

*Faire Enrich( $\mathbf{n}_0$ )*

*%Routine récursive Enrich( $\mathbf{n}$ )*

*Calculer les coefficients d'échelle  $u_{\mathbf{c}^\pm(\mathbf{n})} = (u_{\mathbf{c}^\pm(\mathbf{n}),\alpha})_{\alpha \in \mathcal{P}_\pi}$  (par quadrature)*

Calculer les détails  $d_{\mathbf{n}} = M^{\Psi, -1}u_{c^-(\mathbf{n})} + M^{\Psi, 0}u_{c^+(\mathbf{n})}$   
**si**  $\mathbf{n} = \mathbf{n}_0$  ou  $\|d_{\mathbf{n}}\|_{l^2} > \epsilon_J(\eta)$  **alors**  
 Créer des feuilles aux enfants de  $\mathbf{n} : \mathbb{T}_{J,\eta}^b \leftarrow \mathbb{T}_{J,\eta}^b \cup \{c^\pm(c^\pm(\mathbf{n}))\}$   
 Faire  $\text{Enrich}(c^\pm(\mathbf{n}))$   
**sinon**  
 Supprimer les enfants de  $\mathbf{n} : \mathbb{T}_{J,\eta}^b \leftarrow \mathbb{T}_{J,\eta}^b \setminus \{c^\pm(\mathbf{n})\}$   
**fin si**

Il s'agit d'un algorithme d'enrichissement d'arbre. Il sera adapté dans la section 5.1.4.1 au cas d'un arbre binaire  $\mathbb{T}$  pour lequel on connaît déjà le développement SE de  $U^{\mathbb{T}}$ .

#### 4.3.2.2 Algorithme correspondant à $\mathcal{D}_{J,\eta}^\sharp$

On construit  $\mathbb{T}_{J,\eta}^\sharp$  et les coefficients d'échelle de  $U^{\mathbb{T}_{J,\eta}^\sharp}$  en partant de l'arbre complet  $\mathbb{T}_J$  et des coefficients d'échelle de  $U$  sur  $\mathcal{L}(\mathbb{T}_J)$ . On procède de façon itérative à un élagage de l'arbre ainsi qu'à un seuillage du développement SE de  $U^{\mathbb{T}_J}$  en fonction de la taille de ses détails.

L'algorithme de construction est le suivant. On définit  $\mathcal{L}^-(\mathbb{T})$  comme l'ensemble des feuilles de gauche d'un arbre binaire  $\mathbb{T}$ , *i.e.*

$$\mathcal{L}^-(\mathbb{T}) := \{1 \in \mathcal{L}(\mathbb{T}), 1 = c^-(\mathbf{p}(1))\}, \quad (4.108)$$

et pour toute feuille  $1 \in \mathcal{L}(\mathbb{T})$ , on introduit un indicateur  $\text{verif}(1)$  qui vaut *true* si la feuille doit être testée et *false* sinon. On introduit également un indicateur global  $\text{elag}$  qui vaut *true* tant qu'il peut exister des feuilles à supprimer et *false* sinon.

#### Algorithme 4.2.

*%Initialisation :*  
 $\forall 1 \in \mathcal{L}(\mathbb{T}_J) : \text{Calculer } u_1 = (u_{1,\alpha})_{\alpha \in \mathcal{P}_\pi}$ ,  $\text{verif}(1) = \text{true}$   
 $\mathbb{T}_{J,\eta}^\sharp = \mathbb{T}_J$   
 $\text{elag} = \text{true}$   
**tant que**  $\text{elag} = \text{true}$  **faire**  
 $\text{elag} = \text{false}$   
**pour** toute feuille  $1 \in \mathcal{L}^-(\mathbb{T}_{J,\eta}^\sharp)$  **faire**  
 $\text{Elag}(1)$   
**fin pour**  
**fin tant que**  
*%Routine Elag(1)*  
**si**  $\text{verif}(1) = \text{true}$  et le nœud sœur  $1' = c^+(\mathbf{p}(1))$  est une feuille **alors**  
 Calculer les détails  $d_{\mathbf{p}(1)} = M^{\Psi, -1}u_1 + M^{\Psi, 0}u_{1'}$   
**si**  $\mathbf{n} \neq \mathbf{n}_0$  et  $\|d_{\mathbf{p}(\mathbf{n}^-)}\|_{l^2} \leq \epsilon_J(\eta)$  **alors**  
 $\text{elag} = \text{true}$   
 Supprimer les feuilles  $(1, 1') : \mathbb{T}_{J,\eta}^\sharp \leftarrow \mathbb{T}_{J,\eta}^\sharp \setminus \{1, 1'\}$   
 Calculer les coefficients d'échelle  $u_{\mathbf{p}(1)} = M^{\Phi, -1}u_1 + M^{\Phi, 0}u_{1'}$

```

    verif(p(1)) = true
  sinon
    Conserver 1 et 1'
    verif(1) = false
  fin si
fin si

```

Il s'agit d'un algorithme d'élagage d'arbre. Il sera utilisé tel quel dans la section 5.1.4.1 dans le cas d'une application à un arbre binaire  $T$  (pas forcément complet) pour lequel on connaît déjà le développement SE de  $U^T$ .

L'opérateur multi-échelle  $\mathcal{M}_J$  n'est pas utilisé pour le calcul des coefficients d'échelle et des détails. En fait, on calcule ces quantités localement en utilisant les matrices des coefficients de filtre. Ceci est possible car, dans le cas de l'AMR-MW d'Alpert, les relations d'échelles (4.37) et (4.38) sont localisées sur les éléments stochastiques.

### 4.3.3 Résultats

On teste dans cette section les algorithmes d'enrichissement 4.1 et d'élagage 4.2 sur trois fonctions de régularité différente :

$$f^1(\xi) = \exp(-200(\xi - 0.5)^2), \quad (4.109)$$

$$f^2(\xi) = \begin{cases} -1, & \xi < 1/3, \\ 3(2\xi - 1) & 1/3 \leq \xi \leq 2/3, \\ 1, & \xi > 2/3, \end{cases} \quad (4.110)$$

$$f^3(\xi) = \tanh(100(\xi - 0.4)). \quad (4.111)$$

On fixe le niveau de résolution maximal à  $J = 10$ . En implémentant les algorithmes 4.1 et 4.2 pour différents ordres polynomiaux  $No$  et paramètres de seuillage  $\eta$ , on trouve les mêmes arbres de détails significatifs, *i.e.*  $\mathcal{D}_{J,\eta}^b = \mathcal{D}_{J,\eta} = \mathcal{D}_{J,\eta}^\sharp$ . Dans ces conditions, on pose  $T_{J,\eta} = T_{J,\eta}^b = T_{J,\eta}^\sharp$  et  $G_{J,\eta} = G(T_{J,\eta})$ .

Dans un premier temps, on fixe le paramètre de seuillage à  $\eta = 10^{-3}$  et on fait varier le degré polynomial  $No$ . On présente sur les deux premières lignes des figures 4.8, 4.9 et 4.10 pour chaque fonction  $f^i$ ,  $i \in \{1, 2, 3\}$ , sa projection seuillée sur l'arbre des détails significatifs  $T_{J,\eta}$  pour  $No = 2$ , soit  $(f^i)^{T_{J,\eta}}$ , et les grilles adaptatives  $G(T_{J,\eta})$  pour  $No \in \{0, 1, 2\}$ . Les grilles sont représentées par une fonction constante par morceaux égale au niveau de résolution utilisé localement. On observe que pour la fonction  $f^1$ , qui est infiniment dérivable, l'ordre polynomial  $No = 2$  est intéressant car on a besoin de moins d'éléments stochastiques. En particulier, dans la zone où la dérivée de  $f^1$  est forte (sur l'intervalle  $[0.4, 0.6]$ ), un niveau de résolution 6 est suffisant, alors que pour  $No = 0$  on atteint le niveau de résolution maximal 10 et pour  $No = 1$  le niveau de résolution 8. Pour la fonction  $f^2$ , continue mais dont la dérivée en  $\xi = 1/3$  et  $\xi = 2/3$  est discontinue, l'analyse des détails conduit à les conserver jusqu'au niveau de résolution 10 pour  $No \in \{0, 1\}$  et 9 pour  $No = 2$

aux voisinages de  $\xi = 1/3$  et  $\xi = 2/3$ . Pour  $N_0 = 0$ , un niveau de résolution 2 est suffisant pour  $\xi < 1/3$  et  $\xi > 2/3$ , puisque  $f^2$  est constante dans ces zones. En revanche, les détails jusqu'au niveau de résolution 10 sont conservés dans la zone où  $f^2$  est affine. Pour  $N_0 \in \{1, 2\}$ , les niveaux de résolution 2 et 3 sont suffisants loin des singularités. Pour la fonction  $f^3$ , qui est quasiment discontinue en  $\xi = 0.4$ , on utilise également des niveaux de résolution grands aux voisinages de la discontinuité pour les différents  $N_0$ . En dehors de la zone de singularité, des niveaux de résolution faibles sont suffisants puisque la fonction est constante.

Ensuite, on fait varier l'ordre polynomial et le paramètre de seuillage et on vérifie qu'on obtient une erreur d'approximation  $\|(f^i)^{T_J} - (f^i)^{T_{J,\eta}}\|_{L^2(\Xi)}$  en  $O(\eta)$ , où  $(f^i)^{T_J}$  est la projection de  $f^i$  sur l'arbre complet  $T^J$ . La convergence de l'erreur d'approximation avec le paramètre  $\eta$  est représentée pour les différentes fonctions en bas à gauche sur les figures 4.8, 4.9 et 4.10 pour  $N_0 \in \{0, 1, 2\}$  et  $\eta = 10^{-k}$ ,  $k \in \{0, \dots, 10\}$ . On observe bien le taux de décroissance de l'erreur en  $O(\eta)$ . Pour la fonction  $f^2$ , l'erreur est identiquement nulle pour les petites valeurs de seuillage. En effet, pour  $\eta$  petit, les détails calculés sont significatifs dans les zones de singularité et sont donc tous conservés. Dans les zones où  $f^2$  est constante, la projection seuillée obtenue est exacte pour le niveau de résolution 2 pour les différents  $N_0$ . Pour  $N_0 > 0$ , elle est également exacte dans la zone où  $f^2$  est affine. Pour  $N_0 = 0$ , les détails dans cette zone sont toujours significatifs et sont donc toujours conservés.

Enfin, on représente en bas à droite sur les figures 4.8, 4.9 et 4.10 l'erreur d'approximation  $\|(f^i)^{T_J} - (f^i)^{T_{J,\eta}}\|_{L^2(\Xi)}$  en fonction du nombre de degrés de liberté obtenus en faisant varier  $\eta$ . Le nombre de degrés de liberté est égal à la dimension de l'espace d'approximation stochastique, *i.e.*  $P(T_{J,\eta}) = P_\pi \times \text{card}(\mathcal{L}(T_{J,\eta}))$ . Pour la fonction  $f^1$ , on observe que pour  $\eta \geq 10^{-4}$ , on a besoin de moins de degrés de liberté avec  $N_0 = 1$  qu'avec  $N_0 = 0$  pour obtenir une erreur d'approximation d'ordre  $\eta$ ; cependant la représentation de la solution avec  $N_0 = 1$  et  $\eta < 10^{-5}$  est moins intéressante puisqu'on a alors besoin de plus de degrés de liberté. La représentation de la solution utilisant  $N_0 = 2$  est la plus intéressante pourvu que  $\eta \geq 10^{-6}$ ; au-dessous de cette valeur, le nombre de degrés de liberté devient plus important que pour  $N_0 = 0$ . Pour la fonction  $f^2$ , le plus intéressant est d'utiliser le degré  $N_0 = 1$ ; on pouvait s'y attendre puisque la fonction est linéaire par morceaux et l'utilisation d'un espace d'approximation stochastique de polynômes de degré  $\geq 2$  n'est pas nécessaire. Néanmoins, la représentation utilisant  $N_0 = 2$  reste plus intéressante que celle utilisant  $N_0 = 0$ . Pour la fonction  $f^3$ , les représentations utilisent grossièrement le même nombre de degrés de liberté et atteignent une erreur similaire pour  $\eta \geq 10^{-3}$ . Au dessous de cette valeur, la représentation avec  $N_0 = 0$  est la plus intéressante en termes de nombre de degrés de liberté. Cela est dû au fait que cette fonction est essentiellement constante par morceaux.

Les tests étudiés ci-dessus sont représentatifs de différents cas de figures que l'on est susceptible de rencontrer dans les systèmes hyperboliques stochastiques. La première observation que l'on peut faire est que si l'on souhaite utiliser des degrés  $N_0 > 0$ , il convient de considérer des valeurs de  $\eta$  pas trop petites. Dans ce cas, il est intéressant de prendre un degré polynomial  $N_0 > 0$  pour les fonctions à fortes

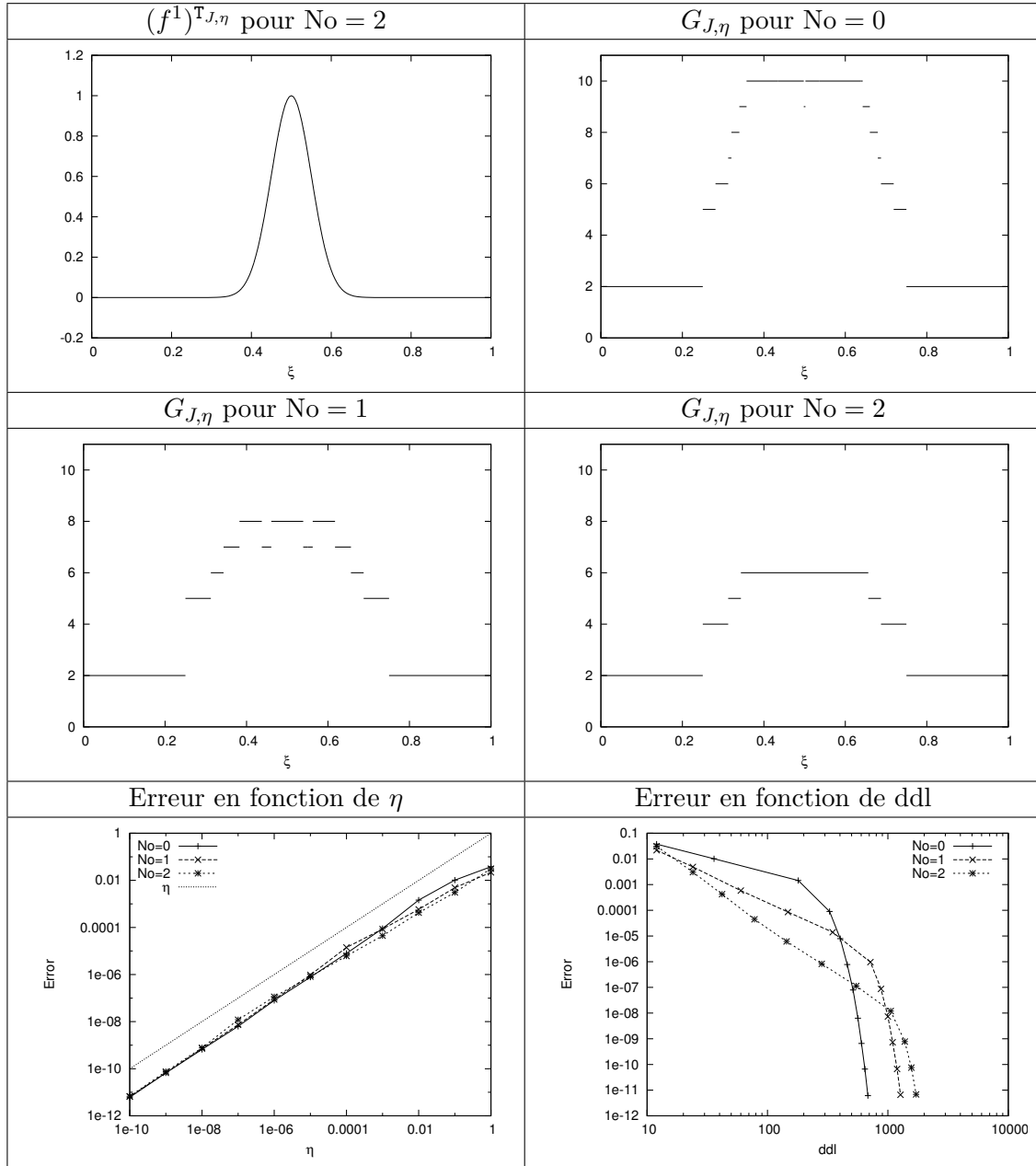


FIGURE 4.8 – Projection seuillée de  $f^1$  et grilles adaptatives correspondantes pour  $No \in \{1, 2, 3\}$ . Calculs avec  $J = 10$  et  $\eta = 10^{-3}$ . Ligne du bas : Erreur d'approximation  $\|(f^1)^{T_J} - (f^1)^{T_{J,\eta}}\|_{L^2(\Xi)}$  en fonction de  $\eta$  (à gauche) et du nombre de degrés de liberté (à droite) pour  $No \in \{0, 1, 2\}$ .



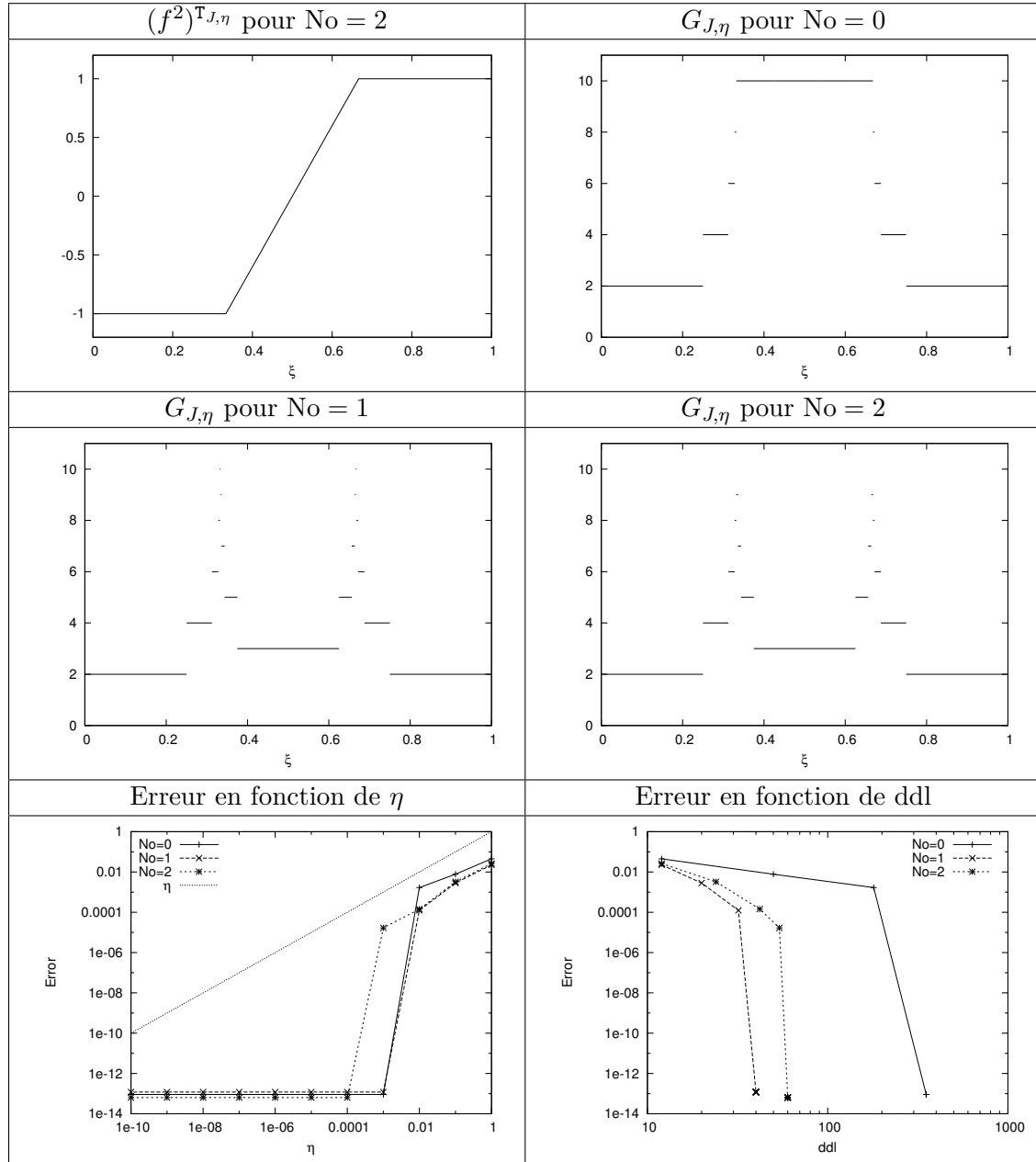


FIGURE 4.9 – Projection seuillée de  $f^2$  et grilles adaptatives correspondantes pour  $No \in \{1, 2, 3\}$ . Calculs avec  $J = 10$  et  $\eta = 10^{-3}$ . Ligne du bas : Erreur d'approximation  $\|(f^2)^{T_J} - (f^2)^{T_{J,\eta}}\|_{L^2(\Xi)}$  en fonction de  $\eta$  (à gauche) et du nombre de degrés de liberté (à droite) pour  $No \in \{0, 1, 2\}$ .

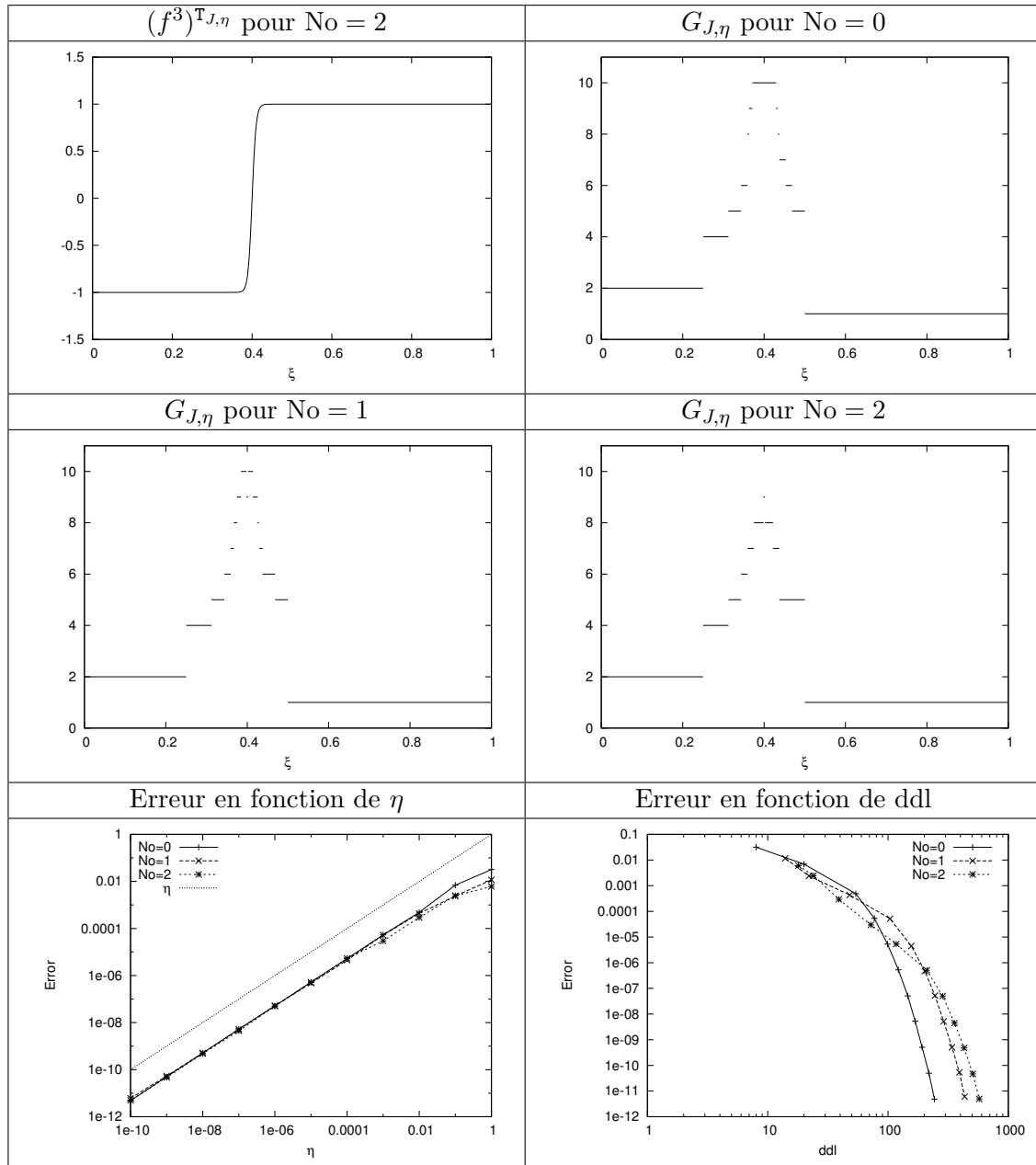


FIGURE 4.10 – Projection seillée de  $f^3$  et grilles adaptatives correspondantes pour  $No \in \{1, 2, 3\}$ . Calculs avec  $J = 10$  et  $\eta = 10^{-3}$ . Ligne du bas : Erreur d'approximation  $\|(f^3)^{T_J} - (f^3)^{T_{J,\eta}}\|_{L^2(\Xi)}$  en fonction de  $\eta$  (à gauche) et du nombre de degrés de liberté (à droite) pour  $No \in \{0, 1, 2\}$ .

pentés et les “détentes” en termes de degrés de liberté. Pour les fonctions constantes par morceaux, les différents ordres polynomiaux sont équivalents. Néanmoins, on rappelle que le schéma de Roe diffuse la solution de Galerkin sur quelques cellules stochastiques, de sorte que la solution obtenue ne sera jamais vraiment discontinue. Ces observations plaident par conséquent pour l'utilisation d'un degré polynomial  $N_0 > 0$  et des valeurs de  $\eta$  pas trop petites.

## 4.4 Synthèse

L'AMR-MW d'Alpert présentée dans ce chapitre va servir de cadre pour la construction du solveur de Roe stochastique adaptatif au chapitre suivant. On résume ici les principaux outils dont on se servira par la suite.

1. Pour un niveau de résolution maximum  $J$  et  $U \in L^2(\Xi)$ , la projection  $L^2(\Xi)$ -orthogonale de  $U$  sur  $S_J^{N_0}$ , notée  $U^J$ , admet deux développements stochastiques équivalents : le premier est le développement SE, cf (4.28) avec  $j = J$ , correspondant au développement de  $U^J$  en coefficients d'échelle  $J$  sur les bases de polynômes de Legendre par morceaux sur les éléments stochastiques ; le second est le développement MW, cf (4.29) avec  $j = J$ , correspondant au développement de  $U^J$  en coefficients d'échelle 0 sur la base de polynômes de Legendre sur  $[0, 1]$  et en coefficients de détails jusqu'au niveau  $J - 1$  sur les multi-ondelettes de Le Maître. Ces développements sont représentés sur une structure d'arbre binaire complet  $T_J$  de profondeur  $J$ , chaque nœud de l'arbre supportant les coefficients d'échelle et les détails.
2. Les multi-ondelettes ont leurs  $(N_0 + 1)$  premiers moments qui s'annulent, cf (4.22), ce qui implique une décroissance des détails avec le niveau de résolution si la fonction  $U$  est régulière. Pour un critère de seuillage fixé  $\eta > 0$ , la projection seuillée  $U^{J,\eta}$  correspond au développement MW de  $U^J$  pour lequel on a gardé l'ensemble des détails significatifs appartenant à l'ensemble  $\mathcal{D}_{J,\eta}$ . L'erreur d'approximation commise par le seuillage est en  $O(\eta)$ .
3. On souhaite représenter les détails significatifs de  $U^{J,\eta}$  sur un arbre binaire de profondeur  $\leq J$ . Pour cela, on a recours à l'un ou l'autre des algorithmes présentés à la section 4.3.2. Dans le chapitre suivant, on pourra utiliser ces algorithmes pour initialiser la solution. On utilisera également ce type d'algorithmes pour seuiller et enrichir les arbres représentant sur chaque cellule spatiale la solution au cours du temps.
4. Enfin, dans le cas de l'AMR-MW d'Alpert, le calcul de la transformation multi-échelles  $\mathcal{M}_J$  n'est pas nécessaire. Les coefficients d'échelle et les détails s'obtiennent localement en utilisant les matrices des coefficients de filtre (4.42) et (4.43).

# Solveur de Roe adaptatif 1D

## Sommaire

<b>5.1 Aspects algorithmiques</b> . . . . .	<b>126</b>
5.1.1 Présentation générale du solveur de Roe adaptatif . . . . .	126
5.1.2 Union d'arbres binaires . . . . .	128
5.1.3 Opérateurs de restriction et de prédiction . . . . .	128
5.1.4 Élagage et enrichissement d'arbres . . . . .	130
<b>5.2 Application aux équations d'Euler</b> . . . . .	<b>132</b>

On s'intéresse à des problèmes où un choc apparaît presque sûrement en temps fini. Ce choc reste localisé aussi bien dans le domaine déterministe  $\Omega \times [0, T]$  pour chaque réalisation des paramètres  $\xi \in \Xi = [0, 1]$ , que dans le domaine stochastique  $\Xi$  pour un point donné  $(x, t) \in \Omega \times [0, T]$ . En d'autres termes, la solution est régulière presque partout sur  $\Omega \times [0, T] \times \Xi$ . Cette observation plaide pour l'utilisation de stratégies adaptatives où l'effort de calcul est concentré au voisinage des chocs, alors qu'une discrétisation plus grossière est utilisée là où la solution est continue. Dans ce mémoire, on considère uniquement une adaptation de la discrétisation stochastique en utilisant un maillage spatial fixé (le pas de temps satisfaisant une condition CFL).

L'objectif de ce chapitre est d'étendre le solveur de Roe stochastique présenté au chapitre 2. Le but est de développer un solveur de Roe adaptatif utilisant une discrétisation stochastique adaptative qui dépend de la variable spatiale  $x$  et du temps  $t$ . Grâce aux outils d'AMR introduits au chapitre précédent, cela revient à faire dépendre de  $(x, t)$  les arbres  $\mathbf{T}$  définissant l'espace stochastique d'approximation  $S^{\text{No}}(\mathbf{T})$ . Dans le contexte d'une approximation VF, on note  $\mathbf{T}_i^n$  l'arbre associé à la cellule  $i$  du maillage spatial au temps discret  $t^n$ , de sorte que la solution stochastique approchée (solution de Galerkin) en la cellule  $i$  et au temps  $t^n$  a pour développements SE et MW

$$U_i^n(\xi) = \sum_{\mathbf{k} \in \text{SE}(\mathbf{T}_i^n)} (u_i^n)_{\mathbf{k}} \Phi_{\mathbf{k}}(\xi) \quad (5.1)$$

$$= \sum_{\mathbf{k} \in \text{MW}(\mathbf{T}_i^n)} (d_i^n)_{\mathbf{k}} \Psi_{\mathbf{k}}(\xi) \in S^{\text{No}}(\mathbf{T}_i^n). \quad (5.2)$$

Pour alléger les notations, on supprime par la suite l'indice No de l'espace stochastique d'approximation puisque l'ordre polynomial est fixé. On note donc  $S(\mathbf{T})$  l'espace stochastique d'approximation associé à un arbre binaire général  $\mathbf{T}$ .

## 5.1 Aspects algorithmiques

Le solveur de Roe adaptatif est une extension du solveur de Roe stochastique présenté au chapitre 2 à des solutions définies sur des espaces d'approximation adaptés, *i.e.*, l'intégration de la solution de Galerkin au cours du temps a lieu sur le vecteur des coefficients d'échelle correspondant aux détails significatifs. Pour cela, on a recours à trois ingrédients essentiels :

- un opérateur d'union d'arbres binaires pour le calcul du flux numérique de Roe aux interfaces entre les couples d'états gauche et droit qui peuvent être définis sur des arbres binaires différents (et donc des espaces d'approximation stochastique différents), ainsi que pour l'intégration en temps sur une cellule de discrétisation spatiale ;
- un opérateur d'enrichissement, qui reprend l'idée de Harten [37] d'enrichir l'arbre et donc l'espace d'approximation stochastique sur lequel on définit la solution avant le calcul des flux pour l'intégration en temps. Ce choix est motivé par le fait que de nouvelles structures (forte variation, singularité, discontinuité, ...) peuvent apparaître dans la solution puisque les fonctions de flux sont souvent non linéaires, de sorte que des espaces stochastiques d'approximation plus grands que ceux utilisés pour la solution peuvent être nécessaires pour l'évaluation précise des flux ;
- un opérateur d'élagage, basé sur l'opérateur de seuillage défini dans le chapitre précédent, utilisé à la fin de chaque pas de temps pour éviter l'apparition d'arbres trop riches là où la solution est régulière en  $\xi$  et peut donc être approchée par un espace d'approximation stochastique de faible dimension.

Les procédures d'enrichissement et d'élagage vont nécessiter deux outils pour naviguer entre des arbres binaires : un opérateur de restriction et un opérateur de prédiction. L'opérateur de restriction consiste à restreindre une fonction  $U^T \in S(T)$  sur un espace d'approximation stochastique plus petit  $S(T^-)$  où  $T^- \subset T$ . On note  $\mathcal{R}_{T \downarrow T^-}$  cet opérateur. Par ailleurs, l'opérateur de prédiction consiste à étendre une fonction  $U^T \in S(T)$  sur un espace d'approximation stochastique plus grand  $S(T^+)$  où  $T \subset T^+$ . On note  $\mathcal{P}_{T \uparrow T^+}$  cet opérateur.

### 5.1.1 Présentation générale du solveur de Roe adaptatif

Partant d'une donnée initiale consistant, pour chaque cellule spatiale, en un arbre  $T_i^0$  de détails significatifs (construit par exemple en utilisant l'algorithme 4.1) et l'approximation correspondante de la condition initiale  $U_i^0 \in S(T_i^0)$ , l'algorithme est le suivant :

I) Boucles sur les interfaces : (calcul du flux)

1. Définir l'arbre union de l'interface :  $T_{i-1/2}^n = T_{i-1}^n \cup T_i^n$
2. Définir l'arbre enrichi de l'interface :  $T_{i-1/2}^* = \mathcal{E} \left( T_{i-1/2}^n \right)$
3. Prédire les états gauche et droit sur  $S(T_{i-1/2}^*)$  :

$$\mathcal{P}_{\mathbb{T}_{i-1}^n \uparrow \mathbb{T}_{i-1/2}^*} U_{i-1}^n \text{ et } \mathcal{P}_{\mathbb{T}_i^n \uparrow \mathbb{T}_{i-1/2}^*} U_i^n$$

4. Pour chaque feuille  $1 \in \mathcal{L}(\mathbb{T}_{i-1/2}^*)$  :
  - (a) Calculer les flux locaux  $f_1(u_{i-1})$  et  $f_1(u_i)$ , où  $u_{i-1}$  et  $u_i$  sont les vecteurs des coefficients d'échelle des états gauches et droits prédits sur  $S(\mathbb{T}_{i-1/2}^*)$ , et calculer la matrice  $[a_{i-1,i}^{\text{Roe}}]_1$
  - (b) Déterminer le polynôme local  $q_{d_{\text{fit}},\{\lambda'_1\}}$
  - (c) Assembler la contribution de la feuille au flux numérique de l'interface  $\varphi_{i-1/2}^{\text{Roe}}$

II) Boucle sur les cellules spatiales : (intégration en temps)

1. Construire l'arbre d'intégration de la cellule :
$$\mathbb{T}_i^* = \mathbb{T}_{i-1/2}^* \cup \mathbb{T}_{i+1/2}^*$$
2. Prédire les flux gauche et droit et l'état initial sur  $S(\mathbb{T}_i^*)$  :  $\mathcal{P}_{\mathbb{T}_{i-1/2}^* \uparrow \mathbb{T}_i^*} \varphi_{i-1/2}$ ,  $\mathcal{P}_{\mathbb{T}_{i+1/2}^* \uparrow \mathbb{T}_i^*} \varphi_{i+1/2}$  et  $\mathcal{P}_{\mathbb{T}_i^n \uparrow \mathbb{T}_i^*} U_i^n$
3. Intégrer pour obtenir  $U_i^* \in S(\mathbb{T}_i^*)$
4. Elaguer l'arbre :  $\mathbb{T}_i^{n+1} = \mathcal{C}(\mathbb{T}_i^*) \subseteq \mathbb{T}_i^*$
5. Restreindre la solution :  $U_i^{n+1} = \mathcal{R}_{\mathbb{T}_i^* \downarrow \mathbb{T}_i^{n+1}} U_i^*$ .

Comme dans la section 4.3, on fixe une profondeur d'arbre maximale, notée  $J$ , qui est la même pour tous les arbres du maillage spatial. Pour le calcul du flux de Roe, on procède comme au Chapitre 2 en appliquant le solveur de Roe indépendamment sur chaque feuille  $1 \in \mathcal{L}(\mathbb{T}_{i-1/2}^*)$ . Plus précisément, en notant  $[a_{i-1,i}^{\text{Roe}}]_1$  la sous-matrice d'ordre  $mP_\pi$  associée à la feuille  $1 \in \mathcal{L}(\mathbb{T}_{i-1/2}^*)$ , on approche sa valeur absolue de la façon suivante :

$$\left| [a_{i-1,i}^{\text{Roe}}]_1 \right| \approx q_{d_{\text{fit}},\{\lambda'_1\}}([a_{i-1,i}^{\text{Roe}}]_1), \quad (5.3)$$

où  $\{\lambda'_1\}$  est le spectre approché de  $[a_{i-1,i}^{\text{Roe}}]_1$  et  $q_{d_{\text{fit}},\{\lambda'_1\}}$  est le polynôme de degré  $\leq d_{\text{fit}}$  minimisant la fonctionnelle quadratique

$$\sum_{\lambda \in \{\lambda'_1\}} \left( |\lambda| - q_{d_{\text{fit}},\{\lambda'_1\}}(\lambda) \right)^2. \quad (5.4)$$

Comme dans la section 2.3.3, on utilise pour  $\{\lambda'_1\}$  l'ensemble des valeurs propres de la matrice Jacobienne stochastique évaluée aux nœuds d'intégration de Gauss sur la feuille 1. Au voisinage de points soniques, ce schéma de Roe nécessite une correction entropique, cf section 2.4.

Enfin, pour assurer la stabilité de l'intégration en temps, le pas de temps  $\Delta t^n$  doit être sélectionné pour satisfaire une condition CFL de la forme (2.41). Dans le contexte adaptatif, la vitesse caractéristique maximale est prise sur l'ensemble du spectre approché (pour chaque feuille de chaque arbre union à chaque interface).

### 5.1.2 Union d'arbres binaires

Pour deux arbres binaires  $T_1$  et  $T_2$ , on définit leur arbre-union  $T_{1\cup 2} := T_1 \cup T_2$  comme l'unique arbre minimal (en terme du nombre de feuilles) tel que pour toute feuille  $1 \in \mathcal{L}(T_{1\cup 2})$ ,

$$\exists ! l_1 \in \mathcal{L}(T_1), \exists ! l_2 \in \mathcal{L}(T_2), S(1) = S(l_1) \cap S(l_2). \quad (5.5)$$

Il est clair que  $S(T_1) \subset S(T_{1\cup 2})$  et  $S(T_2) \subset S(T_{1\cup 2})$ . L'union de deux arbres binaires est illustrée dans la Figure 5.1.

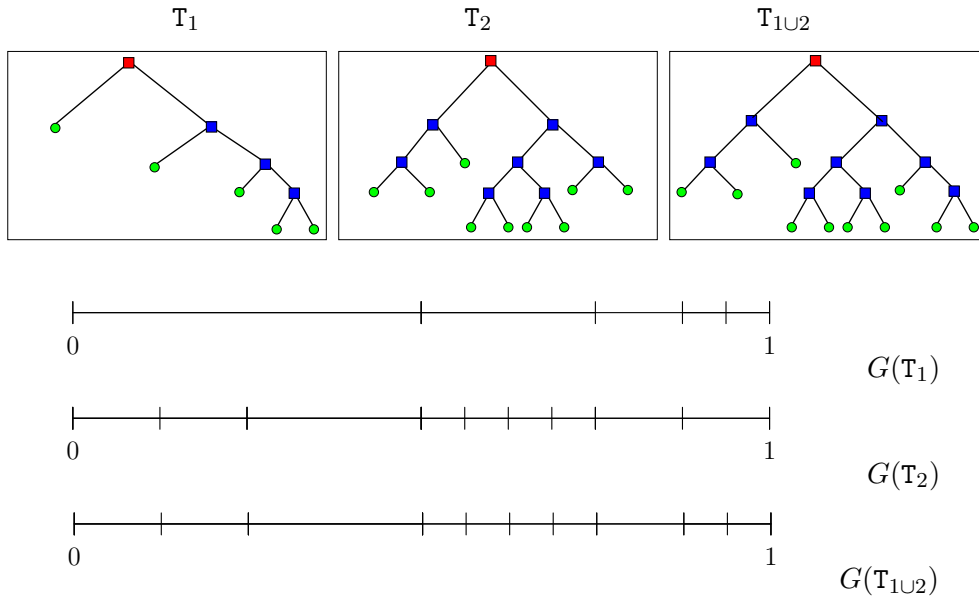


FIGURE 5.1 – Exemple de l'union de deux arbres binaires. Ligne du haut :  $T_1$  (à gauche),  $T_2$  (au centre), et  $T_{1\cup 2}$  (à droite). Lignes du bas : Partitions correspondantes du domaine stochastique.

### 5.1.3 Opérateurs de restriction et de prédiction

#### 5.1.3.1 Opérateur de restriction

Soient deux arbres binaires  $T^- \subset T$  (si bien que  $S(T^-) \subset S(T)$ ) et soit  $U^T \in S(T)$ . On définit la restriction  $\mathcal{R}_{T \downarrow T^-} U^T$  de  $U^T$  dans  $S(T^-)$  comme la projection  $L^2(\Xi)$ -orthogonale de  $U^T$  sur  $S(T^-)$ , *i.e.*,

$$\langle (U^T - \mathcal{R}_{T \downarrow T^-} U^T), \Phi \rangle = 0, \quad \forall \Phi \in S(T^-). \quad (5.6)$$

En termes de détails, l'opération de restriction est directe : les détails conservés de  $\mathcal{R}_{T \downarrow T^-} U^T$  sont les mêmes que ceux de  $U^T$  correspondants. Ainsi, si  $U^T =$

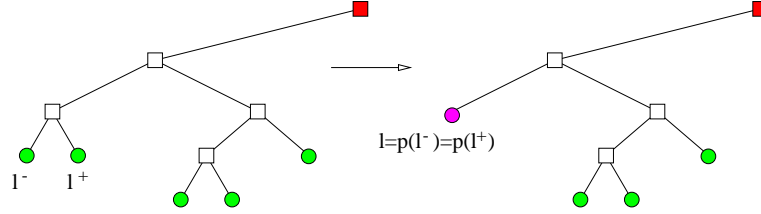


FIGURE 5.2 – Opération de restriction : suppression de deux feuilles sœurs.

$\sum_{\mathbf{n} \in \mathcal{N}^*(\mathbb{T})} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}$ , on a simplement

$$\mathcal{R}_{\mathbb{T} \downarrow \mathbb{T}^-} U^{\mathbb{T}} = \sum_{\mathbf{n} \in \mathcal{N}^*(\mathbb{T}^-)} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}. \quad (5.7)$$

Le calcul des coefficients d'échelle de la restriction s'obtient récursivement. On construit une séquence d'arbres binaires  $\mathbb{T}^{(i)}$  telle que

$$\mathbb{T} = \mathbb{T}^{(0)} \supset \mathbb{T}^{(1)} \supset \dots \supset \mathbb{T}^{(i)} \supset \dots \supset \mathbb{T}^{(l)} = \mathbb{T}^-, \quad (5.8)$$

où deux arbres consécutifs diffèrent seulement d'une génération, *i.e.*, une feuille de  $\mathbb{T}^{(i+1)}$  est soit une feuille soit un nœud avec des enfants feuilles dans  $\mathbb{T}^{(i)}$ . Ainsi, la transition de  $\mathbb{T}^{(i)}$  à  $\mathbb{T}^{(i+1)}$  consiste à supprimer un ensemble de paires de feuilles sœurs pour réduire itérativement l'arbre. Le processus est illustré à la Figure 5.2 dans le cas de la suppression d'une seule paire de feuilles. Dans le cas de la suppression d'une paire de feuilles sœurs (gauche et droite)  $\{1^-, 1^+\} \in \mathcal{L}(\mathbb{T}^{(i)})$ , les coefficients d'échelle de  $U^{\mathbb{T}^{(i)}}$  associés à la feuille  $\mathbf{l} = \mathbf{p}(1^-) = \mathbf{p}(1^+) \in \mathcal{L}(\mathbb{T}^{(i+1)})$  sont

$$u_{\mathbf{l}} = M^{\Phi, -1} u_{1^-} + M^{\Phi, 0} u_{1^+}, \quad (5.9)$$

avec les matrices des coefficients de filtre  $M^{\Phi, s}, s \in \{-1, 0\}$ , définies à la section 4.1.2.3.

### 5.1.3.2 Opérateur de prédiction

Soient deux arbres binaires  $\mathbb{T} \subset \mathbb{T}^+$  (si bien que  $S(\mathbb{T}) \subset S(\mathbb{T}^+)$ ) et soit  $U^{\mathbb{T}} \in S(\mathbb{T})$ . On définit la prédiction  $\mathcal{P}_{\mathbb{T} \uparrow \mathbb{T}^+} U^{\mathbb{T}}$  de  $U^{\mathbb{T}}$  dans  $S(\mathbb{T}^+)$  comme l'injection de  $U^{\mathbb{T}}$  dans  $S(\mathbb{T}^+)$ . Comme pour l'opérateur de restriction, les détails restent inchangés pour les nœuds  $\mathbf{n} \in \mathcal{N}^*(\mathbb{T})$  et on met à zéro les détails des autres nœuds. Ainsi, si  $U^{\mathbb{T}} = \sum_{\mathbf{n} \in \mathcal{N}^*(\mathbb{T})} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}$ , on a simplement

$$\mathcal{P}_{\mathbb{T} \uparrow \mathbb{T}^+} U^{\mathbb{T}} = \sum_{\mathbf{n} \in \mathcal{N}^*(\mathbb{T}^+)} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha}^+ \Psi_{\mathbf{n},\alpha}, \quad (5.10)$$

avec pour tout  $\mathbf{n} \in \mathcal{N}^*(\mathbb{T}^+)$  et tout  $\alpha \in \mathcal{P}_\pi$ ,

$$d_{\mathbf{n},\alpha}^+ = \begin{cases} d_{\mathbf{n},\alpha}, & \text{si } \mathbf{n} \in \mathcal{N}^*(\mathbb{T}), \\ 0, & \text{sinon.} \end{cases} \quad (5.11)$$



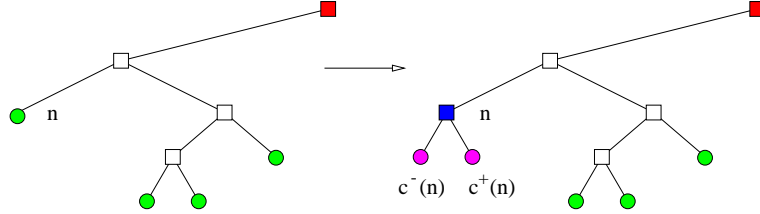


FIGURE 5.3 – Opérateur de prédiction : création des enfants d'un nœud feuille.

Pour les coefficients d'échelle de la prédiction, on peut à nouveau procéder itérativement en utilisant une série d'arbres intermédiaires de plus en plus grands, différenciant entre eux d'une seule génération. Cette fois, l'opération élémentaire consiste à ajouter des enfants à certaines feuilles de l'arbre, comme illustré sur la Figure 5.3. Les coefficients d'échelle associés aux feuilles sont

$$u_{c^-(n)} = (M^{\Phi, -1})^T u_n, \quad u_{c^+(n)} = (M^{\Phi, 0})^T u_n. \quad (5.12)$$

Finalement, pour deux arbres  $T \subset T^+$ , on observe que  $\mathcal{R}_{T+\downarrow T} \circ \mathcal{P}_{T\uparrow T^+} = I_{S(T)}$ , où  $I_{S(T)}$  est l'opérateur identité dans  $S(T)$ , alors qu'en général  $\mathcal{P}_{T\uparrow T^+} \circ \mathcal{R}_{T+\downarrow T} \neq I_{S(T^+)}$ . Une remarque intéressante est qu'on peut d'ailleurs utiliser l'opération  $(I_{S(T_1)} - \mathcal{P}_{T_0\uparrow T_1} \circ \mathcal{R}_{T_1\downarrow T_0})$  sur la grille  $G_1$  pour générer les multi-ondelettes mères de Le Maître  $\{\Psi_\alpha^{No}\}_{\alpha \in \mathcal{P}_\pi}$ . Il suffit pour cela d'appliquer cette opération à l'ensemble des  $P_\pi$  fonctions définies sur  $\Xi = [0, 1]$  pour  $\alpha \in \mathcal{P}_\pi$  par

$$f^\alpha(\xi) = \begin{cases} \Phi_\alpha^{No}(2\xi), & \text{si } 0 \leq \xi \leq 1/2, \\ 0, & \text{sinon,} \end{cases} \quad (5.13)$$

puis d'orthonormaliser les fonctions obtenues.

## 5.1.4 Élagage et enrichissement d'arbres

### 5.1.4.1 Procédure d'élagage

La procédure d'élagage est appliquée après l'intégration en temps (donc en  $t^{n+1}$ ), lorsque la solution sur la cellule  $i$  est connue sur un arbre  $T_i^*$ . L'objectif est de définir le sous-arbre  $T_i^{n+1} \subseteq T_i^*$  associé aux détails significatifs de la solution au temps  $t^{n+1}$ . On procède comme suit. On construit une séquence d'arbres imbriqués, obtenus par la suppression de paires de feuilles sœurs d'un arbre à l'autre. On supprime un couple de feuilles sœurs  $(1^-, 1^+)$  si leur parent  $p(1^-) \in \mathcal{D}_{|T|, \eta}$ , i.e. si le critère

$$\|d_{p(1^-)}\|_{l^2} = \left( \sum_{\alpha \in \mathcal{P}_\pi} (d_{p(1^-), \alpha})^2 \right)^{1/2} \leq \epsilon_J(\eta) \quad (5.14)$$

est satisfait. La procédure s'arrête lorsqu'on a supprimé tous les couples de sœurs possibles. Cette procédure est similaire à l'algorithme 4.2 proposé dans le chapitre précédent pour la construction de l'arbre  $T_{J, \eta}^\sharp$ .

Si la taille  $m$  du système hyperbolique stochastique est supérieure à 1,  $d_{\mathbf{n},\alpha}$  est un vecteur de taille  $m$ , i.e.  $d_{\mathbf{n},\alpha} := (d_{\mathbf{n},\alpha}^k)_{k \in \{1, \dots, m\}}$ , pour tout  $\mathbf{n} \in \mathcal{N}^*(\mathbb{T}_i^*)$  et pour tout  $\alpha \in \mathbb{P}_\pi$ . Le test d'élagage (5.14) est alors réalisé sous la forme suivante :

$$\|d_{\mathbb{P}(1^-)}\|_{(l^2)^m} = \left( \sum_{k=1}^m \sum_{\alpha \in \mathbb{P}_\pi} (d_{\mathbb{P}(1^-),\alpha}^k)^2 \right)^{1/2} \leq \epsilon_J(\eta). \quad (5.15)$$

Par ailleurs, si le système n'est pas adimensionné, il convient de normaliser chaque vecteur de détails  $k$ ,  $k \in \{1, \dots, m\}$ , par exemple par la norme  $L^2$  de la  $k$ -ème composante de la solution.

#### 5.1.4.2 Procédure d'enrichissement

La procédure d'enrichissement est appliquée avant l'intégration en temps. Son objectif est de construire un arbre capable de représenter la solution à la fin du pas de temps. L'idée est de calculer le flux à l'interface dans un espace d'approximation stochastique plus grand que  $S(\mathbb{T}_{i-1/2}^n)$ . Cela revient à enrichir l'arbre union, en divisant certaines de ses feuilles. L'idée intuitive introduite par Harten est que l'ensemble des détails significatifs de la solution numérique évolue "lentement" d'un pas de temps à l'autre, pourvu que le schéma numérique utilisé vérifie une CFL. Plus précisément, soit  $\mathbb{T}_J$  l'arbre binaire complet de profondeur  $J$ , et soient  $\mathbb{T}_{J,\eta}^n$  et  $\mathbb{T}_{J,\eta}^{n+1}$  les arbres des détails significatifs de la solution aux temps  $t^n$  et  $t^{n+1}$  (on omet ici l'indice de la cellule spatiale). On peut résumer l'idée de Harten de la manière suivante :

**Heuristique 5.1.** (*Harten*) On peut agrandir l'arbre  $\mathbb{T}_{J,\eta}^n$  en un arbre  $\mathbb{T}_{J,\eta}^*$  contenant à la fois  $\mathbb{T}_{J,\eta}^n$  et  $\mathbb{T}_{J,\eta}^{n+1}$  de sorte que, si  $(U^{n+1})^{\mathbb{T}_J} = E_{J,\Delta t}(U^n)^{\mathbb{T}_J}$ , avec  $E_{J,\Delta t}$  l'opérateur d'intégration en temps sur l'arbre  $\mathbb{T}_J$ , on a

$$\|(U^n)^{\mathbb{T}_J} - (U^n)^{\mathbb{T}_{J,\eta}^*}\|_{L^2(\Xi)} \leq C\eta \text{ et } \|(U^{n+1})^{\mathbb{T}_J} - (U^{n+1})^{\mathbb{T}_{J,\eta}^*}\|_{L^2(\Xi)} \leq C\eta, \quad (5.16)$$

i.e. l'arbre  $\mathbb{T}_{J,\eta}^*$  est adapté pour décrire la solution à la fois aux temps  $t^n$  et  $t^{n+1}$ .

La stratégie d'enrichissement proposée par Harten consiste à faire grandir l'arbre en fonction de la taille de ses détails, en prenant en compte leur propriété de décroissance [37]. La stratégie que l'on adopte s'en inspire. L'algorithme pour trouver l'arbre enrichi  $\mathbb{T}_{i-1/2}^*$  est le suivant :

#### Algorithme 5.1.

*%Initialisation :*

$$\mathbb{T}_{i-1/2}^* = \mathbb{T}_{i-1/2}^n$$

**pour**  $1 \in \mathcal{L}(\mathbb{T}_{i-1/2}^*)$  **faire**

*%ne pas tester les feuilles dont on a élagué les enfants*

**si** feuille non issue d'un élagage au temps  $t^n$  **alors**

**si**  $\|d_{\mathbb{P}(1)}\|_{l^2} > 2^{(N_0+1)}\epsilon_J(\eta)$  **alors**

```

%créer des feuilles
T*_{i-1/2} ← T*_{i-1/2} ∪ {c^-(1), c^+(1)}
fin si
fin si
fin pour

```

L'idée intuitive est la suivante : on part de l'arbre des détails significatifs  $T_{i-1/2}^n$  et on parcourt les feuilles  $\mathbf{1} \in \mathcal{L}(T_{i-1/2}^n)$ . On rappelle que cet algorithme s'applique immédiatement avec l'élagage de l'arbre à la fin du pas de temps précédent. Aussi, si une feuille  $\mathbf{1}$  est issue d'un élagage à l'étape d'élagage précédente, cela signifie que les détails qui lui étaient associés n'étaient pas significatifs. Dans ce cas, ces feuilles ne feront pas l'objet d'un enrichissement. Sinon, on cherche à estimer le détail de ce nœud. S'inspirant de la décroissance des détails (4.33), on fait l'hypothèse qu'on peut estimer les détails  $d_{\mathbf{1},\alpha}$  par  $d_{\mathbf{p}(\mathbf{1}),\alpha} \times 2^{-(N_o+1)}$ . Par conséquent, on choisit d'enrichir une feuille  $\mathbf{1} \in \mathcal{L}(T_{i-1/2}^*)$  si le critère

$$\|d_{\mathbf{p}(\mathbf{1})}\|_{l^2} = \left( \sum_{\alpha \in P_\pi} (d_{\mathbf{p}(\mathbf{1}),\alpha})^2 \right)^{1/2} > 2^{(N_o+1)} \epsilon_J(\eta) \quad (5.17)$$

est satisfait.

De la même façon que pour le critère d'élagage, si la taille du système stochastique  $m$  est supérieur à 1, le test d'enrichissement (5.17) est choisi sous la forme suivante :

$$\|d_{\mathbf{p}(\mathbf{1})}\|_{(l^2)^m} = \left( \sum_{k=1}^m \sum_{\alpha \in P_\pi} (d_{\mathbf{p}(\mathbf{1}),\alpha}^k)^2 \right)^{1/2} > 2^{(N_o+1)} \epsilon_J(\eta), \quad (5.18)$$

et l'on normalise les  $k$  vecteurs de détails si le système n'est pas adimensionné.

## 5.2 Application aux équations d'Euler

Dans cette section, on teste le solveur de Roe adaptatif sur le cas test de la section 3.2.1. On considère donc le problème du tube à choc de Sod unidimensionnel, où l'écoulement d'un gaz parfait est régi par les équations d'Euler sans point sonique. L'incertitude porte sur le coefficient adiabatique, qui est paramétré par une variable aléatoire uniforme sous la forme (3.18) et les conditions initiales sont données par (3.22). On utilise  $N_c = 200$  cellules pour la discrétisation du domaine spatial, un ordre polynomial  $N_o = 2$  et un degré  $d_{\text{fit}} = 8$  pour l'approximation polynomiale des matrices de décentrement. Enfin, on contraint les arbres de sorte que  $1 \leq |\mathbf{T}| \leq 10$  et on utilise les tests d'élagage (5.15) et d'enrichissement (5.18).

Dans un premier temps, on fixe le paramètre de seuillage  $\eta = 10^{-3}$  et on calcule la solution de Galerkin jusqu'à  $t = 5$ . Pour vérifier la validité du développement stochastique et l'efficacité du solveur de Roe adaptatif, on montre dans la figure 5.4 une reconstruction de la densité de Galerkin à divers instants. Sur le plan  $(x, \xi)$ , on représente la profondeur des feuilles  $|\mathbf{1}|$ . Ces graphes sont à comparer avec ceux

de la figure 3.27, où la solution de Galerkin était calculée avec une discrétisation stochastique uniforme de niveau de résolution  $Nr = 3$  et d'ordre polynomial  $No = 2$ . On observe que les solutions sont similaires et que la méthode adaptative proposée capture donc correctement la dynamique des équations d'Euler. De plus, la résolution stochastique est adaptée en espace et en temps à la régularité locale de la solution dans le domaine stochastique, de sorte qu'une discrétisation plus fine est observée au voisinage du choc. Cette adaptation permet ainsi de contrôler les oscillations non physiques déclenchées par le phénomène de Gibbs qui avaient été observées dans le cas de la discrétisation uniforme avec  $Nr = 3$ . En revanche, pour les points  $(x, \xi)$  pas trop près de la discontinuité, une discrétisation grossière de l'espace stochastique est suffisante. La figure 5.5 présente la distribution du nombre de feuilles en espace et en temps. On remarque que le nombre maximum de feuilles est atteint le long du chemin de l'onde de choc, les valeurs maximales correspondant aux instants où l'onde de choc est réfléchi sur les parois du tube. En comparant la figure 5.5 à la figure 3.26, on remarque que les deux tracés ont un comportement identique, indiquant que le domaine stochastique est davantage raffiné aux endroits où la solution possède un écart type important, *i.e.*, lorsque la solution comporte de fortes variations en  $\xi$ .

Pour observer le comportement de l'erreur stochastique en fonction de  $\eta$ , on définit la mesure d'erreur sur le vecteur d'état  $U(x, t, \xi) = (\rho(x, t, \xi), q(x, t, \xi), E(x, t, \xi))$  par

$$\epsilon_h^2(x_i, t^n) := \frac{1}{M} \sum_{j=1}^M \|(U_i^n)^{T_i^n}(\xi^{(j)}) - U^h(x_i, t^n, \xi^{(j)})\|_{l^2}^2, \quad (5.19)$$

où  $(U_i^n)^{T_i^n}(\xi^{(j)})$  et  $U^h(x_i, t^n, \xi^{(j)})$  sont évaluées pour  $\xi^{(j)}$  appartenant à un échantillon de réalisations respectivement à partir du développement stochastique de la solution de Galerkin sur l'arbre correspondant à la cellule de centre  $x_i$  et au temps discret  $t^n$  et en résolvant le problème d'Euler déterministe discret. On utilise un échantillon de taille  $M = 10^3$ .

Les graphes de la figure 5.6 présentent la distribution spatiale de l'erreur stochastique  $\epsilon_h^2(x_i, t^n)$  aux instants  $t^n = 0.2$  (à gauche) et  $t^n = 0.5$  (à droite) pour la solution de Galerkin adaptative pour différentes valeurs du paramètre de seuillage  $\eta$  et pour la solution de Galerkin obtenue avec une discrétisation stochastique uniforme de niveau de résolution  $Nr = J$  et d'ordre polynomial  $No = 2$ . Pour cette simulation, on calcule la solution pour les deux méthodes avec un pas de temps fixé  $\Delta t$ ; pour simplifier, ce pas de temps est pris égal au minimum de tous les pas de temps obtenus avec le solveur de Roe adaptatif. Le graphe de gauche correspondant à l'instant  $t^n = 0.2$  est à comparer avec le graphe de gauche de la figure 3.30, où la même erreur était représentée pour une discrétisation stochastique uniforme de niveau de résolution  $Nr = 3$  et d'ordre polynomial  $No = 2$ . À cet instant, les ondes ne se sont pas encore réfléchies sur les parois du tube et l'incertitude ne s'est pas encore propagée sur tout le domaine spatial. On distingue trois zones d'erreur correspondant aux voisinages des trois ondes, l'erreur atteignant son maximum au

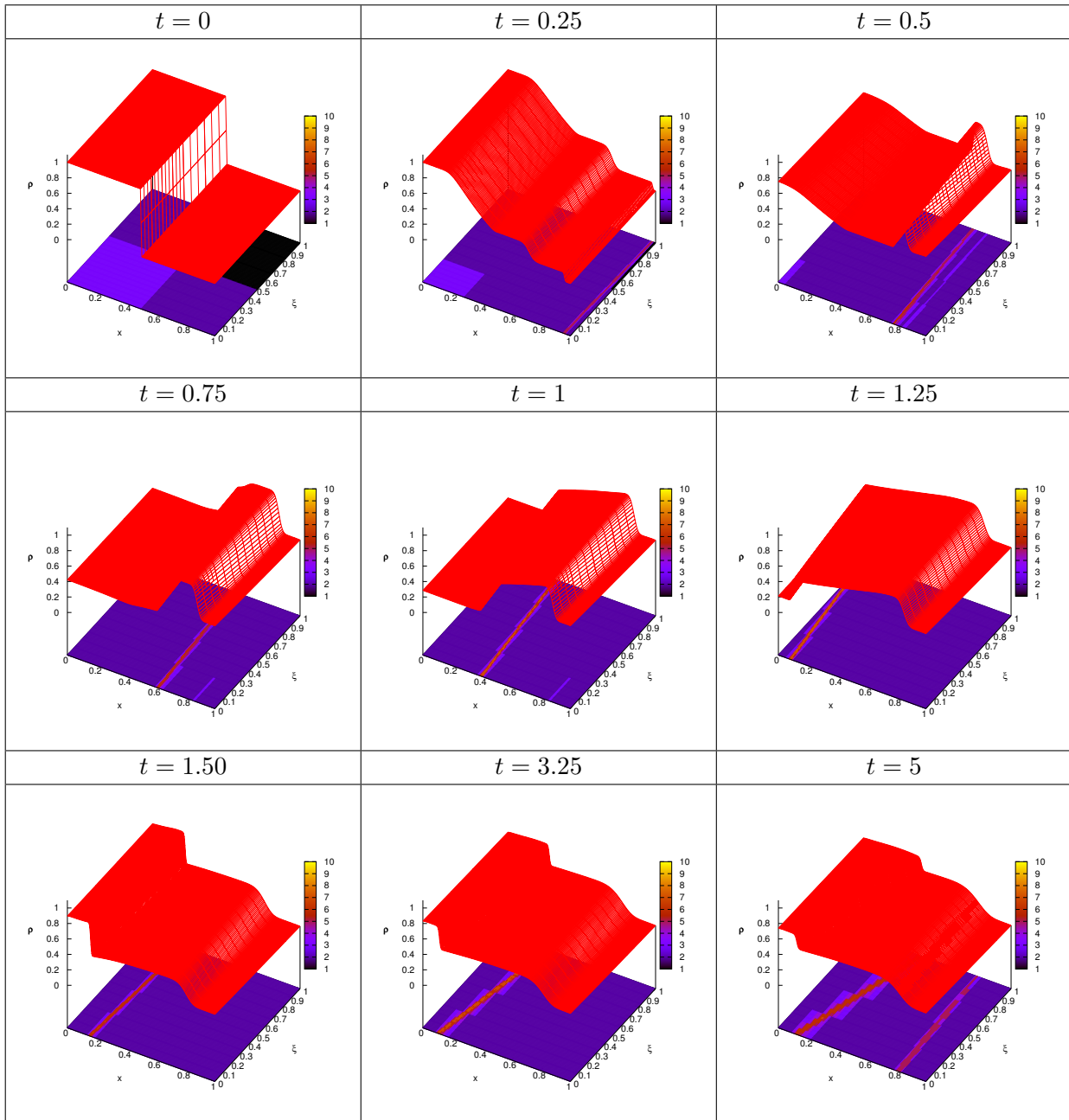


FIGURE 5.4 – Cas test Euler : Reconstruction de la densité de Galerkin à différents instants.

voisinage du choc. Lorsque l'on diminue le critère de seuillage  $\eta$ , l'erreur locale décroît puis finit par stagner à une valeur proche de l'erreur obtenue pour la résolution uniforme avec  $Nr = J$ , du fait du niveau de résolution maximal autorisé. Sur le graphe de droite, l'erreur stochastique est présentée à l'instant  $t^n = 0.5$ , pour lequel l'onde de choc vient de se réfléchir sur la paroi droite du tube. On distingue à nouveau différentes zones d'erreurs correspondant aux voisinages des trois ondes,

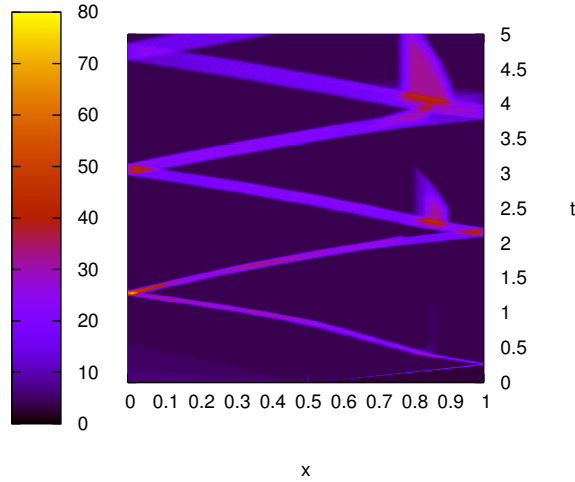


FIGURE 5.5 – Cas test Euler : Distribution espace-temps du nombre de feuilles.

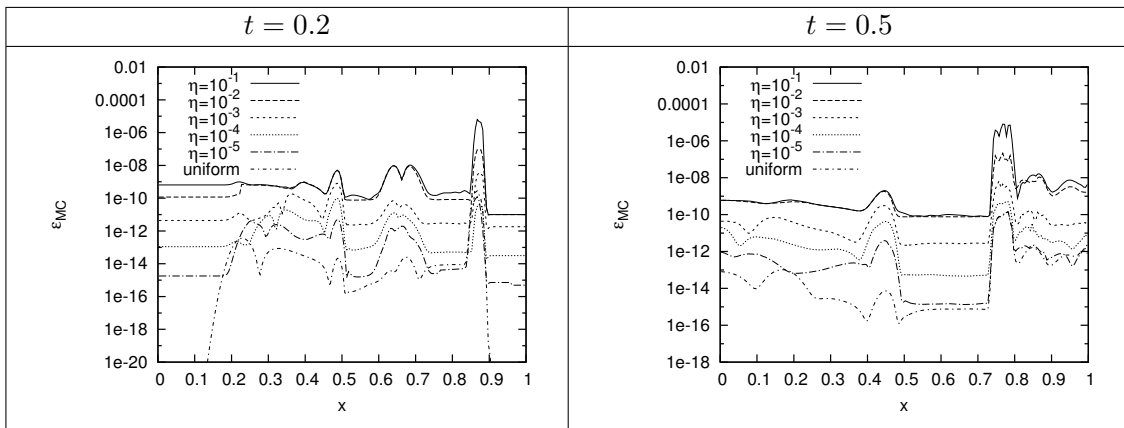


FIGURE 5.6 – Cas test Euler : Erreur stochastique  $\epsilon_h(x_i, t^n)$  pour  $\eta = 10^{-k}$ ,  $k \in \{1, 2, 3, 4, 5\}$  aux instants  $t^n = 0.2$  (à gauche) et  $t^n = 0.5$  (à droite).

l'erreur maximale étant à nouveau atteinte au voisinage du choc. Le comportement de l'erreur est similaire au temps précédent : celle-ci diminue avec le paramètre de seuillage  $\eta$  et finit par stagner à une valeur proche de l'erreur obtenue pour la résolution uniforme avec  $N_r = J$ . Différentes contributions interviennent dans l'erreur stochastique, d'une part le choix d'un niveau de résolution maximum  $J = 10$ , qui conduit à négliger tous les détails de niveau de résolution supérieur ou égal à 10, et d'autre part, les erreurs d'approximation pseudo-spectrales commises sur les flux. L'étude détaillée de ces différentes contributions est en cours d'analyse.

Finalement, on considère l'erreur stochastique intégrée sur le domaine spatial

$$\epsilon_h^2(t^n) := \frac{\Delta x}{M} \sum_{i=1}^{N_c} \sum_{j=1}^M \|(U_i^n)^T(\xi^{(j)}) - U^h(x_i, t^n, \xi^{(j)})\|_{l^2}^2. \quad (5.20)$$

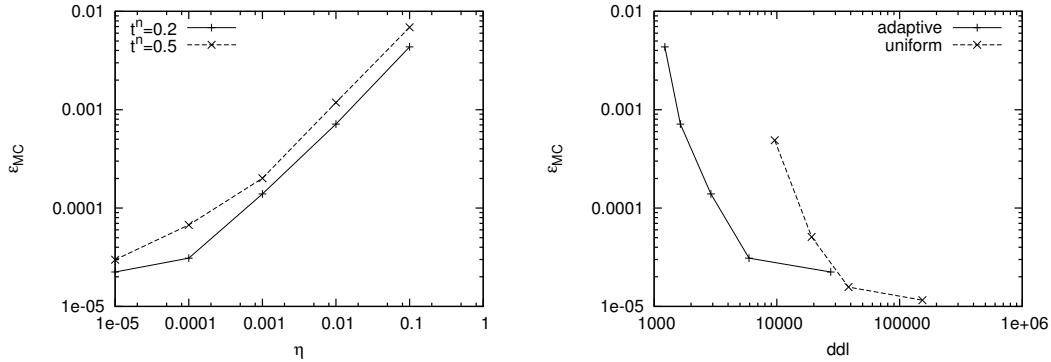


FIGURE 5.7 – Cas test Euler : À droite : Erreur stochastique  $\epsilon_h^2(t^n)$  pour différents  $\eta$  aux instants  $t^n = 0.2$  et  $t^n = 0.5$ . À gauche : Erreur stochastique  $\epsilon_h^2(t^n = 0.2)$  en fonction du nombre de degrés de liberté pour différents  $\eta$  et  $Nr$ .

Cette mesure d'erreur est représentée dans la figure 5.7 sur le graphe de gauche aux instants  $t^n = 0.2$  et  $t^n = 0.5$  pour la solution de Galerkin adaptative en fonction du paramètre de seuillage  $\eta$ . Sur le graphe de droite, on compare la méthode adaptative à la méthode non adaptative en représentant  $\epsilon_h^2(t^n = 0.2)$  en fonction du nombre de degrés de liberté pour  $\eta = 10^{-k}$ ,  $k \in \{1, 2, 3, 4, 5\}$  à  $J = 10$  fixé et pour  $Nr \in \{4, 5, 6, 8\}$ . Puisque le nombre de degrés de liberté pour la méthode adaptative dépend du temps, on utilise par convention le nombre moyen de degrés de liberté sur l'intervalle de temps  $[0, 0.2]$ . Sur le graphe de gauche, on observe que l'erreur pour la méthode adaptative exhibe un taux de décroissance en fonction de  $\eta$  similaire aux deux instants considérés : l'erreur stochastique converge en gros en  $O(\sqrt{\eta})$  pour des valeurs de  $\eta$  pas trop petites. L'erreur à l'instant  $t = 0.5$  est plus importante qu'à l'instant  $t = 0.2$ , suite à l'accumulation des erreurs d'approximations pseudo-spectrales, néanmoins elle reste du même ordre de grandeur. Pour finir, les erreurs pour les méthodes adaptative et non adaptative sur le graphe de gauche exhibent un taux de décroissance similaire, mais la méthode adaptative réalise un niveau d'erreur donné avec beaucoup moins de degrés de liberté, sauf quand  $\eta$  devient très petit, où le gain n'est alors plus significatif.

# Extension au cas multidimensionnel

## Sommaire

<b>6.1 Arbres binaires multiD</b> . . . . .	<b>137</b>
6.1.1 Indicateur directionnel de partition . . . . .	137
6.1.2 Arbres équivalents . . . . .	139
6.1.3 Union d'arbres binaires multiD . . . . .	139
6.1.4 Opérateurs de restriction et de prédiction . . . . .	140
<b>6.2 Elagage et enrichissement</b> . . . . .	<b>140</b>
6.2.1 Elagage anisotrope . . . . .	141
6.2.2 Enrichissement isotrope . . . . .	142
<b>6.3 Application à l'équation de Burgers</b> . . . . .	<b>142</b>
6.3.1 Cas test 1 : vitesse des ondes positives . . . . .	142
6.3.2 Cas test 2 : vitesses des ondes positives et négatives . . . . .	143

On se place à présent dans le cas multidimensionnel  $N > 1$ , de sorte que  $\Xi = [0, 1]^N$ . Le but de ce chapitre est d'étendre les outils d'analyse multi-résolution (AMR) 1D du chapitre 4 et le solveur de Roe adaptatif 1D du chapitre 5 au cas multidimensionnel (multiD). Le solveur de Roe adaptatif multiD est testé sur l'équation de Burgers en deux dimensions stochastiques.

## 6.1 Arbres binaires multiD

Plusieurs possibilités peuvent être envisagées pour construire une AMR dans le cas multiD. Une possibilité, utilisée dans le cas déterministe, repose sur des structures d'arbre où chaque nœud possède  $2^N$  ou aucun enfant (Quad-arbre, Octo-arbre, . . .). Le problème de cette approche est sa limitation intrinsèque à des petites dimensions  $N$ . On choisit ici une approche différente conservant la structure d'arbre binaire où on ajoute un indicateur directionnel de partition.

### 6.1.1 Indicateur directionnel de partition

Tout d'abord, on garde la structure binaire de l'arbre (un nœud a aucun ou deux enfants) et on introduit pour chaque nœud  $\mathbf{n} \in \mathcal{N}(\mathbf{T}) \setminus \mathcal{L}(\mathbf{T})$  un indicateur, noté  $d(\mathbf{n}) \in \{1, \dots, N\}$ , donnant la direction par rapport à laquelle la partition



dyadique est appliquée pour construire ses deux enfants  $c^-(\mathbf{n})$  et  $c^+(\mathbf{n})$ . Pour une feuille  $\mathbf{1}$ , on fixe la convention  $d(\mathbf{1}) = 0$ . Un exemple d'arbre binaire multiD et de la partition de  $\Xi$  correspondante est donné sur la figure 6.1 pour  $N = 2$ .

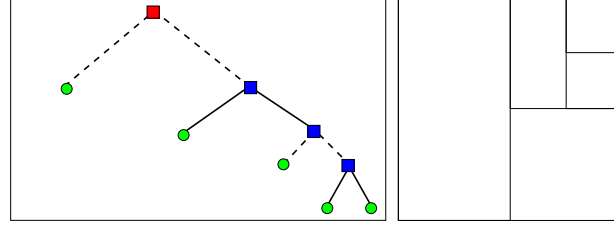


FIGURE 6.1 – À gauche : Exemple d'arbre binaire multiD pour  $N = 2$ . Les segments composés de tirets (respectivement d'un seul trait) représentent une partition par rapport à la première (respectivement deuxième) direction. À droite : Partition correspondante du domaine stochastique.

Ensuite, la base polynomiale d'échelle grossière multidimensionnelle  $\{\Phi_\alpha^{N_0}\}_{\alpha \in \mathcal{P}_\pi}$  est construite par tensorisation de la base d'échelle grossière unidimensionnelle, à savoir la base de Legendre sur  $[0, 1]$ . La tensorisation peut être complète ou partielle ; dans tous les cas, on note  $\Pi_{N_0}^N[\xi]$  l'espace polynomial multidimensionnel correspondant et on continue à noter  $P_\pi$  sa dimension (si bien que  $P_\pi = (N_0 + 1)^N$  dans le cas d'une tensorisation complète). L'espace d'approximation stochastique  $S^{N_0}(\mathbf{T})$  et sa dimension  $P(\mathbf{T})$  sont toujours définis respectivement par (4.84) et (4.85). En outre, les fonctions de base SE sont définies en utilisant des transformations de coordonnées par dimension, de la forme (4.11), menant à un développement SE formellement similaire à (4.86). L'extension de la base MW est moins directe puisqu'il n'est pas possible de procéder par tensorisation de la base unidimensionnelle de  $W_0^{N_0}$  pour obtenir un ensemble "universel" de fonctions multi-ondelettes mères  $\{\Psi_\alpha^{N_0}\}_{\alpha \in \mathcal{P}_\pi}$ . Cela est dû à la construction de l'arbre où chaque nœud correspond à une partition par rapport à une direction  $d(\mathbf{n})$  unique : la base de détail  $\{\Psi_{\mathbf{n},\alpha}\}_{\alpha \in \mathcal{P}_\pi}$  associée à un nœud  $\mathbf{n} \in \mathcal{N}(\mathbf{T}) \setminus \mathcal{L}(\mathbf{T})$  dépend de  $d(\mathbf{n})$ . Cependant, il existe  $N$  ensembles de fonctions de détail anisotropes  $\{\Psi_\alpha^d\}_{\alpha \in \mathcal{P}_\pi}$  qui engendrent les espaces de détail supportant une partition dans la direction  $d \in \{1, \dots, N\}$ .

Ainsi, en utilisant la convention d'ensemble  $\mathcal{N}^*(\mathbf{T})$  définie par (4.89) pour inclure les coefficients d'échelle grossière, le développement MW multidimensionnel pour une fonction  $U \in L^2(\Xi)$  s'écrit sous la forme

$$U(\xi) \approx U^{\mathbf{T}}(\xi) = \sum_{\mathbf{n} \in \mathcal{N}^*(\mathbf{T})} \sum_{\alpha \in \mathcal{P}_\pi} d_{\mathbf{n},\alpha} \Psi_{\mathbf{n},\alpha}^{d(\mathbf{n})}(\xi), \quad (6.1)$$

où  $\Psi_{\mathbf{n},\alpha}^{d(\mathbf{n})}$  est une mise à l'échelle de la fonction de détail anisotrope associée au nœud  $\mathbf{n}$  et le détail directionnel  $d_{\mathbf{n},\alpha}$  est déterminé par  $d_{\mathbf{n},\alpha} = \langle U, \Psi_{\mathbf{n},\alpha}^{d(\mathbf{n})} \rangle$ .

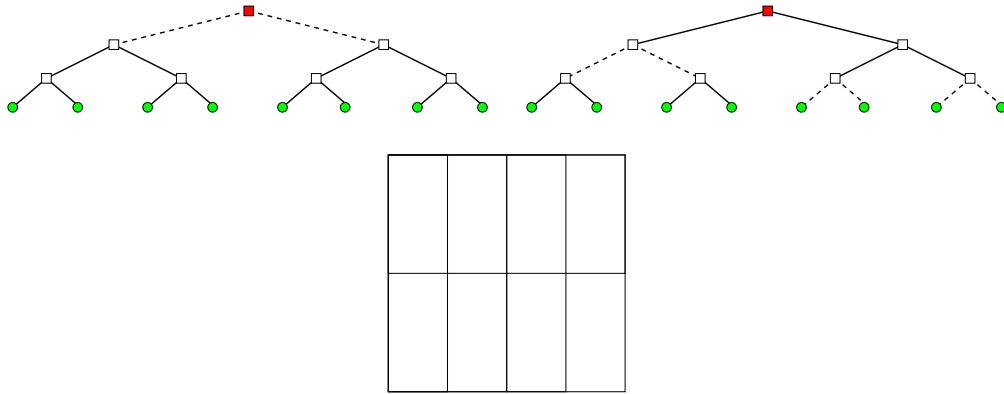


FIGURE 6.2 – Exemple pour  $N = 2$  de deux arbres équivalents (en haut) donnant lieu à la même partition de  $\Xi$  (en bas). Les segments composés de tirets (respectivement d'un seul trait) représentent une partition par rapport à la première (respectivement deuxième) direction.

### 6.1.2 Arbres équivalents

Il existe une différence essentielle entre les arbres binaires 1D et multiD : pour  $N > 1$ , il existe en général plus d'un arbre donnant lieu à la même partition de  $\Xi$ . Ce fait est illustré sur la Figure 6.2 pour  $N = 2$ . Cela conduit à introduire la notion d'arbres équivalents.

**Définition 6.1.** *Deux arbres binaires  $T$  et  $T'$  sont dits équivalents si ils partagent le même ensemble de feuilles,*

$$T \equiv T' \iff \mathcal{L}(T) = \mathcal{L}(T'). \quad (6.2)$$

Deux arbres binaires équivalents fournissent donc le même espace d'approximation stochastique. On étend la relation d'inclusion (4.92) dans le cas multidimensionnel de la façon suivante :

$$T \subset T' \iff \exists T'' \equiv T', \mathcal{N}(T) \subset \mathcal{N}(T''). \quad (6.3)$$

### 6.1.3 Union d'arbres binaires multiD

Pour deux arbres binaires  $T_1$  et  $T_2$ , on définit leur arbre-union  $T_{1 \cup 2} := T_1 \cup T_2$  dans le cas multiD comme l'un des arbres minimaux (en terme du nombre de feuilles) tel que toute feuille  $\mathbf{l} \in \mathcal{L}(T_{1 \cup 2})$  vérifie la propriété (5.5). En effet, l'arbre-union n'est pas unique dès que  $N > 1$ , car différents arbres minimaux peuvent être construits pour satisfaire (5.5). Ces arbres-union sont tous équivalents et fournissent le même espace stochastique d'approximation  $S(T_{1 \cup 2})$ . L'union de deux arbres est illustrée dans la Figure 6.3 pour  $N = 2$ .

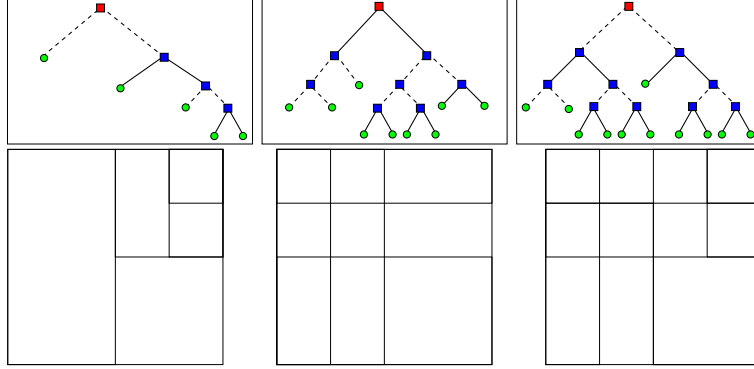


FIGURE 6.3 – Exemple pour  $N = 2$  de l'union de deux arbres binaires. Ligne du haut :  $T_1$  (à gauche),  $T_2$  (au centre), et  $T_{1 \cup 2}$  (à droite). Les segments composés de tirets (respectivement d'un seul trait) représentent une partition par rapport à la première (respectivement deuxième) direction. Ligne du bas : Partitions correspondantes du domaine stochastique.

#### 6.1.4 Opérateurs de restriction et de prédiction

Soient trois arbres binaires  $T^- \subset T \subset T^+$  et soit  $U^T \in S(T)$ . La restriction  $\mathcal{R}_{T \downarrow T^-} U^T$  de  $U^T$  sur  $S(T^-)$  est toujours définie comme la projection  $L^2(\Xi)$ -orthogonale de  $U^T$  sur  $S(T^-)$  par l'équation (5.6). De même, la prédiction  $\mathcal{P}_{T \uparrow T^+} U^T$  de  $U^T$  sur  $S(T^+)$  est toujours définie comme l'injection de  $U^T$  dans  $S(T^+)$ .

Pour déterminer les coefficients d'échelle de la restriction de  $U^T$  sur les feuilles de l'arbre  $T^-$  et ceux de la prédiction de  $U^T$  sur les feuilles de l'arbre  $T^+$ , on procède itérativement comme décrit dans la section 5.1.3, en utilisant cette fois deux familles composées de  $N$  matrices  $\{M_d^{\Phi, -1}\}_{d \in \{1, \dots, N\}}$  et  $\{M_d^{\Phi, 0}\}_{d \in \{1, \dots, N\}}$ , chacune associée à une direction  $d$  du domaine stochastique.

## 6.2 Elagage et enrichissement

En multiD, on définit les niveaux de résolution directionnels d'un nœud, noté  $\{\text{Nr}(\mathbf{n}, d')\}_{d' \in \{1, \dots, N\}}$ , comme le nombre de fois que l'on a coupé dyadiquement le domaine stochastique dans la direction  $d'$  pour obtenir ce nœud. On impose  $\text{Nr}(\mathbf{n}, d') \leq J$ , pour tout nœud et pour tout  $d' \in \{1, \dots, N\}$ , de sorte que  $|\mathbf{n}| = \sum_{d' \in \{1, \dots, N\}} \text{Nr}(\mathbf{n}, d') \leq NJ$ . On définit donc un arbre binaire complet, toujours noté  $T_J$ , de profondeur  $|T_J| \leq NJ$ . On note que cet arbre complet ne sera jamais construit par la suite.

Soient  $T_J$  un arbre binaire complet, une fonction  $U \in L^2(\Xi)$  et  $U^{T_J}$  sa projection  $L^2(\Xi)$ -orthogonale sur l'espace d'approximation stochastique  $S(T_J)$ . On souhaite construire un arbre  $T_{J, \eta}$ , de profondeur maximale  $|T_{J, \eta}| \leq NJ$ , qui représente les nœuds dont les détails  $d_{\mathbf{n}, \alpha}$  sont significatifs. Le développement SE de la projection seuillée  $U^{T_{J, \eta}} := P_{S(T_{J, \eta})} U$  sera alors déterminé sur les feuilles de l'arbre  $T_{J, \eta}$ . Dans

le cas multidimensionnel, l'ensemble des détails significatifs  $\mathcal{D}_{J,\eta}$  s'écrit

$$\mathcal{D}_{J,\eta} := \left\{ \mathbf{n} \in \mathcal{N}^*(\mathbb{T}_J) : \|d_{\mathbf{n}}\|_{l^2} = \left( \sum_{\alpha \in \mathbb{P}_\pi} |d_{\mathbf{n},\alpha}|^2 \right)^{1/2} > \epsilon_J(\eta) \right\}, \quad (6.4)$$

où  $\epsilon_J(\eta) = 2^{-NJ/2}\eta$ .

**Remarque 6.1.** *En pratique, les fonctions de base ne sont pas normalisées (cf remarque 4.3). La valeur de seuillage dépend alors du niveau de résolution directionnel du nœud considéré, i.e. la valeur de seuillage vaut  $\epsilon_{\mathbf{n}}(\eta) = 2^{(|\mathbf{n}|-NJ)/2}\eta$ .*

### 6.2.1 Elagage anisotrope

Grâce au choix de représentation par arbre binaire multiD, l'opérateur de seuillage peut être rendu anisotrope à condition d'y ajouter une procédure de génération d'arbres équivalents. En effet, cet opérateur dépend des directions de partitions successives de  $\Xi$  qui sont choisies pour générer la discrétisation stochastique et son arbre binaire  $\mathbb{T}$  correspondant. En d'autres termes, on pourra mettre à zéro uniquement les détails  $d_{\mathbf{n},\alpha}$  générés par les directions  $d(\mathbf{n})$  imposées par la structure de l'arbre que l'on considère. Ceci est clairement insatisfaisant puisque, pour  $N > 1$ , il existe plusieurs arbres équivalents et on souhaiterait que la procédure de seuillage soit indépendante de tout choix particulier. Pour éviter tout arbitrage, l'arbre correspondant à une séquence de détails est périodiquement remplacé par ses arbres équivalents. Ces arbres équivalents sont construits en cherchant dans l'arbre courant les nœuds  $\mathbf{n}$  dont les enfants  $\mathbf{c}^-(\mathbf{n})$  et  $\mathbf{c}^+(\mathbf{n})$  sont partitionnés consécutivement dans une même direction qui est différente de la direction ayant conduit à ces enfants :  $d(\mathbf{c}^-(\mathbf{n})) = d(\mathbf{c}^+(\mathbf{n})) \neq d(\mathbf{n})$ . Quand un tel nœud est trouvé, les directions de la partition sont échangées,  $d(\mathbf{n}) \leftrightarrow d(\mathbf{c}^-(\mathbf{n})) = d(\mathbf{c}^+(\mathbf{n}))$ , et une permutation des descendants des enfants est effectuée. Cette opération, illustrée dans la Figure 6.4, est appliquée périodiquement et aléatoirement tout au long de la procédure de seuillage.

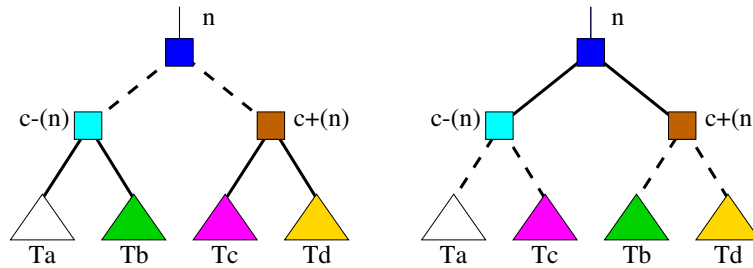


FIGURE 6.4 – Illustration de l'opération élémentaire pour générer des arbres binaires équivalents. à gauche : un nœud avec ses deux enfants divisés selon la même direction. À droite : directions de partition échangées entre les deux générations et permutation des sous-arbres (triangles) des petits enfants réalisée.

### 6.2.2 Enrichissement isotrope

Le dernier outil nécessaire à l'extension du solveur de Roe adaptatif en multiD est la mise en place d'un opérateur d'enrichissement avant le calcul du flux pour chaque interface. Dans les calculs présentés ci-dessous, on décide simplement d'enrichir l'arbre union de l'interface  $\mathbb{T}_{i-1/2}^n$  en divisant toutes ses feuilles dans chacune des  $N$  directions du domaine stochastique. En procédant ainsi, l'arbre enrichi  $\mathbb{T}_{i-1/2}^* = \mathcal{E}(\mathbb{T}_{i-1/2}^n)$  a  $2^N \text{card}(\mathcal{L}(\mathbb{T}_{i-1/2}^n))$  feuilles au total, montrant qu'un tel enrichissement n'est possible en pratique que dans les situations où  $N$  est petit. Il est donc nécessaire de considérer des stratégies plus fines lorsque  $N$  est grand. Des travaux visant à construire un critère d'enrichissement anisotrope basé sur des détails directionnels unidimensionnels de la solution sont actuellement en cours.

## 6.3 Application à l'équation de Burgers

Le solveur de Roe stochastique adaptatif est testé sur les cas tests 1 et 2 du chapitre 3. Il s'agit de l'équation de Burgers en deux dimensions stochastiques avec conditions initiales incertaines sans point sonique. Les bases SE considérées ici correspondent aux polynômes de Legendre complètement tensorisés de degré partiel  $\leq \text{No}$ . Dans les exemples ci-dessous, on contraint les noeuds à un niveau de résolution directionnel  $\text{Nr}(\mathbf{n}, d') \leq J$  pour tout  $d' \in \{1, \dots, N\}$  et une profondeur  $|\mathbf{n}| \leq J$ , de sorte que la profondeur maximale des arbres autorisée est  $J$ .

### 6.3.1 Cas test 1 : vitesse des ondes positives

On considère un domaine spatial unidimensionnel  $\Omega = [0, 1]$  avec des conditions aux bords périodiques. L'équation qui régit le système est donnée sous forme conservative par

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0, \quad F(U) = \frac{U^2}{2}, \quad (6.5)$$

et on considère une condition initiale incertaine  $U^0(x, \xi)$  composés de trois états déterministes constants par morceaux en  $x$ , définie par (3.2) et illustrée dans la figure 3.1 pour  $N_c = 200$  cellules. Pour ce cas test, on contraint les arbres à une profondeur  $1 \leq |\mathbb{T}| \leq 10$  ( $J = 10$ ) et on utilise un paramètre de seuillage  $\eta = 10^{-4}$ . L'ordre polynomial est  $\text{No} = 3$ .

Les réalisations de la solution de Galerkin reconstruites à partir de ses développements stochastiques à différents instants, ainsi que son espérance et son écart-type, exhibent des profils similaires à ceux des figures 3.2 et 3.3. La méthode adaptative proposée capture donc correctement la dynamique de l'équation de Burgers.

La résolution stochastique est adaptée en espace et en temps à la régularité locale de la solution dans le domaine stochastique, de sorte que l'on s'attend à une discrétisation plus fine le long des ondes de choc. Cette caractéristique est observée dans la figure 6.5, où la solution de Galerkin est tracée en fonction de  $(\xi_1, \xi_2)$  pour différents instants  $t \in [0.2, 0.6]$  à un point d'observation mobile  $x_o(t) = 0.25 + 0.5t$

initialement localisé entre les deux chocs stochastiques. Les profils de la solution sont similaires à ceux de la figure 3.4 obtenus avec une discrétisation stochastique uniforme. De plus, on obtient comme on le souhaitait une adaptation locale de la discrétisation stochastique en fonction de la position spatiale et du temps. Ainsi, le domaine stochastique est discrétisé de façon fine au voisinage des discontinuités (en fait dans les parties raides à cause de la diffusivité du schéma de Roe), alors qu'il est grossièrement maillé partout ailleurs.

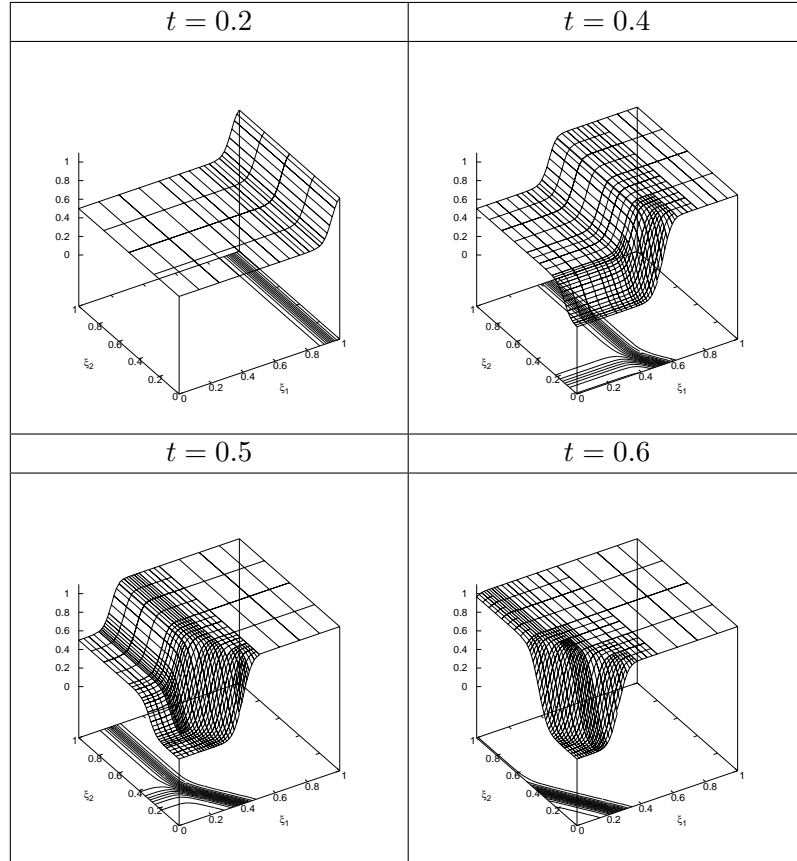


FIGURE 6.5 – Cas test 1 : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  pour un point mobile  $x_o(t) = 0.25 + 0.5t$  et pour différents instants.

### 6.3.2 Cas test 2 : vitesses des ondes positives et négatives

On considère à nouveau l'équation de Burgers mais avec une condition initiale incertaine  $U^0(x, \xi)$  définie en utilisant deux états incertains,  $U^+(\xi_1)$  et  $U^-(\xi_2)$ , le premier presque sûrement positif et le second presque sûrement négatif. On prend pour  $x \in [0, 1]$ , une condition initiale similaire à (3.5) mais avec

$$\begin{aligned} U^+(\xi_1) &= 1 - 0.1(2\xi_1 - 1), & \xi_1 &\sim \mathcal{U}[0, 1] \rightarrow U^+ \sim \mathcal{U}[0.9, 1.1], \\ U^-(\xi_2) &= -1 - 0.05(2\xi_2 - 1), & \xi_2 &\sim \mathcal{U}[0, 1] \rightarrow U^- \sim \mathcal{U}[-1.05, -0.95]. \end{aligned} \quad (6.6)$$

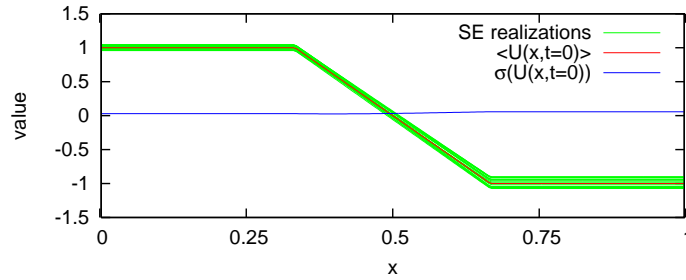


FIGURE 6.6 – Cas test 2 : Échantillon de 20 réalisations de la condition initiale, espérance et écart type.

On résout l'équation de Burgers stochastique avec des conditions aux bords de Dirichlet,  $U = U^+$  en  $x = 0$  et  $U = U^-$  en  $x = 1$ . La condition initiale est illustrée sur la figure 6.6. On utilise  $N_c = 201$  cellules pour la discrétisation spatiale.

Bien qu'initialement continue, la solution de Galerkin va développer en temps fini une discontinuité avec un saut stochastique  $|U^+ - U^-|$  et une vitesse de propagation stochastique  $(U^+ + U^-)/2$ . Pour illustrer la procédure adaptative, on présente dans la figure 6.7 la partition du domaine stochastique  $\Xi$  correspondant à trois

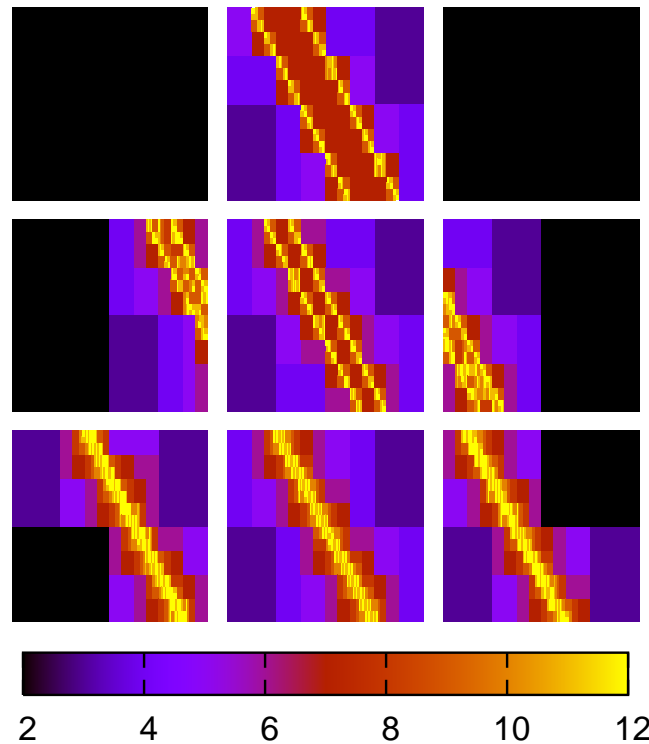


FIGURE 6.7 – Cas test 2 : Partion du domaine stochastique  $\Xi$  aux instants 0.2, 0.4 et 1.5 (de haut en bas) et aux positions spatiales  $x = 0.48, 0.5$  et  $0.52$  (de gauche à droite). L'échelle de couleurs donne la profondeur des feuilles.

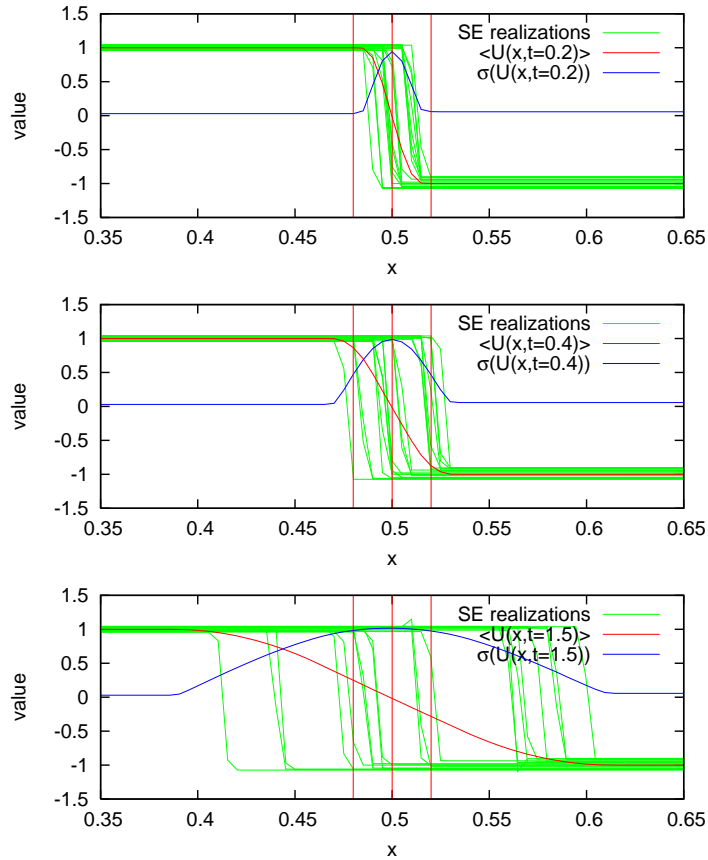


FIGURE 6.8 – Cas test 2 : Solution de Galerkin aux instants  $t = 0.2, 0.4$  et  $1.5$  (de haut en bas). Echantillon de 20 réalisations de la solution de Galerkin, espérance et écart type. Seule la portion du domaine de calcul la plus intéressante est montrée.

positions spatiales différentes et aux instants  $t = 0.2, 0.4$  et  $1.5$ . Pour cette expérience, on contraint les arbres de sorte que  $2 \leq |\mathbb{T}| \leq 12$  et on utilise un paramètre de seuillage  $\eta = 10^{-4}$ . La solution de Galerkin est présentée dans la figure 6.8 aux mêmes instants et sur une partie du domaine spatial  $x \in [0.35, 0.65]$ , où l'espérance et l'écart-type de la solution sont tracés, ainsi qu'un ensemble de réalisations. Les lignes rouges verticales dans la figure 6.8 indiquent les positions spatiales ( $x = 0.5 \pm 0.02$ ) pour lesquelles la partition du domaine stochastique est montrée dans la figure 6.7.

Les figures 6.7 et 6.8 méritent d'être commentées. À  $t = 0.2$ , les arbres en  $x = 0.5 \pm 0.02$  sont tels que  $|\mathbb{T}| = 2$ , la valeur minimale, puisque la solution est toujours continue dans le domaine stochastique pour cet instant et ces positions spatiales. Pour  $x = 0.5$ , des détails stochastiques sont nécessaires pour capturer le choc qui s'est formé presque sûrement et le niveau de résolution maximum  $J$  est atteint pour les arbres à cet instant le long de deux lignes qui traversent  $\Xi$ . Entre ces deux lignes, on observe que la profondeur maximale  $J$  n'est pas atteinte. Cette région correspond à la portion du domaine stochastique où la solution est continue



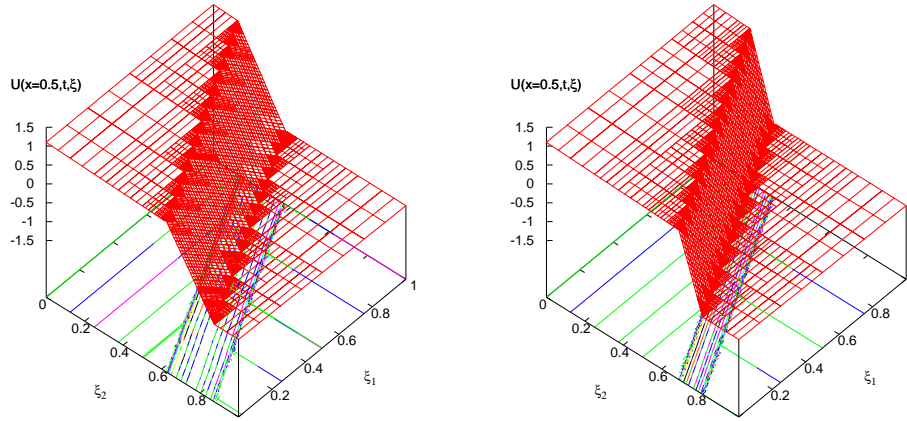


FIGURE 6.9 – Cas test 2 : Solution de Galerkin en fonction de  $(\xi_1, \xi_2)$  en  $x_o = 0.5$  et  $t = 0.2$  (gauche) et  $t = 0.4$  (droite).

à cause de la diffusivité du schéma de Roe. Ce fait est reflété par la solution de Galerkin  $U(\xi_1, \xi_2)$  en  $x = 0.5$  et à  $t = 0.2$  dans le graphe de gauche de la figure 6.9, où on observe des états intermédiaires  $U^- < U(\xi) < U^+$ . Lorsque le temps augmente, la distance entre ces deux lignes correspondant au raffinement maximum diminue, reflétant une solution de Galerkin de plus en plus raide par rapport à  $\xi$  (cf les graphes pour  $t = 0.4$  et  $t = 1.5$  de la figure 6.7 et le graphe de droite de la figure 6.9). À cause de la vitesse du choc  $(U^+ + U^-)/2$ , le choc a atteint les positions  $x = 0.5 \pm 0.02$  avec une probabilité contenue dans  $]0, 1[$ , de sorte que des détails sont à présent nécessaires sur une portion de  $\Xi$  pour prendre en compte la discontinuité de la solution; en fait, les détails sont utiles uniquement sur exactement la moitié du domaine stochastique à ces deux positions pour  $t = 0.4$ . On observe un décalage de la position du choc dans le domaine stochastique à  $x = 0.5 \pm 0.02$  entre  $t = 0.4$  et  $t = 1.5$ . Cela est dû au fait que la condition initiale est presque sûrement non symétrique, de sorte que certaines réalisations des chocs se propagent vers la gauche du domaine et d'autres vers la droite. De plus, on s'attend à ce que pour un temps suffisamment long, on ait en n'importe quelle position spatiale une partition du domaine stochastique  $\Xi$  similaire à la partition en  $x = 0.5$  pour  $t = 1.5$ , et une solution de Galerkin avec un choc séparant deux états  $U^+$  ou  $U^-$  selon le signe de  $2(\xi_1 - 1/2) + (\xi_2 - 1/2)$  puisque la vitesse du choc est  $(U^+ + U^-)/2$ . Cela est reflété dans la ligne du bas de la figure 6.7.

Une caractéristique importante de la stratégie adaptative est que le raffinement permet de contrôler l'amplitude et la mesure de probabilité des overshoots créés par les discontinuités dans la solution de Galerkin. Pour illustrer ce point, on répète l'expérience précédente en faisant varier la profondeur maximale autorisée des arbres  $J$  et on compare les solutions à l'instant  $t = 1.5$ . Dans la figure 6.10, on observe ainsi une décroissance de la fréquence et de l'amplitude des overshoots lorsque l'on augmente  $J$ . La figure 6.11 présente l'évolution en temps du nombre total de

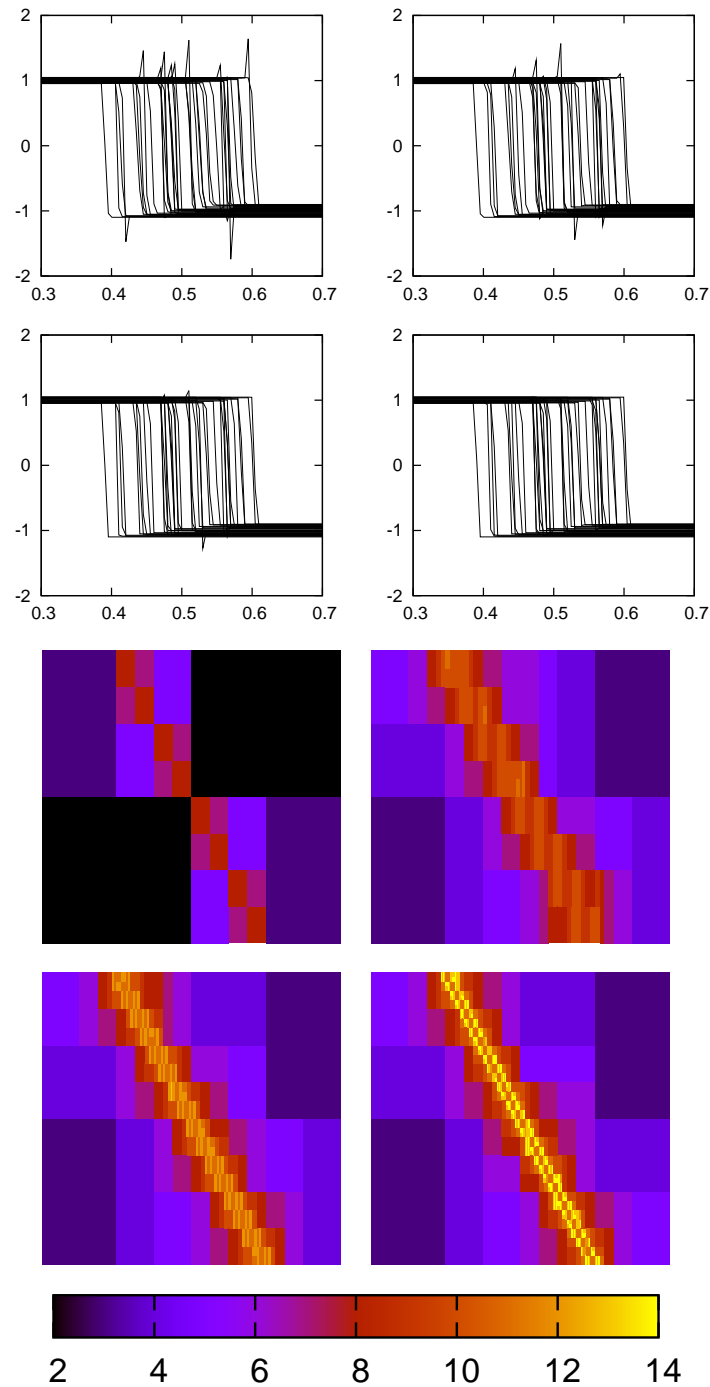


FIGURE 6.10 – Cas test 2 : Échantillon de 30 réalisations de la solution de Galerkin à  $t = 1.5$ , pour une profondeur d'arbres maximale égale à 8 (en haut à gauche), 10 (en haut à droite), 12 (en bas à gauche) et 14 (en bas à droite), et maillages stochastiques correspondants en  $x = 0.5$ . L'échelle de couleurs donne la profondeur des feuilles.

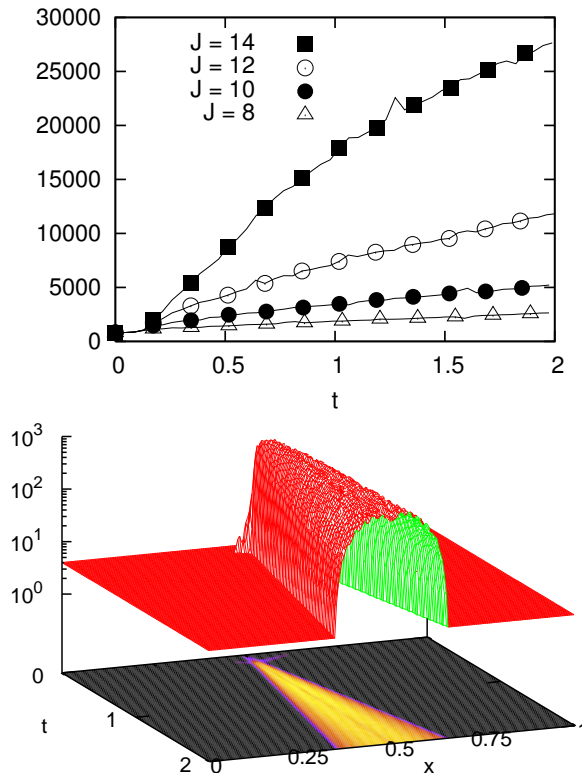


FIGURE 6.11 – Cas test 2. En haut : Évolution avec le temps du nombre total de feuilles pour les solutions de Galerkin utilisant une profondeur maximale égale à 8, 10, 12 et 14. En bas : Distribution espace-temps du nombre de feuilles pour la solution de Galerkin utilisant une profondeur maximale égale à 14.

feuilles pour la solution discrète pour différents  $J$ , ainsi que la distribution des feuilles en espace et en temps (on notera l'échelle logarithmique verticale) pour le cas  $J = 14$ . Pour des temps courts  $t \lesssim 0.15$ , le nombre de feuilles est essentiellement indépendant de  $J$ , puisque le choc ne s'est pas encore formé de sorte que l'adaptation est contrôlée par  $\eta$ . Pour des temps plus longs, le nombre de feuilles augmente en gros linéairement avec  $t$ , reflétant la dépendance linéaire en  $t$  de la portion du domaine spatial  $\Omega$  affecté par le choc stochastique comme observé sur le graphe du bas de la figure 6.11. De manière intéressante, si on augmente  $J$  de 2, le nombre de feuilles est multiplié par moins de 4.

Pour analyser l'efficacité du solveur de Roe adaptatif multiD, on le compare avec le solveur de Roe stochastique utilisant une discrétisation uniforme du domaine stochastique, en utilisant un niveau de résolution  $N_r$  fixé (*i.e.* on utilise un arbre binaire complet  $T_J$  avec  $J = N_r$  dans chaque cellule spatiale). La figure 6.12 présente un échantillon de 30 réalisations de la solution de Galerkin à  $t = 1.5$  obtenu avec des niveaux de résolution uniformes  $N_r = 3$  et 4. Les niveaux de résolution  $N_r = 3$  et 4 ont été choisis de sorte que le nombre total de degré de liberté pour le cas uniforme,

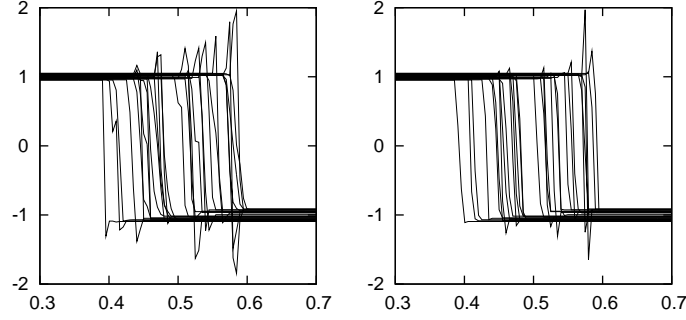


FIGURE 6.12 – Cas test 2 : Échantillon de 30 réalisations de la solution de Galerkin à  $t = 1.5$ , dans le cas d'une discrétisation uniforme du domaine stochastique pour un niveau de résolution  $N_r$  égal à 3 et 4.

*i.e.*,  $\text{dof}_{\text{unif}} = 2^{2N_r}(\text{No}+1)^2 \times N_c$ , est du même ordre de grandeur que le nombre total de degrés de liberté dans le cas adaptatif à  $t^n = 1.5$ , *i.e.*,  $\text{dof}_{\text{adap}} = \sum_{i=1}^{N_c} P(\mathbb{T}_i^n)$ . Par conséquent, ces résultats sont à comparer avec les graphes de la deuxième ligne de la figure 6.10 pour une profondeur maximale des arbres  $J = 12$  et 14. On observe que la fréquence et l'amplitude des overshoots sont considérablement réduites lorsque l'on utilise la méthode adaptative.

Pour évaluer plus quantitativement le contrôle des overshoots, on mesure l'amplitude des overshoots intégrés sur le domaine spatial de la façon suivante :

$$\epsilon_{\text{over}}^2(t^n = 1.5) = \frac{\Delta x}{M} \sum_{i=1}^{N_c} \sum_{j=1}^M \left( (U_i^n(\xi^{(j)}) - U^+(\xi^{(j)}))_+^2 + (U_i^n(\xi^{(j)}) - U^-(\xi^{(j)}))_-^2 \right), \quad (6.7)$$

où  $U_i^n(\xi^{(j)})$  est la solution de Galerkin évaluée en  $\xi^{(j)}$  pour un échantillon de réalisations,  $U^\pm(\xi^{(j)})$  sont les bornes inférieure et supérieure de la solution stochastique en  $\xi^{(j)}$ , et les indices  $\pm$  sont les parties positive et négative d'un nombre réel. On utilise un échantillon de dimension  $M = 10^6$ . La figure 6.13 présente la quantité  $\epsilon_{\text{over}}$  en fonction du nombre de degrés de liberté (on notera les échelles logarithmiques) pour  $N_r = 1, 2, 3, 4, 5, 6$  et  $J = 4, 6, 8, 10, 12, 14$ . Puisque le nombre de degrés de liberté pour la méthode adaptative dépend du temps, on utilise par convention le nombre moyen de degrés de liberté sur l'intervalle de temps  $[0, t^n]$ . On observe que  $\epsilon_{\text{over}}$  décroît plus rapidement pour la méthode adaptative que pour la méthode non adaptative. Par exemple, pour atteindre une valeur de  $10^{-2}$  pour  $\epsilon_{\text{over}}$ , le nombre de degrés de liberté nécessaires pour la méthode adaptative est de deux ordres de grandeur moindre que pour la méthode non adaptative. Ces économies en terme de degrés de liberté réalisent des économies en temps de calcul considérables.

Finalement, on compare la méthode adaptative à la méthode non adaptative

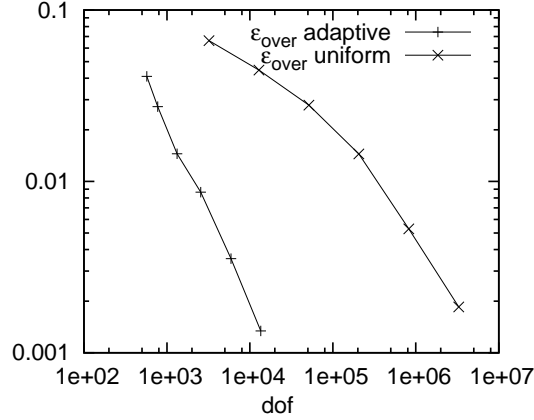


FIGURE 6.13 – Cas test 2 : Mesure  $\epsilon_{\text{over}}(t^n = 1.5)$  pour différents  $N_r$  et  $J$  en fonction du nombre de degrés de liberté.

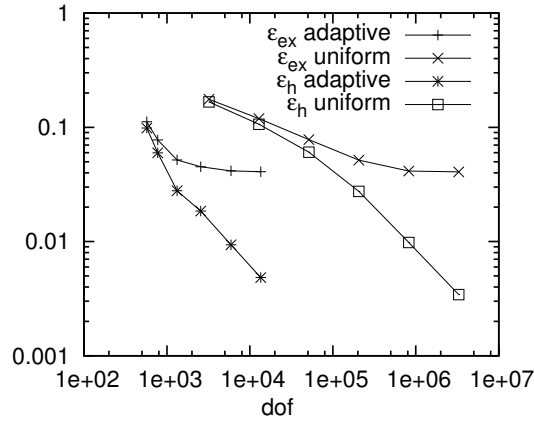


FIGURE 6.14 – Cas test 2 : Erreurs  $\epsilon_{\text{ex}}$  et  $\epsilon_h$  à  $t^n = 1.5$  pour différents  $N_r$  et  $J$  en fonction du nombre de degrés de liberté.

grâce aux mesures d'erreur suivantes :

$$\epsilon_{\text{ex}}^2(t^n = 1.5) := \frac{\Delta x}{M} \sum_{i=1}^{N_c} \sum_{j=1}^M \left( U_i^n(\xi^{(j)}) - U^{\text{ex}}(x_i, t^n, \xi^{(j)}) \right)^2, \quad (6.8)$$

$$\epsilon_h^2(t^n = 1.5) := \frac{\Delta x}{M} \sum_{i=1}^{N_c} \sum_{j=1}^M \left( U_i^n(\xi^{(j)}) - U^h(x_i, t^n, \xi^{(j)}) \right)^2, \quad (6.9)$$

où  $U_i^n(\xi^{(j)})$ ,  $U^{\text{ex}}(x_i, t^n, \xi^{(j)})$ , et  $U^h(x_i, t^n, \xi^{(j)})$  sont évaluées pour  $\xi^{(j)}$  appartenant à un échantillon de réalisations respectivement à partir du développement stochastique de la solution de Galerkin, de la solution exacte du problème de Burgers déterministe dans la cellule de centre  $x_i$  et au temps discret  $t^n$ , et en résolvant le problème de Burgers déterministe discret. On utilise un échantillon de taille  $M = 10^5$ . Ces

deux mesures d'erreur sont représentées dans la figure 6.13 en fonction du nombre de degrés de liberté pour  $N_r \in \{1, 2, 3, 4, 5, 6\}$  et  $J \in \{4, 6, 8, 10, 12, 14\}$ . Les erreurs pour les méthodes adaptative et non adaptative exhibent un taux de décroissance similaire, mais la méthode adaptative réalise un niveau d'erreur donné avec beaucoup moins de degrés de liberté. On remarque également que la mesure d'erreur  $\epsilon_{\text{ex}}$  stagne au-delà d'un certain nombre de degrés de liberté car l'erreur est alors dominée par l'erreur de discrétisation spatiale.



# Conclusion et perspectives

---

Dans ce manuscrit, on a proposé et étudié théoriquement et numériquement une méthode de Galerkin stochastique adaptative pour les systèmes hyperboliques stochastiques de lois de conservation, présentant des discontinuités dans les domaines spatial et stochastique. La discrétisation stochastique est effectuée par la méthode de Galerkin en utilisant un espace de polynômes par morceaux. Dans un premier temps, on a analysé l'hyperbolicité du système de Galerkin stochastique. Puis, on a construit un solveur de Roe avec des matrices de décentrement calculées de manière efficace. On a également proposé une extension du correcteur entropique de Dubois et Mehlman au problème de Galerkin stochastique. Des tests numériques ont été présentés sur les équations de Burgers et d'Euler en deux et une dimensions stochastiques respectivement, sans et avec point sonique, montrant la précision et la robustesse de la méthode.

Enfin, on a proposé une stratégie adaptative innovante basée sur des outils d'analyse multi-résolution afin de construire des espaces d'approximation stochastiques localement raffinés dépendant de l'espace et du temps. Le solveur de Roe adaptatif a été testé sur les équations de Burgers et d'Euler en deux et une dimensions stochastiques respectivement. Les résultats illustrent la capacité de la méthode à traiter les systèmes hyperboliques non linéaires présentant des chocs, tout en réalisant des gains en calculs significatifs grâce à la discrétisation stochastique adaptée. À notre connaissance, il s'agit de la première méthode de Galerkin stochastique adaptative proposant une adaptation de la discrétisation stochastique localisée en espace et en temps, basée de surcroît sur un cadre mathématique (analyse multi-résolution) rigoureux.

La méthode présentée dans ce manuscrit recèle diverses pistes en vue de son amélioration. Tout d'abord, le solveur de Roe stochastique suppose d'avoir à disposition l'expression explicite des valeurs propres et des vecteurs propres du système hyperbolique stochastique. Ce n'est pas une limitation pour les équations de Burgers et d'Euler. Cependant, pour des systèmes de lois de conservation généraux, la connaissance *a priori* des caractéristiques de la solution peut ne pas être possible. Des stratégies alternatives pour traiter de tels problèmes restent une question ouverte.

De plus, le solveur utilisé en espace, bien que contrôlant de manière relativement satisfaisante la diffusion numérique, est seulement d'ordre un, ce qui conduit rapidement à une domination de l'erreur de discrétisation du schéma par l'erreur de discrétisation spatiale lorsque l'espace d'approximation stochastique est suffisam-



ment riche. Il serait donc intéressant de construire un solveur d'ordre supérieur en espace, en utilisant par exemple des techniques de type WENO ou DG (assorties de limiteurs adéquats). Une alternative serait de contrôler la discrétisation spatiale en utilisant également une méthode adaptative.

Ensuite, des techniques d'enrichissement plus avancées ont besoin d'être développées pour traiter le cas de plus grandes dimensions stochastiques. Elles font l'objet de travaux en cours. Une analyse des temps de calcul est également à prévoir dans le cas de grandes dimensions stochastiques.

Enfin, il serait souhaitable de comparer notre méthode avec les autres méthodes existantes pour propager les incertitudes dans les systèmes hyperboliques stochastiques, par exemple avec celles de Abgrall [1], Poette *et al.* [69, 70] et Mishra et Schwab [61]. Dans cet objectif, un benchmark pourrait être proposé.

# Bibliographie

- [1] R. ABGRALL – « A simple, flexible and generic deterministic approach to uncertainty quantifications in non linear problems : application to fluid flow problems », Tech. Report 00325315, INRIA, 2008, soumis à *J. Comput. Phys.* (Cité en pages 28 et 154.)
- [2] M. ABRAMOWITZ & I. STEGUN – Handbook of Mathematical Functions, Dover, 1970. (Cité en pages 9 et 12.)
- [3] R. AGARWAL & P. WONG – Error inequalities in polynomial interpolation and their applications, vol. 262, Kluwer, 1993. (Cité en page 19.)
- [4] B. ALPERT – « A class of bases in L2 for the sparse representation of integral operators », SIAM J. Math. Anal. **24** (1993), p. 246–262. (Cité en pages 101 et 102.)
- [5] R. ASKEY & J. WILSON – « Some basic hypergeometric polynomials that generalize Jacobi polynomials », Memoirs of the American Mathematical Society **54** (1985), p. 1–55. (Cité en page 12.)
- [6] I. BABUŠKA, F. NOBILE & R. TEMPONE – « A stochastic collocation method for elliptic partial differential equations with random input data », SIAM J. Numer. Anal. **45** (2007), no. 3, p. 1005–1034. (Cité en page 20.)
- [7] A. BARTH, C. SCHWAB & N. ZOLLINGER – « Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients », Tech. Report 18, Seminar for Applied Mathematics, ETH Zürich, 2010. (Cité en page 6.)
- [8] T. J. BARTH – « On the propagation of statistical model parameter uncertainty in cfd calculations », J. Theoretical and Computational Fluid Dynamics (2011), Accepted 11/05/2010, In Press. (Cité en page 28.)
- [9] M. BERVEILLER, B. SUDRET & M. LEMAIRE – « Stochastic finite element : a non intrusive approach by regression », Eur. J. Comput. Mech. **15** (2006), p. 81–92. (Cité en page 19.)
- [10] R. CAMERON & W. MARTIN – « The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals », Ann. Math. **48** (1947), p. 385–392. (Cité en pages 7 et 8.)
- [11] C. CLENSHAW & A. CURTIS – « A method for numerical integration on an automatic computer », Numerische Mathematik **2** (1960), p. 197–205. (Cité en page 18.)
- [12] A. COHEN – Numerical analysis of wavelet methods, Studies in Mathematics and its Applications, vol. 32, North-Holland Publishing Co., Amsterdam, 2003. (Cité en pages 109 et 111.)
- [13] A. COHEN, S. KABER, S. MÜLLER & M. POSTEL – « Fully adaptive multiresolution finite volume schemes for conservation laws », Math. Comp. **72** (2001), no. 241, p. 183–225 (electronic). (Cité en pages 98 et 110.)

- [14] T. CRESTAUX – « Méthode adaptative d'intégration multi-dimensionnelle et sélection d'une base de polynômes de chaos », Thèse, Université Paris 13, 2011. (Cit  en page 18.)
- [15] S. DAS, R. GHANEM & J. SPALL – « Asymptotic sampling distribution for Polynomial Chaos representation of data : a maximum entropy and Fisher information approach », SIAM J. Sci. Comp. **30** (2008), no. 5, p. 2207–2234. (Cit  en page 4.)
- [16] M. DEB, I. BABUSKA & J. ODEN – « Solution of stochastic partial differential equations using Galerkin finite element techniques », Comput. Methods Appl. Mech. Engrg. **190** (2001), p. 6359–6372. (Cit  en pages 13 et 29.)
- [17] B. DEBUSSCHERE, H. NAJM, A. MATTA, O. KNIO, R. GHANEM & O. LE MAÎTRE – « Protein labeling reactions in electrochemical microchannel flow : Numerical prediction and uncertainty propagation », Physics of Fluids **15** (2003), no. 8, p. 2238–2250. (Cit  en page 21.)
- [18] B. DEBUSSCHERE, H. NAJM., P. PÉBAY, O. KNIO, R. GHANEM & O. LE MAÎTRE – « Numerical challenges in the use of polynomial chaos representations for stochastic processes », SIAM J. Sci. Comput. **26** (2004), no. 2, p. 698–719. (Cit  en pages 28, 37 et 79.)
- [19] P. DEGOND, P. PEYRARD, G. RUSSO & P. VILLEDIEU – « Polynomial upwind schemes for hyperbolic systems », C. R. Acad. Sci. Paris Sér. I Math. **328** (1999), no. 6, p. 479–483. (Cit  en page 46.)
- [20] G. DEODATIS – « Weighted integral method », J. Engrg. Math. **117** (1991), no. 8, p. 1851–1864. (Cit  en page 6.)
- [21] G. DEODATIS & M. SHINOZUKA – « Weighted integral method », J. Engrg. Math. **117** (1991), no. 8, p. 1865–1877. (Cit  en page 6.)
- [22] C. DESCELIERS, R. GHANEM & C. SOIZE – « Maximum likelihood estimation of stochastic chaos representations from experimental data », Int. J. Num. Meth. Eng. **66** (2006), no. 6, p. 978–1001. (Cit  en page 4.)
- [23] C. DESCELIERS, C. SOIZE & R. GHANEM – « Identification of chaos representations of elastic properties of random media using experimental vibration tests », Comp. Mech. **39** (2007), no. 6, p. 831–838. (Cit  en page 4.)
- [24] B. DESPRÉS – Lois de Conservations Eulériennes, Lagrangiennes et Méthodes Numériques, Springer-Verlag, 2010. (Cit  en pages 22 et 43.)
- [25] D. A. DI PIETRO & A. ERN – Mathematical Aspects of Discontinuous Galerkin Methods, Springer-Verlag, 2011. (Cit  en page 25.)
- [26] O. DITLEVSEN & N. TARP-JOHANSEN – « Choice of input fields in stochastic finite-elements », Probabilistic Engineering Mechanics **14** (1999), p. 63–72. (Cit  en page 4.)
- [27] F. DUBOIS & G. MEHLMAN – « A non-parameterized entropy correction for Roe's approximate Riemann solver », Numer. Math. **73** (1996), no. 2, p. 169–208. (Cit  en pages 29, 49 et 50.)

- [28] A. ERN & J.-L. GUERMOND – Theory and Practice of Finite Elements, Springer-Verlag, 2004. (Cit  en page 20.)
- [29] B. GANAPATHYSUBRAMANIAN & N. ZABARAS – « Sparse grid collocation schemes for stochastic natural convection problems », J. Comput. Phys. **225** (2007), p. 652–685. (Cit  en page 20.)
- [30] L. GE, K. CHEUNG & M. KOBAYASHI – « Stochastic solution for uncertainty propagation in nonlinear shallow-water Equations », Journal of Hydraulic Engineering **Dec. 2008** (2008), p. 1732–1743. (Cit  en pages 28, 29 et 37.)
- [31] T. GERSTNER & M. GRIEBEL – « Numerical integration using sparse grids », Numer. Algorithms **18** (1998), p. 209–232. (Cit  en page 18.)
- [32] R. GHANEM & P. SPANOS – Stochastic Finite Elements : A Spectral Approach, Dover, 2003, 2nd edition. (Cit  en pages 7, 10, 20 et 21.)
- [33] E. GODLEWSKI & P. RAVIART – Numerical Approximation of Hyperbolic Systems of Conservation Laws, Applied Mathematical Sciences, vol. 118, Springer-Verlag, New York, 1996. (Cit  en pages 22 et 43.)
- [34] D. GOTTLIEB & D. XIU – « Galerkin method for wave equations with uncertain coefficients », Commun. Comput. Phys. **3** (2008), no. 2, p. 505–518. (Cit  en page 28.)
- [35] M. GRIGORIU – Stochastic Calculus, Applications in Sciences and Engineering, Birkh user, 2002. (Cit  en page 6.)
- [36] A. HARTEN – « High resolution schemes for hyperbolic conservation laws », J. Comput. Phys. **49** (1983), no. 3, p. 357–393. (Cit  en page 49.)
- [37] — , « Multiresolution algorithms for the numerical solution of hyperbolic conservation laws », Comm. Pure Appl. Math. **48** (1995), no. 12, p. 1305–1342. (Cit  en pages 126 et 131.)
- [38] — , « Multiresolution representation of data : a general framework », SIAM J. Numer. Anal. **33** (1996), no. 3, p. 1205–1256. (Cit  en pages 30 et 98.)
- [39] J. S. HESTHAVEN & T. WARBURTON – Nodal discontinuous Galerkin methods : algorithms, analysis, and applications, Springer-Verlag, 2008. (Cit  en page 25.)
- [40] N. HOVHANNISYAN, S. M LLER & R. SCH FER – « Adaptive multiresolution discontinuous galerkin schemes for conservation laws », Tech. report, 2010. (Cit  en page 97.)
- [41] A. KEESE & H. MATTHIES – « Numerical methods and Smolyak quadrature for nonlinear stochastic partial differential equations », Tech. report, Institute of Scientific Computing TU Braunschweig Brunswick, 2003. (Cit  en page 18.)
- [42] F. KEINERT – Wavelets and multiwavelets, Studies in Advanced Mathematics, Chapman & Hall/CRC, Boca Raton, FL, 2004. (Cit  en page 97.)
- [43] M. KLEIBER & T. HIEN – The Stochastic Finite Element Method, John Wiley and Sons, 1992. (Cit  en page 5.)

- [44] O. KNIO & O. LE MAÎTRE – « Uncertainty propagation in CFD using polynomial chaos decomposition », Fluid Dyn. Res. **38** (2006), p. 616–640. (Cité en page 21.)
- [45] O. LE MAÎTRE & O. KNIO – Spectral Methods for Uncertainty Quantification, Scientific Computation, Springer-Verlag, New York, 2010, With applications to computational fluid dynamics. (Cité en pages 7 et 21.)
- [46] O. LE MAÎTRE, O. KNIO, H. NAJM & R. GHANEM – « A stochastic projection method for fluid flow. I. Basic formulation », J. Comput. Phys. **173** (2001), p. 481–511. (Cité en page 21.)
- [47] O. LE MAÎTRE, O. KNIO, H. NAJM & R. GHANEM – « Uncertainty propagation using Wiener-Haar expansions », J. Comput. Phys. **197** (2004), no. 1, p. 28–57. (Cité en pages 13 et 27.)
- [48] O. LE MAÎTRE, H. NAJM, R. GHANEM & O. KNIO – « Multi-resolution analysis of Wiener-type uncertainty propagation schemes », J. Comput. Phys. **197** (2004), no. 2, p. 502–531. (Cité en pages 13, 29, 35, 97, 101 et 102.)
- [49] O. LE MAÎTRE, H. NAJM, P. PÉBAY, R. GHANEM & O. KNIO – « Multi-resolution-analysis scheme for uncertainty quantification in chemical systems », SIAM J. Sci. Comput. **29** (2007), no. 2, p. 864–889. (Cité en pages 13 et 35.)
- [50] O. LE MAÎTRE, M. REAGAN, H. NAJM, R. GHANEM & O. KNIO – « A stochastic projection method for fluid flow. II. Random process », J. Comput. Phys. **181** (2002), p. 9–44. (Cité en page 21.)
- [51] O. LE MAÎTRE – « A Newton method for the resolution of steady stochastic Navier-Stokes equations », Comput. & Fluids **38** (2009), no. 8, p. 1566–1579. (Cité en page 87.)
- [52] G. LIN, C.-H. SU & G. KARNIADAKIS – « Predicting shock dynamics in the presence of uncertainties », J. Comput. Phys. **217** (2006), no. 1, p. 260–276. (Cité en page 29.)
- [53] G. LIN, C.-H. SU & G. KARNIADAKIS – « Stochastic modeling of random roughness in shock scattering problems : theory and simulations », Comput. Methods Appl. Mech. Engrg. **197** (2008), no. 43-44, p. 3420–3434. (Cité en page 28.)
- [54] J. LIU – Monte Carlo Strategies in Scientific Computing, Springer-Verlag, 2001. (Cité en page 6.)
- [55] X. LIU, S. OSHER & T. CHAN – « Weighted essentially non-oscillatory schemes », J. Comput. Phys. **115** (1994), p. 200–212. (Cité en page 25.)
- [56] M. LOÈVE – Probability Theory, Springer-Verlag, 1977. (Cité en page 4.)
- [57] X. MA & N. ZABARAS – « An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations », J. Comput. Phys. **228** (2009), p. 3084–3113. (Cité en page 20.)
- [58] L. MATHELIN & M. HUSSAINI – « A stochastic collocation algorithm for uncertainty analysis », Tech. Report NASA/CR-2003-212153, NASA Langley Research Center, 2003. (Cité en page 20.)

- [59] L. MATHELIN, M. HUSSAINI & T. ZANG – « Stochastic approaches to uncertainty quantification in CFD simulations », Numer. Algorithms **38** (2005), no. 1-3, p. 209–236. (Cité en page 28.)
- [60] M. MCKAY, W. CONOVER & R. BECKMAN – « A comparison of three methods for selecting values of input variables in the analysis of output from a computer code », Technometrics **21** (1979), p. 239–245. (Cité en pages 6 et 17.)
- [61] S. MISHRA & C. SCHWAB – « Sparse tensor multi-level Monte Carlo Finite Volume methods for hyperbolic conservation laws with random initial data », Tech. Report 24, Seminar for Applied Mathematics, ETH Zürich, 2010. (Cité en pages 6, 26 et 154.)
- [62] W. MOROKOFF & R. CAFLISCH – « Quasi-Monte Carlo integration », J. Comput. Phys. **122** (1995), no. 2, p. 218–230. (Cité en page 17.)
- [63] S. MÜLLER – Adaptive multiscale schemes for conservation laws, Lecture Notes in Computational Science and Engineering, vol. 27, Springer-Verlag, Berlin, 2003. (Cité en pages 98 et 110.)
- [64] H. NAJM – « Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics », Ann. Rev. Fluid Mech. **41** (2009), p. 35–52. (Cité en page 21.)
- [65] H. NAJM, B. DEBUSSCHERE, Y. MARZOUK, S. WIDMER & O. LE MAÎTRE – « Uncertainty quantification in chemical systems », Internat. J. Numer. Methods Engrg. **80** (2009), no. 6-7, p. 789–814. (Cité en page 21.)
- [66] M. NDJINGA – « Computing the matrix sign and absolute value functions », C. R. Math. Acad. Sci. Paris **346** (2008), no. 1-2, p. 119–124. (Cité en page 45.)
- [67] F. NOBILE, R. TEMPONE & C. WEBSTER – « An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data », SIAM J. Numer. Anal. **46** (2008), no. 5, p. 2411–2442. (Cité en page 20.)
- [68] — , « A sparse grid stochastic collocation method for partial differential equations with random input data », SIAM J. Numer. Anal. **46** (2008), no. 5, p. 2309–2345. (Cité en page 20.)
- [69] G. POETTE, B. DESPRÉS & D. LUCOR – « Uncertainty quantification for systems of conservation laws », J. Comput. Phys. **228** (2009), no. 7, p. 2443–2467. (Cité en pages 28, 37 et 154.)
- [70] — , « Treatment of uncertain material interfaces in compressible flows », Comput. Methods Appl. Mech. Engrg. **200** (2011), p. 284–308. (Cité en pages 28, 37 et 154.)
- [71] F. PUKELSHEIM – Optimal Design of Experiments, Classics in Applied Mathematics, vol. 50, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006, Reprint of the 1993 original. (Cité en page 19.)
- [72] M. REAGAN, H. NAJM, R. GHANEM & O. KNIO – « Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection », Combustion and Flame **132** (2003), p. 545–555. (Cité en page 18.)

- [73] P. ROE – « Some contributions to the modelling of discontinuous flows », in Large-Scale Computations in Fluid Mechanics, Part 2 (La Jolla, Calif., 1983), Lectures in Appl. Math., vol. 22, Amer. Math. Soc., Providence, RI, 1985, p. 163–193. (Cité en pages 25 et 49.)
- [74] M. SHINOZUKA & G. DEODATIS – « Response variability of stochastic finite element systems », J. Engrg. Math. **114** (1988), no. 3, p. 499–519. (Cité en page 6.)
- [75] S. SMOLYAK – « Quadrature and interpolation formulas for tensor products of certain classes of functions », Dokl. Akad. Nauk SSSR **4** (1963), p. 240–243. (Cité en page 18.)
- [76] C. SOIZE – « A comprehensive overview of a non-parametric probabilistic approach of model uncertainty for predictive models in structural dynamics », J. Sound and Vibrations **288** (2005), p. 623–652. (Cité en page 2.)
- [77] — , « Construction of probability distributions in high dimension using the maximum entropy principle : Applications to stochastic processes, random fields and random matrices », Int. J. Num. Meth. Eng. **76** (2008), no. 10, p. 1583–1611. (Cité en page 4.)
- [78] C. SOIZE & R. GHANEM – « Reduced chaos decomposition with random coefficients of vector-valued random variables and random fields », Comp. Meth. App. Mech. Eng. **198** (2009), no. 21–26, p. 1926–1934. (Cité en page 4.)
- [79] J. STOER & R. BULIRSCH – Introduction to Numerical Analysis, 3ème éd., Texts in Applied Mathematics, vol. 12, Springer-Verlag, New York, 2002, Translated from the German by R. Bartels, W. Gautschi and C. Witzgall. (Cité en pages 12, 17, 18 et 46.)
- [80] E. TORO – Riemann solvers and numerical methods for fluid dynamics : A practical introduction, 2ème éd., Springer-Verlag, Berlin, 1999. (Cité en pages 22 et 43.)
- [81] J. TRYOEN, O. LE MAÎTRE & A. ERN – « Adaptive Anisotropic Stochastic Discretization Schemes for Uncertain Conservation Laws », in Proceedings of the ASME 2010 3rd Joint US-European Fluids Engineering Summer Meeting, 2010. (Cité en page 30.)
- [82] J. TRYOEN, O. LE MAÎTRE, M. NDJINGA & A. ERN – « Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems », J. Comput. Phys. **229** (2010), no. 18, p. 6485–6511. (Cité en page 29.)
- [83] — , « Roe solver with entropy corrector for uncertain hyperbolic systems », J. Comput. Appl. Math. **235** (2010), no. 2, p. 491–506. (Cité en page 29.)
- [84] J. WALDVOGEL – « Fast construction of the Fejér and Clenshaw–Curtis quadrature rules », BIT Num. Math. **43** (2003), no. 1, p. 001–018. (Cité en page 18.)
- [85] X. WAN & G. KARNIADAKIS – « Multi-element generalized polynomial chaos for arbitrary probability measures », SIAM J. Sci. Comp. **28** (2006), no. 3, p. 901–928. (Cité en pages 13 et 29.)

- 
- [86] N. WIENER – « The Homogeneous Chaos », Amer. J. Math. **60** (1938), p. 897–936. (Cité en page 7.)
- [87] D. XIU – Numerical Methods for Stochastic Computations : A Spectral Method Approach, Princeton University Press, 2010. (Cité en page 7.)
- [88] D. XIU & J. HESTHAVEN – « High-order collocation methods for differential equations with random inputs », SIAM J. Sci. Comput. **27** (2005), no. 3, p. 1118–1139. (Cité en page 20.)
- [89] D. XIU & G. KARNIADAKIS – « The Wiener-Askey Polynomial Chaos for stochastic differential equations », SIAM J. Sci. Comput. **24** (2002), p. 619–644. (Cité en page 12.)
- [90] F. YAMAZAKI, M. SHINOZUKA & G. DASGUPTA – « Neumann expansion for stochastic finite element analysis », J. Engrg. Math. **114** (1988), no. 8, p. 1335–1354. (Cité en page 6.)