



**THÈSE / UNIVERSITÉ DE RENNES 1**  
*sous le sceau de l'Université Européenne de Bretagne*

En Cotutelle Internationale avec

**l'Université Ibn Tofaïl de Kenitra, Maroc**

pour le grade de

**DOCTEUR DE L'UNIVERSITÉ DE RENNES 1**

*Mention : Mathématiques appliquées*

**Ecole doctorale MATISSE**

présentée par

**Mestapha OUMOUNI**

Préparée à l'INRIA Rennes - Bretagne Atlantique

---

**Analyse numérique  
de méthodes performantes  
pour les EDP stochastiques  
modélisant l'écoulement  
et le transport en milieux  
poreux**

**Thèse soutenue à Kenitra  
le 05 Juin 2013**

Composition du jury :

**Allami BENYAICHE**

Professeur à l'université Ibn Tofaïl, Kenitra /  
Président

**Abdellatif CHARAFI**

Professeur à l'université Al Akhawayn, Ifrane /  
Rapporteur

**Olivier LE MAÎTRE**

Chargé de recherche CNRS au LIMSI /  
Rapporteur

**Olivier PIRONNEAU**

Professeur à l'université Pierre et Marie Curie  
Paris / Rapporteur

**Anthony NOUY**

Professeur à l'école Centrale de Nantes /  
Examineur

**Jocelyne ERHEL**

Directrice de recherche, INRIA/  
directrice de thèse

**Zoubida MGHAZLI**

Professeure à l'université Ibn Tofaïl, Kenitra /  
directrice de thèse



# Remerciements

Mes premiers remerciements vont naturellement à mes deux directeurs de thèse, Mmes Jocelyne Erhel et Zoubida Mghazli, qui m'ont permis de poursuivre le travail initié lors de mémoire du master par une thèse en co-tutelle entre l'Université Ibn Tofaïl et l'Université de Rennes 1. Je les remercie pour l'aide compétente qu'ils m'ont apportée, pour leurs encouragements et disponibilités, leurs qualités humaines et scientifiques.

Je tiens à adresser mes profonds remerciements à tous les membres du jury : M. Benyaaich Alami, (président du jury), Mmes. Jocelyne Erhel et Zoubida Mghazli (directeurs de thèse), M. Anthony Nouy (Examinateur) et en particulier aux rapporteurs, Messieurs Olivier Pironeau, Olivier Lemaître et Alami Talibi, que je souhaite leur exprimer toute ma gratitude, d'abord, d'avoir accepté de rapporter ce travail, malgré leurs emplois du temps assez chargés. Ainsi pour leur lecture attentive du manuscrit et leurs remarques constructives. Je remercie également tous les projets qui m'ont financé mes séjours au sein de l'équipe sage dans le cadre de ma thèse en co-tutelle, en particulier le projet européen Coadvice qui m'a financé un séjour de 9 mois, l'université de Rennes1 qui m'a accordé une bourse de mobilité pour un séjour de 3 mois, et enfin, les projets HYDRINV et PHC Volubilis N°MA/10/225 qui m'ont financé un mois de stage en France.

Je remercie également et chaleureusement tous les membres de l'équipe SAGE, d'abord pour leurs accueils agréables durant tous mes séjours à l'IRISA, ainsi que les grandes motivations que j'ai eu pour l'avancement de mon travail. d'abord les permanents : Jocelyne Erhel, Bernard Philippe, Édouard Canot et Géraldine Pichot pour leurs qualités humaines et scientifiques. Ensuite, les autres membres passés et présents, Mohamad Muhieddine, Babtist Poirriez, Julia Charrier, Nadir Soualem, Souhila Sabit, Sinda Khalfallah, Louis-Bernard Nguenang, Marwen ben Refifa, Salwa Mansour, Lionel Lenôtre, Grégoire Lecourt et Aurélien Le Gentil m'ont tous, d'une façon explicite ou implicite, présenté une aide scientifique ou simplement leur partage d'amitié. Je profite aussi à remercier les professeurs Arnaud Debauche et Jean-Raynald De Dreuzy d'avoir accepté à assister à mes séminaires et leurs remarques constructives. Je remercie également tous les membres de l'équipe "courir à l'IRISA", qui m'ont fait découvrir l'esprit de la compétition en course à pieds.

Je remercie aussi tous les membres de l'équipe I.M.A de l'Université Ibn tofaïl de Kenitra, professeurs et doctorants, pour la bonne humeur et leurs conseils.

Enfin, mes pensées vont à mes parents, à mes frères et ma soeur, qui m'ont toujours soutenu et encouragé, ainsi que mes amis, pour tout ce qu'ils m'avaient apporté.



# Table des matières

Table des matières	vi
Liste des tableaux	vii
Table des figures	ix
<b>1 Problème modèle et méthodes numériques d'approximation</b>	<b>5</b>
1.1 Formulation du problème modèle . . . . .	5
1.1.1 Le bruit de dimension finie . . . . .	6
1.1.2 Erreur de troncature sur les paramètres . . . . .	8
1.1.3 Formulation faible et formulation forte . . . . .	11
1.1.4 Régularité stochastique . . . . .	12
1.1.5 Espace d'approximation déterministe . . . . .	13
1.2 Méthode d'échantillonnage de Monte-Carlo . . . . .	14
1.3 Méthodes de type approximation locale . . . . .	16
1.4 Méthode des éléments finis stochastiques . . . . .	16
1.5 Méthode de collocation stochastique . . . . .	18
1.5.1 Interpolation par une tensorisation pleine . . . . .	19
1.5.2 Construction de Smolyak . . . . .	21
1.5.3 Construction de Smolyak anisotrope . . . . .	25
<b>2 Calcul de l'espérance de la solution d'une EDP stochastique unidimensionnelle à l'aide d'une base réduite</b>	<b>29</b>
2.1 Formulation du problème modèle . . . . .	30
2.2 Construction d'une base stochastique réduite . . . . .	31
2.2.1 Algorithme . . . . .	34
2.2.2 Assemblage de la matrice de rigidité . . . . .	34
2.3 Exemple numérique . . . . .	35
2.3.1 Comparaison avec la méthode de Monte Carlo . . . . .	37
2.3.2 Comparaison avec la méthode de collocation stochastique . . . . .	38
2.4 Conclusion . . . . .	40
<b>3 Méthode des grilles clairsemées adaptative par un indicateur d'erreur</b>	<b>41</b>

3.1	Motivation . . . . .	41
3.2	Estimation de l'erreur et adaptation dans l'espace stochastique . . . . .	43
3.3	Adaptation des grilles clairsemées . . . . .	46
3.3.1	Estimation des poids d'anisotropie . . . . .	46
3.3.2	Estimation du niveau $w$ convenable . . . . .	51
3.4	Tests numériques . . . . .	53
3.4.1	Cas du paramètre $k$ non linéaire . . . . .	53
3.4.2	Cas du paramètre $k$ linéaire . . . . .	59
3.5	Conclusion . . . . .	60
<b>4</b>	<b>Méthode adaptative des grilles clairsemées via l'extrapolation de Richardson</b>	<b>63</b>
4.1	Estimation de l'erreur . . . . .	64
4.1.1	Cas d'un ensemble stochastique borné . . . . .	65
4.1.2	Cas d'un ensemble stochastique non borné . . . . .	70
4.2	Approche des grilles clairsemées adaptative . . . . .	77
4.3	Application à un problème elliptique stochastique . . . . .	81
4.4	Conclusion . . . . .	84
<b>I</b>	<b>Analyse numérique du couplage de l'équation d'écoulement et de l'équation d'advection-diffusion</b>	<b>87</b>
<b>5</b>	<b>Analyse numérique de l'équation d'advection-diffusion dans un milieu aléatoire</b>	<b>88</b>
5.1	Problème physique . . . . .	89
5.1.1	Équation de l'écoulement . . . . .	89
5.1.2	Équation d'advection-diffusion . . . . .	89
5.1.3	Quantités d'intérêts . . . . .	90
5.2	Description de l'approche numérique . . . . .	90
5.2.1	Méthodes particulières stochastiques . . . . .	90
5.2.2	Approximation dans un milieu aléatoire . . . . .	93
5.3	Analyse de l'approche numérique . . . . .	94
5.3.1	Approximation de l'équation de l'écoulement . . . . .	94
5.3.2	Erreur de discrétisation en temps . . . . .	96
5.3.3	Erreur de la discrétisation spatiale . . . . .	100
5.3.4	Erreur sur l'extension moyenne . . . . .	102
5.3.5	Erreur sur la dispersion moyenne . . . . .	105
5.4	Erreur de la troncature . . . . .	108
5.5	Exemples numériques . . . . .	111
5.6	Conclusion . . . . .	111

5.7	Développement de Karhunen-Loève pour les champs aléatoires . . . . .	116
5.7.1	Formulation du problème . . . . .	116
5.8	Champ Gaussian de covariance exponentielle . . . . .	118
	<b>Bibliographie</b>	<b>121</b>

# Liste des tableaux

2.1	Comparaison de l'approche de projection sur la base réduite avec la méthode de Monte Carlo. . . . .	38
2.2	Comparaison de l'approche de projection sur la base réduite avec la méthode de collocation stochastique . . . . .	39
3.1	$N = 11$ Premières valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori. . . . .	55
3.2	Les 11 premières valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori. . . . .	58
3.3	Les 10 premières valeurs des poids d'anisotropie obtenues par une approche a posteriori. . . . .	60
4.1	Les $N = 10$ valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori. . . . .	80
4.2	Le cardinal de la grille clairsemée anisotrope $\#\mathcal{H}_\alpha(w, N)$ , les poids d'anisotropie sont donnés dans le tableau 4.1 . . . . .	80
4.3	Indicateur sur l'approximation de la moyenne $\mathbb{E}[Z]$ utilisant la grille clairsemée anisotrope (GCA) et l'extrapolation de Richardson (ER). . . . .	81
4.4	Indicateur sur l'approximation de la variance $\text{Var}[Z]$ utilisant la grille clairsemée anisotrope (GCA) et l'extrapolation de Richardson (ER). . . . .	81
4.5	L'ordre de troncature $N$ dans la série (4.48) par rapport à la longueur de corrélation $l_c$ et la tolérance souhaitée $\text{tol}_N$ . . . . .	83
4.6	Les $N = 21$ valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori, $N = 10$ pour le cas $l_c = 1$ et $N = 21$ pour le cas $l_c = 1/2$ . . . . .	83
4.7	Comparaison du taux de convergence amélioré $\mu$ comme dans (4.49) avec celui $\tilde{\mu}$ comme dans (4.50). . . . .	84



# Table des figures

1.1	Pour $\Gamma = [-1, 1]^2$ et $w = 5$ , grilles clairsemées isotropes avec des points de Clenshaw-Curtis (à gauche) et des points de Gauss (au milieu). À droite, grille de produit tensoriel complet, avec $m(i_1) = m(i_2) = 2^5$ . . .	22
1.2	Grille clairsemée isotrope (à gauche), grille clairsemée anisotrope $\frac{\alpha_1}{\alpha_2} = \frac{3}{2}$ (au milieu) et $\frac{\alpha_1}{\alpha_2} = 2$ (à droite), pour $\Gamma = [-1, 1]^2$ et $w = 6$ , utilisant des points de Clenshaw-Curtis. . . . .	26
2.1	Les dix premières valeurs et vecteurs propres $\{(\hat{\lambda}_n, \hat{\varphi}_n)\}_{n=1}^N$ de la covariance de $G$ . . . . .	37
2.2	Les dix premières valeurs et vecteurs propres $\{(\lambda_n, \varphi_n)\}_{n=1}^N$ de la covariance de $k^{-1}$ . . . . .	38
2.3	L'erreur relative de la moyenne par projection et Monte Carlo . . . . .	39
2.4	La moyenne par projection et Monte Carlo . . . . .	40
3.1	Comparaison de la grille $\#\mathcal{H}_{\hat{\alpha}}(w, N)$ avec $\#\mathcal{H}_{\alpha}(w, N)$ , (à gauche $l_c = 0.5$ , et à droite $l_c = 0.2$ ). . . . .	55
3.2	Comparaison du cardinal de la grille clairsemée $\mathcal{H}_{\alpha}(w, N)$ avec celui de la grille $\mathcal{H}_{\hat{\alpha}}(w, N)$ , (à gauche $l_c = 0.5$ , à droite $l_c = 0.2$ ). . . . .	56
3.3	Comparaison de l'erreur $C\ \mathbb{E}[\lambda_{ref} - \mathcal{A}_{\alpha}(w, N)\lambda]\ $ avec celle commise sur $u$ (à gauche $l_c = 0.5$ , à droite $l_c = 0.2$ ). . . . .	57
3.4	L'espérance approchée $\mathbb{E}[\mathcal{A}_{\alpha}(5, 41)]$ et la variance approchée $\text{var}[\mathcal{A}_{\alpha}(5, 41)]$ . 57	
3.5	Comparaison du cardinal de la grille clairsemée $\mathcal{H}_{\alpha}(w, N)$ avec celui de la grille $\mathcal{H}_{\hat{\alpha}(w, N)}$ , (gauche $l_c = 0.5$ , droite $l_c = 0.2$ ). . . . .	58
3.6	Comparaison de l'erreur $C\ \mathbb{E}[\lambda_{ref} - \mathcal{A}_{\alpha}(w, N)\lambda]\ $ avec celle sur $u$ (à gauche $l_c = 0.5$ , à droite $l_c = 0.2$ ). . . . .	59
3.7	L'espérance approchée $\mathbb{E}[\mathcal{A}_{\alpha}(4, 41)]$ et la variance $\text{var}[\mathcal{A}_{\alpha}(4, 41)]$ , le cas des variables bornés. . . . .	59
3.8	À gauche : Comparaison de l'erreur $C_w\ \mathbb{E}[\lambda_{ref} - \mathcal{A}_{\alpha}(w, N)\lambda]\ _{\infty}$ avec celle sur $u$ . À droite : Comparaison du cardinal de la grille $\mathcal{H}_{\alpha}(w, N)$ avec celui de la grille $\mathcal{H}_{\hat{\alpha}}(w, N)$ . . . . .	61
3.9	L'espérance approchée $\mathbb{E}[\mathcal{A}_{\alpha}(6, 51)]$ et la variance $\text{var}[\mathcal{A}_{\alpha}(6, 51)]$ . . . .	61

4.1	Comparaison de l'indicateur classique avec l'indicateur basé sur l'extrapolation de Richardson dans le cas $l_c = 1$ avec $N = 10$ . À gauche l'indicateur de l'espérance et à droite celui de la variance . . . . .	85
4.2	Comparaison de l'indicateur classique avec celui basé sur l'extrapolation de Richardson dans le cas $l_c = 0.5$ avec $N = 21$ . À gauche l'indicateur de l'espérance et à droite celui de la variance. . . . .	85
4.3	L'espérance $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$ et la variance $\text{Var}[\mathcal{A}_\alpha(w, N)u]$ , le cas $l_c = \frac{1}{2}$ avec $N = 21$ et $w = 4$ . . . . .	86
5.1	L'extension (à gauche) et la dispersion explicite (à droite) avec un pas de temps $\delta t = 1/50$ et $M = 10^6$ de simulations . . . . .	112
5.2	La dispersion approchée par une différence finie. . . . .	112
5.3	Valeurs propres de la covariance exponentielle, $l_c = 1$ , $\sigma = 1$ . . . . .	119
5.4	Vecteurs propres de la covariance exponentielle, $l_c = 1$ , $\sigma = 1$ . . . . .	120

# Introduction

Les dernières décennies ont connu un énorme progrès dans la résolution des équations aux dérivées partielles. La simulation numérique de telles équations est un outil essentiel, et est devenue de plus en plus prometteuse et rapide grâce aux progrès technologiques et à la puissance des calculateurs. Cependant, les résultats numériques peuvent être affectés par différents types d'erreurs, comme les erreurs sur les modèles et les erreurs dues à la simulation numérique. Ces dernières peuvent être contrôlées et réduites utilisant diverses techniques numériques sophistiquées, telles que des techniques d'adaptation et les estimations d'erreur a posteriori [51, 63]. Cependant, les erreurs sur les modèles sont loin d'être aussi bien contrôlées. En particulier les modèles mathématiques classiques souffrent de plusieurs sources d'incertitudes, qui réduisent la fiabilité des simulations numériques. Elles sont données généralement par deux classes, les incertitudes épistémiques et les incertitudes aléatoires. Les incertitudes aléatoires sont associées aux variabilités intrinsèques de la nature, comme l'hétérogénéité dans les milieux ou les matériaux. Quant aux incertitudes épistémiques, elles apparaissent dans les modèles en raison du manque de connaissances sur les propriétés des milieux, les erreurs sur les mesures ou le manque et l'insuffisance des mesures sur les données d'entrée, telles que les coefficients du modèle, les termes de forçage, les conditions aux bords, la géométrie du milieu, etc... Par exemple, dans l'étude des écoulements souterrains, les propriétés du sous-sol telles que la porosité et la perméabilité de l'aquifère sont extrapolées à partir des mesures prises seulement sur un nombre fini de points spatiaux.

Dans le but d'augmenter la précision des prédictions numériques et obtenir des prévisions assez fiables sur le modèle étudié, les simulations doivent prendre en compte ces différents types d'erreurs. La compréhension et la quantification sont cruciales pour comprendre et quantifier les différences entre les prédictions numériques et le comportement du système réel.

L'incertitude peut être traitée par plusieurs techniques, par exemple : la théorie des ensembles flous, la théorie de l'évidence, la théorie des probabilités, etc. (cf [24],[55],[34], [5], et les références qui s'y trouvent). Dans la théorie des probabilités, les incertitudes sont modélisées comme des champs ou des processus aléatoires. Cette théorie ayant une base mathématique solide, elle constitue un outil efficace pour décrire et quantifier la propagation des incertitudes. La quantification d'incertitude dans ce cas consiste généralement à considérer un système stochastique gouverné par une équation aux dérivées partielles stochastique (EDPS) ou une équation différentielle stochastique (EDS). Les entrées de ce système sont modélisées comme des fonctions stochastiques.

L'approche traditionnelle pour résoudre les EDPS ou les EDS est la méthode de Monte-Carlo (MC) [27]. Cette approche donne une approximation des statistiques, telles que l'espérance, la variance ou des moments de certaines quantités d'intérêt de la solution. L'inconvénient principal de cette approche est sa convergence lente, qui engendre un effort de calcul généralement exorbitant, en particulier pour des systèmes physiques complexes et de grande taille.

Récemment, des méthodes spectrales ont été développées, telles que la méthode de Galerkin et de collocation stochastiques [34, 5, 6]. Elles offrent un outil robuste pour la résolution des problèmes d'EDPS. Ces méthodes approchent la réponse du modèle comme fonction stochastique, avec une grande précision et une convergence exponentielle. Cependant elles exigent la résolution d'une série de problèmes déterministes de grande taille qui sont parfois totalement couplés, limitant leur application à des faibles dimensions stochastiques. Ces méthodes engendrent un coût de calcul qui augmente de façon exponentielle avec la dimension stochastique et leur mise en œuvre peut devenir parfois impraticable. Il est intéressant donc de bien chercher à minimiser le coût et l'effort de la résolution, tout en choisissant avec soin les degrés de liberté nécessaires.

Cette thèse a pour objectif principal l'étude et l'analyse des méthodes efficaces pour la résolution d'un problème stochastique dans le but de la quantification des incertitudes. On s'intéresse d'abord à la résolution de l'équation de l'écoulement stochastique en milieux poreux, par une méthode d'interpolation des grilles clairsemées adaptatives, ce qui constitue la première partie. La deuxième partie concerne la quantification du transport d'un polluant dans un milieu aléatoire. Une approche numérique de résolution est présentée avec une analyse mathématique de l'erreur a priori.

La suite est organisée de la façon suivante : dans le premier chapitre, nous introduisons le problème modèle traité dans la première partie, il s'agit du problème de diffusion linéaire avec des coefficients et données stochastiques,

$$\begin{cases} -\operatorname{div}(k\nabla u) = f & \text{dans } \Omega \times D \\ u|_{\partial D} = 0. \end{cases} \quad (1)$$

où le triplet  $(\Omega, \mathcal{F}, d\mathbb{P})$  est un espace de probabilité complet et  $D$  un domaine physique borné et régulier. On fait des hypothèses sur les paramètres aléatoires pour que le problème soit bien posé, ainsi que pour le transformer en une formulation déterministe nécessaire pour l'approximation numérique. Nous donnons également un rappel des différentes méthodes numériques d'approximation de tels problèmes, ainsi qu'un rappel sur leurs estimations d'erreur montrant la qualité de la convergence.

Dans le deuxième chapitre, nous proposons une méthode de réduction de modèle qui permet d'approcher l'espérance  $\mathbb{E}[u]$ , solution du problème d'écoulement unidimension-

nel. L'approche est basée sur la décomposition de Karhunen-Loève (K-L) de l'inverse du paramètre de diffusion et une projection similaire à celle de Galerkin, moins coûteuse. La projection de la solution sur un espace à dimension réduite est construite de manière que l'espérance de la solution et celle du projeté coïncident. Généralement cette projection ne coïncide pas avec la solution. L'estimation a priori d'erreur prouve l'efficacité et la convergence de l'approche unidimensionnelle considérée.

Malheureusement, nous n'avons pas pu étendre l'idée du deuxième chapitre pour des problèmes en dimensions supérieures. Cependant, dans le même esprit d'extension à 2d ou 3d de l'espace physique, nous proposons d'utiliser dans le chapitre 3 l'inverse de diffusion comme un indicateur d'erreur  $\lambda = k^{-1}f$ , pour déterminer le nombre total des degrés de liberté de la solution. On montre que l'erreur commise par une interpolation polynomiale de Smolyak  $\mathcal{A}_\alpha(w, N)$  sur cet indicateur est une borne supérieure de l'erreur commise sur  $u$ .

Dans l'approche classique [50], les poids d'anisotropie  $\alpha$  sont choisis comme étant le taux de convergence de  $u$  suivant chaque direction. Cela engendre un effort important sous forme de problèmes déterministes, alors qu'un grand nombre de ces problèmes ne contribue pas à la construction de l'approximation  $\mathcal{A}_\alpha(w, N)u$ . L'utilisation de l'indicateur  $\lambda$  pour l'approximation de  $\alpha$  exige seulement la résolution de problèmes d'interpolation. On utilise aussi cet indicateur pour déterminer un niveau convenable pour l'interpolation  $\mathcal{A}_\alpha(w, N)u$ . Cette approche permet de réduire un coût de calcul significatif, et ceci sans impact sur la précision.

Dans le chapitre 4 nous démontrons une amélioration de l'erreur a priori de la méthode des grilles clairsemées anisotrope.

La comparaison avec l'estimation de l'erreur donnée dans [50] montre une compatibilité de cette amélioration avec le cas isotrope [49]. Nous avons aussi appliqué une technique d'adaptation donnée dans [11, 48] dans le cas de la méthode isotrope. Elle est basée sur l'extrapolation de Richardson et consiste à utiliser l'information de l'erreur a priori précédente et une combinaison linéaire des résultats calculés auparavant pour prédire une amélioration des résultats. Ceci permet d'éviter un effort de calcul considérable à un niveau supplémentaire.

Dans la deuxième partie, au chapitre 5, on s'intéresse à l'approximation numérique de la solution du couplage entre l'équation d'écoulement dans un milieu poreux aléatoire et l'équation d'advection-diffusion. Plus précisément, l'approximation de la valeur

moyenne de l'extension d'un soluté introduit au temps initial, et la valeur moyenne de la macro-dispersion qui représente la vitesse avec laquelle le soluté se propage.

On fait le choix de discrétiser l'équation de l'écoulement par la méthode des éléments finis mixtes, afin d'approcher efficacement la vitesse  $v$  (l'erreur commise sur la vitesse dans  $L^2$  est optimale). Ensuite, l'équation d'advection-diffusion est vue comme une équation de Fokker-Planck, et sa solution est donc approchée par une méthode particulière probabiliste. L'extension et la dispersion moyenne peuvent en effet être exprimées comme l'espérance d'une fonction d'un processus  $X_t$ , solution de l'équation différentielle stochastique correspondante. Les deux sont calculées par un schéma d'Euler pour l'EDS et une méthode de Monte-Carlo. Une estimation de l'erreur faible de

la discrétisation en temps a été démontrée dans [47], nous présentons une nouvelle démonstration de cette estimation aussi fine que [15] mais qui reste valable pour un bruit multiplicatif. On montre aussi que l'erreur de discrétisation spatiale est d'ordre  $h$ , et finalement, on donne l'estimation de l'erreur globale commise sur l'extension moyenne. Cette estimation dépend de la régularité de la perméabilité  $k$ .

La dispersion est calculée explicitement comme une espérance d'une fonction de  $X_t$ , son approximation est ainsi donnée en fonction des réalisations d'une discrétisation du processus. L'estimation de l'erreur globale commise sur la dispersion est établie lorsque la perméabilité est de classe  $C^{1,\alpha}$ .

# Chapitre 1

## Problème modèle et méthodes numériques d'approximation

Dans ce chapitre, nous introduisons le problème modèle traité dans la première partie. Il s'agit du problème de diffusion linéaire avec coefficients et données stochastiques. On y ajoute des hypothèses sur les paramètres aléatoires pour que le problème soit bien posé et pour le transformer en une formulation déterministe, nécessaire pour l'approximation numérique. Ici nous rappelons également les différentes méthodes numériques d'approximation de tels problèmes, ainsi que leurs estimations d'erreur caractérisant la qualité de la convergence.

### 1.1 Formulation du problème modèle

On considère un triplet  $(\Omega, \mathcal{F}, d\mathbb{P})$ , espace de probabilité complet, où  $\Omega$  est l'ensemble des résultats,  $\mathcal{F}$  la  $\sigma$ -algèbre des événements de  $\Omega$ ,  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  une mesure de probabilité. Soient  $D$  un ouvert de  $\mathbb{R}^d$  de frontière  $\partial D$  assez-régulière,  $k(\cdot, \cdot)$  et  $f(\cdot, \cdot)$  deux champs aléatoires sur  $\Omega \times D$ . L'objectif est la résolution du problème elliptique suivant :

$$\begin{cases} -\operatorname{div}(k\nabla u) = f & \text{dans } \Omega \times D \\ u|_{\partial D} = 0. \end{cases} \quad (1.1)$$

Nous supposons les hypothèses suivantes sur les paramètres aléatoires  $k$  et  $f$ , pour que notre problème soit bien posé :

**Hypothèse 1.1.**

- **H1** :  $f$  est un champ aléatoire du second ordre, c'est-à-dire  $f \in L^2(\Omega) \otimes L^2(D)$ .
- **H2** :  $k$  est un champ aléatoire uniformément borné et strictement positif (p.s) sur  $\Omega$ , il existe  $0 < k_{min} < k_{max}$ , telles que :

$$\mathbb{P}\left\{\omega, k_{min} \leq k(\omega, \cdot) \leq k_{max}\right\} = 1. \quad (1.2)$$

Si on définit l'espace de Hilbert  $L^2(\Omega) \otimes H_0^1(D)$  par :

$$L^2(\Omega) \otimes H_0^1(D) := \left\{v : \Omega \longrightarrow H_0^1(D); \int_{\Omega} \|v\|_{H_0^1(D)}^2 d\mathbb{P} < \infty\right\},$$

alors, le problème variationnel associé à  $(\mathcal{P}_s)$  est donné par la formulation faible suivante :

$$\begin{cases} \text{Trouver } u \in L^2(\Omega) \otimes H_0^1(D) \text{ tel que} \\ \mathbb{E}[(k \cdot \nabla u, \nabla \mu)_{L^2(D)}] = \mathbb{E}[(f, \mu)_{L^2(D)}] \quad \forall \mu \in L^2(\Omega) \otimes H_0^1(D). \end{cases} \quad (1.3)$$

Sous l'Hypothèse **H2**, la forme bilinéaire  $a(\cdot, \cdot) = \mathbb{E}[(k \nabla \cdot, \nabla \cdot)_{L^2(D)}]$  est continue et coercive sur  $L^2(\Omega) \otimes H_0^1(D) \times L^2(\Omega) \otimes H_0^1(D)$ , et par l'Hypothèse **H1**, la forme linéaire  $l(\cdot) = \mathbb{E}[(f, \cdot)_{L^2(D)}]$  est continue sur  $L^2(\Omega) \otimes H_0^1(D)$ , alors par le théorème de Lax-Milgram, le problème variationnel (1.3) admet une solution unique dans l'espace  $L^2(\Omega) \otimes H_0^1(D)$ .

**1.1.1 Le bruit de dimension finie**

Dans de nombreuses applications, les paramètres aléatoires sont estimés par un nombre fini de variables aléatoires non corrélées, parfois indépendantes, comme dans le cas de la troncature de Karhunen-Loève [44, 45, 5]. Ainsi on peut supposer que la diffusion  $k$  et la donnée  $f$  dépendent d'une suite de  $N$  variables aléatoires  $\{Y_n\}_{n=1}^N$ . Le nombre  $N$  est parfois appelé la dimension stochastique.

**Exemple 1.1.** (Champ aléatoire constant par morceaux)

On décompose le domaine physique en  $N$  zones disjointes  $D = \cup_{n=1}^N T_n$ , dans chaque zone  $T_n$  le paramètre de diffusion est caractérisé par une variable aléatoire de moyenne non nulle, comme suit :

$$k(\omega, x) = k_{min} + \sum_{n=1}^N \sigma_n \chi_{T_n}(x) Y_n(\omega),$$



où  $\chi_{T_n}$  est la fonction caractéristique de la zone  $T_n$  et  $k_{min}$ ,  $\sigma_n$  sont des constantes positives, et la famille  $\{Y_n\}_{n=1}^N$  sont des variables aléatoires positives et indépendantes et de variance unité.

**Exemple 1.2.** *Troncature de Karhunen-Loève*

La décomposition de Karhunen-Loève est un outil puissant pour discrétiser n'importe quel champ aléatoire du second ordre  $k(\cdot, \cdot) \in L^2(\Omega) \otimes L^2(D)$ . Elle est composée d'une partie déterministe (sa moyenne ou son espérance), et une partie aléatoire, décomposée sur la base des valeurs propres et des fonctions propres de sa fonction de covariance  $cov[k]$  comme suit :

$$k(\omega, x) = \mathbb{E}[k](x) + \sum_{n=1}^{\infty} \sqrt{\lambda_n} \phi_n(x) Y_n(\omega), \quad (1.4)$$

où  $\{\lambda_j\}_{j=1}^{\infty}$  et  $\{\phi_j\}_{j=1}^{\infty}$  sont respectivement les valeurs et les fonctions propres solutions du problème aux valeurs propres suivant :

$$\int_D cov[k](x, y) \phi(y) dy = \lambda \phi(x).$$

La fonction de covariance  $cov[k]$  est donnée par :

$$cov[k](x, y) := \mathbb{E}[k(\cdot, x)k(\cdot, y)] - \mathbb{E}[k(\cdot, x)]\mathbb{E}[k(\cdot, y)],$$

et la suite  $\{Y_j\}_{j=1}^{\infty}$  est une suite de variables aléatoires décorrélées données par :

$$Y_n(\omega) = \frac{1}{\sqrt{\lambda_n}} \int_D (k(\omega, x) - \mathbb{E}[k](x)) \phi_n(x) dx. \quad (1.5)$$

On définit la troncature de Karhunen-Loève  $k_N$  comme la troncature de la série (1.4) à l'ordre  $N$  comme suit :

$$k_N(\omega, x) = \mathbb{E}[k](x) + \sum_{n=1}^N \sqrt{\lambda_n} \phi_n(x) Y_n(\omega).$$

Cette troncature est la décomposition optimale du processus  $k(\cdot, \cdot)$  au sens de  $L^2(\Omega) \otimes L^2(D)$  sur l'ensemble de toutes les décompositions de type  $\sum_{n=1}^N v_n \zeta_n$ , telles que  $(\zeta_n, v_n) \in L^2(\Omega) \times L^2(D)$ .

**Exemple 1.3.** *(Coefficients non linéaires)*

Dans certains cas, afin d'assurer quelques propriétés sur les données, le coefficient  $k$  est donné par une transformation non linéaire d'un autre champ aléatoire gaussien  $G$ ,  $k(\omega, x) = g(G(\omega, x))$ , où  $g$  est une fonction régulière et strictement positive. Une telle

transformation peut garantir la positivité ou la bornitude de  $k$ , ainsi que l'indépendance des variables aléatoires qui le décrivent.

Une transformation classique qu'on trouve dans beaucoup d'applications consiste à prendre  $g(t) = e^t$ ; elle permet de garantir que le paramètre  $k = e^G$  soit strictement positif pour tout  $x \in D$ , (p.s) sur  $\Omega$ . Le champ aléatoire  $G$  est donné par exemple par sa discrétisation de (K.L) et ainsi  $k$  est discrétisé de la façon suivante :

$$k(\omega, x) = \exp \left( b_0(x) + \sum_{n=1}^N b_n(x) Y_n(\omega) \right). \quad (1.6)$$

Si  $G$  est un champ Gaussien, alors les variables aléatoires  $Y_1, \dots, Y_N$  sont des gaussiennes centrées réduites, ainsi dans ce cas, non seulement elles sont décorrélées mais de plus elles sont indépendantes. En particulier le champ  $k$  ne satisfait pas l'hypothèse H2, puisque nous avons :

$$k_{min} = e^{-|b_0| - \sum_{n=1}^N |b_n|_\infty |Y_n(\omega)|} \quad \text{et} \quad k_{max} = e^{|b_0| + \sum_{n=1}^N |b_n|_\infty |Y_n(\omega)|}.$$

Cependant on peut montrer dans ce cas que  $u \in L^p(\Omega) \otimes H_0^1(D)$  puisque  $k_{min}$ ,  $k_{min}^{-1}$  et  $k_{max}$  sont dans  $L^p(\Omega)$  pour tout  $p > 1$ .

### Remarque 1.1.

Il est intéressant, si possible, dans le cas où la donnée  $f$  est stochastique, de la développer sur le même système de variables aléatoires  $(Y_n)_{n=1}^N$  décrivant le coefficient de diffusion  $k$ . Cela permet de réduire la dimension stochastique du problème, c'est-à-dire le nombre des variables aléatoires décrivant la solution  $u$ .

## 1.1.2 Erreur de troncature sur les paramètres

En pratique les informations sur les paramètres du modèle  $k$  et  $f$  dont on dispose sont limitées. Par exemple, la moyenne et la fonction de corrélation (covariance) des champs peuvent être connues. Les champs aléatoires sont considérés dans un espace de dimension infinie, ainsi une approximation de ces paramètres dans un espace de faible dimension stochastique est une étape cruciale pour l'approximation du problème (1.1).

L'erreur commise par une perturbation des entrées du modèle est indépendante de la méthode numérique d'approximation utilisée tout en affectant le processus d'approximation. La quantification de l'erreur commise par une troncature des paramètres dans

le cas d'un champ aléatoire de dimension infinie est un point important et crucial pour rendre la quantification des incertitudes assez fiable.

La proposition suivante donne l'erreur commise par une troncature du paramètre  $k$  et de la donnée  $f$  dans la norme de l'espace  $L^2(\Omega; H_0^1(D))$ .

**Proposition 1.1.** ([5])

*Considérons  $u$  la solution de (1.1), et soit  $u_N$  la solution du même problème avec un champ de diffusion tronqué  $k_N$  et une donnée tronquée  $f_N$ . Alors il existe une constante  $C$  indépendante de l'ordre de troncature  $N$ , telle que l'estimation suivante soit vérifiée :*

$$\|u - u_N\|_{L^2(\Omega; H_0^1(D))} \leq C \left( \|k - k_N\|_{L^\infty(\Omega; L^\infty(D))} + \|f - f_N\|_{L^2(\Omega; L^2(D))} \right). \quad (1.7)$$

### Cas d'une troncature de Karhunen-Loève

L'estimation (1.7) exprime la continuité du problème (1.1) par rapport aux coefficients  $k$  et  $f$ . La quantité  $\eta = \|k - k_N\|_{L^\infty(\Omega; L^\infty(D))} + \|f - f_N\|_{L^2(\Omega; L^2(D))}$  est une borne supérieure de l'erreur  $u - u_N$ . Cette borne supérieure peut être explicitement contrôlée dans le cas d'une troncature de Karhunen-Loève. En effet, considérons le développement (K-L) donné dans (1.4) pour  $k$  et  $f$  :

$$k(\omega, x) = \mathbb{E}[k](x) + \sum_{n=1}^{\infty} \sqrt{\lambda_n} \phi_n(x) Y_n(\omega), \quad (1.8)$$

$$f(\omega, x) = \mathbb{E}[f](x) + \sum_{n=1}^{\infty} \sqrt{\hat{\lambda}_n} \hat{\phi}_n(x) Z_n(\omega), \quad (1.9)$$

où les couples  $\{(\lambda_n, \phi_n)\}_{n=1}^{\infty}$ ,  $\{(\hat{\lambda}_n, \hat{\phi}_n)\}_{n=1}^{\infty}$  sont respectivement le spectre de l'opérateur de covariance de  $k$  et  $f$ . Par ailleurs, supposons que les variables aléatoires  $\{Y_n\}_{n=1}^{\infty}$  soient uniformément bornées, nous avons l'estimation suivante sur  $k$  :

$$\|k - k_N\|_{L^\infty(\Omega; L^\infty(D))} \leq C_k \left( \sum_{n=N+1}^{\infty} \sqrt{\lambda_n} \|\phi_n\|_{\infty} \right);$$

et l'estimation sur  $f$  est donnée par :

$$\|f - f_N\|_{L^2(\Omega; L^2(D))} \leq \sqrt{\sum_{n=N+1}^{\infty} \hat{\lambda}_n},$$

La décroissance des valeurs propres de l'opérateur de covariance dépend étroitement de la régularité du noyau de l'opérateur. La proposition qui suit montre que la décroissance est algébrique quand la covariance a une régularité de Sobolev  $H^p$ , et qu'elle est exponentielle pour une covariance analytique [29, 56].

**Proposition 1.2.** [29, 56]

Soit  $\{(\lambda_n, \phi)\}_{n=1}^\infty$  le spectre de l'opérateur de covariance de  $k$  donné par (1.8),

1) Il existe deux constantes positives  $c_1$  et  $c_2$ , telle que pour tout  $n \geq 1$  les  $\lambda_n$  vérifient, selon la régularité de la covariance :

– lorsque le noyau de covariance est analytique par morceaux :

$$\lambda_n \leq c_1 e^{-c_2 n^{\frac{1}{d}}}.$$

– lorsque le noyau de covariance est  $H^p$  par morceaux :

$$\lambda_n \leq c_1 n^{-\frac{p}{d}}.$$

2) Pour tout  $s > 0$  il existe  $c_s > 0$  telle que pour tout  $n > 1$ , les fonctions  $\phi_n$  vérifient

$$\|\phi_n\|_\infty \leq c_s (\lambda_n)^{-s}.$$

Par conséquent, connaissant la régularité de la fonction de covariance des champs aléatoires  $k$  et  $f$ , la proposition précédente donne une estimation globale de l'erreur de troncature de Karhunen-Loève. En particulier lorsque la donnée  $f$  a une covariance analytique, l'erreur de troncature  $\|u - u_N\|$  est dominée par celle du champ de diffusion  $\|k - k_N\|$  selon la régularité de la covariance de  $k$  comme le montrent les estimations suivantes.

1. Lorsque le noyau de covariance est analytique par morceaux. Alors pour tout  $s < \frac{1}{2}$ , il existe  $c_s > 0$  telle que pour tout  $N > 1$ , nous avons :

$$\|u - u_N\|_{L^2(\Omega; H_0^1(D))} \leq C_s e^{(-c_2(\frac{1}{2}-s)N^{\frac{1}{d}})},$$

2. Lorsque le noyau de covariance est  $H^p$  par morceaux. Alors pour tout  $s < 1$ , il existe  $c_s > 0$  telle que pour tout  $N > 1$ , nous avons :

$$\|u - u_N\|_{L^2(\Omega; H_0^1(D))} \leq c_s N^{1-p\frac{(1-s)}{d}},$$

Dans le cas d'un champ log-normal et une donnée déterministe, on trouve dans [17] des estimations de l'erreur forte qui est de l'ordre  $O(N^{\frac{1}{2}-s})$  avec  $s < \frac{1}{2}$ , et une estimation de l'erreur faible commise sur la moyenne de l'ordre  $O(N^{-1})$ .

Dans le prochain paragraphe, nous allons transformer le problème (1.1) en un autre qui est déterministe. Pour cela, nous introduisons quelques notations.

**Notations :**

Dans la suite, on note par  $\Gamma_n = Y_n(\Omega)$  l'image de  $\Omega$  par  $Y_n$  pour chaque  $n = 1, \dots, N$ , et par  $\Gamma = \prod_{n=1}^N \Gamma_n \subset \mathbb{R}^N$  l'image  $\mathbf{Y}(\Omega)$ . On suppose que le vecteur aléatoire  $\mathbf{Y}$  possède une densité de probabilité conjointe connue  $\varrho \in L^\infty(\Gamma)$ . On suppose aussi que la famille des variables aléatoires  $\{Y_n\}_{n=1}^N$  sont indépendantes, ainsi dans ce cas la densité de probabilité se factorise comme  $\varrho = \prod_{n=1}^N \rho_n$ , où chaque  $\rho_n$  est la densité de  $Y_n$ .

**1.1.3 Formulation faible et formulation forte**

Le champ aléatoire  $k$  et la donnée  $f$  sont paramétrisés par le vecteur aléatoire  $\mathbf{Y}$  par  $k(\omega, x) = k(Y_1, \dots, Y_N; x)$  et  $f(\omega, x) = f(Y_1, \dots, Y_N; x)$ . Par le lemme de Doob-Dynkin, la solution  $u$  du problème (1.1) est aussi paramétrisée par le même vecteur aléatoire  $\mathbf{Y}$  tel que  $u(\omega, x) = u(Y_1, \dots, Y_N; x)$ . Par conséquent le problème variationnel faible (1.3) est équivalent à la formulation faible et déterministe suivante :

$$\left\{ \begin{array}{l} \text{Trouver } u \in L^2(\Gamma) \otimes H_0^1(D) \text{ tel que} \\ \int_{\Gamma} (k \cdot \nabla u, \nabla \mu)_{L^2(D)} \varrho dy = \int_{\Gamma} (f, \mu)_{L^2(D)} \varrho dy, \quad \forall \mu \in L^2(\Gamma) \otimes H_0^1(D). \end{array} \right. \quad (1.10)$$

La preuve de l'existence et l'unicité du problème (1.10) s'appuie sur l'hypothèse **H2**, qui affirme que la diffusion  $k$  est strictement positive et bornée uniformément sur  $\Omega$ , c'est-à-dire que les deux constantes  $k_{min}$  et  $k_{max}$  sont indépendantes de tous les événements de  $\Omega$ . Une application directe du théorème de Lax-Milgram, montre que (1.10) est bien posé.

Pour certaines classes de méthodes d'approximation du problème (1.1), telles que la méthode de Monte-Carlo et la méthode de collocation stochastique, on introduit le problème faible-spatial obtenu en projetant (1.1) dans l'espace  $H_0^1(D)$ , donné par la formulation forte (problème fort-stochastique) suivante :

$$\left\{ \begin{array}{l} \text{Trouver } u(y, \cdot) \in H_0^1(D) \text{ tel que} \\ \int_D k(y) \nabla u \cdot \nabla v dx = \int_D f v dx \quad \forall v \in H_0^1(D), \text{ pour p.s } y \in \Gamma. \end{array} \right. \quad (1.11)$$

Une simple application du théorème de Lax-Milgram permet d'affirmer que pour chaque

$y \in \Gamma$  (p.s), le problème (1.11) est bien posé. La solution  $u$  satisfait l'estimation suivante :

$$\|u(y, \cdot)\|_{H_0^1(D)} \leq C_D k_{min}^{-1} \|f(y, \cdot)\|_{L^2(D)}, \quad (1.12)$$

où  $C_D$  est la constante de Poincaré.

On conclut que, sous l'estimation (1.12), la solution  $u$  est dans  $L_q^p(\Gamma) \otimes H_0^1(D)$  si  $f$  est dans  $L_q^p(\Gamma) \otimes L^2(D)$ .

Dans le cas où  $k_{min}$ , la borne inférieure de  $k$ , dépend de  $\omega$ , alors  $u$  est dans  $L_q^p(\Gamma) \otimes H_0^1(D)$  si  $f$  est déterministe et  $k_{min}^{-1}$  est dans  $L_q^p(\Gamma)$ .

Dans le cas où les deux fonctions  $f$  et  $k_{min}$  sont aléatoires, alors pour tout  $p > 1$ ,  $u \in L_q^p(\Gamma) \otimes H_0^1(D)$  si  $f \in L_q^{2p}(\Gamma) \otimes L^2(D)$  et  $k_{min}^{-1} \in L^{2p}(\Omega)$ . En effet :

$$\mathbb{E}[\|u\|_{H_0^1(D)}^p] \leq C_D \mathbb{E}[k_{min}^{-p} \|f\|_{L^2(D)}^p] \leq C_D \sqrt{\mathbb{E}[(k_{min}^{-1})^{2p}] \mathbb{E}[\|f\|_{L^2(D)}^{2p}]}.$$

Nous avons aussi le même résultat si les deux champs aléatoires  $k$  et  $f$  sont indépendants dans  $L^p$ .

### 1.1.4 Régularité stochastique

L'approximation d'une fonction  $\varphi$  (quelconque) par des polynômes globaux est une démarche naturelle que l'on rencontre dans divers contextes de l'analyse. Lorsque la fonction  $\varphi$  est assez régulière (fonction analytique), cette approximation polynomiale est plus précise.

La convergence et la compétitivité des méthodes spectrales, en particulier celles de collocation stochastique, sont basées sur une régularité stochastique de la solution du problème (1.1). Cette régularité concerne l'analyticité de la fonction  $u$  suivant la variable stochastique  $y$  [6, 49, 50].

Ainsi dans la technique d'adaptation des méthodes de collocation que nous allons considérer, nous supposons que  $u$  peut être prolongeable analytiquement dans une région du plan complexe  $\mathbb{C}$ .

D'une manière similaire à [49, 50], nous allons nous appuyer dans la suite sur une analyse unidimensionnelle suivant chaque direction  $y_n$  pour chaque  $n = 1, \dots, N$ . Pour

ce faire, nous introduisons la notation suivante : pour chaque  $n = 1, \dots, N$ , on pose  $\Gamma_n^* = \prod_{j \neq n}^N \Gamma_j$ , et tout élément  $y \in \Gamma$ , peut s'écrire en  $y = (y_n, y_n^*)$  où  $y_n \in \Gamma_n$  et  $y_n^* \in \Gamma_n^*$ .

On suppose que la solution de (1.1) satisfait l'hypothèse de régularité suivante :

**Hypothèse 1.2.** *Pour chaque  $y_n \in \Gamma_n$ , il existe un réel  $\tau_n > 0$  pour lequel la fonction  $u(y_n, y_n^*, \cdot)$  comme fonction de  $y_n$ ,  $u : \Gamma_n \rightarrow \mathcal{C}^0(\Gamma_n^*; H_0^1(D))$  admet un prolongement analytique  $u(z, y_n^*, x)$ , pour tout  $z$  dans l'ensemble complexe  $\Sigma_n(\Gamma_n; \tau_n) \in \mathbb{C}$  donné par :*

$$\Sigma_n(\Gamma_n; \tau_n) = \{z \in \mathbb{C}, \text{dist}(z, \Gamma_n) \leq \tau_n\}.$$

De plus, il existe  $\eta > 0$  qui ne dépend pas de  $n$ , tel que pour tout  $z \in \Sigma_n(\Gamma_n; \tau_n)$ ,

$$\|u(z)\|_{\mathcal{C}^0(\Gamma_n^*; H_0^1(D))} \leq \eta.$$

L'hypothèse précédente peut être vérifiée pour les champs aléatoires présentés dans la section précédente, ainsi que lorsque le paramètre de diffusion  $k$  et la donnée  $f$  sont des fonctions analytiques par rapport à chaque variable  $y_n$ . En effet, si  $k$  et  $f$  sont analytiques dans un disque de rayon  $r_n$ , alors il existe deux nombres  $M > 0$ ,  $\gamma_n > 0$  finis, tels que :

$$\|\partial_{y_n}^m k(y, \cdot)\|_{L^\infty(D)} \leq M \gamma_n^m m! \quad \text{et} \quad \|\partial_{y_n}^m f(y, \cdot)\|_{L^2(D)} \leq M \gamma_n^m m! \quad (1.13)$$

Les deux bornes (1.13) sont exactement les hypothèses données dans (cf [6] Lemme 5) pour que la solution  $u$  satisfasse l'hypothèse de régularité 1.2. La convergence des méthodes spectrales pour la résolution de (1.1) repose sur la régularité de la fonction  $u$ , cette convergence est exponentielle et son taux dépend du diamètre  $\tau_n$  [6, 5, 49, 50].

### 1.1.5 Espace d'approximation déterministe

L'approximation du problème (1.1) est basée sur deux approximations, l'une par rapport à la variable stochastique et l'autre suivant la variable déterministe. Généralement deux grandes classes de méthodes sont utilisées pour l'approximation de (1.1), les méthodes intrusives et les méthodes non-intrusives [41], selon la manière dont l'approximation dans l'espace stochastique  $L^2(\Gamma)$  est considérée. L'approximation dans l'espace déterministe  $H_0^1(D)$  reste la même pour ces méthodes.

Nous considérons dans la suite un sous-espace d'éléments finis standard  $V_h(D) \subset H_1^0(D)$ , basés sur les fonctions d'approximation linéaire par morceaux définies sur une triangulation régulière  $\mathcal{T}_h$  de pas de discrétisation  $h > 0$ .

Nous avons l'estimation suivante qui concerne l'erreur commise en approchant une fonction  $v \in H_0^1(D)$  par  $v_h$  dans  $V_h(D)$  ([18]).

$$\|v - v_h\|_{H_0^1(D)} \leq Ch\|v\|_{H^2}, \quad (1.14)$$

où la constante  $C$  ne dépend pas du pas  $h$ .

Dans la suite de ce chapitre, nous décrivons les méthodes d'approximation de (1.1), données par les deux approches intrusives et non-intrusives.

- Méthodes non-intrusives : La résolution est basée sur une suite de problèmes déterministes, correspondant à des réalisations (aléatoires comme Monte Carlo ou déterministes comme la collocation) judicieusement choisies du vecteur  $\mathbf{Y}$ . Leur avantage principal est que les codes déterministes existants ne nécessitent aucune adaptation particulière et peuvent être utilisés comme une boîte noire, ce qui apporte une grande flexibilité de la mise en œuvre en parallèle pour ces méthodes.
- Méthodes intrusives : Contrairement aux méthodes non-intrusives basées sur une batterie de réalisations de  $\mathbf{Y}$  pour déterminer les quantités d'intérêt du modèle, les méthodes intrusives de type 'Galerkin' sont basées sur un formalisme proche de celui des codes déterministes. Elles cherchent une approximation de la réponse du modèle à la fois dans l'espace stochastique et déterministe, ainsi une adaptation des codes déterministes existants s'impose pour ces méthodes suivant la nature des problèmes physiques.

## 1.2 Méthode d'échantillonnage de Monte-Carlo

La méthode d'échantillonnage de Monte-Carlo est la méthode la plus attractive pour une approximation des statistiques d'intérêt d'un problème stochastique, notamment le problème (1.1). La méthode présente à la fois une grande simplicité et une généralité pour n'importe quel problème stochastique, et naturellement bien adaptée à la parallélisation. Bien que la méthode dépende légèrement de la dimension stochastique, elle n'exploite pas la régularité des données du modèle pour produire une convergence rapide des quantités approchées.

Pour l'approximation de la moyenne des quantités d'intérêt  $\mathbb{E}[Q(u)]$ , l'approche classique de Monte-Carlo consiste d'abord à choisir  $M$  réalisations  $\{\mathbf{Y}(\omega_j)\}_{j=1}^M$  iid (indépendantes et identiquement distribuées) du vecteur aléatoire  $\mathbf{Y}$ . Ensuite pour chaque réalisation  $\mathbf{Y}(\omega_j)$ , on considère  $u_h(\omega_j, \cdot)$  solution Éléments Finis standard du problème



déterministe suivant :

$$\begin{cases} -\operatorname{div}(k(\omega_j, \cdot) \nabla u(\omega_j, \cdot)) = f(\omega_j, \cdot) & \text{dans } D \\ u(\omega_j, \cdot) |_{\partial D} = 0, \end{cases} \quad (1.15)$$

$\mathbb{E}[Q(u)]$  est alors approchée par la moyenne empirique suivante :

$$\mathcal{A}(Q(u_h); M) := \frac{1}{M} \sum_{j=1}^M Q(u_h)(\omega_j, \cdot).$$

L'erreur de la méthode est naturellement donnée par deux parties, l'une concerne l'erreur d'échantillonnage et l'autre concerne l'erreur de discrétisation spatiale :

$$\begin{aligned} \mathbb{E}[Q(u)] - \mathcal{A}(Q(u_h); M) &= (\mathbb{E}[Q(u)] - \mathbb{E}[Q(u_h)]) + (\mathbb{E}[Q(u_h)] - \mathcal{A}(Q(u_h); M)) \\ &:= \mathcal{E}_h + \mathcal{E}_{h,M}. \end{aligned}$$

Dans cette erreur, le pas de discrétisation spatiale  $h$  contrôle l'erreur de discrétisation  $\mathcal{E}_h$ , tandis que le nombre de simulations  $M$  contrôle l'erreur d'échantillonnage  $\mathcal{E}_{h,M}$ . Par exemple un résultat d'estimation qu'on peut trouver dans la référence [5], concernant l'approximation de la moyenne  $\mathbb{E}[u]$ , est donné dans la proposition suivante :

**Proposition 1.3.** ([5])

*Considérons  $u$  solution de (1.1). Pour chaque  $j = 1, \dots, M$  soit  $u_h(\omega_j)$  la projection de (1.15) sur  $V_h(D)$ , alors nous avons l'estimation suivante :*

$$\left\| \mathbb{E}[u] - \mathcal{A}(u_h; M) \right\|_{H_0^1(D)} \leq C \left( h + \frac{1}{\sqrt{M}} \right).$$

La constante  $C > 0$  est indépendante de  $h$  et  $M$ .

On constate alors que l'inconvénient principal de l'approche de Monte Carlo est sa convergence d'ordre  $O(\frac{1}{\sqrt{M}})$ , donc assez lente, le coût global d'approximation étant  $M$  fois celui du problème déterministe (1.15).

De nombreuses versions de la méthode de Monte-Carlo ont été proposées pour améliorer la convergence, et réduire ainsi le coût de calcul global. Par exemple la méthode de type 'Quasi Monte-Carlo' [14, 59] dont l'erreur de convergence est de l'ordre  $O\left(\frac{\log(M)^N}{M}\right)$ . Cependant, cet avantage d'accélération pourrait se détériorer lorsque  $N$ , le nombre des variables aléatoires de paramétrisation du problème est important. On note aussi que dans ce cas, la construction des suites de points de discrédance faible devient de plus en plus compliquée.

### 1.3 Méthodes de type approximation locale

Les méthodes de type d'approximation locale comme la méthode de perturbation ([38]) offrent une alternative efficace de la méthode de Monte Carlo. Elles permettent d'approcher avec un moindre effort, les deux premiers moments de la solution du problème stochastique. La méthode de type petites perturbations utilise un développement en série de Taylor de la fonction  $u$  autour de la moyenne de  $\mathbf{Y}$ ,  $\mu_{\mathbf{Y}} = \mathbb{E}[\mathbf{Y}]$ . La série de Taylor de  $u$  n'est considérée que pour le premier ou le deuxième ordre :

$$u(\mathbf{Y}, \cdot) \approx u(\mu_{\mathbf{Y}}, \cdot) + \langle \mathbf{Y} - \mu_{\mathbf{Y}}, \frac{\partial u}{\partial \mathbf{Y}} \rangle + \frac{1}{2} \langle (\mathbf{Y} - \mu_{\mathbf{Y}}) \cdot (\mathbf{Y} - \mu_{\mathbf{Y}})', \frac{\partial^2 u}{\partial^2 \mathbf{Y}} \rangle, \quad (1.16)$$

puisque pour des ordres plus élevés, la méthode nécessite le calcul des dérivées par rapport à la variable stochastique, qui est généralement très lourdement coûteux. En substituant le développement (1.16) dans le problème (1.3), on peut obtenir les coefficients de (1.16) comme solution d'une suite de  $(N + N^2 + 1)$  problèmes d'EDP déterministes..

Babuška et Chatzipantelidis dans [4] proposent une méthode de perturbation avec des approximations successives dont la convergence est basée sur la coercivité uniforme du coefficient de diffusion  $k$ .

Ces méthodes de perturbation sont généralement applicables uniquement pour des problèmes dont les coefficients ont des petites variations et avec peu d'incertitudes i.e le cas avec des petites variances.

### 1.4 Méthode des éléments finis stochastiques

La méthode des éléments finis stochastiques pour un problème d'EDP stochastique, notamment (1.1), suit le même principe que l'approche de Galerkin pour les problèmes déterministes. Elle définit dans l'espace stochastique une approximation de  $u$  représentée sur une certaine base d'approximation (polynômes de Chaos, base d'ondelettes, polynômes par morceaux) [34, 43, 5, 21].

La méthode discrétise le problème faible (1.10) et traite la variable stochastique et déterministe séparément, en introduisant un espace Élément Fini construit par le tensoriel  $W_p(\Gamma) \otimes V_h(D)$ , et une discrétisation de (K.L) des paramètres  $k$  et  $f$ .

Le sous-espace  $V_h(D)$  est un sous-espace d'approximation d'Éléments Finis classique

de l'espace  $H_0^1(D)$ , et  $W_p(\Gamma)$  est un espace d'approximation acceptable de l'espace stochastique  $L^2(\Gamma)$ . Dans le cas des polynômes de Chaos de degré total  $p$ , l'espace  $W_p(\Gamma)$  est construit comme suit :

$$W_p(\Gamma) := \text{Vect} \left\{ \psi_\beta(\mathbf{Y}) := \left( \prod_{n=1}^N \psi_{\beta_n}(Y_n), \psi_{\beta_n} \in P_{\beta_n}(\Gamma_n) \right) \right\},$$

où  $\beta \in \mathbb{N}^N$  est tel que  $|\beta| := \sum_{n=1}^N \beta_n \leq p$ , et  $P_{\beta_n}(\Gamma_n)$  l'espace des polynômes sur  $\Gamma_n$  de degré au plus égal à  $\beta_n$ , et les polynômes  $\{\psi_{j_n}\}_{j_n=0}^{\beta_n}$  sont orthogonaux suivant la densité de probabilité  $\rho_n$  de chaque variable aléatoire  $Y_n$ .

En projetant le problème (1.10) dans l'espace  $W_p(\Gamma) \otimes V_h(D)$ , on obtient une approximation de  $u$  sous la forme :

$$u_{p,h}(\omega, x) = \sum_{|\beta| \leq p} u_\beta^h(x) \psi_\beta(Y),$$

comme solution de la projection suivante :

$$\int_\Gamma \int_D k(y, x) \nabla u_{p,h}(y, x) \nabla v(y, x) \varrho(y) dx dy = \int_\Gamma \int_D f(y, x) v(y, x) \varrho(y) dx dy.$$

Chaque coefficient  $u_j^h$  est une fonction dans  $V_h(D)$  et le système  $\{\psi_j(Y)\}_{j=1}^{\mathcal{J}}$  forme une base de l'espace  $W_p(\Gamma)$ . Cette projection conduit à la résolution d'un système linéaire de grande taille dont la matrice de rigidité  $A$  et le vecteur second membre  $F$  sont donnés par les expressions tensorielles suivantes :

$$A = G_0 \otimes K_0 + \sum_{n=1}^N G_n \otimes K_n, \quad (1.17)$$

$$F = g_0 \otimes f_0 + \sum_{n=1}^N g_n \otimes f_n, \quad (1.18)$$

les éléments  $(G_m)_{i,j}$  de chaque matrice stochastique  $G_m$  sont donnés par  $(G_m)_{i,j} = \mathbb{E}[Y_m \psi_i(Y) \psi_j(Y)]$ , où chaque  $\psi_i$  est la  $i$ ème fonction de base de  $W_p(\Gamma)$ . Les matrices  $K_m$  sont les matrices Éléments Finis déterministes, les éléments de chaque vecteur  $g_n$  sont donnés par  $(g_n)_l = \mathbb{E}[Y_n \psi_l(Y)]$ .

On constate qu'en général, le problème variationnel (1.10) conduit à un système linéaire entièrement couplé, dont la dimension est  $\dim(W_p(\Gamma)) \times \dim(V_h)$ . La méthode ainsi exige des techniques très efficaces et des calculs parallèles pour la résolution numérique des systèmes linéaires résultants [42].

Toujours pour un coefficient  $k$  linéaire par rapport à  $y$ , Babuška et al. proposent dans [5] un choix particulier des fonctions de base de l'espace  $W_p(\Gamma)$ , appelée base de

polynômes bi-orthogonale. Ce choix permet de découpler le système linéaire en  $P$  systèmes déterministes. La base proposée est obtenue en résolvant une suite de problèmes aux valeurs propres, (pour plus de détails voir [5, 6]).

Sous des hypothèses légèrement plus fortes que (H1), la méthode présente une convergence géométrique, donnée par l'estimation suivante :

**Théorème 1.1.** ([5])

Soit  $u$  solution de (1.1), et  $u_{p,h}$  la projection de  $u$  dans l'espace  $W_p(\Gamma) \otimes V_h(D)$ . Alors il existe une constante  $C > 0$  telle que pour tout  $\tau \in ]0; 1[$ , il existe  $(r_n)_{n=1}^N \in ]0; 1[$  tels que pour tout  $h$  et  $p = (p_1, \dots, p_N) \in \mathbb{N}^N$  on ait :

$$\|u - u_{p,h}\|_{L^2(\Gamma, H_0^1(D))} \leq C \left( h + \frac{1}{\tau} \sum_{n=1}^N (r_n)^{p_n+1} \right).$$

## 1.5 Méthode de collocation stochastique

L'idée de la méthode de collocation stochastique consiste à calculer des réalisations de la solution en résolvant (1.11) sur une grille de points déterministes et prédéterminés (points de collocation). Une approximation de la solution est ainsi construite par l'interpolation polynomiale de Lagrange.

La méthode est complètement parallélisable et en contraste avec Monte Carlo, elle exploite la régularité des paramètres  $k$  et  $f$  par rapport à la variable  $y$  pour obtenir des taux de convergence beaucoup plus rapides.

L'approche de collocation construit une approximation  $u_{h,p}$  de la solution  $u$  de (1.1) dans l'espace d'éléments finis  $W_{p,h}$ . Cet espace est donné par le produit tensoriel des sous-espaces  $W_{p,h} := \mathcal{P}_p(\Gamma) \otimes V_h(D)$ , où  $V_h(D)$  l'espace classique des éléments finis de  $H_0^1(D)$  et le sous-espace  $\mathcal{P}_p(\Gamma)$  est formé par les polynômes à  $N$  variables de degré au plus  $p = (p_1, \dots, p_N)$ .

On introduit l'approximation semi-discrète  $u_h : \Gamma \rightarrow V_h(D)$ , obtenue par la projection de (1.11) dans  $V_h(D)$ , (p.s) pour tout  $y \in \Gamma$  :

$$\begin{cases} \text{trouver } u_h(y, \cdot) \in V_h(D) \text{ telle que} \\ \int_D k(y, \cdot) \nabla u_h(y, x) \cdot \nabla v_h(x) dx = \int_D f(y, x) v_h(x) dx, \quad \forall v_h \in V_h(D). \end{cases} \quad (1.19)$$

On résout (1.19) pour des points convenables d'un ensemble  $\{y_j\}_j$  dans  $\Gamma$ . On obtient ainsi une famille de réalisations  $\{u_h(y_j, \cdot)\}_j$ , permettant de construire  $u_{h,p}$  par une interpolation polynomiale de Lagrange. Cette approximation peut être obtenue soit par

une tensorisation complète, soit en utilisant une interpolation sur des grilles clairsemées basées sur la construction de Smolyak [58, 62].

### 1.5.1 Interpolation par une tensorisation pleine

L'interpolation par un produit tensoriel complet construit la grille des points de collocation par un produit cartésien complet de tous les points unidimensionnels. Ainsi, ce produit tensoriel complet conduit à une approximation polynomiale de plus haut degré suivant chaque dimension  $y_n$ .

Pour chaque  $n = 1, \dots, N$ , soit  $i_n \in \mathbb{N}$  un niveau d'interpolation, et soit  $\mathcal{X}^{i_n} := \{y_{n,1}, \dots, y_{n,m(i_n)}\} \subset \Gamma_n$  l'ensemble des points de collocation suivant la direction  $y_n$ .

Pour  $v \in \mathcal{C}^0(\Gamma_n; H_0^1(D))$ , on introduit la suite des opérateurs d'interpolation unidimensionnelle de Lagrange  $\mathcal{U}^{i_n} : \mathcal{C}^0(\Gamma_n; H_0^1(D)) \rightarrow \mathcal{P}_{p_n}(\Gamma_n) \otimes H_0^1(D)$  définis par :

$$\mathcal{U}^{i_n}(v)(z, x) = \sum_{j=1}^{m(i_n)} v(y_{n,j}, x) \ell_{n,j}^{i_n}(z) \quad \forall v \in \mathcal{C}^0(\Gamma_n; H_0^1(D)), \quad (1.20)$$

où chaque  $\ell_{n,j}^{i_n} \in \mathcal{P}_{p_n}(\Gamma_n)$  est le polynôme de Lagrange de degré  $p_n = m(i_n) - 1$  :

$$\ell_{n,j}^{i_n}(z) = \prod_{k=1, k \neq j}^{m(i_n)} \frac{z - y_{n,k}}{y_{n,j} - y_{n,k}} \quad z \in \Gamma_n.$$

On définit l'opérateur d'interpolation de Lagrange de produit tensoriel complet par

$$\begin{aligned} \mathcal{A}_i : \mathcal{C}^0(\Gamma; H_0^1(D)) &\longrightarrow \mathcal{P}_p(\Gamma) \otimes H_0^1(D) \\ \mathcal{A}_i v(y, x) &:= \mathcal{U}^{i_1} \otimes \dots \otimes \mathcal{U}^{i_N} v(y, x) = \sum_{\mathbf{j} \leq \mathbf{m}(\mathbf{i})} v(y_{\mathbf{j}}, x) L_{\mathbf{j}}^{\mathbf{i}}(y), \end{aligned} \quad (1.21)$$

où  $p = (p_1, \dots, p_N)$ ,  $\mathcal{P}_p(\Gamma) = \otimes_{n=1}^N \mathcal{P}_{p_n}(\Gamma_n)$ , et nous avons posé  $\mathbf{j} = (j_1, \dots, j_N)$ ,  $\mathbf{m}(\mathbf{i}) = (m(i_1), \dots, m(i_N))$  et les points d'interpolation dans  $\Gamma$ ,  $y_{\mathbf{j}} = (y_{1,j_1}, \dots, y_{n,j_n}, \dots, y_{N,j_N})$  avec  $y_{n,j_n} \in \Gamma_n$  l'un des nœuds dans  $\mathcal{X}^{i_n}$ . Pour chaque  $y = (y_1, \dots, y_N) \in \Gamma$ , nous avons posé  $L_{\mathbf{j}}^{\mathbf{i}}(y) = \otimes_{n=1}^N \ell_{n,j_n}^{i_n}(y_n)$  le polynôme multi-dimensionnel de Lagrange.

On remarque que le calcul de l'interpolation  $\mathcal{A}_i v$  impose l'évaluation de  $v$  sur la grille des nœuds :

$$\mathcal{H}_{\infty}^{\mathbf{i}} = \mathcal{X}^{i_1} \times \dots \times \mathcal{X}^{i_N},$$

de cardinal  $M = \prod_{n=1}^N m(i_n)$ .

Le choix des points de collocation s'avère primordial pour la qualité des approximations et les propriétés de convergence, en particulier lorsque le nombre des points de collocation augmente ([1]). L'utilisation des points de Gauss et de Clenshaw-Curtis est préférée à celle des points équidistants (phénomène de Runge), car ils offrent des meilleures propriétés d'interpolation. Les points de Clenshaw-Curtis présentent également l'avantage d'être une suite de points imbriqués ; lorsqu'on double l'ordre, cela peut être intéressant pour une approximation adaptative.

En utilisant la formule d'interpolation (1.21) appliquée à la solution  $u_h \in \mathcal{C}^0(\Gamma; V_h(D))$  de (1.19), on peut approcher efficacement l'espérance  $\mathbb{E}[u]$  et la variance  $\text{var}[u]$  de la fonction  $u$ , de la manière suivante :

$$\begin{aligned}\mathbb{E}[u] &\approx \bar{u}_h := \sum_{\mathbf{j} \leq m(\mathbf{i})} u_h(\mathbf{y}_{\mathbf{j}}, \cdot) w_{\mathbf{j}}, \\ \text{var}[u] &\approx \sum_{\mathbf{j} \leq m(\mathbf{i})} u_h(\mathbf{y}_{\mathbf{j}}, \cdot)^2 w_{\mathbf{j}} - (\bar{u}_h)^2,\end{aligned}$$

où les poids d'intégration multidimensionnelle  $w_{\mathbf{j}}$  sont donnés par la tensorisation :

$$w_{\mathbf{j}} = \prod_{n=1}^N \int_{\Gamma_n} \ell_{n, j_n}^{i_n}(y) \rho_n(y) dy.$$

L'erreur commise en approchant la solution  $u$  par l'interpolation  $\mathcal{A}_{\mathbf{i}} u_h$  est naturellement donnée par deux parties, la première concerne l'erreur d'interpolation polynomiale dans  $L^2_{\varrho}(\Gamma)$  et la deuxième concerne l'erreur classique de la discrétisation spatiale.

$$\|u - \mathcal{A}_{\mathbf{i}} u_h\|_{L^2_{\varrho}(\Gamma; H^1_0(D))} \leq \|u - \mathcal{A}_{\mathbf{i}} u\|_{L^2_{\varrho}(\Gamma; H^1_0(D))} + \|\mathcal{A}_{\mathbf{i}} u - \mathcal{A}_{\mathbf{i}} u_h\|_{L^2_{\varrho}(\Gamma; H^1_0(D))}. \quad (1.22)$$

Le théorème suivant fournit un résultat de convergence de la méthode de collocation [6]. Sous l'hypothèse 1.2 de la régularité et avec l'utilisation des points de Gauss, l'estimation du premier terme  $\|u - \mathcal{A}_{\mathbf{i}} u\|_{L^2_{\varrho}(\Gamma; H^1_0(D))}$  dans (1.22) montre que la convergence est exponentielle.

**Théorème 1.2.** ([6])

*Soit  $u$  la solution exacte de (1.1) et  $\mathcal{A}_{\mathbf{i}} u$  son approximation par la formule (1.21). On suppose que  $u$  admet une extension analytique suivant chaque direction dans la région complexe  $\Sigma_n(\Gamma_n; \tau_n) := \{z \in \mathbb{C}, \text{dist}(z, \Gamma_n) \leq \tau_n\}$ , de taille  $\tau_n > 0$ .*

*Alors il existe une constante  $C > 0$  indépendante de chaque  $p_n = m(i_n) - 1$ , telle que :*

$$\|u - \mathcal{A}_{\mathbf{i}} u\|_{L^2_{\varrho}(\Gamma) \otimes H^1_0(D)} \leq C \sum_{n=1}^N \beta_n(p_n) e^{-\nu_n p_n^{\varrho_n}}, \quad (1.23)$$

où

$$\begin{aligned}
 & \bullet \text{ si } \Gamma_n \text{ est borné } \begin{cases} \theta_n = \beta_n = 1 \\ \nu_n = \log \left( \frac{2\tau_n}{|\Gamma_n|} \left( 1 + \sqrt{1 + \frac{|\Gamma_n|^2}{4\tau_n^2}} \right) \right). \end{cases} \\
 & \bullet \text{ si } \Gamma_n \text{ est non borné } \begin{cases} \theta_n = 1/2, \beta_n = O(\sqrt{p_n}) \\ \nu_n = \tau_n/2. \end{cases}
 \end{aligned} \tag{1.24}$$

L'estimation de l'erreur du théorème 1.2 montre la convergence exponentielle de la méthode. Chaque terme  $\beta_n(p_n)e^{-\nu_n p_n^{\theta_n}}$  de l'estimation représente la contribution de la direction  $y_n$  à l'erreur globale. Cette contribution est primordiale pour le calcul des poids d'anisotropie et l'estimation de l'erreur de la méthode anisotrope des grilles clairsemées.

### 1.5.2 Construction de Smolyak

La construction de Smolyak fournit un outil efficace pour construire des fonctions de plusieurs variables par une interpolation de Lagrange. Elle a été introduite par Smolyak [58] pour le calcul des intégrales multidimensionnelles, puis développée pour l'interpolation des fonctions de plusieurs variables, [11, 54, 8].

L'algorithme utilise des produits tensoriels creux des formules d'interpolation unidimensionnelle, où les indices de niveaux de chaque dimension sont choisis dans l'ensemble de contraintes d'un niveau total  $w$ . Ainsi cette approche de Smolyak est donnée par une combinaison linéaire de produits tensoriels des règles d'interpolation unidimensionnelles, utilisant des approximations de faibles ordres. Ceci fournit une stratégie d'interpolation avec un nombre réduit de points de collocation, de telle manière que l'erreur d'interpolation soit la même que pour la tensorisation de produit plein.

On considère une famille d'opérateurs d'interpolation unidimensionnelle de Lagrange  $(\mathcal{U}^l)_l$ , et on définit  $\Delta_l$  l'opérateur de différence par :

$$\Delta_l = \mathcal{U}^l - \mathcal{U}^{l-1}.$$

Soit  $v$  une fonction continue d'une seule variable, alors  $\lim_{K \rightarrow \infty} \sum_{l=1}^K \Delta_l v = \lim_{K \rightarrow \infty} \mathcal{U}^K v = v$ .

Dans le cas multidimensionnel, afin de construire une approximation d'une fonction à plusieurs variables, on prend le produit tensoriel de ces opérateurs de différence, tronqués à un certain niveau acceptable.

On définit l'opérateur d'interpolation de Smolyak  $\mathcal{A}(w, N)$  de niveau d'approximation  $w \in \mathbb{N}$  par la formule suivante :

$$\mathcal{A}(w, N) = \sum_{\mathbf{i} \in X(w, N)} (\Delta_{i_1} \otimes \dots \otimes \Delta_{i_N}), \quad (1.25)$$

où  $X(w, N)$  est l'ensemble des multi-indices dans  $\mathbb{N}^N$  :

$$X(w, N) := \left\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1}, \sum_{n=1}^N (i_n - 1) \leq w \right\}. \quad (1.26)$$

Une formule équivalente à celle donnée dans (1.25) mais plus pratique ([62]) est donnée par :

$$\mathcal{A}(w, N) = \sum_{\mathbf{i} \in Y(w, N)} (-1)^{w+N-|\mathbf{i}|_1} \binom{N-1}{w+N-|\mathbf{i}|_1} \mathcal{U}^{i_1} \otimes \dots \otimes \mathcal{U}^{i_N}, \quad (1.27)$$

où  $|\mathbf{i}|_1 = \sum_{n=1}^N i_n$ , l'ensemble des indices  $Y(w, N)$  est donné par,

$$Y(w, N) := \left\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1}, w+1 \leq \sum_{n=1}^N i_n \leq w+N \right\}.$$

Le calcul de l'opérateur d'interpolation de Smolyak  $\mathcal{A}(w, N)$  donné dans (1.27) appliquée à  $v$ , nécessite une évaluation de la fonction  $v$  sur la grille clairsemée des nœuds suivante :

$$\mathcal{H}_1(w, N) = \bigcup_{\mathbf{i} \in Y(w, N)} \{\mathcal{X}^{i_1} \times \dots \times \mathcal{X}^{i_N}\}.$$

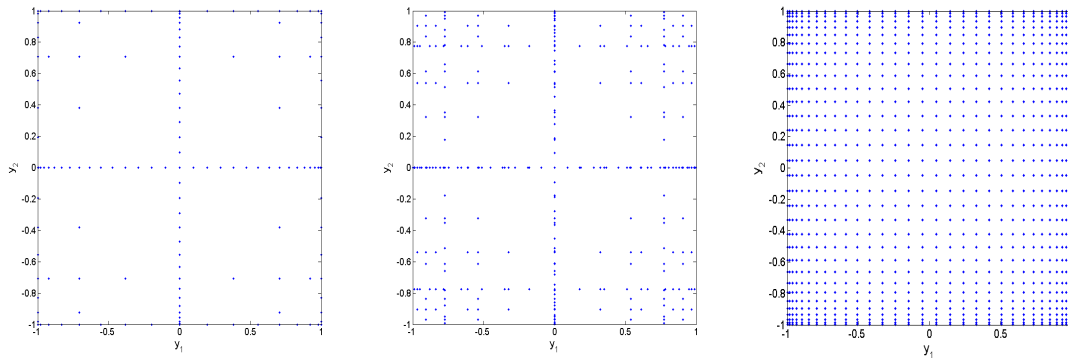


FIGURE 1.1 – Pour  $\Gamma = [-1, 1]^2$  et  $w = 5$ , grilles clairsemées isotropes avec des points de Clenshaw-Curtis (à gauche) et des points de Gauss (au milieu). À droite, grille de produit tensoriel complet, avec  $m(i_1) = m(i_2) = 2^5$ .



L'interpolation par tensorization de Smolyak est particulièrement intéressante lorsque les nœuds de collocation sont imbriqués, et c'est le cas lorsque les points de collocation unidimensionnels  $\{\mathcal{X}^{i_n}\}_{i_n=1}^{m_n}$  utilisés sur chaque dimension  $y_n$  vérifient l'imbrication  $\mathcal{X}^i \subset \mathcal{X}^{i+1}$ . Cela conduit à une réduction significative du nombre de points d'interpolation de la grille clairsemée  $\mathcal{H}_1(w, N)$ . En effet, on a alors

$$\mathcal{H}_1(w, N) = \bigcup_{|\mathbf{i}|_1=w+N} \{\mathcal{X}^{i_1} \times \dots \times \mathcal{X}^{i_N}\}.$$

Ceci est l'intérêt majeur de ces formules imbriquées permettant de réduire le coût d'une procédure d'adaptation, puisque lorsque l'on augmente  $w$ , les calculs déjà effectués sur des grilles de niveau inférieur peuvent être réutilisés. Plusieurs ensemble de points de collocation imbriqués sont disponibles selon leur niveau d'exactitude et leur cardinal : Newton-Cotes, Clenshaw-Curtis, Gauss-Patterson,...

Nous pouvons utiliser la formule d'approximation donnée dans (1.27) pour approcher l'espérance  $\mathbb{E}[u]$  et la variance  $\text{var}[u]$  de la fonction  $u$ , comme suit :

$$\begin{aligned} \mathbb{E}[u] &\approx \mathbb{E}[\mathcal{A}(w, N)u_h] = \sum_{\mathbf{i} \in Y(w, N)} \sum_{\mathbf{j} \leq \mathbf{m}(\mathbf{i})} u_h(y_{\mathbf{j}}, \cdot) \tilde{w}_{\mathbf{j}}^{\mathbf{i}}, \\ \text{var}[u] &\approx \text{var}[\mathcal{A}(w, N)u_h] = \sum_{\mathbf{i} \in Y(w, N)} \sum_{\mathbf{j} \leq \mathbf{m}(\mathbf{i})} u_h(y_{\mathbf{j}}, \cdot)^2 \tilde{w}_{\mathbf{j}}^{\mathbf{i}} - (\mathbb{E}[\mathcal{A}(q, N)u_h])^2, \end{aligned}$$

où  $\{y_{\mathbf{j}}\}_{\mathbf{j} \leq \mathbf{m}(\mathbf{i})}$  est l'ensemble des points d'interpolation dans  $\Gamma$  qui correspondent au niveau  $\mathbf{i}$ , chaque poids d'intégration multidimensionnelle  $\tilde{w}_{\mathbf{j}}^{\mathbf{i}}$  est donné par l'intégrale suivante :

$$\tilde{w}_{\mathbf{j}}^{\mathbf{i}} := c_{\mathbf{i}} \int_{\Gamma} L_{\mathbf{j}}^{\mathbf{i}}(y) \varrho(y) dy = c_{\mathbf{i}} \prod_{n=1}^N \int_{\Gamma_n} \ell_{n, j_n}^{i_n}(z) \varrho_n(z) dz.$$

Les coefficients de combinaison  $c_{\mathbf{i}}$  sont définis par :

$$c_{\mathbf{i}} = (-1)^{w+N-|\mathbf{i}|_1} \binom{N-1}{w+N-|\mathbf{i}|_1}. \quad (1.28)$$

La méthode des grilles clairsemées isotropes converge sous l'hypothèse 1.2, au moins algébriquement, par rapport au cardinal des points de collocation de la grille isotrope  $M = \#\mathcal{H}_1(w, N)$ . Considérons  $\nu = \min_n \nu_n$ , où chaque  $\nu_n$  est le taux de convergence donné dans le Théorème 1.2. L'estimation suivante est donnée dans [49] pour le cas où l'ensemble  $\Gamma$  est borné. Elle concerne la convergence de la formule (1.27) utilisant des points de Clenshaw-Curtis ou de Gauss.

### Théorème 1.3.

Considérons la fonction  $u \in L^2_{\varrho}(\Gamma; H^1_0(D))$  satisfaisant l'hypothèse 1.2, alors la formule d'interpolation de Smolyak (1.27) satisfait l'estimation suivante :

## 1. Points de Clenshaw-Curtis :

$$\|u - \mathcal{A}(w, N)u\|_{L^2_\varrho(\Gamma; H_0^1(D))} \leq C_1(\nu, N)M^{-\mu}, \quad (1.29)$$

$$\mu := \frac{\nu}{\log(2) + 1 + \log(N)}. \quad (1.30)$$

## 2. Points de Gauss :

$$\|u - \mathcal{A}(w, N)u\|_{L^2_\varrho(\Gamma; H_0^1(D))} \leq C_2(\nu, N)M^{-\tilde{\mu}}, \quad (1.31)$$

$$\tilde{\mu} := \frac{e \log(2)\nu}{\zeta + \log(N)}, \quad (1.32)$$

où  $\zeta \approx 2.1$  et les deux constantes  $C_1, C_2$  ne dépendent pas de  $M$ .

Sous les hypothèses du théorème précédent, pour de grandes valeurs de niveau  $w$ , une estimation plus précise montrant une convergence sous-exponentielle de la méthode est vérifiée. Cependant, cette situation de convergence n'est pas pratique, à cause du problème de la grande dimensionalité.

## Points d'interpolation

1. **Les abscisses de Clenshaw-Curtis** Les points de Clenshaw-Curtis sont les extréma des polynômes de Chebyshev, soit  $m(i) \in \mathbb{N}$  :

$$z_j^i = -\cos\left(\frac{\pi(j-1)}{m(i)-1}\right), \quad j = 1, \dots, m(i),$$

avec pour  $m = 1$ ,  $z_0 = 0$ . Ces points de Clenshaw-Curtis sont optimaux par rapport à la norme de  $L^\infty$ . La suite de points est emboîtée si on choisit l'ordre de croissance  $m(i) = 2^{i-1} + 1, i > 1$ . Cependant, le degré d'exactitude de ces points est seulement  $m(i) - 1$ .

2. **Les abscisses de Gauss**

Les points de Gauss sont donnés comme racines des polynômes orthogonaux suivant  $\varrho$ , la densité de probabilité de  $\mathbf{Y}$ . Ces points de Gauss sont optimaux par rapport à la norme de  $L^2$ , avec un degré polynomial d'exactitude maximal  $2m(i) - 1$ . Cependant, ces points ne sont pas en général imbriqués, contrairement aux points de Clenshaw-Curtis.

D'autres points ont été proposés, qui cherchent à préserver et combiner la qualité d'imbrication et d'optimalité, par exemple les points de Patterson ( voir [19] pour plus de détails sur la construction).

### 1.5.3 Construction de Smolyak anisotrope

Dans la construction de Smolyak isotrope des grilles clairsemées, toutes les directions sont traitées équitablement, car toute permutation d'un multi-idice  $\mathbf{i} \in X(w, N)$ , reste aussi dans  $X(w, N)$ . Cependant dans plusieurs problèmes pratiques, la solution présente une certaine anisotropie dans le sens où ses directions ont différents comportements et variations. Dans ce cas l'utilisation de la méthode isotrope des grilles clairsemées peut dégrader la qualité de la convergence, à cause de la grande dimension et de la mauvaise répartition des points de collocation.

Par une analogie avec l'adaptation déterministe, il est avantageux d'utiliser plus de points de collocation dans les directions qui contribuent plus à l'erreur globale et moins de points dans les autres qui contribuent moins à l'erreur, tout en gardant de hauts niveaux de précision.

Considérons  $N$  réels strictement positifs  $\alpha_1, \dots, \alpha_N$ ,  $w \in \mathbb{N}$  et posons  $\underline{\alpha} = \min_n \alpha_n$ . Suivant ([32, 49]), on décrit l'interpolation de Smolyak des grilles clairsemées anisotropes par la formule suivante :

$$\mathcal{A}_\alpha(w, N) = \sum_{\mathbf{i} \in X_\alpha(w, N)} (\Delta_{i_1} \otimes \dots \otimes \Delta_{i_N}). \quad (1.33)$$

où l'ensemble des multi-indices  $X_\alpha(w, N)$  est donné par :

$$X_\alpha(w, N) := \left\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1}, \sum_{n=1}^N (i_n - 1)\alpha_n \leq w\underline{\alpha} \right\}. \quad (1.34)$$

Une formule équivalente de (1.33) sous la forme d'une combinaison de produits tensoriels des interpolations unidimensionnelles est donnée par :

$$\mathcal{A}_\alpha(w, N) = \sum_{\mathbf{i} \in Y_\alpha(w, N)} c_\alpha(\mathbf{i}) \mathcal{U}^{i_1} \otimes \dots \otimes \mathcal{U}^{i_N}. \quad (1.35)$$

Les coefficients de combinaison  $c_\alpha(\mathbf{i})$  sont définis par :

$$c_\alpha(\mathbf{i}) = \sum_{j \in \{0,1\}^N, \mathbf{i}+j \in X_\alpha(w, N)} (-1)^{|\mathbf{j}|_1},$$

et  $Y_\alpha(w, N)$ , ensemble des indices, est défini par :

$$Y_\alpha(w, N) = X_\alpha(w, N) \setminus X_\alpha(w - \frac{|\alpha|_1}{\underline{\alpha}}, N),$$

ou encore par :  $Y_\alpha(w, N) = \left\{ \mathbf{i} \geq \mathbf{1} : w\underline{\alpha} < \sum_{n=1}^N \alpha_n i_n \leq w\underline{\alpha} + |\alpha|_1 \right\}$ , avec  $|\alpha|_1 = \sum_{n=1}^N \alpha_n$ . Comme dans le cas isotrope (1.27), pour calculer  $\mathcal{A}_\alpha(w, N)u$ , nous sommes amenés à

évaluer la fonction  $u$  sur la grille clairsemée anisotrope :

$$\mathcal{H}_\alpha(w, N) = \bigcup_{\mathbf{i} \in Y_\alpha(w, N)} \{\mathcal{X}^{i_1} \times \dots \times \mathcal{X}^{i_N}\}.$$

L'introduction des poids d'anisotropie  $\{\alpha_n\}_{n=1}^N$  nous permet d'approcher la solution  $u$  d'une façon adaptative, telle que, quand la direction  $y_n$  contribue à l'erreur plus que la direction  $y_m$ , alors on aura  $\alpha_n < \alpha_m$  et ainsi plus de points seront attribués suivant  $y_n$  que suivant la direction  $y_m$ . La quantité  $\frac{\alpha_n}{\underline{\alpha}}$  mesure le degré d'importance de chaque direction  $y_n$ .

La stratégie de cette approche repose sur la construction du vecteur de poids  $\alpha > 0$  et le niveau global d'approximation  $w \in \mathbb{N}$ . L'idée principale présentée dans [50] consiste à relier chaque poids d'anisotropie  $\alpha_n$  avec la vitesse de convergence exponentielle suivant la direction correspondante  $y_n$ . Suivant le cas où  $\Gamma$  est borné ou non borné, et une fonction satisfaisant l'Hypothèse 1.2, la convergence est donnée dans le Théorème 1.2.

Un choix optimal pour ces poids d'anisotropie dans le cas où  $\Gamma$  est borné consiste à choisir  $\alpha_n = \nu_n$ , pour chaque  $n = 1, \dots, N$ , où les taux  $\nu_n$  sont donnés dans le Théorème 1.2. Ainsi la détermination du vecteur  $\alpha$  peut être faite soit par une procédure *a priori* ou par une information *a posteriori* [6, 50].

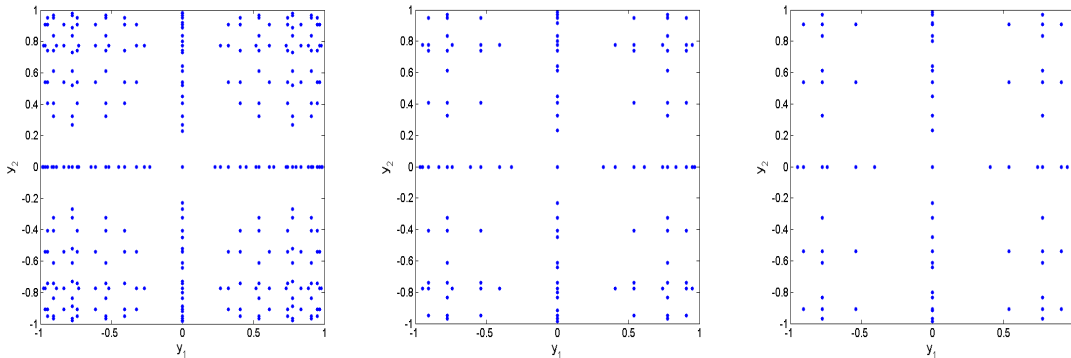


FIGURE 1.2 – Grille clairsemée isotrope (à gauche), grille clairsemée anisotrope  $\frac{\alpha_1}{\alpha_2} = \frac{3}{2}$  (au milieu) et  $\frac{\alpha_1}{\alpha_2} = 2$  (à droite), pour  $\Gamma = [-1, 1]^2$  et  $w = 6$ , utilisant des points de Clenshaw-Curtis.

Nous notons que la méthode de Smolyak des grilles clairsemées isotropes est un cas particulier des grilles anisotropes. Ceci peut être observé en prenant les composantes du vecteur poids tous identiques  $\alpha_1 = \alpha_2 = \dots = \alpha_N = 1$ .

Les coefficients d'anisotropie assurent une partition adaptative des points de col-

location entre les directions. Le choix de ces coefficients d'anisotropie est pris comme étant le taux de convergence exponentiel de  $u$  [50], l'approximation par une approche a posteriori, exige généralement, un grand nombre de résolution de systèmes linéaires. Dans le chapitre 3 on propose un indicateur pour les déterminer avec moins d'effort.

Si de plus le niveau total  $w$  est bien choisi, le nombre des indices de l'ensemble  $Y_\alpha(w, N)$  sera réduit et l'effort de l'approximation sera nettement réduit, puisque le cardinal  $\#\mathcal{H}_\alpha(w, N)$  augmente rapidement voire exponentiellement avec  $w$ .

Pour déterminer un niveau  $w$  acceptable, un estimateur d'erreur classique [32] consiste à calculer deux approximations successives, pour  $w \geq 0$ ,

$$E(w, N) := \mathcal{A}_\alpha(w, N)u - \mathcal{A}_\alpha(w + 1, N)u.$$

Notons que cette procédure pour de grandes dimensions  $N$  peut être très coûteuse, puisque pour chaque test de l'erreur, une grande suite de résolution de systèmes linéaires de taille  $\dim(V_h)$  est exigée. Dans le chapitre 4 nous proposons une méthodologie d'adaptation efficace pour réduire nettement le coût de calcul exigé pour une approximation convenable  $\mathcal{A}_\alpha(w, N)u$ .

Le théorème suivant fournit une estimation d'erreur de la méthode de la grille anisotrope. Il montre la convergence suivant le cardinal des points de collocation de la grille anisotrope. Ce résultat de convergence est démontré dans [50] dans le cas où  $\Gamma$  est borné.

Notons par  $M = \#\mathcal{H}_\alpha(w, N)$  le cardinal des points de la grille clairsemée de la formule d'interpolation anisotrope  $\mathcal{A}_\alpha(w, N)$ . Les poids d'anisotropie  $\alpha_n$  sont choisis égaux aux taux de convergence  $\nu_n$  pour chaque  $n = 1, \dots, N$ .

**Théorème 1.4.** [50]

Considérons la fonction  $u \in L^2_\varrho(\Gamma; H^1_0(D))$  satisfaisant l'hypothèse 1.2. Si on suppose que  $w$  vérifie  $w \leq \frac{\sum_{n=1}^N \alpha_n}{\alpha \log(2)}$ ; alors la formule d'interpolation anisotrope  $\mathcal{A}_\alpha(w, N)$ , donnée dans (1.33) satisfait l'estimation suivante :

– Points de Clenshaw Curtis : il existe  $C_1(\alpha, N)$  indépendante de  $M$  telle que :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{L^2_\varrho(\Gamma; H^1_0(D))} \leq C_1(\alpha, N)M^{-\mu}, \quad (1.36)$$

$$\mu := \frac{\alpha(\log(2)e - \frac{1}{2})}{\log(2) + \sum_{n=1}^N \frac{\alpha}{\alpha_n}}. \quad (1.37)$$

- *Poins de Gauss* : il existe  $C_2(\alpha, N)$  indépendante de  $M$  telle que

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{L^2_\Gamma(\Gamma; H^1_0(D))} \leq C_2(\alpha, N)M^{-\tilde{\mu}}, \quad (1.38)$$

$$\tilde{\mu} := \frac{\underline{\alpha} \log(2)e}{2 \log(2) + \sum_{n=1}^N \frac{\alpha}{\alpha_n}}. \quad (1.39)$$

**Remarque 1.2.**

- *Sous les mêmes hypothèses du théorème 1.4*, pour de grandes valeurs de niveau  $w$  qui vérifie  $w > \frac{\sum_{n=1}^N \alpha_n}{\underline{\alpha} \log(2)}$ , une estimation plus précise montrant une convergence sous-exponentielle de la méthode anisotrope est montrée dans [50]. À cause du problème de la grande dimensionnalité, dans la pratique, on ne s'intéresse qu'à l'estimation algébrique.
- *En supposant que les poids satisfont  $\alpha_n \approx \underline{\alpha}$ , pour  $n = 1, \dots, N$ , la formule d'interpolation anisotrope devient isotrope  $\mathcal{A}_\alpha(w, N) \approx \mathcal{A}(w, N)$ . On peut remarquer que dans ce cas, les estimations d'erreur du théorème 1.4 ne sont pas compatibles avec celles de Théorème 1.3. Dans le chapitre 4 nous allons établir des estimations a priori de la méthode anisotrope des grilles clairsemées plus consistantes et compatibles avec celles de la méthode isotrope. Elles seront utiles pour la construction d'un indicateur d'erreur efficace pour la méthode.*
- *La convergence de la méthode anisotrope des grilles clairsemées, dans le cas où  $\Gamma$  est non borné, est traitée au chapitre 4.*

## Chapitre 2

# Calcul de l'espérance de la solution d'une EDP stochastique unidimensionnelle à l'aide d'une base réduite

Dans ce Chapitre, on présente un résultat concernant l'approximation de l'espérance  $\mathbb{E}[u]$ , où  $u$  est la solution d'un problème d'EDP elliptique unidimensionnelle avec un paramètre et une donnée aléatoires.

Dans les méthodes de Galerkin et de collocation stochastiques, le coût de l'approximation de la réponse, peut être dramatique. L'effort de calcul et le stockage de mémoire engendrés limitent leur application à des faibles dimensions stochastiques.

Nous proposons une méthodologie de réduction de modèle qui permet d'approcher l'espérance de la solution. L'approche présentée est basée sur la décomposition de Karhunen-Loève (K-L) de l'inverse du paramètre de diffusion. Une projection de la solution sur un espace à dimension réduite est construite de telle manière que l'espérance de la solution et celle du projeté de la solution coïncident. L'estimation a priori d'erreur prouve l'efficacité et la convergence de l'approche considérée.

## 2.1 Formulation du problème modèle

Considérons  $(\Omega, \mathcal{F}, d\mathbb{P})$  un espace de probabilisé complet.

Soient  $I = ]a, b[$  un ouvert borné de  $\mathbb{R}$ ,  $k$  et  $f$  deux champs aléatoires sur  $\Omega \times I$ . L'objectif est de calculer la moyenne de la solution du problème unidimensionnel suivant :

$$(\mathcal{P}_s) \begin{cases} -(k.u')' = f & \text{dans } \Omega \times I \\ u(\cdot, a) = 0, k u'(\cdot, b) = \zeta(\cdot). \end{cases} \quad (2.1)$$

### Hypothèse 2.1.

On suppose que les champs  $k$  et  $f$  vérifient les Hypothèses (1.1) et que  $\zeta \in L^2(\Omega)$ .

On suppose aussi que  $k$  et  $f$  sont des bruits finis, et qu'il existe  $N \in \mathbb{N}$  tel que les décompositions (K-L), s'écrivent :

$$k(\omega, x) = \bar{k}(x) + \sum_{n=1}^N k_n(x) Y_n(\omega) \quad (2.2)$$

$$f(\omega, x) = \bar{f}(x) + \sum_{n=1}^N f_n(x) Y_n(\omega) \quad (2.3)$$

où  $\{Y_n\}_{n=1}^N$  sont des variables aléatoires uniformément bornées.

Le problème variationnel associé à  $(\mathcal{P}_s)$  est donné par la formulation :

$$(\mathcal{P}_v) \begin{cases} \text{Trouver } u \in L^2(\Omega) \otimes V(I) \text{ telle que} \\ \mathbb{E}[(k.u', \mu')_{L^2(I)}] = \mathbb{E}[(f, \mu)_{L^2(I)} + \zeta(\cdot)\mu(\cdot, b)] \quad \forall \mu \in L^2(\Omega) \otimes V(I) \end{cases} \quad (2.4)$$

où l'espace  $V(I) := \{v \in H^1(I); v(a) = 0\}$ . Il admet une solution unique d'après le théorème de Lax-Milgram.

La méthode de Galerkin stochastique approche  $u$  dans  $L^2(\Omega)$  sur la base des polynômes orthogonaux multidimensionnels, de cardinal  $\mathcal{P} = \frac{(p+N)!}{p!N!}$ , conduisant à la résolution d'un système de très grande taille. L'effort de calcul et de stockage en mémoires engendrés limitent leur application à des faibles dimensions stochastiques.

On donne dans la suite, une méthodologie pour calculer une approximation de  $\mathbb{E}[u]$ . Elle consiste à projeter  $(\mathcal{P}_s)$  sur un sous-espace engendré par un nombre réduit de variables aléatoires. Ce projeté n'est pas une approximation de  $u$ , mais on montre que l'espérance de ce projeté coïncide avec  $\mathbb{E}[u]$ .



## 2.2 Construction d'une base stochastique réduite

Le développement de Karhunen-Loève (K-L) permet de décomposer tout champ aléatoire en une partie déterministe (la moyenne ou l'espérance), et une partie aléatoire décomposée sur la base des valeurs propres et des fonctions propres de la covariance du champ aléatoire. Le paramètre  $k$  est strictement positif et borné, alors son inverse,  $k^{-1}$  est du second ordre. On peut alors déterminer sa moyenne  $\mathbb{E}[k^{-1}]$ , sa covariance  $\text{cov}[k^{-1}]$  et par conséquent son développement K-L, noté par :

$$k^{-1}(\omega, x) = \mathbb{E}[k^{-1}](x) + \sum_{n=1}^{\infty} \sqrt{\hat{\lambda}_n} b_n(x) Z_n(\omega), \quad (2.5)$$

où  $\{\hat{\lambda}_j\}_{j=1}^{\infty}$  et  $\{b_j\}_{j=1}^{\infty}$  sont respectivement les valeurs et les fonctions propres de  $\text{cov}[k^{-1}]$ , et  $\{Z_j\}_{j=1}^{\infty}$  sont des variables aléatoires décorélées données par :

$$Z_n(\omega) = \frac{1}{\sqrt{\hat{\lambda}_n}} \int_I (k^{-1}(\omega, x) - \mathbb{E}[k^{-1}](x)) b_n(x) dx. \quad (2.6)$$

Posons  $\mathcal{S}_{\infty} := \overline{\text{Vect}\{Z_j; j \geq 0\}}$ ,  $Z_0 = 1$ , et notons par  $\mathcal{P}u$  la projection de  $u$  sur  $\mathcal{S}_{\infty} \otimes V(I)$ , définie par :

$$\mathbb{E} \left[ (k \cdot (\mathcal{P}u)', \mu')_{L^2(I)} \right] = \mathbb{E} \left[ (k \cdot u', \mu')_{L^2(I)} \right], \quad \forall \mu \in \mathcal{S}_{\infty} \otimes V(I). \quad (2.7)$$

La proposition suivante montre que les deux fonctions stochastiques  $u$  et  $\mathcal{P}u$  ont la même espérance.

### Proposition 2.1.

*Soit  $u$  la solution de  $(\mathcal{P}_s)$  et  $\mathcal{P}u$  donnée par (2.7), alors nous avons*

$$\mathbb{E}[u](x) = \mathbb{E}[\mathcal{P}u](x), \quad \forall x \in I.$$

### Preuve

Pour toute  $v \in V(I)$ , considérons la fonction test  $\mu = \int_a^x \frac{v'(y)}{k(\omega, y)} dy$ . Nous avons, pour chaque  $x$  fixé, la variable aléatoire  $\mu(\cdot, x)$  est dans  $\mathcal{S}_{\infty}$ , et pour chaque  $\omega \in \Omega$ , la fonction  $\mu(\omega, \cdot)$  est dans  $V(I)$ , on en déduit alors que  $\mu(\cdot, \cdot) \in \mathcal{S}_{\infty} \otimes V(I)$  telle que  $k\mu' = v'$ , pour tout  $v \in V(I)$ . Ainsi, en substituant  $\mu$  dans la projection (2.7) on obtient pour tout  $v$  dans  $V(I)$  :  $\mathbb{E}[(\mathcal{P}u)', v']_{L^2(I)} = \mathbb{E}[(u', v')_{L^2(I)}]$ , ou encore :

$$\left( (\mathbb{E}[u])' - (\mathbb{E}[(\mathcal{P}u)])', v' \right)_{L^2(I)} = 0 \quad \forall v \in V(I),$$

et puisque les deux fonctions  $\mathbb{E}[u]$  et  $\mathbb{E}[(\mathcal{P}u)]$  sont dans  $V(I)$ , alors on en déduit que  $\mathbb{E}[u](\cdot) = \mathbb{E}[\mathcal{P}u](\cdot)$ . ■

**Remarque 2.1.**

- Le cas où les conditions sont de Dirichlet-Robin :  $u(\cdot, a) = 0$  et  $u(\cdot, b) + k(\cdot, b)u'(\cdot, b) = \gamma$ , où  $\gamma \in \mathbb{R}$ , la proposition reste vraie avec la même preuve.
- dans le cas où les conditions sont de Dirichlet homogène, la preuve de la proposition 1 n'est pas valable, car il y a une difficulté à trouver une fonction test  $\mu \in \mathcal{S}_\infty \otimes H_0^1(I)$  telle que  $k\mu' = v'$ ,  $\forall v \in H_0^1(I)$ .

Quand la dimension stochastique  $N$  est grande, l'approximation de  $u$  peut être très coûteuse. Ainsi, si on est intéressé plus par  $\mathbb{E}[u]$  que par  $u$ , en tronquant la série (2.5), jusqu'à un ordre  $M$  "convenable", et en projetant  $u$  dans  $\text{Vect}\{Z_j; j \leq M\} \otimes V(I)$ , on obtient une bonne approximation de  $\mathbb{E}[u]$ , nécessitant un degré de liberté stochastique minimal.

Notons par  $\mathcal{S}_M := \text{Vect}\{Z_j; 0 \leq j \leq M\}$  et  $V_h := \{\phi \in \mathcal{C}^0(\bar{I}) \mid \phi|_T \in P_1(T), \forall T \in \mathcal{T}_h, \phi(a) = 0\}$ , l'espace d'approximation par des éléments finis de degré 1, où  $(\mathcal{T}_h)$  est une triangulation régulière de  $I$ . La proposition suivante donne une estimation, dans les deux normes  $|\cdot|_{1,I}$  et  $\|\cdot\|_{0,I}$ , de l'erreur commise en approchant  $\mathbb{E}[u]$  par  $\mathbb{E}[\mathcal{P}_{M,h}u]$ , où  $\mathcal{P}_{M,h}u$  est la projection de  $u$  dans l'espace  $\mathcal{S}_M \otimes V_h$ , définie par :

$$\mathbb{E}[(k \cdot (u' - (\mathcal{P}_{M,h}u)'), \mu'_{M,h})_{L(I)}] = 0, \quad \forall \mu_{M,h} \in \mathcal{S}_M \otimes V_h. \quad (2.8)$$

**Proposition 2.2.**

Soit  $u$  la solution de  $(\mathcal{P}_s)$  et  $\mathcal{P}_{M,h}u$  sa projection sur  $\mathcal{S}_M \otimes V_h$ . Supposons que les fonctions propres  $(b_n)_{n \geq 1}$  données dans (2.5) sont uniformément bornées, alors on a l'estimation a priori suivante :

$$\begin{cases} |\mathbb{E}[u] - \mathbb{E}[\mathcal{P}_{M,h}u]|_{1,I} \leq C_1 \left( h + \sqrt{\sum_{j \geq M+1} \hat{\lambda}_j} \right) \\ \|\mathbb{E}[u] - \mathbb{E}[\mathcal{P}_{M,h}u]\|_{0,I} \leq C_0 \left( h^2 + \sqrt{\sum_{j \geq M+1} \hat{\lambda}_j} \right) \end{cases} \quad (2.9)$$

où  $C_1$  et  $C_0$  sont des constantes positives indépendantes de  $h$  et de  $M$ .

### Preuve

L'idée de la démonstration consiste à décomposer l'erreur en deux quantités, dont l'une dépend de  $M$  et l'autre de  $h$ .

Introduisons  $\mathcal{P}_M u$  la projection de  $u$  dans  $\mathcal{S}_M \otimes V(I)$ , donnée par l'équation suivante :

$$\mathbb{E}[(k \cdot (\mathcal{P}u - \mathcal{P}_M u)', \mu'_M)_{L^2(I)}] = 0 \quad \forall \mu_M \in \mathcal{S}_M \otimes V(I). \quad (2.10)$$

Nous avons pour  $s = 0$  et  $s = 1$ ,

$$\| \mathbb{E}[u] - \mathbb{E}[\mathcal{P}_{M,h}u] \|_{s,I} \leq \| \mathbb{E}[\mathcal{P}u] - \mathbb{E}[\mathcal{P}_M u] \|_{s,I} + \| \mathbb{E}[\mathcal{P}_M u] - \mathbb{E}[\mathcal{P}_{M,h}u] \|_{s,I}. \quad (2.11)$$

Puisque  $\mathcal{P}_{M,h}u$  est une approximation de  $\mathcal{P}_M u$  dans  $V_h$ , nous avons l'estimation donnée par les approximations déterministes classiques :

$$\| \mathcal{P}_M u - \mathcal{P}_{M,h}u \|_{s,I} \leq c_s h^{2-s} \| \mathcal{P}_M u \|_{H^2(I)}$$

où  $c_s > 0$  est indépendante de  $h$ , ainsi par l'inégalité de Jensen, on déduit l'estimation du premier terme de (2.11) :

$$\| \mathbb{E}[\mathcal{P}_M u - \mathcal{P}_{M,h}u] \|_{s,I} \leq \mathbb{E}[\| \mathcal{P}_M u - \mathcal{P}_{M,h}u \|_{s,I}] \leq c_s h^{2-s} \mathbb{E}[\| \mathcal{P}_M u \|_{H^2(I)}]. \quad (2.12)$$

On choisit dans (2.10) comme fonction test :

$$\mu_M = \int_a^x \mathbb{E}[(\mathcal{P}u - \mathcal{P}_M u)'] (k^{-1} - \sum_{j \geq M+1} \sqrt{\hat{\lambda}_j} b_j Z_j) dy.$$

On obtient l'égalité suivante :

$$\| \mathbb{E}[(\mathcal{P}u - \mathcal{P}_M u)'] \|_{0,I}^2 = \mathbb{E} \left[ \left\| \left( k \cdot (\mathcal{P}u - \mathcal{P}_M u)', \sum_{j \geq M+1} \sqrt{\hat{\lambda}_j} b_j Z_j \mathbb{E}[(\mathcal{P}u - \mathcal{P}_M u)'] \right) \right\|_{L^2(I)}^2 \right],$$

par l'inégalité de Cauchy-Schwarz et le fait que  $\mathbb{E}[Z_j Z_i] = \delta_{ij}$ , on déduit :

$$\| \mathbb{E}[(\mathcal{P}u - \mathcal{P}_M u)'] \|_{0,I}^2 \leq k_{max} \| b \|_{\infty} \| (\mathcal{P}u - \mathcal{P}_M u)' \|_{0,\Omega \otimes I} \| \mathbb{E}[(\mathcal{P}u - \mathcal{P}_M u)'] \|_{0,I} \sqrt{\sum_{j \geq M+1} \hat{\lambda}_j}.$$

Par conséquent :

$$\| \mathbb{E}[(\mathcal{P}u - \mathcal{P}_M u)'] \|_{0,I} \leq 2k_{max} \| b \|_{\infty} \| (\mathcal{P}u)' \|_{0,\Omega \otimes I} \sqrt{\sum_{j \geq M+1} \hat{\lambda}_j}. \quad (2.13)$$

Par l'inégalité de Poincaré, nous avons de même l'estimation de l'erreur dans  $L^2(I)$ . En combinant les deux bornes (2.12) et (2.13), on déduit l'estimation de la proposition ■

**Remarque 2.2.**

- L'approximation  $\mathcal{P}_{M,h}u$  est obtenue en résolvant le problème éléments finis suivant :

$$\mathbb{E}[(k \cdot (\mathcal{P}_{M,h}u)', \mu'_{M,h})_{L^2(I)}] = \mathbb{E}[(f, \mu_{M,h})_{L^2(I)} + \xi \mu_{M,h}(\cdot, 1)] \quad \forall \mu_{M,h} \in \mathcal{S}_M \otimes V_h. \quad (2.14)$$

- Dans le cas où le champ de diffusion  $k$  est approché par une troncature  $k_N$ , nous devons considérer  $\|k - k_N\|_\infty$  l'erreur commise par cette troncature dans l'estimation de la proposition (2.2).
- Quand le second membre  $f$  est une fonction déterministe et que  $\xi$  est une constante, la projection  $\mathcal{P}u$  donnée par (2.7) coïncide avec la solution  $u$ . Dans le cas contraire elles ne coïncide pas et  $u$  peut être approché dans l'espace engendré par les variables aléatoires  $\{Y_n Z_j\}_{n,j=1}^N$ .
- On peut aussi utiliser n'importe quel développement en série de l'inverse  $k^{-1}$ , comme par exemple un système orthonormé de  $L^2(I)$ . Le fait de proposer d'utiliser le développement de K.L, permet d'assurer l'optimalité de la troncature (le nombre  $M$  est minimal).

**2.2.1 Algorithme**

L'approximation de la moyenne  $\mathbb{E}[u]$  revient à approcher  $\mathbb{E}[\mathcal{P}u]$  par  $\mathbb{E}[\mathcal{P}_{M,h}u]$ . La première étape consiste à choisir le nombre de termes  $M$  dans le développement (2.5), ensuite calculer  $\mathcal{P}_{M,h}u$  en utilisant la projection (2.14), puis en déduire la moyenne  $\mathbb{E}[\mathcal{P}_{M,h}u]$ , selon l'algorithme suivant :

1. Calculer  $\text{cov}[k^{-1}]$  la covariance de  $k^{-1}$ ,
2. Pour  $\epsilon > 0$  donné, déterminer les  $M$  valeurs et vecteurs propres de  $\text{cov}[k^{-1}]$ , tels que  $\|k^{-1} - k_M^{-1}\|_{L^2(\Omega; L^2(D))} \leq \epsilon$
3. Calculer la matrice de rigidité (2.15) et le vecteur du second membre (2.16).
4. Résoudre le système linéaire résultant.

**2.2.2 Assemblage de la matrice de rigidité**

Comme dans le cas de la méthode de Galerkin stochastique, cette approche conduit à une suite de systèmes linéaires couplés. Mais la taille est beaucoup plus petite que pour la méthode de Galerkin. On choisit le système de variables aléatoires  $\{Z_j\}_{j=0}^M$ , et on considère une base de l'espace élément fini  $V_h(D)$ ,  $\{\phi_i\}_{i=1}^{N_D}$ . Le problème (2.14)

conduit au système linéaire  $A\mathbf{w} = \mathbf{F}$ , où la matrice  $A$  et le vecteur  $\mathbf{F}$  ont une structure de  $M + 1$  blocs :

$$A = S_0 \otimes K_0 + \sum_{m=1}^M S_m \otimes K_m \quad (2.15)$$

$$\mathbf{F} = b_0 \otimes F_0 + \sum_{m=1}^M b_m \otimes F_m \quad (2.16)$$

où  $S_m$  est la matrice stochastique donnée par :

$$(S_0)_{i,j} = \mathbb{E}[Z_i Z_j] = \delta_{i,j} \quad (2.17)$$

$$(S_m)_{i,j} = \mathbb{E}[Y_m Z_i Z_j] \quad i, j = 0, \dots, M. \quad (2.18)$$

et  $K_m$  est la matrice élément fini déterministe donnée par :

$$(K_0)_{i,j} = \int_I \mathbb{E}[k] \nabla \phi_i \nabla \phi_j dx \quad (2.19)$$

$$(K_m)_{i,j} = \int_I k_m \nabla \phi_i \nabla \phi_j dx \quad i, j = 0, \dots, N_D. \quad (2.20)$$

Le second membre  $\mathbf{F}$  est donné par :

$$(b_0)_i = \mathbb{E}[Z_i] \quad i = 0, \dots, M, \quad (2.21)$$

$$(F_0)_j = \int_D f_0 \phi_j dx \quad j = 1, \dots, N_I. \quad (2.22)$$

$$(b_m)_i = \mathbb{E}[Y_m Z_i] \quad i = 0, \dots, M, \quad (2.23)$$

$$(F_m)_j = \int_I f_m \phi_j dx \quad j = 1, \dots, N_I. \quad (2.24)$$

## 2.3 Exemple numérique

Nous considérons le problème elliptique unidimensionnel suivant :

$$(\mathcal{P}_1) \begin{cases} -(k.u')' = f & \text{dans } \Omega \times I = \Omega \times ]0, 1[ \\ u(0, \omega) = 0, \quad (ku')(1) = \gamma. \end{cases} \quad (2.25)$$

Le champ aléatoire  $k$  est log-normal donné par :  $k(\omega, x) = \exp(G(x, \omega))$ , avec  $G(x, \omega)$  est un champ gaussien de moyenne nulle et de variance  $\sigma^2$ , satisfaisant une covariance gaussienne donnée par :  $\text{cov}[G](x_1, x_2) = \sigma^2 \exp(-\frac{(x_1 - x_2)^2}{l^2})$ , où  $l$  représente la longueur de corrélation. On vérifie aisément que les deux champs  $k$  et  $k^{-1}$  vérifient les statistiques suivantes :

– La moyenne :

$$\mathbb{E}[k] = \mathbb{E}[k^{-1}] = e^{\sigma^2/2}$$

– La variance :

$$\text{Var}[k] = \text{Var}[k^{-1}] = e^{2\sigma^2} - e^{\sigma^2}$$

– la fonction covariance :

$$\text{cov}[k](x, y) = \text{cov}[k^{-1}](x, y) = e^{\sigma^2} \exp \left\{ \sigma^2 e^{-\frac{(x-y)^2}{l^2}} \right\} - e^{\sigma^2}.$$

On prend  $f = 1$ ,  $\gamma = 1$ ,  $l = 1$  et  $\sigma = 1$ . On utilise les fonctions constantes par morceaux pour approcher le spectre de l'opérateur de covariance, ainsi l'approximation de la troncature du K-L, avec  $N = 10$  pour  $G$ ,  $k$  et  $k^{-1}$ , donnée par :

$$\begin{aligned} G(\omega, x) &\approx \sum_{n=1}^N \sqrt{\hat{\lambda}_n} \hat{\varphi}_n(x) \xi_n(\omega), \\ k(\omega, x) &\approx e^{\sigma^2/2} + \sum_{n=1}^N \sqrt{\lambda_n} \varphi_n(x) Y_n(\omega), \\ k^{-1}(\omega, x) &\approx e^{\sigma^2/2} + \sum_{n=1}^N \sqrt{\lambda_n} \varphi_n(x) Z_n(\omega), \end{aligned}$$

où  $\{(\hat{\lambda}_n, \hat{\varphi}_n)\}_{n=1}^N$  (figure 2.1) et  $\{(\lambda_n, \varphi_n)\}_{n=1}^N$  (figure 2.2), sont les  $N$  premières valeurs et vecteurs propres de la covariance  $\text{cov}[G](x, y)$  et  $\text{cov}[k^{-1}](x, y)$  respectivement, et les variables aléatoires  $(\xi_n)$ ,  $(Y_n)$  et  $(Z_n)$  sont données par :

$$\xi_n(\omega) = \frac{1}{\sqrt{\hat{\lambda}_n}} \int_I G(\omega, x) \hat{\varphi}_n(x) dx, \quad (2.26)$$

$$Y_n(\omega) = \frac{1}{\sqrt{\lambda_n}} \int_I (k(\omega, x) - e^{\sigma^2/2}) \varphi_n(x) dx, \quad (2.27)$$

$$Z_n(\omega) = \frac{1}{\sqrt{\lambda_n}} \int_I (k^{-1}(\omega, x) - e^{\sigma^2/2}) \varphi_n(x) dx. \quad (2.28)$$

Les variables aléatoires  $\xi_n$  sont gaussiennes puisque le champ aléatoire  $G$  est gaussien. La projection nécessite le calcul des matrices stochastiques  $S_m$  dont les éléments sont donnés par les espérances (2.18). Comme dans le cas de la méthode de Galerkin, ces coefficients peuvent être approchés par une quadrature ou une approximation de Monte Carlo. Dans cet exemple, on utilise les expressions (2.27) et (2.28), ainsi que le fait que les variables  $\xi_n$  sont gaussiennes pour les calculer explicitement sous forme d'intégrale déterministe sur  $I$ .

On considère pour la discrétisation spatiale une approximation par éléments finis  $P_1$  avec  $n_x = 40$  degrés de liberté.

**Remarque 2.3.** Les valeurs propres  $\hat{\lambda}_n$  et  $\lambda_n$  décroissent exponentiellement vers zéros, puisque les deux fonctions covariance de  $G$  et  $k^{-1}$  sont analytiques (voir [56]). On remarque aussi que  $\hat{\lambda}_n < \lambda_n$  du fait que  $\text{Var}[G] < \text{Var}[k^{-1}]$ , on pourrait prendre moins de termes dans le développement de  $G$  que dans ceux de  $k$  et  $k^{-1}$ .

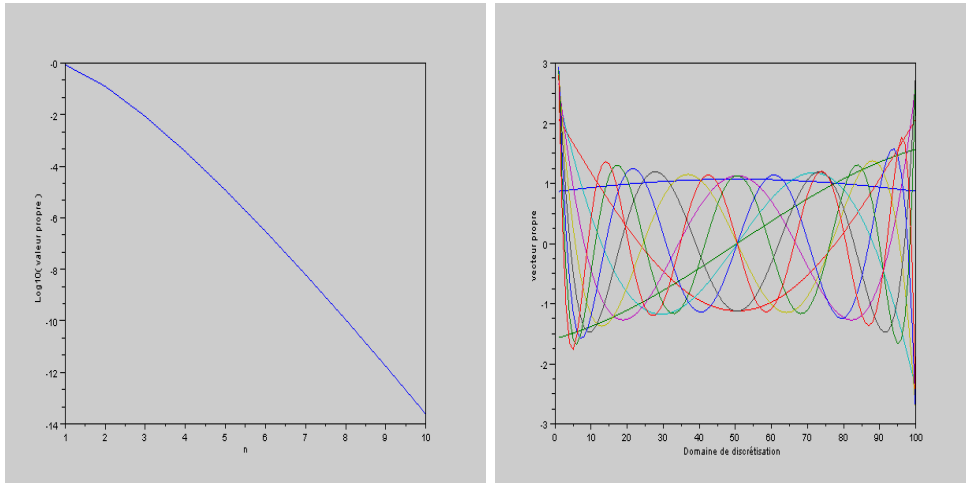


FIGURE 2.1 – Les dix premières valeurs et vecteurs propres  $\{(\hat{\lambda}_n, \hat{\varphi}_n)\}_{n=1}^N$  de la covariance de  $G$ .

### 2.3.1 Comparaison avec la méthode de Monte Carlo

Pour comparer notre approche avec l'approche Monte Carlo, nous considérons des réalisations du champ  $k^{-1}$  à partir d'un échantillonnage du vecteur gaussien  $\xi = (\xi_1, \dots, \xi_{10})$ . Nous considérons respectivement  $M = 10^4$  et  $M = 10^5$  réalisations de  $\xi$ .

Le tableau (2.1) nous donne une comparaison entre l'approche de projection sur la base réduite et celle de Monte Carlo (MMC), pour approcher l'espérance  $\mathbb{E}[u]$ .

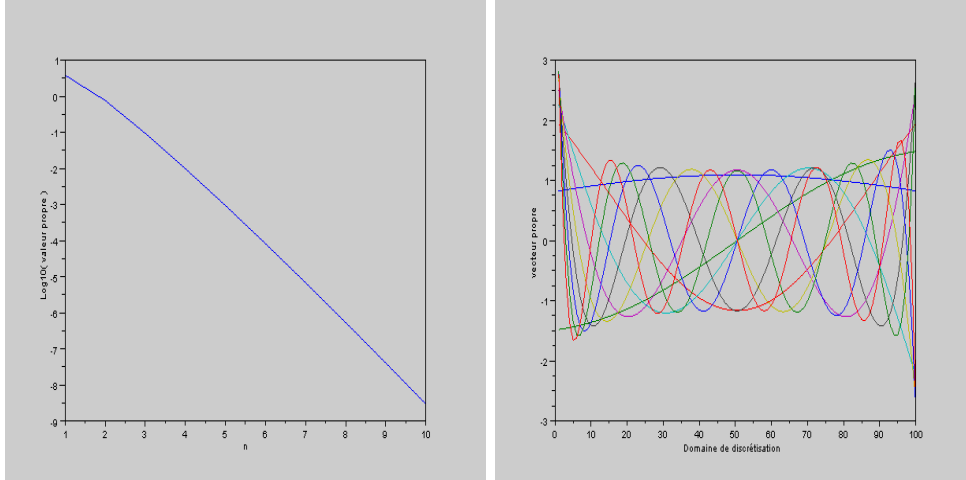


FIGURE 2.2 – Les dix premières valeurs et vecteurs propres  $\{(\lambda_n, \varphi_n)\}_{n=1}^N$  de la covariance de  $k^{-1}$ .

Méthodes	$\ \mathbb{E}[u] - \mathbb{E}[u_h]\ _{0,I}$	Temps de calcul en s
MMC, ( $10^4$ )	0.009	18.04s
MMC, ( $10^5$ )	0.0043	145.56s
Projection	0.0000002	5.065s

TABLE 2.1 – Comparaison de l’approche de projection sur la base réduite avec la méthode de Monte Carlo.

### 2.3.2 Comparaison avec la méthode de collocation stochastique

La méthode de Galerkin stochastique introduit un espace d’élément fini  $\mathcal{S}_p$  de  $L^2(\Omega)$ , donné par les polynômes de Chaos qui sont orthogonaux suivant la densité conjointe des paramètres. Le nombre d’éléments de cette base croît exponentiellement avec la dimension stochastique  $N$  et le degré polynomial  $p$ . Dans notre exemple, nous avons pris  $N = 10$  et si on prend comme degré d’approximation  $p = 4$ , la dimension de  $\mathcal{S}_p$  sera  $P = 1001$ . La matrice de rigidité possède la même structure que (2.15) avec une dimension 40040, tandis que la dimension de la matrice dans notre projection est seulement  $\dim(A) = 440$ . D’autre part, dans le cas de la méthode de Galerkin stochastique, on doit déterminer la densité de chaque variable aléatoire  $Y_n$  donnée par (2.27) pour construire les polynômes orthogonaux suivant la densité conjointe de  $\mathbf{Y} = (Y_1, \dots, Y_N)$ , sous l’hypothèse que les variables aléatoires  $Y_n$  sont indépendantes (ce qui n’est pas vrai, sauf si chaque  $Y_n$  suit une loi normale).



Méthodes	$  \mathbb{E}[u] - \mathbb{E}[u_h]  _{0,I}$	Temps de calcul en s
Collocation $p = 3$	0.007	110.23s
Collocation $p = 4$	0.0004	1764.41s
Projection	0.0000002	5.065s

TABLE 2.2 – Comparaison de l'approche de projection sur la base réduite avec la méthode de collocation stochastique

La méthode de collocation stochastique possède la même précision que la méthode de Galerkin. Elles sont équivalentes lorsque le champ  $k$  est linéaire suivant  $y$  ([6]). La méthode de collocation utilise des réalisations déterministes du champ aléatoire  $k$ . Les points de collocation sont donnés par exemple par les points de Gauss, qui sont les zéros des polynômes d'Hermite.

On compare dans le tableau 2.2 la méthode de collocation stochastique avec notre approche, en utilisant  $p = 3$  et  $p = 4$  points de collocation suivant chaque direction  $n$ .

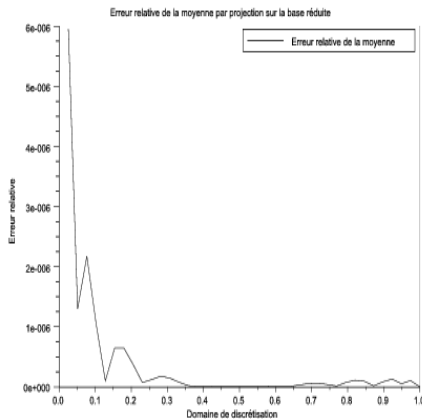


FIGURE 2.3 – L'erreur relative de la moyenne par projection et Monte Carlo

La figure 2.3 montre l'erreur relative commise sur la moyenne, en utilisant une projection sur une base réduite de dimension  $M = 10$  et par Monte Carlo avec deux nombres de simulations  $10^4$  et  $10^5$ .

On change la condition au limite au point  $x = 1$  par une condition non homogène de Dirichlet  $u(\cdot, 0) = 0$  et  $u(\cdot, 1) = 2$ . Dans la figure 2.4 on trace les deux courbes de

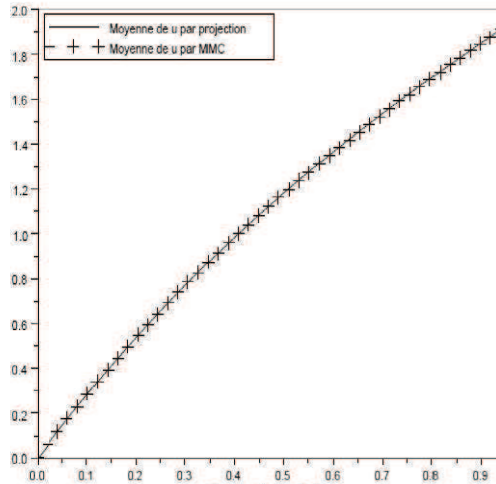


FIGURE 2.4 – La moyenne par projection et Monte Carlo

l'approximation de la moyenne  $\mathbb{E}[u]$ , avec ces conditions non homogènes de Dirichlet, par projection sur la base réduite et Monte-Carlo avec  $10^6$  simulations. Elle montre aussi que la projection sur la base réduite donne une bonne approximation de la moyenne.

## 2.4 Conclusion

Nous avons proposé une méthodologie pour approcher l'espérance de la solution d'un problème elliptique unidimensionnel avec des entrées stochastiques. Elle consiste à projeter le problème sur une base de variables aléatoires bien choisies et en nombre minimal. Ce projeté de la solution n'approche pas la solution elle-même, mais on montre qu'elles ont la même espérance.

En comparaison avec la méthode de Monte Carlo, l'approximation de la moyenne sur cette base minimale semble très efficace, même dans le cas de grande dimension stochastique. Les avantages de cette construction sont multiples, elle nous permet une grande réduction du stockage mémoire et du temps de calcul exigés par les méthodes stochastiques spectrales.

Il serait intéressant d'étendre cette méthodologie aux problèmes de dimension supérieure avec des conditions limites quelconques.

# Chapitre 3

## Méthode des grilles clairsemées adaptative par un indicateur d'erreur

Dans ce chapitre, on s'intéresse à une approximation du problème (1.1) par une méthode de collocation adaptative des grilles clairsemées. On cherche à maintenir un certain niveau de précision de l'approximation, tout en réduisant le nombre de degrés de liberté stochastiques et l'effort de calcul, par le biais d'un indicateur d'erreur satisfaisant une borne supérieure de l'erreur. Sous une hypothèse d'analyticité de la solution  $u$ , on démontre une estimation d'erreur efficace, qui permet d'estimer les poids d'anisotropie convenables de la méthode pour un ensemble stochastique borné ou non borné.

### 3.1 Motivation

Le coût de calcul augmente exponentiellement avec la dimension stochastique, parfois ce coût peut être prohibitif. Il est donc intéressant de chercher un moyen approprié utilisant des techniques d'adaptation efficaces pour réduire le coût de calcul global. En choisissant soigneusement le nombre total des degrés de liberté stochastiques, la réduction de l'effort serait maintenue sans perte de précision.

Dans les approches d'adaptation classiques, lorsqu'on s'intéresse à approcher la solution  $u$  avec une interpolation de produit tensoriel complet  $\mathcal{A}_i u$ , donnée par (1.21), une analyse unidimensionnelle consiste à traiter l'interpolation direction par direction.

Dans les références [5] et [6], suivant chaque direction, on commence par une interpolation polynomiale de degré zéro et on augmente le degré jusqu'à atteindre une précision souhaitée avant de passer à autre direction.

Cette approche semble très coûteuse, car à chaque itération, il faut résoudre une série de problèmes d'EDP déterministes. Le nombre d'itérations exigées augmente avec  $N$ , la dimension stochastique.

L'approximation de  $u$  avec une interpolation des grilles clairsemées  $\mathcal{A}_\alpha(w, N)u$  donnée dans (1.33) exige, d'abord l'estimation des poids d'anisotropie  $\{\alpha_n\}_{n=1}^N$ , ensuite  $w$ , le niveau total de la précision de la formule  $\mathcal{A}_\alpha(w, N)$ . On utilise généralement une approche basée sur une analyse a posteriori dans le calcul des poids d'anisotropie [50]. Cette approche demande la résolution d'une série d'EDP déterministes, alors que les résultats de calcul ne sont pas utilisés dans la construction de l'interpolation  $\mathcal{A}_\alpha(w, N)u$ . Le niveau global satisfait un critère d'arrêt sur des approximations successives de  $u$ , conduisant aussi à des calculs additionnels très coûteux.

Nous considérons le problème (1.1), où le coefficient de diffusion  $k$  et la donnée  $f$  satisfont l'hypothèse 1.1. On peut décomposer (1.1) en deux équations de premier ordre suivantes :

$$\begin{cases} -\nabla u = k^{-1}p \\ \operatorname{div}(p) = f \end{cases}$$

Le comportement de chaque dimension stochastique pour  $u$  est celui de  $k^{-1}p$ , puisque les deux fonctions  $\nabla u$  et  $u$  ont le même comportement par rapport à la variable stochastique  $y$ . On peut espérer de la deuxième équation que  $p$  et  $f$  auront le même comportement suivant  $y$ . Par conséquent, nous proposons de combiner ces deux paramètres pour obtenir un indicateur  $\lambda = k^{-1}f$ , qui va nous permettre de prévoir le comportement des différentes directions et d'identifier ainsi, le nombre total des degrés de liberté de  $u$  avec un moindre effort.

Dans le chapitre 2 (cf [25]), pour le cas d'un problème elliptique unidimensionnel, nous avons calculé un développement de Karhunen-Loève de l'inverse  $k^{-1}$ , et utilisé les variables aléatoires du développement comme base pour construire une projection de la fonction  $u$  qui nous permet d'obtenir une bonne approximation de l'espérance  $\mathbb{E}[u]$ . Dans la suite du chapitre, nous proposons d'utiliser une approche similaire pour approcher la solution  $u$  dans l'espace  $L^2(\Gamma)$ . Elle consiste à développer d'abord l'indicateur  $\lambda$  sur la base des polynômes de Lagrange multidimensionnels comme dans (1.21) et (1.35) avec une tolérance donnée, puis de réutiliser cette base pour estimer la solution  $u$ .

## 3.2 Estimation de l'erreur et adaptation dans l'espace stochastique

Nous commençons par énoncer la proposition suivante qui montre que la quantité  $\|\lambda - \mathcal{A}_i \lambda\|_{L^2_\varrho(\Gamma; L^2(D))}$ , l'erreur commise sur  $\lambda = k^{-1}f$ , est une borne supérieure de l'erreur  $\|u - \mathcal{A}_i u\|_{L^2_\varrho(\Gamma; H^1_0(D))}$ . Ainsi la précision satisfaite par une approximation de l'indicateur  $\lambda$  est également maintenue par celle de la fonction  $u$ .

Les deux erreurs  $\|u - \mathcal{A}_i u\|_{L^2_\varrho(\Gamma; H^1_0(D))}$  et  $\|\lambda - \mathcal{A}_i \lambda\|_{L^2_\varrho(\Gamma; L^2(D))}$ , dépendent de la régularité de  $u$  et  $\lambda$ , telle que la taille de la région d'analyticité suivant chaque direction (cf Théorème 1.2). Par conséquent, pour comparer ces deux quantités, il suffit de comparer la région d'analyticité suivant chaque dimension pour  $u$  et  $\lambda$ .

Pour cela, nous adoptons les notations suivantes : pour chaque  $n = 1, \dots, N$ , on pose  $\Gamma_n^* = \prod_{j \neq n}^N \Gamma_j$ ,  $y = (y_n, y_n^*)$  où  $y_n \in \Gamma_n$  et  $y_n^* \in \Gamma_n^*$ . On suppose que la donnée  $f$  est polynomiale par rapport à la variable  $y$ .

### Proposition 3.1.

*Considérons  $u$  la solution exacte de (1.1),  $\lambda = k^{-1}f$ ,  $\mathcal{A}_i u$  et  $\mathcal{A}_i \lambda$  leurs approximations par l'interpolation de produit tensoriel complet (1.21). Alors nous avons l'estimation suivante :*

$$\|u - \mathcal{A}_i u\|_{L^2_\varrho(\Gamma; H^1_0(D))} \leq C \|\lambda - \mathcal{A}_i \lambda\|_{L^2_\varrho(\Gamma; L^2(D))}.$$

où la constante  $C > 0$  peut décroître avec  $M$ , le nombre des points de collocation.

### Preuve

1) Premièrement, on traite le cas où  $k$  est linéaire par rapport à  $y$  :

$$k = b_0 + \sum_{n=1}^N b_n y_n$$

Puisque la donnée  $f$  est polynomiale en  $y$ , alors on compare la région d'analyticité de  $u$  avec celle de  $k^{-1}$ . La fonction  $u(\cdot, y_n^*, x)$  comme fonction de la variable  $y_n$ , peut être prolongé analytiquement dans le disque  $\{|z - y_n| < \tau_n(u)\}$ , et par la règle de d'Alembert, nous avons :

$$\frac{1}{\tau_n(u)} = \limsup_{m \rightarrow \infty} \frac{\frac{\|\sqrt{k} \partial_{y_n}^m \nabla u\|_{L^2(D)}}{m!}}{\frac{\|\sqrt{k} \partial_{y_n}^{m-1} \nabla u\|_{L^2(D)}}{(m-1)!}}.$$

D'autre part :  $k^{-1}(z, y_n^*, x) = \frac{k^{-1}(y_n, y_n^*, x)}{1 + \frac{(z-y_n)b_n(x)}{k(y_n, y_n^*, x)}}$ , alors l'inverse  $k^{-1}(\cdot, y_n^*, x)$  peut être aussi prolongée analytiquement dans le disque  $\{|z - y_n| < \tau_n(k^{-1})\}$ , où :

$$\frac{1}{\tau_n(k^{-1})} = \left\| \frac{b_n}{k} \right\|_{\infty}.$$

En dérivant  $m$ -fois le problème variationnel (1.11) suivant chaque direction  $y_n$ , on obtient pour tout  $v \in H_0^1(D)$  :

$$\partial_{y_n}^m \int_D k \nabla u \nabla v dx = \int_D \sum_{l=0}^m \binom{m}{l} \partial_{y_n}^l k \partial_{y_n}^{m-l} \nabla u \nabla v dx = \partial_{y_n}^m \int_D f v dx,$$

et puisque pour  $l > 1$  on a  $\partial_{y_n}^l k = 0$  et  $\partial_{y_n} k = b_n$ , alors :

$$\int_D k \frac{\partial_{y_n}^m \nabla u}{m!} \nabla v dx = - \int_D b_n \frac{\partial_{y_n}^{m-1} \nabla u}{(m-1)!} \nabla v dx + \int_D \frac{\partial_{y_n}^m f}{m!} v dx. \quad (3.1)$$

Prenons  $v = \partial_{y_n}^m \nabla u$  dans (3.1) et par l'inégalité de Cauchy-Schwarz :

$$\frac{\|\sqrt{k} \partial_{y_n}^m \nabla u\|_{L^2(D)}}{m!} \leq \left\| \frac{b_n}{k} \right\|_{\infty} \frac{\|\sqrt{k} \partial_{y_n}^{m-1} \nabla u\|_{L^2(D)}}{(m-1)!} + \frac{C_D}{\sqrt{k_{\min}(y)}} \frac{\|\partial_{y_n}^m f\|_{L^2(D)}}{m!}.$$

La fonction  $f$  est polynomiale en  $y$ , alors :  $\limsup_{m \rightarrow \infty} \frac{\|\partial_{y_n}^m f\|_{L^2(D)}}{m!} = 0$ , on en déduit :

$$\limsup_{m \rightarrow \infty} \frac{\frac{\|\sqrt{k} \partial_{y_n}^m \nabla u\|_{L^2(D)}}{m!}}{\frac{\|\sqrt{k} \partial_{y_n}^{m-1} \nabla u\|_{L^2(D)}}{(m-1)!}} \leq \left\| \frac{b_n}{k} \right\|_{\infty}.$$

Par conséquent, nous avons ainsi obtenu  $\tau_n(u) \geq \tau_n(k^{-1})$ , qui montre que la fonction  $u$  est plus régulière que  $\lambda$ , il existe  $C(u, \lambda) > 0$  (qui peut décroître avec  $M$  le nombre des nœuds de collocation) telle que

$$\|u - \mathcal{A}_i u\|_{L^2(\Gamma; H_0^1(D))} \leq C \|\lambda - \mathcal{A}_i \lambda\|_{L^2(\Gamma; L^2(D))}.$$

2) Maintenant on suppose que  $k$  est non linéaire suivant  $y$  comme dans le cas  $k(y, x) = g(b_0(x) + \sum_{n=1}^N b_n(x) y_n)$ , où  $g$  est analytique et strictement positive.

Remarquons que la fonction  $k$  satisfait ici une hypothèse qui assure l'analyticité de  $u$  (cf [6]).

Considérons  $\{\mathcal{V}_l\}_l$ , une partition de l'ensemble  $\Gamma$  de pas  $h_l$ , telle que sur chaque élément  $\mathcal{V}_l$  nous avons l'approximation suivante pour  $k$  :

$$k(y, x) \approx \sum_{n=1}^N a_n(y^l, x) y_n \quad \forall y \in \mathcal{V}_l,$$

où  $y^l$  est le centre de l'élément  $\mathcal{V}_l$ . Considérons  $\mathcal{A}_1^l u$  l'interpolation de Lagrange de la restriction de  $u$  dans chaque élément  $\mathcal{V}_l$ . D'après le résultat précédent, l'estimation suivante est vérifiée :

$$\|u - \mathcal{A}_1^l u\|_{L^2(\mathcal{V}_l; H_0^1(D))} \leq \tilde{c} \|\lambda - \mathcal{A}_1^l \lambda\|_{L^2(\mathcal{V}_l; L^2(D))}. \quad (3.2)$$

D'autre part, si on pose  $\psi_l(y, \cdot) = u(y, \cdot) - \mathcal{A}_1 u(y, \cdot)$  pour chaque  $y \in \mathcal{V}_l$ , alors nous avons :

$\mathcal{A}_1^l \psi_l = \mathcal{A}_1^l u - \mathcal{A}_1^l(\mathcal{A}_1 u) = \mathcal{A}_1^l u - \mathcal{A}_1 u$ , puisque  $\mathcal{A}_1 u$  est un polynôme. Selon les estimations classiques de la convergence  $h_p$  de l'approximation par morceaux [18, 5], l'estimation suivante est vérifiée :

$$\|\psi_l - \mathcal{A}_1^l \psi_l\|_{L^2(\mathcal{V}_l; H_0^1(D))} \approx \hat{c} h_l^p \|\psi_l\|_{L^2(\mathcal{V}_l; H_0^1(D))}, \quad (3.3)$$

où  $p$  est le plus petit degré polynomial d'interpolation suivant les directions de  $\mathcal{A}_1^l$ . On en déduit alors :

$$\|u - \mathcal{A}_1^l u\|_{L^2(\mathcal{V}_l; H_0^1(D))} \approx \hat{c} h_l^p \|u - \mathcal{A}_1 u\|_{L^2(\mathcal{V}_l; H_0^1(D))}. \quad (3.4)$$

Avec les mêmes arguments, nous avons :

$$\|\lambda - \mathcal{A}_1^l \lambda\|_{L^2(\mathcal{V}_l; H_0^1(D))} \approx \hat{c} h_l^p \|\lambda - \mathcal{A}_1 \lambda\|_{L^2(\mathcal{V}_l; L^2(D))}. \quad (3.5)$$

Par conséquent, en combinant (3.4) et (3.5) avec (3.2) nous obtenons dans chaque élément  $\mathcal{V}_l$  :

$$\|u - \mathcal{A}_1 u\|_{L^2(\mathcal{V}_l; H_0^1(D))} \leq \tilde{C} \|\lambda - \mathcal{A}_1 \lambda\|_{L^2(\mathcal{V}_l; L^2(D))}.$$

Ainsi l'estimation est vérifiée sur  $\Gamma$ . ■

### Remarque 3.1.

*Le résultat de la proposition reste vraie pour l'interpolation de Smolyak des grilles clairsemées  $\mathcal{A}_\alpha(N, w)$ , puisque l'analyse de cette méthode repose sur l'analyticité de la fonction interpolée (cf, Théorème 1.3 et 1.4).*

*Lorsque le paramètre  $k$  est linéaire par rapport à  $y$  la constante  $C$  décroît avec le cardinal de la grille  $\mathcal{H}_\alpha(w, N)$ , puisque dans ce cas la région d'analyticité de  $u$  est plus grande que celle de l'indicateur  $\lambda$ .*

### Remarque 3.2. Choix des points de référence

*Lorsque nous tentons de déterminer le nombre de points d'interpolation  $m(i_n)$  selon la direction  $y_n$  pour l'approximation  $\mathcal{A}_1$ , nous fixons les autres directions à des points  $y_n^*$  (points de référence) et nous cherchons le nombre de point  $m(i_n)$  de telle sorte que deux approximations successives de type  $\mathcal{U}^{i_n}$  soient proches.*

Chaque degré  $m(i_n) - 1$  peut dépendre du point de référence choisi  $y_n^*$ , ainsi, il doit être choisi avec soin. Pour les champs aléatoires discutés dans la section (1.1.1), le point de référence  $y_n^* = 0$  est convenable, puisque pour ces exemples de champs, il n'y a pas d'interaction entre les dimensions stochastiques.

Dans la section qui suit, comme application de la Proposition 3.1, nous utilisons l'indicateur ( $\lambda = k^{-1}f$ ) pour déterminer les degrés de liberté stochastiques appropriés pour calculer l'opérateur d'interpolation  $\mathcal{A}_\alpha(w, N)$ .

L'intérêt d'utiliser cet indicateur, est que la recherche des paramètres  $\alpha$  et  $w$  de la formule  $\mathcal{A}_\alpha(w, N)\lambda$  exige seulement l'évaluation de problèmes d'interpolation, tandis que l'approche classique nécessite en plus, la résolution d'un grand nombre de systèmes linéaires, dont une grande partie de l'effort n'est pas exploitée dans le calcul de  $\mathcal{A}_\alpha(w, N)u$ .

### 3.3 Adaptation des grilles clairsemées

Le vecteur poids  $\alpha$  et le niveau global  $w$  sont les principaux ingrédients de la méthode de collocation sur la grille clairsemée anisotrope, donnée par la formule  $\mathcal{A}_\alpha(w, N)$ . Chaque composante  $\alpha_n$  du vecteur  $\alpha$  caractérise l'importance de la direction  $y_n$  de l'approximation de  $u$ , et le niveau global  $w$  détermine le nombre total des nœuds de la méthode. Le calcul de ces ingrédients exige généralement un effort excessif et inutile. Ainsi, en choisissant  $\alpha$  et  $w$  avec soin et avec moins d'efforts, la méthode des grilles clairsemées peut être assez efficace.

#### 3.3.1 Estimation des poids d'anisotropie

L'idée principale donnée dans [50], consiste à choisir chaque poids d'anisotropie  $\alpha_n$  comme  $\nu_n$ , le taux de convergence exponentiel suivant la direction  $y_n$  comme dans le Théorème 1.2. Selon ce choix, les directions qui ont une large région d'analyticité auront moins de points de collocation, et vice-versa.

L'erreur globale commise par une interpolation polynomiale  $\mathcal{A}_1 u$  de la fonction  $u$  peut être divisée en  $N$  parties, comme suit :

$$\|u - \mathcal{A}_1 u\|_{L^2_\varrho(\Gamma, H^1_0(D))} \leq \sum_{n=1}^N \varepsilon_n,$$



où chaque partie  $\varepsilon_n$  est une contribution de l'erreur commise suivant la direction  $y_n$ , qui est estimée, quand  $\Gamma_n$  est borné, par :

$$\varepsilon_n \approx c_n e^{-\nu_n(m(i_n)-1)}, \quad \forall n = 1, \dots, N, \quad (3.6)$$

et quand  $\Gamma_n$  est non borné, elle est estimée par :

$$\varepsilon_n \approx O(\sqrt{m(i_n) - 1}) e^{-\nu_n \sqrt{m(i_n)-1}}, \quad \forall n = 1, \dots, N, \quad (3.7)$$

où  $m(i_n)$  est le nombre de points de collocation utilisés par  $\mathcal{U}^{i_n}$ .

En général, dans de nombreux problèmes, la connaissance a priori des taux de convergence  $\{\nu_n\}_{n=1}^N$  n'est pas disponible. Ainsi une approche basée sur l'information a posteriori permet de déterminer numériquement ces taux de convergence, ainsi que les poids d'anisotropie.

Un problème est fortement anisotrope si ses dimensions représentent beaucoup de variations. Dans la pratique, beaucoup de problèmes présentent un certain degré d'anisotropie entre les directions. Un tel degré d'anisotropie suivant une direction  $y_n$  peut être quantifié par le rapport  $\frac{\alpha_n}{\underline{\alpha}}$ , où  $\underline{\alpha} = \min_{1 \leq n \leq N} \alpha_n$ .

Dans le cas d'un problème fortement anisotrope, nous observerons (voir section 3.4) que le vecteur de poids obtenu numériquement à partir de la borne (3.7) n'est pas assez convenable, dans le sens où la grille clairsemée  $\mathcal{H}_\alpha(w, N)$  pour un niveau donné  $w$  est un peu dense en comparaison avec la grille qui correspond aux poids obtenus par la borne (3.6). Il semble que la propriété d'anisotropie des poids peut être affectée par la racine carrée qui apparaît dans (3.7).

Afin de remédier à cette situation, on propose d'utiliser une estimation de l'erreur comme dans (3.6). Nous supposons, dans le cas d'un ensemble  $\Gamma$  non borné, que la densité de probabilité  $\varrho$  est à décroissance gaussienne. Ainsi chaque densité  $\rho_n$  de la variable aléatoire  $Y_n$  satisfait l'estimation suivante pour tout  $z \in \mathbb{R}$  :

$$\rho_n(z) \leq C_n e^{-\delta z^2}, \quad C_n > 0 \text{ et } \delta > 0. \quad (3.8)$$

Nous notons par  $H_p(t) \in \mathcal{P}_p(\mathbb{R})$  les polynômes d'Hermite normalisés,

$$H_p(t) = \frac{1}{\sqrt{p! 2^p \pi^{\frac{1}{2}}}} (-1)^p e^{-t^2} \left( \frac{d^p}{dt^p} \right) \mu(t).$$

Nous rappelons que les polynômes d'Hermite forment une base orthonormale complète (cf, [60]) de l'espace à poids  $L_\mu^2(\mathbb{R})$ , où  $\mu(t) = e^{-t^2}$ . Notons aussi par  $h_p(t) = \sqrt{\mu(t)} H_p(t)$ , les fonctions d'Hermite associées.

La proposition suivante donne une estimation de l'erreur commise par l'opérateur d'interpolation  $\mathcal{U}^m v$ , pour une fonction  $v(z)$  entière, possédant une croissance au plus exponentielle, satisfaisant (3.8). La preuve est basée sur une estimation des coefficients de Fourier-Hermite ([10]) pour une fonction entière et qui décroît comme un noyau gaussien.

**Proposition 3.2.**

Considérons une fonction  $\kappa$  dans  $L^2_\rho(\mathbb{R}; L^2(D))$ , où la densité  $\rho$  vérifie (3.8), et soit  $\mathcal{U}^i \kappa$  l'interpolée de Lagrange de  $\kappa$ , donnée par  $m(i)$  points de Gauss-Hermite. Supposons que pour chaque  $t \in \mathbb{R}$ , la fonction  $\kappa(t, \cdot)$  peut être prolongée analytiquement sur tout le plan complexe, et qu'il existe  $\nu > 0$  tel que

$$\|e^{-\nu|z|} \kappa(z)\|_{L^2(D)} \leq C_\nu < \infty, \quad \forall z \in \mathbb{C}. \quad (3.9)$$

Alors, il existe  $g > 0$  telle que l'estimation suivante soit vérifiée :

$$\|\kappa(t, \cdot) - \mathcal{U}^i \kappa(t, \cdot)\|_{L^2_\rho(\mathbb{R}; L^2(D))} \leq C_g e^{-gm(i)}, \quad (3.10)$$

où la constante  $C_g$  ne dépend pas de  $m(i)$  le nombre des points de Gauss pour un niveau donné  $i > 0$ .

**Preuve**

On considère la fonction  $\tilde{\kappa}(t, x) = \kappa(w(t), x)$  où  $w(t) = \frac{\sqrt{2}t}{\delta}$ . Le développement sur la base des polynômes d'Hermite de  $\tilde{\kappa}$  est donné comme suit :

$$\tilde{\kappa}(t, x) = \sum_{p=0}^{\infty} \kappa_p(x) H_p(t) \quad \text{où} \quad \kappa_p(x) = \int_{\mathbb{R}} \tilde{\kappa}(t, x) H_p(t) e^{-t^2} dt. \quad (3.11)$$

Définissons la fonction  $F(t, x) = \tilde{\kappa}(t, x) e^{-\frac{t^2}{2}}$ . Le développement en série de Fourier-Hermite de la fonction  $F$  est donné par :

$$F(t, x) = \sum_{p=0}^{\infty} F_p(x) h_p(t) \quad \text{où} \quad F_p(x) = \int_{\mathbb{R}} F(t, x) h_p(t) dt. \quad (3.12)$$

On en déduit que les coefficients d'Hermite de  $F$  coïncident avec les coefficients  $\kappa_p$  (3.11) de la fonction  $\tilde{\kappa}$ . En effet,

$$F_p(x) = \int_{\mathbb{R}} F(t, x) h_p(t) dt = \int_{\mathbb{R}} \tilde{\kappa}(t, x) e^{-t^2} H_p(t) dt = \kappa_p.$$

La fonction  $F$  est entière comme produit de deux fonctions entières. D'autre part, puisque  $\kappa$  satisfait la borne (3.9), alors  $F$  décroît comme un noyau gaussien puisque,

$$\|\tilde{\kappa}(t, x)\| e^{-t^2/2} \leq C_\nu e^{\sqrt{2}\frac{\nu}{\delta}t} e^{-t^2/2} \leq C_\nu e^{-\frac{1}{2}(t-\sqrt{2}\frac{\nu}{\delta})^2} e^{\frac{\nu^2}{\delta^2}} \leq d e^{-qt^2},$$

où les deux constantes  $d$  et  $q$  sont positives et dépendent de  $\nu$  et  $\delta$ .

On en déduit que les coefficients de Fourier-Hermite  $F_p$  décroissent géométriquement [10] de la façon suivante :

$$F_p \leq K e^{-gp} \quad \forall p \in \mathbb{N},$$

où le taux  $g > 0$  et la constante  $K > 0$  est indépendante de  $p$ .

D'autre part, l'erreur peut être majorée comme suivant (cf [6], Lemme 7) :

$$\|\kappa(t, \cdot) - \mathcal{U}^i \kappa(t, \cdot)\|_{L^2_\rho(\mathbb{R}; L^2(D))} \leq \min_{v \in \mathcal{P}_{p(i)} \otimes L^2(D)} \max_{t \in \mathbb{R}} \left\| \|\kappa - v\|_{L^2(D)} e^{-\frac{(td)^2}{4}} \right\|.$$

En prenant  $v = \Pi_{m(i)} \kappa$  comme la troncature du développement d'Hermite de  $\kappa$  donné dans (3.11) jusqu'à l'ordre  $m(i)$ , et  $\tilde{v} = \Pi_{m(i)}(\tilde{\kappa})$  celui de  $\tilde{\kappa}$ , on en déduit que :

$$\begin{aligned} \|\kappa(t, \cdot) - \mathcal{U}^i \kappa(t, \cdot)\|_{L^2_\rho(\mathbb{R}; L^2(D))} &\leq \max_{t \in \mathbb{R}} \left\| \|\kappa - \Pi_{m(i)}(\kappa)\|_{L^2(D)} e^{-\frac{(td)^2}{4}} \right\| \\ &\leq \max_{t \in \mathbb{R}} \left\| \|\tilde{\kappa} - \Pi_{m(i)}(\tilde{\kappa})\|_{L^2(D)} e^{-\frac{t^2}{2}} \right\| \\ &\leq \sum_{p=m(i)+1}^{\infty} \|k_p\|_{L^2(D)} \max_{t \in \mathbb{R}} \left| H_p(t) e^{-\frac{t^2}{2}} \right| \\ &\leq C_1 \sum_{p=m(i)+1}^{\infty} e^{-gp} \max_{t \in \mathbb{R}} |h_p(t)|. \end{aligned}$$

D'autre part nous avons  $\max_{t \in \mathbb{R}} |h_p(t)| \leq 1$  (cf, [10]), ainsi, on en déduit que :

$$\|\kappa(t, \cdot) - \mathcal{U}^i \kappa(t, \cdot)\|_{L^2_\rho(\mathbb{R}; L^2(D))} \leq C_1 \sum_{p=m(i)+1}^{\infty} e^{-gp} \leq C_2 e^{-gm(i)}.$$

■

### Remarque 3.3.

Quand le champ de diffusion  $k$  est exponentiel,  $k = e^G$ , alors  $\lambda = e^{-G} f$  est analytique sur tout le plan  $\mathbb{C}$  suivant chaque direction  $y_n$ . Ainsi dans le cas des variables aléatoires non bornées, on peut utiliser l'estimation (3.10) pour calculer le taux de convergence suivant chaque direction de l'indicateur  $\lambda$ .

De nombreux tests numériques montrent aussi que le taux de décroissance donné dans (3.10) donne une convergence de la formule  $\mathcal{A}_\alpha(w, N)$  et ceci pour des champs non linéaires qui sont seulement analytiques.

Afin de déterminer le vecteur des poids  $\alpha$ , nous utilisons l'estimation d'erreur (3.10) de la Proposition 3.2, pour un ensemble  $\Gamma$  borné ou non borné. Ainsi nous utilisons une

approche a posteriori pour estimer chaque poids  $\alpha_n = g_n$ , pour chaque  $n = 1, \dots, N$ . On choisit la règle de croissance  $m(i_n) = i_n$ , l'algorithme suivant permet de calculer le vecteur poids  $\alpha$  pour  $\lambda$ .

**Algorithme pour le calcul de  $\alpha_n$**

- On fixe  $y_n^* = 0$  comme point de référence, i.e  $y_j = 0$  pour tout  $j \neq n$ .
- On considère le nombre  $\tilde{m}_n$  tel que  $\tilde{\lambda}(y_n, 0) := \mathcal{U}^{\tilde{m}_n} \lambda(y_n, 0)$  est considérée comme approximation de référence de  $\lambda$  suivant la direction  $y_n$ .
- pour chaque  $p = 1, \dots, \tilde{m}_n - 1$ , on calcule l'erreur :

$$\text{err}_n(p) = \|\tilde{\lambda}(y_n, 0) - \mathcal{U}^p \lambda(y_n, 0)\|$$

- On utilise l'approximation des moindres carrés pour estimer  $\alpha_n$  et  $c_n$  solutions du problème de minimisation suivant :

$$\min_{\alpha_n, c_n} \|\log(\text{err}_n) - \log(c_n) + P\alpha_n\|_2^2$$

où  $\text{err}_n$  est le vecteur dont chaque composante  $\text{err}_n(p)$ , représente l'erreur commise par  $\mathcal{U}^p \lambda$ . Les composantes du vecteur  $P$  sont données par le nombre de points de collocation  $1, \dots, \tilde{m}_n - 1$ . ■

On remarque que le calcul des poids  $\alpha_1, \dots, \alpha_N$ , avec cet algorithme, exige seulement l'évaluation d'une suite de problèmes d'interpolation. Le nombre total de problème est de cardinal  $S = \sum_{n=1}^N \frac{\tilde{m}_n(\tilde{m}_n+1)}{2}$ . Ainsi le coût de calcul est nettement réduit, en comparaison avec l'approche utilisant  $u$  qui exige la résolution de même nombre  $S$  de systèmes linéaires de grande taille, chacun est une approximation de (1.19). L'algorithme qui approche les poids de  $u$  [50] est donné par les étapes suivantes

- On fixe  $y_n^* = 0$  comme point de référence, i.e  $y_j = 0$  pour tout  $j \neq n$ .
- On considère le nombre  $\tilde{m}_n$  tel que  $\tilde{u}_h(y_n, 0) := \mathcal{U}^{\tilde{m}_n} u_h(y_n, 0)$  est considérée comme approximation de référence de  $u$  suivant la direction  $y_n$ .
- pour chaque  $p = 1, \dots, \tilde{m}_n - 1$ , on calcule l'erreur :

$$\text{err}_n(p) = \|\tilde{u}_h(y_n, 0) - \mathcal{U}^p u_h(y_n, 0)\|$$

- On utilise l'approximation des moindres carrés pour estimer  $\alpha_n$  et  $c_n$  solutions du problème de minimisation suivant :

$$\min_{\alpha_n, c_n} \|\log(\text{err}_n) - \log(c_n) + P\alpha_n\|_2^2,$$

pour chaque points d'interpolation  $u_h$  est solution du problème (1.19),  $\text{err}_n$  est le vecteur dont chaque composante  $\text{err}_n(p)$ , représente l'erreur commise par  $\mathcal{U}^p u_h$ . Les composantes du vecteur  $P$  sont données par le nombre de points de collocation  $1, \dots, \tilde{m}_n - 1$ . ■

### 3.3.2 Estimation du niveau $w$ convenable

Dans les approches classiques, lorsque l'on s'intéresse à l'approximation de  $u$  par l'interpolation  $\mathcal{A}_\alpha(w, N)u$ , le niveau convenable  $w$  de l'approximation satisfait la condition d'arrêt  $\|\mathcal{A}_\alpha(w, N)u - \mathcal{A}_\alpha(w+1, N)u\| \approx \varepsilon$ , où  $\varepsilon$  est une tolérance donnée. Généralement l'effort de l'approximation  $\mathcal{A}_\alpha(w+1, N)u$  au niveau  $w+1$  est comparable, parfois supérieur à tous les efforts combinés des niveaux antérieurs, même dans le cas où les points de la grille sont imbriqués.

On constate alors, que ce critère d'arrêt standard pour l'approximation de  $u$  peut être très couteux, en particulier, quand la dimension  $N$  est grande.

Le résultat de la Proposition 3.1 montre que l'erreur commise sur l'indicateur  $\lambda$  est une borne supérieure de l'erreur sur  $u$ . Ainsi nous pouvons utiliser la borne supérieure  $C\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\|$  afin de trouver un niveau approprié pour l'approximation  $\mathcal{A}_\alpha(w, N)u$ . L'approche d'adaptation consiste à choisir  $w \in \mathbb{N}$  tel que, pour  $\varepsilon > 0$ , une tolérance désirable, l'interpolation  $\mathcal{A}_\alpha(w, N)\lambda$  satisfait l'estimation suivante :

$$\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\| \leq \frac{\varepsilon}{C}, \quad (3.13)$$

où la constante  $C$  est positive, donnée dans la Proposition 3.1. Elle peut dépendre du niveau  $w$  dans le cas où le champ aléatoire est linéaire par rapport à  $y$ . Cette dépendance se traduit par la décroissance de  $C$  en fonction du cardinal  $M$  des points de la grille clairsemée. Par conséquent, l'estimation de  $C$  peut être faite suivant le cas où le champ  $k$  est linéaire ou non linéaire par rapport à la variable  $y$ .

#### Cas où $k$ est non linéaire suivant $y$

Dans ce cas, on suppose que la constante  $C$  ne dépend pas de  $w$ , en particulier qu'elle ne décroît pas en fonction de  $M$ , le cardinal de la grille clairsemée. Le théorème 1.4 montre que l'erreur commise par l'approximation  $\mathcal{A}_\alpha(w, N)u$  satisfait l'estimation suivante :

$$\|u - \mathcal{A}_\alpha(w, N)u\| \approx C_1\|u\|M^{-\mu_1}, \quad (3.14)$$

où le taux de convergence  $\mu_1$  dépend de la taille de la région d'analyticité de  $u$ . Nous avons la même estimation pour l'approximation  $\mathcal{A}_\alpha(w, N)\lambda$  pour  $\lambda$  :

$$\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\| \approx C_1\|\lambda\|M^{-\mu_2}. \quad (3.15)$$

Ainsi, si dans ce cas on suppose que  $\mu_1 \approx \mu_2$ , alors la constante  $C$  est estimée par :

$$C \approx \frac{\|u\|}{\|\lambda\|} \approx \frac{\|\mathcal{A}_\alpha(0, N)u\|}{\|\mathcal{A}_\alpha(0, N)\lambda\|}. \quad (3.16)$$

### Cas où $k$ est linéaire suivant $y$

Dans ce cas la constante  $C$  décroît avec le nombre de points de la grille, puisque la région d'analyticité de  $\lambda$  est contenue dans celle de  $u$ , comme le montre la preuve de la Proposition 3.1. L'erreur commise suivant chaque direction par une interpolation polynomiale de  $\lambda$  est plus grande que celle commise par une interpolation de  $u$  puisque nous avons  $\mu_1 \geq \mu_2$ .

Le taux  $\mu_1$  dépend de  $\alpha_{min}$  le plus petit poids d'anisotropie de l'indicateur  $\lambda$  et  $\mu_2$  dépend du minimum de ceux de la fonction  $u$ . Puisque la contribution de l'erreur est grande suivant la direction de plus petit taux de convergence, alors pour chaque niveau  $w$ , on estime la constante  $C$  par l'estimation suivante :

$$C \approx C_w = \frac{\|\mathcal{A}_\alpha(0, N)u\|}{\|\mathcal{A}_\alpha(0, N)\lambda\|} M_w^{-\gamma}, \quad (3.17)$$

où l'exposant  $\gamma > 0$  est calculé numériquement de la façon suivante :

Lorsque nous estimons le vecteur poids  $\alpha$  pour l'indicateur  $\lambda$ , nous considérons la direction  $y_{\tilde{n}}$  dont le poids correspondant satisfait le minimum *i.e.*,  $\alpha_{\tilde{n}} = \min_{1 \leq n \leq N} \alpha_n$ . Nous calculons à nouveau le poids  $\tilde{\alpha}_{\tilde{n}}$  de la solution  $u$  seulement suivant la direction  $n = \tilde{n}$ . Ainsi nous posons  $\gamma = \tilde{\alpha}_{\tilde{n}} - \alpha_{\tilde{n}}$ .

L'utilisation de l'indicateur  $\lambda$  pour déterminer le niveau  $w$  et le fait de le réutiliser pour calculer l'interpolation  $\mathcal{A}_\alpha(w, N)u$ , nous permet de réduire et parfois d'éviter des calculs inutiles. En particulier, les problèmes déterministes calculés à partir des points de collocation dont les indices  $i_n$  ne sont pas admissibles *i.e.*,  $\mathbf{i} \notin X_\alpha(w, N)$ . Le calcul des poids d'anisotropie ainsi que la formule  $\mathcal{A}_\alpha(w, N)\lambda$  n'exige qu'une suite de problèmes d'interpolation, au lieu de problèmes d'EDP exigés lorsqu'on essaye d'adapter directement la fonction  $u$ .

### 3.4 Tests numériques

Cette section illustre la convergence et l'efficacité de l'approche adaptative proposée pour le problème stochastique elliptique linéaire. Nous considérons un problème bidimensionnel (à deux dimensions spatiales). Nous considérons deux exemples de paramètre de diffusion  $k$  paramétrisé par des variables aléatoires bornées et non bornées. Les résultats des calculs présentés ici sont réalisés avec le logiciel Matlab.

$$\begin{cases} -\operatorname{div}(k\nabla u) = f & \text{dans } \Omega \times D := \Omega \times ]0, 1]^2 \\ u|_{\partial D} = 0 \end{cases} \quad (3.18)$$

où la donnée  $f$  est déterministe,  $f(\omega, x, z) = \cos(x) \sin(z)$ , pour  $x, z \in [0, 1]$ .

#### 3.4.1 Cas du paramètre $k$ non linéaire

On considère dans un premier exemple, une diffusion  $k_N$  qui dépend seulement de  $x \in [0, 1]$  et qui est non linéaire suivant les variables stochastiques comme suit ([50]) :

$$k_N(\omega, x) = 0.5 + \exp\left(Y_1(\omega) \left(\frac{\sqrt{\pi}l_c}{2}\right)^{1/2} + \sum_{n=2}^N \zeta_n \varphi_n(x) Y_n(\omega)\right), \quad (3.19)$$

les réels  $(\zeta_n)_{n=1}^N$  sont donnés par :

$$\zeta_n := \left(\sqrt{\pi}l_c\right)^{1/2} \exp\left(\frac{-\left(\lfloor \frac{n}{2} \rfloor \pi l_c\right)^2}{8}\right), \quad \text{si } n > 1$$

et les fonctions  $\{\varphi_n\}_{n=1}^N$  sont définies par :

$$\varphi_n(x) := \begin{cases} \sin\left(\frac{\lfloor \frac{n}{2} \rfloor \pi x}{2}\right) & \text{si } n \text{ est pair} \\ \cos\left(\frac{\lfloor \frac{n}{2} \rfloor \pi x}{2}\right) & \text{si } n \text{ est impair.} \end{cases}$$

Le paramètre  $l_c$  représente la longueur de corrélation du paramètre  $k$  et l'ensemble des variables aléatoires  $\{Y_n\}_{n=1}^N$  sont de même loi, indépendantes, de moyenne nulle et de variance unité.

#### Cas des variables non bornées

Pour étudier la convergence et l'efficacité de l'approche proposée pour la construction de Smolyak anisotrope, nous considérons le cas où les variables  $Y_1, \dots, Y_N$  sont des

variables aléatoires gaussiennes, et nous considérons les deux cas où la longueur de corrélation  $l_c = 1/2$  avec  $N = 11$  et  $l_c = 0,2$  avec  $N = 41$ .

Nous allons approcher  $u$  par les formules  $\mathcal{A}_\alpha(w, N)u$  et calculer, à la fois,  $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$  et  $\text{var}[\mathcal{A}_\alpha(w, N)u]$  comme des approximations de l'espérance et de la variance de  $u$  respectivement.

Nous déterminons d'abord le vecteur des poids appropriés  $\alpha$  pour l'indicateur  $\lambda = k^{-1}f$  avec une approche a posteriori, puis nous cherchons un niveau convenable  $w$  telle que l'interpolation  $\mathcal{A}_\alpha(w, N)u$  satisfait un seuil donné pour l'espérance  $\mathbb{E}[u]$  et la variance  $\text{var}[u]$ . Notons que si l'espérance approchée  $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$  satisfait une tolérance espérée pour un certain  $w$ , i.e  $\|\mathbb{E}[u] - \mathbb{E}[\mathcal{A}_\alpha(w, N)u]\| \approx \text{tol}$ , il en est de même pour l'approximation de la variance  $\|\text{var}[u] - \text{var}[\mathcal{A}_\alpha(w, N)u]\| \approx \text{tol}$ , avec le même niveau  $w$ . En effet la fonction  $(u - \mathbb{E}[u])^2$  est analytique, avec un diamètre de la région d'analyticité au moins égal à celui de  $u$ .

Ici nous avons  $\Gamma = \mathbb{R}^N$ , et les points de collocation dans ce cas sont les racines de polynômes d'Hermite. L'espace d'éléments finis pour la discrétisation spatiale est donné par une combinaison des fonctions linéaires par morceaux sur un maillage uniforme de  $D$  avec  $n_d = 712$  inconnues, et le pas de discrétisation spatiale  $h = 0.05$ .

Pour déterminer le vecteur poids  $\alpha$  dans ce cas ( $\Gamma$  non borné), nous utilisons l'estimation d'erreur donnée dans (3.10) de la Proposition 3.2 au lieu de l'estimation (3.7).

Nous désignons par  $\alpha$  et  $\tilde{\alpha}$  les deux vecteurs poids obtenus par une information a posteriori en utilisant la borne (3.10), respectivement pour l'indicateur  $\lambda$  et pour la fonction  $u$ . Notons aussi par  $\hat{\alpha}$  le vecteur poids pour  $\lambda$  obtenu en utilisant dans l'algorithme la borne  $e^{-\text{qrt}\alpha_n m_n}$ . Dans le Tableau 3.1, on compare ces différents poids d'anisotropie, dont le cas  $N = 41$ , on présente les 11 premières valeurs.

On peut remarquer, à partir du Tableau 3.1, que les composantes de vecteur poids  $\alpha$  sont comparables et proches de celles de  $\tilde{\alpha}$ . Cependant, celles du vecteur  $\hat{\alpha}$  sont complètement différentes, car elles sont obtenues avec deux bornes différentes (3.10) et (3.7). En comparant les valeurs de  $\alpha$  avec celles de  $\hat{\alpha}$ , on remarque que :

$$\frac{\alpha_n}{\alpha_{\min}} > \frac{\hat{\alpha}_n}{\hat{\alpha}_{\min}}, \quad \forall n = 1, \dots, N.$$

Ainsi, on conclut que la borne (3.7) utilisée pour le calcul de vecteur  $\hat{\alpha}$  conduit à une perte d'anisotropie entre les directions.

La figure 3.1 donne le cardinal des grilles anisotropes clairsemées  $\mathcal{H}_\alpha(w, N)$  et



$l_c \backslash \alpha$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$	$\alpha_{11}$
1/2	2.57	2.26	2.27	4.31	4.31	8.18	8.18	12.50	12.51	20.58	20.59
1/5	3.13	2.55	2.55	2.88	2.88	3.4	3.4	4.34	4.35	5.92	5.92
$l_c \backslash \alpha$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\alpha}_4$	$\tilde{\alpha}_5$	$\tilde{\alpha}_6$	$\tilde{\alpha}_7$	$\tilde{\alpha}_8$	$\tilde{\alpha}_9$	$\tilde{\alpha}_{10}$	$\tilde{\alpha}_{11}$
1/2	2.57	2.36	2.09	3.58	3.51	8.31	7.04	11.73	10.8	21.48	21.77
1/5	3.13	2.88	2.50	2.33	2.16	2.56	2.62	3.55	3.27	4.22	4.18
$l_c \backslash \alpha$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\alpha}_6$	$\hat{\alpha}_7$	$\hat{\alpha}_8$	$\hat{\alpha}_9$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$
1/2	9.47	7.91	7.92	15.14	15.14	21.87	21.87	30.16	30.17	42.96	42.95
1/5	10.98	9.25	9.27	10.24	10.24	12.03	12.04	15.34	15.34	16.17	16.17

TABLE 3.1 –  $N = 11$  Premières valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori.

$\mathcal{H}_{\hat{\alpha}}(w, N)$ . Nous observons que les poids obtenus par la borne d'erreur (3.10), comme indiqué dans la proposition 3.2, sont bien meilleurs que ceux obtenus en utilisant l'estimation (3.7), puisque  $\#\mathcal{H}_{\alpha}(w, N) \ll \#\mathcal{H}_{\hat{\alpha}}(w, N)$ .

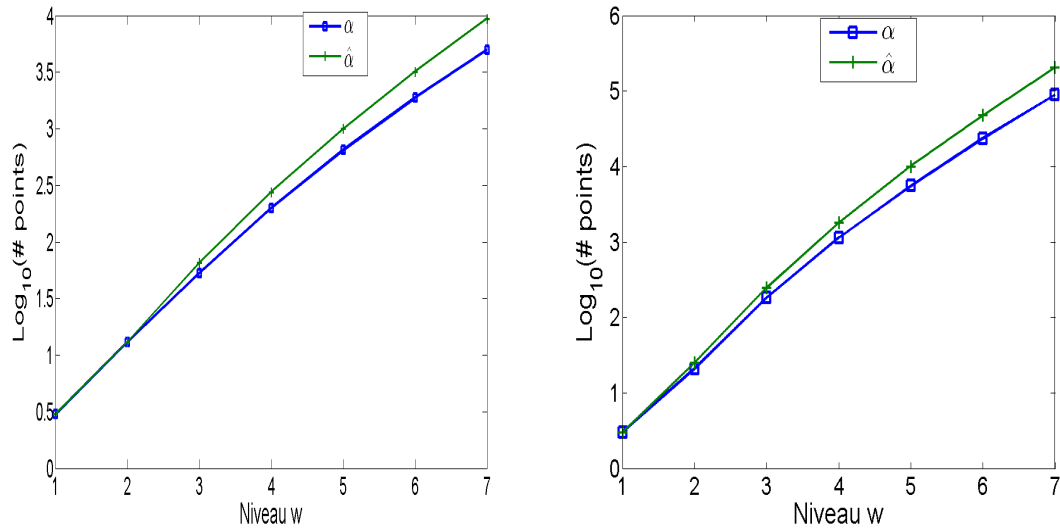


FIGURE 3.1 – Comparaison de la grille  $\#\mathcal{H}_{\hat{\alpha}}(w, N)$  avec  $\#\mathcal{H}_{\alpha}(w, N)$ , (à gauche  $l_c = 0.5$ , et à droite  $l_c = 0.2$ ).

Dans la Figure (3.2), nous comparons le cardinal de la grille clairsemée  $\mathcal{H}_{\alpha}(w, N)$

avec celui de la grille  $\mathcal{H}_{\tilde{\alpha}}(w, N)$ . Nous observons qu'ils sont très comparables, en effet  $\#\mathcal{H}_{\alpha}(w, N) \approx \#\mathcal{H}_{\tilde{\alpha}}(w, N)$ .

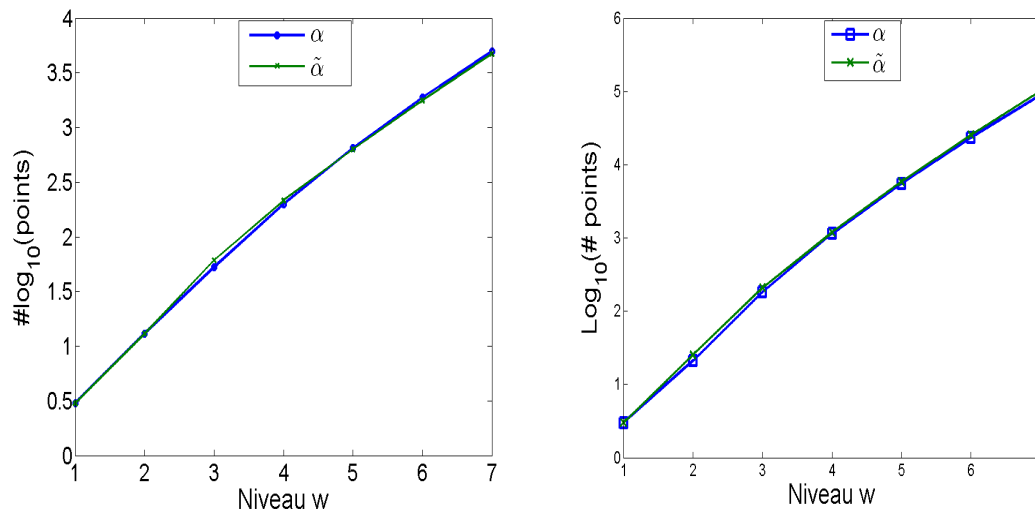


FIGURE 3.2 – Comparaison du cardinal de la grille clairsemée  $\mathcal{H}_{\alpha}(w, N)$  avec celui de la grille  $\mathcal{H}_{\tilde{\alpha}}(w, N)$ , (à gauche  $l_c = 0.5$ , à droite  $l_c = 0.2$ ).

Par la proposition 3.1, l'indicateur  $\lambda = k^{-1}f$  montre que l'erreur  $C\|\lambda - \mathcal{A}_{\alpha}(w, N)\lambda\|$  est une borne supérieure de l'erreur  $\|u - \mathcal{A}_{\alpha}(w, N)u\|$ , ainsi nous l'utilisons pour déterminer un niveau convenable  $w$  satisfaisant un critère désiré  $\|u - \mathcal{A}_{\alpha}(w, N)u\| \approx tol$ . Pour cela, nous cherchons  $w$  qui vérifie  $C\|\lambda - \mathcal{A}_{\alpha}(w, N)\lambda\| \approx tol$ , où la constante  $C$  est estimée par (3.16).

Pour la comparaison des quantités de l'erreur, nous avons pris  $\mathcal{A}_{\alpha}(9, 11)$  et  $\mathcal{A}_{\alpha}(7, 41)$  quand  $l_c = 0.5$ , respectivement  $l_c = 0.2$ , comme solution de référence pour  $u$  et  $\lambda$ .

La figure 3.3 montre que les deux quantités  $\|u - \mathcal{A}_{\alpha}(w, N)u\|$  et  $C\|\lambda - \mathcal{A}_{\alpha}(w, N)\lambda\|$  ont le même comportement.

### Cas des variables bornées

Nous considérons dans ce cas le même exemple du paramètre  $k$  avec les variables aléatoires  $Y_1, \dots, Y_N$  bornées et uniformément réparties dans  $[-1, 1]$ . Les points de collocation sont donnés comme les racines des polynômes de Legendre ( Points de

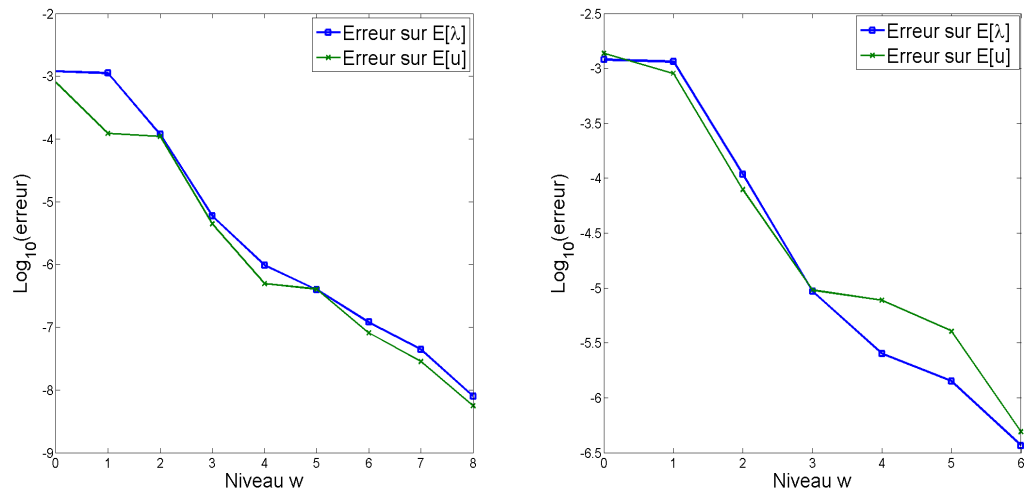


FIGURE 3.3 – Comparaison de l'erreur  $C\|\mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w, N)\lambda]\|$  avec celle commise sur  $u$  (à gauche  $l_c = 0.5$ , à droite  $l_c = 0.2$ ).

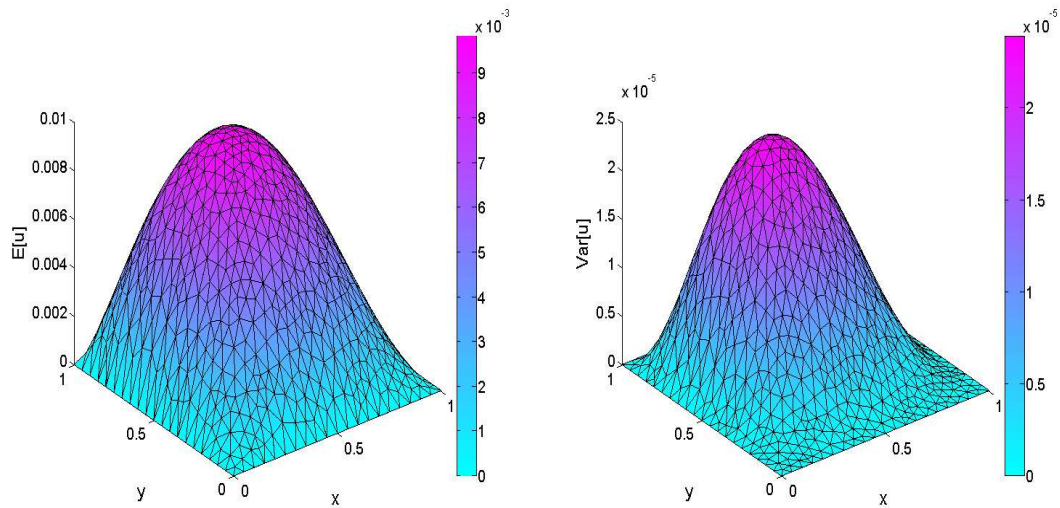


FIGURE 3.4 – L'espérance approchée  $\mathbb{E}[\mathcal{A}_\alpha(5, 41)]$  et la variance approchée  $\text{var}[\mathcal{A}_\alpha(5, 41)]$ .

Gauss-Legendre) ou les points Clenshaw-Curtis.

Pour déterminer le vecteur poids  $\alpha$ , nous utilisons les points de Gauss-Legendre et la borne d'erreur (3.6).

Le tableau 3.2 donne les deux vecteurs poids :  $\alpha$  pour l'indicateur  $\lambda$  et  $\tilde{\alpha}$  pour

la fonction  $u$ . On remarque qu'ils sont très proches et assez comparables. En outre, comme le montre la Figure 3.5,  $\#\mathcal{H}_\alpha(w, N)$  est plus petit que  $\#\mathcal{H}_{\tilde{\alpha}}(w, N)$  et c'est un autre bénéfice. Dans le cas  $l_c = 0.5$  où  $N = 11$ , nous prenons  $\mathcal{A}_\alpha(8, N)$  comme solution de référence pour les deux fonctions  $u$  et  $\lambda$ . Pour le cas  $l_c = 0.2$  où  $N = 41$ , nous considérons  $\mathcal{A}_\alpha(6, N)$  comme solution de référence. La figure 3.6 permet de comparer l'erreur sur  $u$  avec celle sur  $\lambda$ .

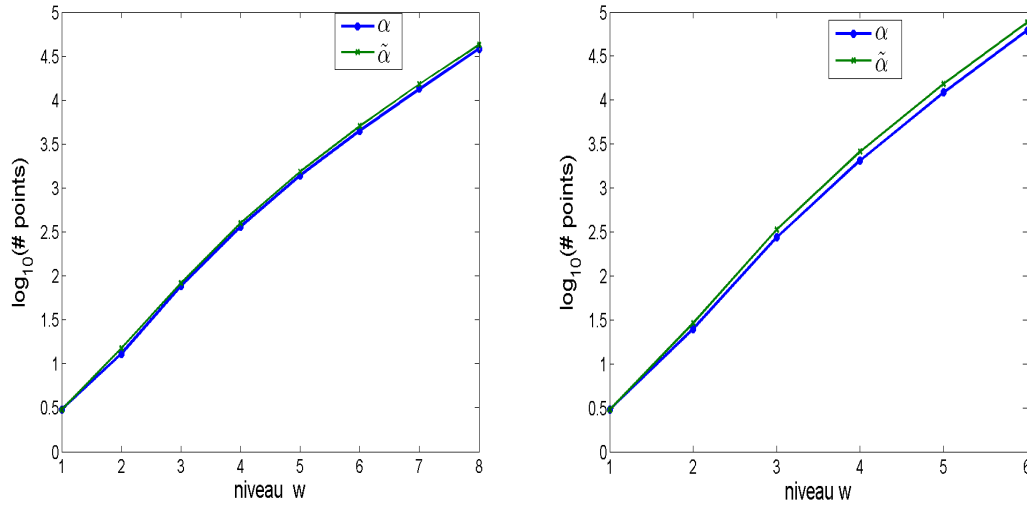


FIGURE 3.5 – Comparaison du cardinal de la grille clairsemée  $\mathcal{H}_\alpha(w, N)$  avec celui de la grille  $\mathcal{H}_{\tilde{\alpha}}(w, N)$ , (gauche  $l_c = 0.5$ , droite  $l_c = 0.2$ ).

$l_c \backslash \alpha$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$	$\alpha_{11}$
1/2	4.83	4.86	4.86	7.21	7.21	10.19	10.20	14.52	14.52	21.68	21.68
1/5	5.72	5.24	5.25	5.54	5.54	6.56	6.56	7.24	7.25	8.13	8.13
$l_c \backslash \alpha$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\alpha}_4$	$\tilde{\alpha}_5$	$\tilde{\alpha}_6$	$\tilde{\alpha}_7$	$\tilde{\alpha}_8$	$\tilde{\alpha}_9$	$\tilde{\alpha}_{10}$	$\tilde{\alpha}_{11}$
1/2	4.83	5.19	4.53	5.92	6.14	10.33	9.5	13.75	12.81	22.58	22.87
1/5	5.72	5.58	4.92	4.82	4.73	5.12	4.97	5.9	5.42	6.44	6.26

TABLE 3.2 – Les 11 premières valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori.

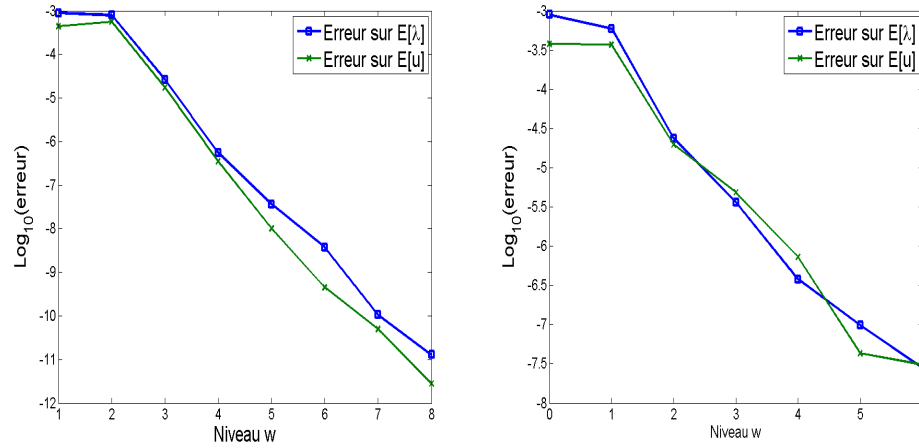


FIGURE 3.6 – Comparaison de l’erreur  $C\|\mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w, N)\lambda]\|$  avec celle sur  $u$  (à gauche  $l_c = 0.5$ , à droite  $l_c = 0.2$ ).

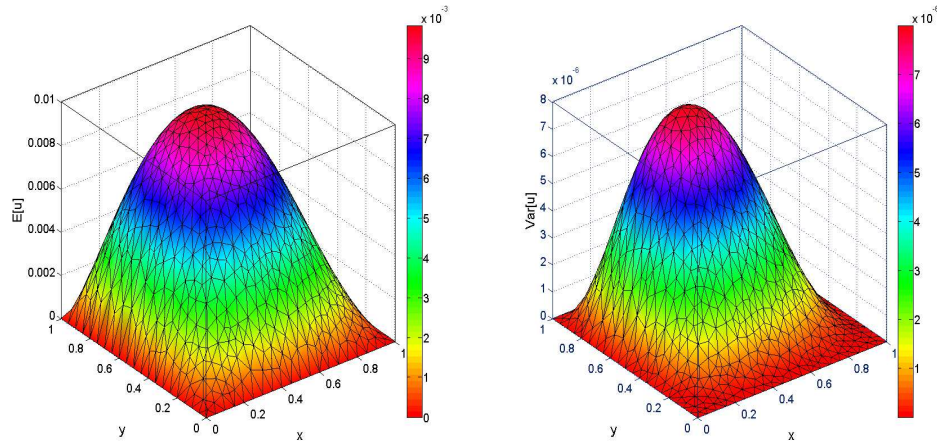


FIGURE 3.7 – L’espérance approchée  $\mathbb{E}[\mathcal{A}_\alpha(4, 41)]$  et la variance  $\text{var}[\mathcal{A}_\alpha(4, 41)]$ , le cas des variables bornés.

### 3.4.2 Cas du paramètre $k$ linéaire

Dans le second exemple, nous considérons un paramètre de diffusion  $k_N$  qui est linéaire par rapport à la variable stochastique  $y \in \Gamma$ , où  $\Gamma$  est un ensemble borné dans  $\mathbb{R}^N$ , comme dans le cas du développement de Karhunen-Loève (K.L) :

$$k(\omega, x) = \frac{\pi^2}{6} + \sum_{n=1}^N \frac{1}{1+n^2} \cos(n\pi x_1) \sin(n\pi x_2) Y_n(\omega), \quad x \in [0, 1]^2.$$

et la donnée  $f(\omega, x) = 0.1 + \sin(x_1)x_2Y_1(\omega)$ . Les variables aléatoires  $Y_1, \dots, Y_N$  sont supposées indépendantes et identiquement distribuées selon une loi uniforme dans  $[-1, 1]$ .

Nous prenons  $N = 51$  et nous calculons  $\alpha$  en utilisant une approche a posteriori. Nous considérons  $\mathcal{A}_\alpha(9, N)$  comme une approximation de référence pour  $u$  et  $\lambda$ .

Le Tableau 3.3 montre les dix premiers coefficients de pondération pour l'indicateur  $\lambda = k^{-1}f$  et la fonction  $u$ . On remarque que les composantes du vecteur  $\alpha$  sont légè-

	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$
$\alpha$ pour $\lambda$	3.89	5.77	7.19	8.26	9.11	9.87	10.48	11.01	11.48	11.9
$\tilde{\alpha}$ pour $u$	4.24	6.44	8.00	9.32	10.19	10.92	11.53	12.06	12.53	12.96

TABLE 3.3 – Les 10 premières valeurs des poids d'anisotropie obtenues par une approche a posteriori.

rement inférieures aux composantes de vecteur  $\#\mathcal{H}_{\tilde{\alpha}}(w, N)$ . Ceci est cohérent avec la théorie, car quand  $k$  est linéaire suivant  $y$ , nous avons montré que la région analytique de la fonction  $u$  est plus grande que celle de l'indicateur  $\lambda$ .

La figure 3.8 (à gauche) montre que le cardinal de la grille clairsemée  $\#\mathcal{H}_\alpha(w, N)$  est légèrement plus petit que la grille  $\#\mathcal{H}_{\tilde{\alpha}}(w, N)$ . Cela peut être expliqué numériquement par le fait que nous avons

$$\frac{\alpha}{\min(\alpha)} \geq \frac{\tilde{\alpha}}{\min(\tilde{\alpha})}.$$

Ainsi, en utilisant le vecteur poids  $\alpha$  pour calculer l'interpolation  $\mathcal{A}_\alpha(w, N)u$ , on résout moins de systèmes linéaires que si on utilise le vecteur  $\tilde{\alpha}$ .

La figure 3.8 (à droite) compare l'erreur sur  $u$  avec  $C_w \|\mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w, N)\lambda]\|$ , où  $C_w$  diminue avec  $M_w$  le cardinal de la grille de niveau  $w$ ,  $\mathcal{H}_\alpha(w, N)$ . On l'estime par  $C_w \approx \frac{\|\mathcal{A}_\alpha(0, N)u\|}{\|\mathcal{A}_\alpha(0, N)\lambda\|} M_w^\mu$ , où  $\mu = \min(\tilde{\alpha}) - \min(\alpha) = \tilde{\alpha}_1 - \alpha_1$ .

## 3.5 Conclusion

Dans ce chapitre, nous avons présenté une stratégie adaptative de la méthode de collocation sur des grilles clairsemées pour l'approximation d'un problème elliptique avec des données stochastiques. L'approximation du problème est donnée par une méthode de collocation anisotrope dans l'espace stochastique, et une approximation d'éléments finis dans l'espace physique.

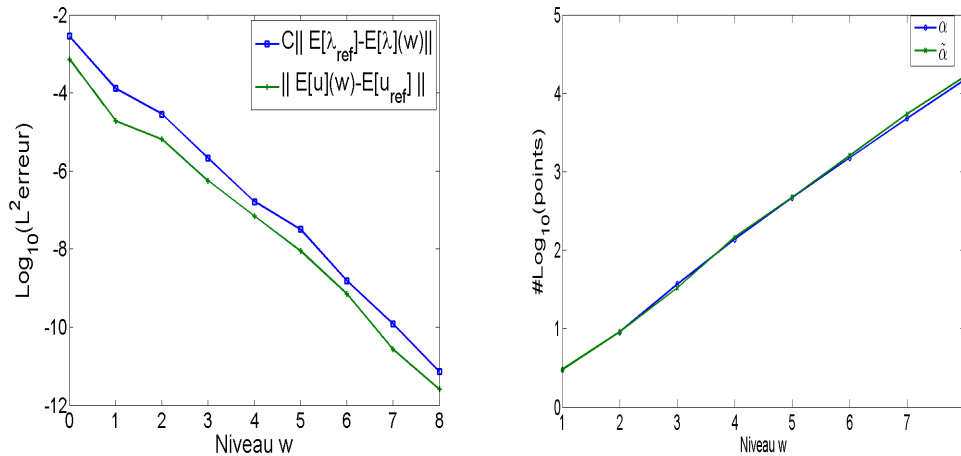


FIGURE 3.8 – À gauche : Comparaison de l’erreur  $C_w \| \mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w, N)\lambda] \|_\infty$  avec celle sur  $u$ . À droite : Comparaison du cardinal de la grille  $\mathcal{H}_\alpha(w, N)$  avec celui de la grille  $\mathcal{H}_{\tilde{\alpha}}(w, N)$ .

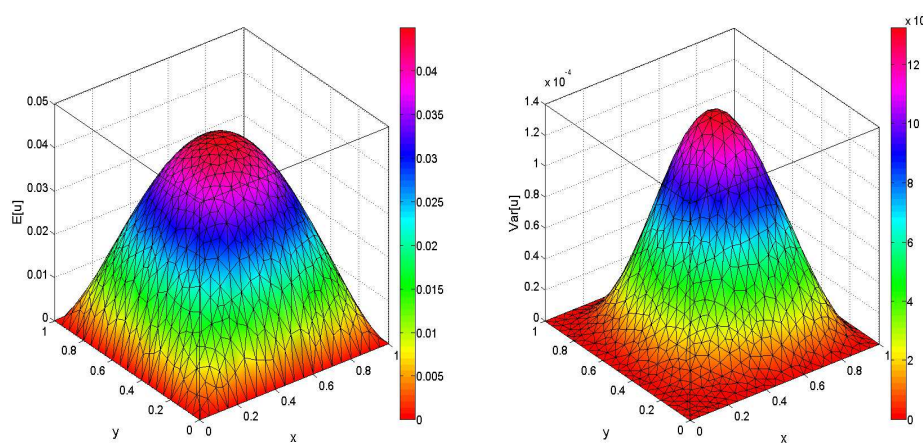


FIGURE 3.9 – L’espérance approchée  $\mathbb{E}[\mathcal{A}_\alpha(6, 51)]$  et la variance  $\text{var}[\mathcal{A}_\alpha(6, 51)]$ .

L’objectif était d’estimer les coefficients d’anisotropie et un niveau convenable de la méthode de collocation. Nous avons mis en place un indicateur d’erreur en fonction de l’inverse du paramètre de la diffusion, permettant ainsi de définir une borne supérieure de l’erreur commise dans l’espace stochastique.

La stratégie a consisté à utiliser cet indicateur, d’abord pour estimer un vecteur des poids d’anisotropie, comme le taux de convergence exponentielle suivant chaque direction, dans le cas d’un ensemble stochastique borné ou non borné. Ensuite pour déterminer un niveau global de la formule des grilles clairsemées anisotropes. L’approche

utilisée montre une réduction significative de l'effort global, puisque l'estimation des paramètres de la méthode exige seulement la résolution de problèmes d'interpolation, au lieu de la résolution d'un grand nombre d'EDP déterministes comme dans les approches classiques.



# Chapitre 4

## Méthode adaptative des grilles clairsemées via l'extrapolation de Richardson

Dans le chapitre précédent, nous avons présenté une méthode adaptative des grilles clairsemées pour l'approximation des statistiques de la solution de (1.1), en particulier sa moyenne  $\mathbb{E}[u]$  et sa variance  $\text{var}[u]$ . Elle consiste à déterminer les paramètres de la formule d'interpolation  $\mathcal{A}_\alpha(w, N)u$  en utilisant un indicateur d'erreur, qui satisfait une borne de l'erreur d'interpolation commise sur  $u$ . Cependant, dans certains cas, (comme le cas d'un champ de diffusion linéaire suivant  $y$ ), la constante de la borne supérieure peut décroître avec le cardinal des points de la grille, ainsi il se peut que le niveau  $w$  estimé pour la formule  $\mathcal{A}_\alpha(w, N)u$  soit légèrement surestimé (généralement à une unité de plus).

Dans ce chapitre nous démontrons une amélioration de l'erreur a priori de la méthode des grilles clairsemées, suivant le nombre total des points de collocation. Une approche adaptative de la méthode est aussi établie, et basée sur les deux points suivants :

- le vecteur des poids d'anisotropie  $\alpha$  est estimé en utilisant l'indicateur  $\lambda = k^{-1}f$  ;
- le niveau convenable  $w$  est déterminé en utilisant un indicateur d'erreur basé sur l'extrapolation de Richardson.

L'extrapolation de Richardson est une technique d'accélération de la convergence des suites numériques, fréquemment utilisée dans l'analyse numérique. Elle consiste à utiliser l'information de l'erreur a priori et une combinaison linéaire des résultats calculés auparavant pour prédire une amélioration des résultats.

On considère le problème (1.1) où le paramètre  $k$  et la donnée  $f$  peuvent être paramétrés par des variables aléatoires bornées ou non bornées. On s'intéresse à l'approximation  $\mathcal{A}_\alpha(w, N)u$  utilisant un niveau  $w$  approprié. On utilisera ainsi l'extrapolation de Richardson pour éviter un effort de calcul considérable à un niveau supplémentaire. Pour cela, on commence par établir des estimations a priori plus consistantes de l'erreur  $\|u - \mathcal{A}_\alpha(w, N)u\|$ , dans le cas d'un ensemble stochastique borné ou non borné.

Dans tout ce qui suit,  $g_n$  désigne le taux de convergence dans la direction  $n$  et  $\underline{g} = \min_{1 \leq n \leq N} g_n$ . Les poids  $\alpha_n$  sont choisis égaux aux taux  $g_n$  et  $\underline{\alpha} = \min_{1 \leq n \leq N} \alpha_n$ . Le nombre de points de la grille est noté par  $M$ .

## 4.1 Estimation de l'erreur

La méthode des grilles clairsemées isotrope décrite dans le chapitre 1 a pour but de réduire le problème de grande dimensionnalité. Effectivement, la méthode des grilles pleines (produit tensoriel complet (1.21)), converge approximativement comme  $C(g, N)e^{-gp}$ , où  $p$  est le degré polynomial (Théorème 1.2). Le nombre de points de colocation vérifie  $M = (1 + p)^N$  et  $\log(M) \leq pN$ , ainsi cette erreur peut être reformulée asymptotiquement de la manière suivante :

$$\|u - \mathcal{A}_i u\| \approx M^{-\frac{g}{N}}, \quad (4.1)$$

L'erreur asymptotique commise par la méthode isotrope des grilles clairsemées  $\mathcal{A}(w, N)$  est donnée dans le théorème 1.3, et peut être réécrite sous la forme suivante :

$$\|u - \mathcal{A}(w, N)u\| \approx M^{-\frac{g}{\log(N)}}. \quad (4.2)$$

Ainsi, une comparaison des deux estimations (4.1) et (4.2) montre que la méthode des grilles clairsemées isotrope permet de réduire le défaut de la grande dimension.

La méthode des grilles clairsemées anisotrope, qui est une généralisation du cas isotrope, permet de répartir les points de colocation entre les directions. Elle conduit, naturellement, à une réduction de l'effort et une amélioration des résultats en comparaison avec le cas isotrope. L'erreur de la méthode anisotrope [50] est donnée dans le théorème 1.4 :

$$\|u - \mathcal{A}_\alpha(w, N)u\| \approx M^{-\frac{g}{\log(2) + \sum_{n=1}^N \frac{g}{g_n}}}. \quad (4.3)$$

On suppose qu'on cherche à approcher une fonction  $v$  par la formule  $\mathcal{A}_\alpha(w, N)v$  et que

les coefficients  $g_n$ , associés à  $v$ , soient donnés par :

$$g_n = \frac{\underline{g}}{1 - \frac{1}{(n+1)^2}}, \quad \forall n = 1, \dots, N.$$

où  $\underline{g} > 0$ , alors on en déduit que :

$$\log(2) + \sum_{n=1}^N \frac{\underline{g}}{g_n} = \log(2) + N - \sum_{n=1}^N \frac{1}{(n+1)^2} \approx N,$$

ainsi, en remplaçant dans (4.3), on obtient pour cet exemple ( où  $\alpha = \mathbf{g}$ ) une approximation de l'erreur de la forme :

$$\|u - \mathcal{A}_\alpha(w, N)u\| \approx M^{-\frac{\underline{g}}{N}} \quad (4.4)$$

On peut remarquer que l'estimation de l'erreur (4.4) n'est pas compatible avec celle donnée dans (4.2), puisque dans cet exemple, la méthode est asymptotiquement isotrope. En d'autres termes l'estimation (4.4) pour cet exemple ne montre pas l'avantage de réduction du problème de grande dimension établie dans le cas de la méthode isotrope.

Dans la suite on va établir des estimations d'erreurs de la méthode des grilles clairsemées anisotrope, qui sont plus consistantes et compatibles avec celles données dans le cas de la méthode isotrope, pour le cas d'un ensemble stochastique borné ou non borné.

### 4.1.1 Cas d'un ensemble stochastique borné

Le théorème suivant donne une estimation de l'erreur commise par la formule  $\mathcal{A}_\alpha(w, N)$ . Cette estimation est compatible avec le cas isotrope et montre une amélioration par rapport à celle donnée dans le théorème 1.4. Elle montre que la méthode anisotrope converge algébriquement par rapport au nombre total des points de la grille lorsque le niveau total d'interpolation  $w$  satisfait :

$$w \leq \frac{\sum_{n=1}^N \alpha_n}{\underline{\alpha} \log(2)}. \quad (4.5)$$

#### **Théorème 4.1.**

*Considérons la fonction  $u \in L^2_\circ(\Gamma; H_0^1(D))$  satisfaisant l'hypothèse 1.2. On suppose que  $w$  vérifie (4.5), alors la formule d'interpolation anisotrope (1.33) satisfait les estimations suivantes :*

– pour les points de Clenshaw-Curtis,

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{L^2_\varrho(\Gamma; H^1_0(D))} \leq C_1(\alpha, N)M^{-\mu}, \quad (4.6)$$

$$\mu := \frac{\underline{\alpha}(\log(2)e - 1/2)}{\log(2) + \log(1 + \sum_{n=1}^N \underline{\alpha}/\alpha_n)} \quad (4.7)$$

– pour les points de Gauss :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{L^2_\varrho(\Gamma; H^1_0(D))} \leq C_2(\alpha, N)M^{-\tilde{\mu}} \quad (4.8)$$

$$\tilde{\mu} := \frac{\underline{\alpha} \log(2)e}{2 \log(2) + \log(1 + \sum_{n=1}^N \underline{\alpha}/\alpha_n)} \quad (4.9)$$

Les deux constantes  $C_1$  et  $C_2$  dépendent seulement de  $N$  et  $\alpha$  et non du cardinal de la grille  $M = \#\mathcal{H}_\alpha(w, N)$ .

Avant de donner la preuve de ce théorème, on commence par établir un résultat qui sera utile dans cette preuve. Il concerne la relation entre le cardinal  $M$  de la grille des points de collocation  $\mathcal{H}_\alpha(w, N)$  et le niveau  $w$ .

**Lemme 4.1.**

*En utilisant la formule d'interpolation des grilles clairsemées (1.33), et selon les points d'interpolation utilisés, Clenshaw-Curtis ou Gauss, le nombre total des points de collocation de la grille au niveau  $w$  satisfait la borne suivante :*

– pour les points de Clenshaw-Curtis ( avec  $m(i)=2^{i-1} + 1$ ) :

$$2^{w-1} \leq M \leq 2^w \left( 1 + \sum_{n=1}^N \frac{\alpha}{\alpha_n} \right)^w, \quad (4.10)$$

– pour les points de Gauss ( avec  $m(i)=2^{i-1} + 1$ ) :

$$2^w \leq M \leq 2^{2w} \left( 1 + \sum_{n=1}^N \frac{\alpha}{\alpha_n} \right)^w. \quad (4.11)$$

**Preuve :**

Nous allons d'abord établir une borne du nombre total de l'ensemble des indices  $\#X_\alpha(w, N)$ , donné dans (1.34). Nous utilisons une preuve par récurrence pour montrer la borne suivante :

$$\#X_\alpha(w, N) \leq \left( 1 + \sum_{n=1}^N \frac{\alpha}{\alpha_n} \right)^w. \quad (4.12)$$

Remarquons d'abord que le cardinal  $\#X_\alpha(w, N)$  est invariant par n'importe quelle permutation des coefficients d'anisotropie  $\{\alpha_n\}_{n=1}^N$  (les points seront différents, mais

leur nombre reste le même par toute permutation donnée des poids  $(\alpha_n)_{n=1}^N$ , ainsi, on peut supposer que ces poids  $\{\alpha_n\}_{n=1}^N$ , forment une suite décroissante de nombres positifs, et on peut alors poser  $\alpha_N = \min_{1 \leq n \leq N} \alpha_n = \underline{\alpha}$ .

Pour  $N = 1$ , nous avons  $\#X_\alpha(w, 1) = w + 1 \leq 2^w$ , admettons alors que la borne est satisfaite par  $\#X_\alpha(w, N)$  jusqu'à l'ordre  $N - 1$ .

Nous avons pour tout  $\mathbf{i} \in X_\alpha(w, N)$  :

$$\sum_{n=1}^N (i_n - 1)\alpha_n \leq \underline{\alpha}w,$$

par conséquent,

$$\sum_{n=1}^{N-1} (i_n - 1)\alpha_n \leq \left( w - (i_N - 1)\frac{\alpha_N}{\underline{\alpha}} \right) \underline{\alpha}.$$

On en déduit la relation suivante :

$$\begin{aligned} \#X_\alpha(w, N) &= \sum_{j_N=1}^{\lfloor w \frac{\alpha}{\alpha_N} + 1 \rfloor} \#X_\alpha(w - (j_N - 1)\frac{\alpha_N}{\underline{\alpha}}, N - 1) \\ &= \sum_{j=0}^w \#X_\alpha(w - j, N - 1), \quad (\text{puisque } \alpha_N = \underline{\alpha}). \end{aligned}$$

Ainsi, par l'hypothèse de récurrence (4.12), nous avons :

$$\#X_\alpha(w, N) \leq \sum_{j=0}^w \left( 1 + \sum_{n=1}^{N-1} \frac{\alpha}{\alpha_n} \right)^{w-j} = \sum_{j=0}^w \left( \sum_{n=1}^N \frac{\alpha}{\alpha_n} \right)^j. \quad (4.13)$$

En remarquant que  $\left( 1 + \sum_{n=1}^N \frac{\alpha}{\alpha_n} \right)^w = \sum_{j=0}^w \binom{w}{j} \left( \sum_{n=1}^N \frac{\alpha}{\alpha_n} \right)^j$ , et le fait que  $\binom{w}{j} \geq 1$  pour tout  $j = 1, \dots, w$ , on déduit que  $\#X_\alpha(w, N)$  satisfait la borne (4.12).

Maintenant, à partir de la formule (1.33), le cardinal  $M$  de la grille  $\mathcal{H}_\alpha(w, N)$  peut être calculé de la façon suivante, dans le cas des points de Clenshaw-Curtis :

$$M = \sum_{\mathbf{i} \in X_\alpha(w, N)} \prod_{n=1}^N \hat{m}(i_n),$$

du fait que les points de Clenshaw-Curtis sont imbriqués lorsqu'on double les points dans chaque niveau qui suit,  $\hat{m}(i_n)$  est donné par :

$$\hat{m}(i) = \begin{cases} 1 & \text{si } i = 1 \\ 2 & \text{si } i = 2 \\ 2^{i-2} & \text{si } i > 2. \end{cases} \quad (4.14)$$

Ainsi, chaque  $\hat{m}(i_n)$  vérifie  $2^{i_n-2} \leq \hat{m}(i_n) \leq 2^{i_n-1}$ , en remarquant que,

$$|\mathbf{i} - 1| = \sum_{n=1}^N (i_n - 1) \leq \frac{1}{\underline{\alpha}} \sum_{n=1}^N \alpha_n (i_n - 1) \leq w, \text{ on obtient :}$$

$$M \leq \sum_{\mathbf{i} \in X_\alpha(w, N)} 2^{|\mathbf{i}-1|} \leq 2^w \#X_\alpha(w, N).$$

D'autre part, si on suppose que  $\underline{\alpha} = \alpha_j$ , le multi-indice  $\hat{\mathbf{i}}$  dont les composantes sont égaux à  $\hat{i}_n = 1$  sauf  $\hat{i}_j = w + 1$  est admissible et  $\hat{m}(\hat{i}_j) = 2^{w-1}$  points. Par conséquent :

$$2^{w-1} \leq M \leq \sum_{\mathbf{i} \in X_\alpha(w, N)} 2^{|\mathbf{i}-1|} \leq 2^w \#X_\alpha(w, N).$$

En utilisant (4.12) on conclut l'estimation dans le cas des points de Clenshaw-Curtis.

Dans le cas des points de Gauss, le nombre total  $M$  vérifie :

$$M = \sum_{\mathbf{i} \in Y_\alpha(w, N)} \prod_{n=1}^N \tilde{m}(i_n), \text{ avec } \tilde{m}(i) = \begin{cases} 1 & \text{si } i = 1 \\ 2^{i-1} + 1 & \text{si } i > 1 \end{cases} \quad (4.15)$$

Si on suppose que  $\underline{\alpha} = \alpha_j$ , le multi-indice  $\tilde{\mathbf{i}}$  dont les composantes sont égaux à  $\tilde{i}_n = 1$  sauf  $\tilde{i}_j = w + 1$  est admissible et  $\tilde{m}(\tilde{i}_j) = 2^w + 1$ , par conséquent  $M \geq 2^w$ .

D'autre part, de l'estimation suivante,

$$\prod_{n=1}^N \tilde{m}(i_n) \leq \prod_{n=1}^N (2^{i_n-1} + 1) \leq \prod_{n=1}^N (2^{2(i_n-1)}),$$

On en déduit l'estimation :

$$2^w \leq M \leq \sum_{\mathbf{i} \in Y_\alpha(w, N)} 2^{2w} \leq 2^{2w} \#Y_\alpha(w, N).$$

■

**Preuve du théorème 4.1 :**

1) CAS DES POINTS DE CLENSHAW-CURTIS (C-C) :

Supposons que  $w$  vérifie (4.5). L'erreur commise dans ce cas (voir [50]) en fonction du niveau  $w$  est donnée par :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{L^2_\alpha(\Gamma; H^1_0(D))} \leq C(\alpha, N) e^{-\tilde{\lambda}(w, N)} \quad (4.16)$$

où la constante  $C(\alpha, N)$  ne dépend pas de  $w$ , et la fonction  $\tilde{\lambda}$  est donnée dans le cas des points de C-C par :

$$\tilde{\lambda}(w, N) := w\underline{\alpha} \left( \log(2)e - \frac{1}{2} \right). \quad (4.17)$$

D'autre part, d'après l'estimation (4.10), nous avons :

$$w \geq \frac{\log(M)}{\log(2) + \log(1 + \sum_{n=1}^N \frac{\underline{\alpha}}{\alpha_n})}.$$

On en déduit que :

$$\tilde{\lambda} \geq \frac{\underline{\alpha}(\log(2)e - 1/2) \log(M)}{\log(2) + \log(1 + \sum_{n=1}^N \frac{\underline{\alpha}}{\alpha_n})}.$$

En substituant ce dernier résultat dans (4.16) on obtient la première estimation du théorème.

## II) CAS DES POINTS DE GAUSS :

Dans ce cas la fonction  $\tilde{\lambda}$  dans (4.16) est définie par :

$$\tilde{\lambda}(w, N) := w\underline{\alpha} \log(2)e, \quad (4.18)$$

ainsi, d'après l'estimation (4.11), nous avons :

$$w \geq \frac{\log(M)}{2 \log(2) + \log(1 + \sum_{n=1}^N \frac{\underline{\alpha}}{\alpha_n})}.$$

On en déduit que :

$$\tilde{\lambda} \geq \frac{\log M \underline{\alpha} \log(2)e}{2 \log(2) + \log(1 + \sum_{n=1}^N \frac{\underline{\alpha}}{\alpha_n})},$$

et en substituant cela dans (4.16), on obtient la deuxième estimation du théorème. ■

### Remarque 4.1.

- L'estimation (4.11) reste vraie, en particulier celle de la majoration de  $M$ , si on utilise un ordre de croissance de points de collocation linéaire ( $m(i) = i$  ou  $m(i) = 2i - 1 \dots$ ), puisque le nombre total des indices  $\#X_\alpha(w, N)$  est indépendant du choix de cet ordre.
- Le résultat de convergence du théorème 4.1 est beaucoup plus consistant que celui donné dans le théorème 1.4 (cf.[50]), il est compatible avec celui de la méthode des grilles clairsemées isotrope, quand les poids  $\{\alpha_n\}_n^N$  ne diffèrent pas beaucoup.

- Même dans le cas d'un problème fortement anisotrope où  $\sum_{n=1}^{\infty} \frac{g}{g_n} < \infty$ , l'estimation (4.6) et (4.8) est une amélioration de celle du théorème 1.4, puisque nous avons pour tout  $t > 0$  :

$$\log(1+t) \leq t.$$

- On pourrait redémontrer de la même façon l'estimation qui montre une convergence sous-exponentielle. Cependant ce cas de convergence n'est vérifié que pour des grandes valeurs de  $w$ , et vu le problème des grande dimensionnalité, on ne s'intéresse dans la pratique qu'à l'estimation algébrique.

### 4.1.2 Cas d'un ensemble stochastique non borné

Dans cette partie, on suppose que la fonction  $\varrho$ , densité conjointe de vecteur aléatoire  $\mathbf{Y}$ , décroît exactement comme le noyau gaussien  $e^{-\sum_{n=1}^N \delta_n y_n^2}$ , où  $\delta_n > 0$ . On suppose aussi que la fonction  $u$  est analytique et que de plus, pour chaque point  $(y_n, y_n^*) \in \Gamma = \mathbb{R}^N$ , la fonction  $v(y_n) = u(y_n, y_n^*)$  comme fonction de  $y_n$  est dans l'espace à poids  $\mathcal{C}_{\sigma_n}^0(\mathbb{R}; H_0^1(D))$  défini par :

$$\mathcal{C}_{\sigma_n}^0(\mathbb{R}; H_0^1(D)) := \left\{ \phi(t, x) : \mathbb{R} \rightarrow H_0^1(D) \text{ continue en } t, \max_{t \in \mathbb{R}} \|e^{-\beta_n t} \phi(t, \cdot)\|_{H_0^1(D)} < \infty \right\},$$

où on a défini le poids  $\sigma_n(t) = e^{-\beta_n |t|}$  pour  $\beta_n > 0$ . On suppose aussi que  $v(y_n)$  satisfait l'hypothèse du lemme suivant (voir [6]) :

#### Lemme 4.2.

Considérons une fonction  $v \in \mathcal{C}_{\sigma_n}^0(\mathbb{R}; H_0^1(D))$ , et supposons qu'elle admet une extension analytique dans la région du plan complexe  $\Sigma = \{z \in \mathbb{C}, \text{dist}(z, \mathbb{R}) \leq \tau_n\}$ , où  $\tau_n > 0$ ,

$$\forall z = y + i\theta \in \Sigma \quad \sigma_n(y) \|v(z)\|_{H_0^1(D)} \leq C_v(\tau_n).$$

Alors, il existe une fonction  $\Theta(m_n)$  telle que

$$\|v - \mathcal{U}^{i_n} v\|_{L_{\rho_n}^2(\mathbb{R}; H_0^1(D))} \leq \Theta(m(i_n) - 1) e^{-g_n \sqrt{2(m(i_n) - 1)}}, \quad (4.19)$$

où on a posé  $g_n = \frac{\tau_n \delta_n}{\sqrt{2}}$  et  $\Theta(m(i_n) - 1) \leq E \sqrt{m(i_n) - 1}$ ,  $m(i_n)$  est le nombre de points de Gauss utilisés par l'opérateur  $\mathcal{U}^{i_n}$  et la constante  $E$  ne dépendant pas de  $m(i_n)$ .



Le théorème 4.3 concerne l'estimation de l'erreur commise par la formule  $\mathcal{A}_\alpha(w, N)$ , il montre deux types de convergence, une algébrique et l'autre sous-exponentielle lorsque la formule utilise des valeurs assez grandes de  $w$ .

Nous allons suivre, pour la preuve de ce théorème, la même démarche que dans [50], pour le cas d'un ensemble  $\Gamma$  borné donné avec quelques adaptations techniques pour prendre en compte la racine dans l'exposant de l'estimation (4.19), et nous utiliserons le résultat du lemme 4.1.

On note dans la suite pour des raisons de simplification  $\mathcal{H} := L^2_\varrho(\Gamma; H_0^1(D))$ , et par  $I_d$  l'opérateur identité des fonctions qui dépendent de  $d$  variables. Pour tout  $n = 1, \dots, N$ , on note par  $I_1^{(n)}$ , l'opérateur unidimensionnel identité  $I_1^{(n)} : L^2_{\rho_n}(\mathbb{R}; H_0^1(D)) \rightarrow L^2_{\rho_n}(\mathbb{R}; H_0^1(D))$ . On suppose aussi qu'on utilise un ordre de croissance des points de collocation de la forme  $m(i_n) = 2^{i_n-1} + 1$ .

En utilisant l'estimation (4.19), nous avons :

$$\begin{aligned} \|\Delta^{i_n} u\|_{L^2_{\rho_n}(\mathbb{R}; H_0^1(D))} &= \|\mathcal{U}^{i_n} u - \mathcal{U}^{i_n-1} u\|_{L^2_{\rho_n}(\mathbb{R}; H_0^1(D))} \\ &\leq \|\mathcal{U}^{i_n} - u\|_{L^2_{\rho_n}(\mathbb{R}; H_0^1(D))} + \|\mathcal{U}^{i_n-1} - u\|_{L^2_{\rho_n}(\mathbb{R}; H_0^1(D))} \\ &\leq E \left( 2^{\frac{i_n-1}{2}} e^{-g_n 2^{\frac{i_n}{2}}} + 2^{\frac{i_n-2}{2}} e^{-g_n 2^{\frac{(i_n-1)}{2}}} \right) \\ &\leq \tilde{E} 2^{\frac{i_n-1}{2}} e^{-g_n \frac{(i_n-1)}{2}}, \end{aligned} \quad (4.20)$$

où la constante  $\tilde{E} = E(\frac{1}{\sqrt{2}} + e^{-g_n(\sqrt{2}-1)})$ .

Définissons les quantités suivantes :

$$\tilde{X}_\alpha(w, d) := \left\{ \mathbf{i} \in \mathbb{N}^d, \mathbf{i} > \mathbf{1} : \underline{\alpha} w \leq \sum_{n=1}^d (i_n - 1) \alpha_n \leq \alpha_d + \underline{\alpha} w \right\}$$

$$R(w, d) = \sum_{\mathbf{i} \in \tilde{X}_\alpha(w, d)} \tilde{E}^d \left( \prod_{n=1}^d 2^{\frac{(i_n-1)}{2}} \right) e^{-h(w, d)} \quad (4.21)$$

$$h(i, d) = \sum_{n=1}^d g_n 2^{\frac{(i_n-1)}{2}}. \quad (4.22)$$

Le lemme qui suit concerne une estimation récursive de l'erreur  $\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}}$ .

**Lemme 4.3.**

*Soit  $u \in \mathcal{H}$  satisfaisant l'hypothèse de lemme 4.2, alors la formule des grilles clairsemées (1.33) basée sur les points de Gauss vérifie l'estimation suivante :*

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}} \leq \sum_{n=1}^N R(w, d), \quad (4.23)$$

avec  $R(w, d)$  est défini par (4.21).

**Preuve du lemme 4.3 :**

En développant par rapport au  $N^{\text{ième}}$  opérateur  $\Delta^{i_N}$ , la formule anisotrope peut être donnée par la représentation équivalente suivante :

$$\begin{aligned} \mathcal{A}_\alpha(w, N) &= \sum_{\mathbf{i} \in X_\alpha(w, N)} \bigotimes_{n=1}^N \Delta^{i_n} = \sum_{\mathbf{i} \in X_\alpha(w, N-1)} \bigotimes_{n=1}^{N-1} \Delta^{i_n} \otimes \sum_{\mathbf{j}=1}^{\mathbf{s}} \Delta^j \\ &= \sum_{\mathbf{i} \in X_\alpha(w, N-1)} \bigotimes_{n=1}^{N-1} \Delta^{i_n} \otimes \mathcal{U}^{\mathbf{s}}. \end{aligned}$$

où nous avons posé  $\mathbf{s} = \lfloor 1 + w \frac{\alpha}{\alpha_N} - \sum_{n=1}^{N-1} (i_n - 1) \frac{\alpha_n}{\alpha_N} \rfloor$ , le symbole  $\lfloor \cdot \rfloor$  désigne la partie entière d'un nombre réel.

On en déduit par une simple récurrence que :

$$\begin{aligned} I_N - \mathcal{A}_\alpha(w, N) &= I_N - \sum_{\mathbf{i} \in X_\alpha(w, N-1)} \bigotimes_{n=1}^{N-1} \Delta^{i_n} \otimes (\mathcal{U}^{\mathbf{s}} - I_1^N) - \sum_{\mathbf{i} \in X_\alpha(w, N-1)} \bigotimes_{n=1}^{N-1} \Delta^{i_n} \otimes I_1^{(N)} \\ &= \sum_{\mathbf{i} \in X_\alpha(w, N-1)} \bigotimes_{n=1}^{N-1} \Delta^{i_n} \otimes (I_1^N - \mathcal{U}^{\mathbf{s}}) + (I_{N-1} - \mathcal{A}_\alpha(w, N-1)) \otimes I_1^{(N)} \\ &= \sum_{d=2}^N \left( \tilde{R}(w, d) \bigotimes_{n=d+1}^N I_1^{(n)} \right) + \left( I_1^{(1)} - \mathcal{A}_\alpha(w, 1) \right) \bigotimes_{n=2}^N I_1^{(n)}, \quad (4.24) \end{aligned}$$

où, nous avons défini  $\tilde{R}(w, d) = \sum_{\mathbf{i} \in X_\alpha(w, d-1)} \bigotimes_{n=1}^{d-1} \Delta^{i_n} \otimes \left( I_1^{(d)} - \mathcal{U}^{\hat{i}_d} \right)$  et pour tout  $(i_1, \dots, i_{d-1}) \in X_\alpha(w, d-1)$ , nous avons posé  $\hat{i}_d = \lfloor 2 + w \frac{\alpha}{\alpha_d} - \sum_{n=1}^{d-1} (i_n - 1) \frac{\alpha_n}{\alpha_d} \rfloor$ . Remarquons que dans cette définition, l'indice  $d$ -dimensionnel  $(i_1, \dots, i_{d-1}, \hat{i}_d)$  est dans l'ensemble des indices  $\tilde{X}_\alpha(w, d)$ .

En utilisant les deux bornes (4.19) et (4.20), le terme  $\tilde{R}(w, d)$  peut être majoré de la façon suivante :

$$\begin{aligned} \|\tilde{R}(w, d)(u)\|_{\mathcal{H}} &\leq \sum_{\mathbf{i} \in X_\alpha(w, d-1)} \prod_{n=1}^{d-1} \|\Delta^{i_n}(u)\|_{\mathcal{H}} \|(I_1^{(d)} - \mathcal{U}^{\hat{i}_d})(u)\|_{\mathcal{H}} \\ &\leq \sum_{\mathbf{i} \in X_\alpha(w, d-1)} \tilde{E}^{d-1} 2^{\sum_{n=1}^{d-1} \frac{(i_n-1)}{2}} e^{-\sum_{n=1}^{d-1} g_n 2^{\frac{i_n-1}{2}}} E 2^{\frac{\hat{i}_d-1}{2}} e^{-g_d 2^{\frac{\hat{i}_d-1}{2}}} \\ &\leq \sum_{\mathbf{i} \in \tilde{X}_\alpha(w, d)} \tilde{E}^d 2^{\sum_{n=1}^d \frac{(i_n-1)}{2}} e^{-h(w, d)} = R(w, d). \quad (4.25) \end{aligned}$$

Ainsi, en utilisant la relation (4.24) et la borne (4.25), l'erreur d'interpolation de la méthode des grilles clairsemées anisotrope peut être bornée par :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}} \leq \sum_{d=2}^N R(w, d) + \|I_1(1) - \mathcal{A}_\alpha(w, 1)\|_2.$$

Le second terme peut être majoré aussi de la même façon par  $R(w, 1)$ , puisque l'ensemble  $\tilde{X}_\alpha(w, 1)$  contient seulement l'indice  $i_1 = \lfloor 2 + \frac{\alpha w}{\alpha_1} \rfloor$  :

$$\begin{aligned} \|(I_1^{(1)} - \mathcal{A}_\alpha(w, 1))u\|_{\mathcal{H}} &= \|(I_1^{(1)} - \mathcal{U}^{\lfloor 1 + \frac{\alpha w}{\alpha_1} \rfloor})u\|_{\mathcal{H}} \\ &\leq E 2^{\frac{\lfloor 1 + \frac{\alpha w}{\alpha_1} \rfloor}{2}} e^{-g_1 \frac{\lfloor 1 + \frac{\alpha w}{\alpha_1} \rfloor}{2}} \\ &\leq \sum_{i_1 \in \tilde{X}_\alpha(w, 1)} \tilde{E} 2^{\frac{i_1 - 1}{2}} e^{-g_1 \frac{i_1 - 1}{2}} := R(w, 1). \end{aligned}$$

Ainsi, l'estimation récursive du lemme est achevée. ■

Le lemme suivant concerne une borne supérieure du terme  $R(w, d)$ .

**Lemme 4.4.**

Soit  $u \in \mathcal{H}$  satisfaisant l'hypothèse de lemme 4.2, et la formule des grilles clairsemées (1.33) utilisant des points de Gauss avec le choix  $\alpha_n = g_n$  où les  $g_n$  sont donnés dans le lemme (4.2). Alors pour tout  $d = 1, \dots, N$  le terme  $R(w, d)$  défini par (4.21) satisfait l'estimation suivante :

$$R(w, d) \leq C(d, \mathbf{g}) \exp \left\{ -\mathcal{S}(d) 2^{w \frac{g}{2\mathcal{S}(d)}} + \frac{w \log(2)}{2} \right\}, \quad (4.26)$$

où nous avons posé  $\mathcal{S}(d) := \sum_{n=1}^d g_n$ .

**Preuve du lemme 4.4**

Considérons la fonction  $h(y, d)$  définie par (4.22). Son développement de Taylor à l'ordre 2 au voisinage du point  $\mathbf{i}^* = (1 + \frac{wg}{\mathcal{S}(d)}, \dots, 1 + \frac{wg}{\mathcal{S}(d)})$  donne pour tout  $\mathbf{z} \in \mathbb{R}^d$  :

$$h(\mathbf{z}, d) = h(\mathbf{i}^*, d) + \nabla h(\mathbf{i}^*, d) \cdot (\mathbf{z} - \mathbf{i}^*) + \frac{1}{2} (\mathbf{z} - \mathbf{i}^*)' \cdot \nabla^2 h(\xi, d) \cdot (\mathbf{z} - \mathbf{i}^*) \quad (4.27)$$

où  $\xi = t\mathbf{z} + (1-t)\mathbf{i}^*$  avec  $t \in [0, 1]$ .

Le terme linéaire est positif pour tout  $\mathbf{z} = \mathbf{i} \in \tilde{X}_\alpha(w, d)$ , en effet :

$$\begin{aligned} \nabla h(\mathbf{i}^*, d) \cdot (\mathbf{i} - \mathbf{i}^*) &= \frac{\log(2)}{2} 2^{w \frac{g}{2\mathcal{S}(d)}} \sum_{n=1}^d g_n \left( i_n - 1 - w \frac{g}{\mathcal{S}(d)} \right) \\ &= \frac{\log(2)}{2} 2^{w \frac{g}{2\mathcal{S}(d)}} \left( \sum_{n=1}^d (i_n - 1) g_n - wg \right) \geq 0. \end{aligned}$$

et le terme quadratique peut être minoré par :

$$\begin{aligned} (\mathbf{i} - \mathbf{i}^*)' \cdot \nabla^2 h(\xi, d) \cdot (\mathbf{i} - \mathbf{i}^*) &= \frac{\log(2)^2}{4} \sum_{n=1}^d g_n 2^{\frac{\xi_n - 1}{2}} (i_n - i_n^*)^2 \\ &\geq \frac{\log(2)^2}{4\sqrt{2}} \sum_{n=1}^d g_n (i_n - i_n^*)^2. \end{aligned}$$

On en déduit que pour tout  $\mathbf{i} \in \tilde{X}_\alpha(w, d)$  :

$$h(\mathbf{i}, d) \geq h(\mathbf{i}^*, d) + \sum_{n=1}^d \frac{(i_n - i_n^*)^2}{2\sigma_n^2}, \quad (4.28)$$

où on a posé  $\sigma_n^2 = \frac{4\sqrt{2}}{g_n \log(2)^2}$ .

La combinaison de (4.21) avec (4.28) donne :

$$\begin{aligned} R(w, d) &\leq \tilde{E}^d e^{-h(\mathbf{i}^*, d)} \sum_{\mathbf{i} \in \tilde{X}_\alpha(w, d)} \left( \prod_{n=1}^d 2^{\frac{i_n - 1}{2}} \exp \left\{ -\frac{(i_n - (1 + \frac{wg}{\mathcal{S}(d)}))^2}{2\sigma_n^2} \right\} \right) \\ &\leq \tilde{E}^d e^{-h(\mathbf{i}^*, d)} \sum_{\mathbf{i} \in \tilde{X}_\alpha(w, d)} \prod_{n=1}^d 2^{\frac{wg}{2\mathcal{S}(d)}} \exp \left\{ -\frac{(i_n - (1 + \frac{wg}{\mathcal{S}(d)} + \frac{\log(2)}{2}\sigma_n^2))^2}{2\sigma_n^2} + \frac{\log(2)^2}{8}\sigma_n^2 \right\} \\ &\leq \tilde{E}^d e^{-h(\mathbf{i}^*, d)} 2^{\frac{wg}{2}} \prod_{n=1}^d \exp \left( \frac{\log(2)^2}{8}\sigma_n^2 \right) T_n, \end{aligned} \quad (4.29)$$

où nous avons posé  $T_n := \sum_{i_n \geq 1} \exp \left\{ -\frac{(i_n - (1 + \frac{wg}{\mathcal{S}(d)} + \frac{\log(2)}{2}\sigma_n^2))^2}{2\sigma_n^2} \right\}$ . Le terme  $T_n$  peut être facilement majoré indépendamment de  $w$  comme suit :

$$T_n \leq 1 + \sqrt{2\pi}\sigma_n.$$

En combinant cette dernière borne avec (4.29) et sachant que  $h(\mathbf{i}^*, d) = \sum_{n=1}^d g_n 2^{\frac{g}{2\mathcal{S}(d)}}$  nous obtenons :

$$R(w, d) \leq C(d, \mathbf{g}) \exp \left\{ -\mathcal{S}(d) 2^{w \frac{g}{2\mathcal{S}(d)}} + \frac{w \log(2)}{2} \right\},$$

où on a posé :

$$C(d, \mathbf{g}) = \tilde{E}^d \prod_{n=1}^d \left( \left( 1 + \frac{4\sqrt{2\pi}}{\sqrt{g_n} \log(2)} \right) \exp \left( \frac{\sqrt{2}}{2g_n} \right) \right).$$

■

**Théorème 4.2.**

Considérons la fonction  $u \in \mathcal{H}$  satisfaisant l'hypothèse du lemme 4.2, où les coefficients de convergence  $\alpha_n = g_n$  sont donnés dans le lemme 4.2, alors la formule d'interpolation (1.33) avec des points de Gauss, satisfait l'estimation suivante :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}} \leq \tilde{C}(N, \mathbf{g})e^{\frac{w \log(2)}{2} - \psi(w, N)}, \quad (4.30)$$

où

$$\psi(w, N) := \begin{cases} w \frac{\log(2)}{2} e & \text{si } 0 \leq w \leq \frac{2\mathcal{S}(N)}{g \log(2)} \\ \mathcal{S}(N) 2^{w \frac{g}{2\mathcal{S}(N)}} & \text{si } w > \frac{2\mathcal{S}(N)}{g \log(2)} \end{cases} \quad (4.31)$$

et la fonction  $\tilde{C}$  ne dépend pas du cardinal  $M$ .

**Preuve :**

En substituant l'estimation du lemme 4.3 dans celle récursive du Lemme 4.4, nous obtenons la borne suivante pour l'erreur d'interpolation :

$$\|u - \mathcal{A}_\alpha(w, N)\|_{\mathcal{H}} \leq \sum_{d=1}^N C(d, \mathbf{g}) \exp \left\{ -\mathcal{S}(d) 2^{w \frac{g}{2\mathcal{S}(d)}} + \frac{w \log(2)}{2} \right\}.$$

La fonction  $\psi(w, N)$  définie par (4.31) est donnée comme le minimum de la fonction  $\tilde{\psi}(w, d) = \mathcal{S}(d) 2^{w \frac{g}{2\mathcal{S}(d)}}$  pour tout  $1 \leq d \leq N$ . En effet, définissons pour tout  $t \in [1, \mathcal{S}(N)]$ , la fonction  $G$  par :

$$G(t) := t 2^{w \frac{g}{2t}}.$$

Nous avons  $\frac{dG}{dt} = \left(1 - \frac{\log(2)wg}{2t}\right) 2^{w \frac{g}{2t}}$ , ainsi  $G$  décroît sur  $[1, t^*]$  et croît sur  $[t^*, \mathcal{S}(N)]$ , où  $t^*$  est le minimum de la fonction  $G$ ,  $t^* = \frac{\log(2)}{2} gw$ .

Par conséquent :

$$G(t) \geq G(t^*) = \frac{\log(2)}{2} w \underline{g} e.$$

Deux situations seront présentées, ou bien le niveau  $w$  est choisi de sorte que  $0 < t^* \leq \mathcal{S}(N)$ , i.e  $0 \leq w \leq \frac{2\mathcal{S}(N)}{g \log(2)}$ , alors dans ce cas nous avons :

$$\psi(w, N) = \frac{\log(2)}{2} w \underline{g} e,$$

ou bien pour des niveaux  $w$  suffisamment grands  $w \geq \frac{\mathcal{S}(N)}{\underline{g} \log(2)}$ , le minimum  $t^*$  est en dehors de  $[1, \mathcal{S}(N)]$ , et puisque la fonction  $G$  est décroissante sur l'intervalle  $[1, \mathcal{S}(N)]$ , alors on en déduit dans ce cas que :

$$\psi(w, N) = \min_{1 \leq d \leq N} \tilde{\psi}(w, d) = \mathcal{S}(N) 2^{w \frac{\underline{g}}{2\mathcal{S}(N)}}.$$

Le résultat de l'estimation du théorème est ainsi vérifiée en prenant :

$$\tilde{C}(\mathbf{g}, N) = \sum_{d=1}^N C(\mathbf{g}, d). \quad (4.32)$$

■

Le théorème suivant concerne le résultat de convergence de la méthode des grilles clairsemées pour des variables aléatoires non bornées. Il montre une estimation de l'erreur commise en fonction du nombre total des points de collocation  $M = \#\mathcal{H}_\alpha(w, N)$ , les points étant ceux de Gauss.

### Théorème 4.3.

Considérons la fonction  $u \in L^2_\rho(\Gamma; H^1_0(D))$  satisfaisant l'hypothèse du Lemme 4.2, où les coefficients de convergence  $\alpha_n = g_n$  sont donnés dans le lemme 4.2. Alors, sous l'hypothèse que  $\underline{g} \geq e^{-1}$ , la formule d'interpolation anisotrope (1.33) avec des points de Gauss, satisfait l'estimation suivante :

– convergence algébrique ( $0 \leq w \leq \frac{2\mathcal{S}(N)}{\underline{g} \log(2)}$ ) :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}} \leq \tilde{C}(\mathbf{g}, N) M^{-\mu}, \quad (4.33)$$

$$\mu := \frac{(\underline{g}e - 1) \log(2)}{4 \log(2) + 2 \log(1 + \sum_{n=1}^N \underline{g}/g_n)}. \quad (4.34)$$

– convergence sous-exponentielle ( $w > \frac{2\mathcal{S}(N)}{\underline{g} \log(2)}$ ) :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}} \leq \tilde{C}(\mathbf{g}, N) \sqrt{M} \exp\left(-\mathcal{S}(N) M^{\frac{\tilde{\mu}}{2}}\right), \quad (4.35)$$

$$\tilde{\mu} := \frac{\log(2) \underline{g}}{\mathcal{S}(N) \left(2 \log(2) + \log(1 + \sum_{n=1}^N \underline{g}/g_n)\right)}. \quad (4.36)$$

La constante  $\tilde{C}$  ne dépend pas du cardinal  $M$  et est donnée par (4.32).

**Preuve :**

La preuve est similaire à celle du cas de l'ensemble stochastique borné, elle consiste à combiner l'estimation du théorème 4.2 avec l'estimation donnant le nombre des points de collocation de Gauss en fonction du niveau  $w$  déduite du lemme 4.1 :

$$\begin{aligned} w &\geq \frac{\log(M)}{2 \log(2) + \log(1 + \sum_{n=1}^N \underline{g}/g_n)} \\ 2^{\frac{w}{2}} &\leq \sqrt{M}. \end{aligned}$$

■

**Remarque 4.2.**

Lorsque le champ aléatoire est exponentiel de la forme  $k = e^G$ , il est une fonction entière par rapport à la variable  $y$ . On conjecture (d'après la comparaison qu'on avait établi entre l'erreur sur  $u$  et l'erreur sur l'indicateur  $k^{-1}f$ ), que la fonction  $u$  est aussi entière suivant  $y$ . L'erreur unidimensionnelle de l'interpolation  $\mathcal{U}^{i_n}$  est de la forme  $O(e^{-g_n m(i_n)})$ , avec  $g_n > 0$ , et ceci même si  $\Gamma$  est non borné comme le montre la proposition 3.2, ainsi on en déduit que l'erreur de la méthode des grilles clairsemées est aussi donnée par :

$$\|u - \mathcal{A}_\alpha(w, N)u\|_{\mathcal{H}} \leq C_1(\mathbf{g}, N)M^{-\mu}, \quad (4.37)$$

$$\mu := \frac{\underline{g}}{2 \log(2) + \log(1 + \sum_{n=1}^N \underline{g}/g_n)}. \quad (4.38)$$

**Remarque 4.3.**

- On observe que le taux de convergence sub-exponentielle est satisfait lorsque  $w > \frac{2S(N)}{\underline{g} \log(2)}$  (des valeurs assez grandes de  $w$ , deux fois plus grandes que dans le cas où  $\Gamma$  est borné). Cependant, cette estimation a peu d'intérêt dans la pratique. Avoir des niveaux  $w$  assez grands est rarement envisageable, à cause du problème de grande dimensionnalité.
- La condition  $\underline{g}e - 1 > 0$  dans le théorème 4.3 est nécessaire pour établir une estimation de l'erreur pour tout  $w \geq 0$ , dans le cas contraire, la convergence est assurée par les grandes valeurs de  $w$ .

## 4.2 Approche des grilles clairsemées adaptative

La méthode des grilles clairsemées anisotrope permet de répartir les points de collocation entre les directions, en prenant en compte leur variation. En comparaison avec

la méthode isotrope, elle permet de réduire significativement l'effort d'approximation. Cependant quand la dimension stochastique  $N$  est grande, le coût du calcul augmente de façon exponentielle avec la dimension, puisque le nombre total des points de collocation augmente avec le niveau  $w$ . Il est donc important de chercher à bien choisir ce niveau  $w$ .

La construction hiérarchique de l'algorithme des grilles clairsemées suggère immédiatement qu'une sorte d'extrapolation peut être utilisée pour prédire la valeur de l'approximation au niveau qui suit en utilisant les résultats obtenus aux niveaux précédents ([11],[48]). Nous proposons d'utiliser l'extrapolation de Richardson pour améliorer l'interpolation en utilisant les résultats antérieurs, et réduire ainsi nettement le coût de l'approximation.

On s'intéresse à l'approximation de la moyenne d'une quantité d'intérêt de la fonction  $u$ , qu'on note par  $\mathbb{E}[Q(u)]$  (par exemple l'espérance et la variance de  $u$ ). On l'approche par l'approximation des grilles clairsemées anisotrope :

$$\mathbb{E}\left[Q(u)\right] \approx \mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w)u\right)\right]. \quad (4.39)$$

Pour déterminer  $\mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w)u\right)\right]$ , la méthode usuelle part de  $w = 0$  et on augmente le niveau par 1 jusqu'à satisfaire le critère d'arrêt suivant :

$$\left\|\mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w)u\right)\right] - \mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w + 1)u\right)\right]\right\|_{H_0^1(D)} \approx tol. \quad (4.40)$$

Le théorème 4.1 donne une estimation de l'erreur commise en approchant  $u$  par l'interpolation  $\mathcal{A}_\alpha(N, w)u$ , en fonction de  $M_w$ , le nombre total des points de collocation de la grille  $\mathcal{H}_\alpha(N, w)$ . Cette dernière estimation reste aussi vérifiée pour  $Q(u)$  (si  $Q$  est analytique, ou si  $Q$  est à dérivée bornée). Avec cette connaissance du comportement de l'erreur en fonction du nombre de points de la grille, nous pouvons appliquer l'extrapolation de Richardson classique de la manière qui suit.

Notons pour simplifier  $Q_w = \mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w)u\right)\right]$ , et supposons que nous avons calculé une estimation de la moyenne  $\mathbb{E}[Q(u)]$  à deux niveaux différents  $w - 1$  et  $w$ , on peut ainsi écrire approximativement selon les deux estimations suivantes :

$$Q_w = \mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w)u\right)\right] = \mathbb{E}\left[Q(u)\right] + C(\alpha, N)M_w^{-\tau} \quad (4.41)$$

$$Q_{w-1} = \mathbb{E}\left[Q\left(\mathcal{A}_\alpha(N, w-1)u\right)\right] = \mathbb{E}\left[Q(u)\right] + C(\alpha, N)M_{w-1}^{-\tau} \quad (4.42)$$



où le taux de décroissance  $\tau$  est donné par  $\tau = \frac{\alpha}{\log(2) + \log(1 + \sum_{n=1}^N \frac{\alpha}{\alpha_n})}$  ( pour  $\Gamma$  borné ). Une combinaison linéaire de (4.41) et (4.42), permet de construire l'opérateur de l'extrapolation de Richardson de la quantité  $\mathbb{E}[Q(u)]$  au niveau  $w + 1$  de la manière suivante :

$$\mathcal{R}\left(\mathbb{E}[Q(u)]\right) = \frac{Q_w M_{w-1}^{-\tau} - Q_{w-1} M_w^{-\tau}}{M_{w-1}^{-\tau} - M_w^{-\tau}}. \quad (4.43)$$

L'extrapolation  $\mathcal{R}\left(\mathbb{E}[Q(u)]\right)$  permet une meilleure approximation de la moyenne  $\mathbb{E}[Q(u)]$  que celle au niveau  $w$ , c'est-à-dire  $Q_w$ . Elle est peut être comparable avec celle obtenue au niveau  $w + 1$ .

$$\begin{aligned} \mathcal{R}\left(\mathbb{E}[Q(u)]\right) - Q_w &= \frac{Q_w M_{w-1}^{-\tau} - Q_{w-1} M_w^{-\tau}}{M_{w-1}^{-\tau} - M_w^{-\tau}} - Q_w \\ &= \frac{(Q_w - Q_{w-1}) M_w^{-\tau}}{M_{w-1}^{-\tau} - M_w^{-\tau}}, \end{aligned}$$

ainsi en utilisant (4.41) et (4.42), on déduit que :

$$\mathcal{R}\left(\mathbb{E}[Q(u)]\right) - Q_w = C(\alpha, N) M_w^{-\tau}. \quad (4.44)$$

D'autre part nous avons :

$$\begin{aligned} \|Q_w - Q_{w+1}\| &\leq \|\mathbb{E}[Q(u)] - Q_w\| + \|\mathbb{E}[Q(u)] - Q_{w+1}\| \\ &\leq C(\alpha, N)(M_w^{-\tau}) + C(\alpha, N)(M_{w+1}^{-\tau}), \end{aligned} \quad (4.45)$$

ainsi les deux estimations (4.45) et (4.44) montrent que les deux erreurs  $\|Q_w - Q_{w+1}\|$  et  $\|\mathcal{R}\left(\mathbb{E}[Q(u)]\right) - Q_w\|$  sont majorées par la même borne. Cela permet de considérer le critère d'arrêt sur l'erreur  $\left\|\mathcal{R}\left(\mathbb{E}[Q(u)]\right) - Q_w\right\|$  au lieu de l'erreur  $\|Q_w - Q_{w+1}\|$ .

Dans notre approche, on ne calcule pas le prochain niveau  $Q_{w+1}$  dans la méthode des grilles clairsemées si le résultat d'extrapolation est proche du résultat au niveau courant  $Q_w$  par rapport à une tolérance donnée.

$$\left\|Q_w - \mathcal{R}\left(\mathbb{E}[Q(u)]\right)\right\|_{H_0^1(D)} \approx tol. \quad (4.46)$$

Ceci est grandement avantageux par rapport à l'approche classique qui compare la solution à deux niveaux successifs, ce qui entraîne souvent à calculer un niveau supplémentaire à un coût substantiel. Généralement le coût d'un nouveau niveau est comparable au coût cumulé de tous les niveaux précédents.

Au lieu de cela, nous utilisons l'extrapolation de Richardson pour estimer notre quantité d'intérêt à un niveau supérieur, par conséquent, une réduction considérable de l'effort est maintenue sans impact sur les précisions.

Pour tester la validité de cette approche comme un moyen d'accélération de la méthode des grilles clairsemées anisotrope, nous considérons l'exemple simple de la variable aléatoire suivante :

$$Z = \exp \left\{ \sum_{n=1}^N c_n Y_n \right\},$$

les  $Y_n$  sont indépendantes et suivent une loi normale  $N(0, 1)$ , et  $c_n = \frac{1}{n^2}$ , pour tout  $n = 1, \dots, N$ . On s'intéresse à l'approximation de la moyenne et la variance de  $Z$  par :

$$\begin{aligned} \mathbb{E}[Z] &\approx \mathbb{E}[\mathcal{A}_\alpha(N, w)Z] \\ \text{Var}[Z] &\approx \mathbb{E}[(\mathcal{A}_\alpha(N, w)Z)^2] - (\mathbb{E}[\mathcal{A}_\alpha(N, w)Z])^2. \end{aligned}$$

La moyenne et la variance exactes de  $Z$  sont données par :

$$\begin{aligned} \mathbb{E}[Z] &= e^{1/2 \sum_{n=1}^N c_n^2} \\ \text{Var}[Z] &= e^{2 \sum_{n=1}^N c_n^2} - e^{\sum_{n=1}^N c_n^2} \end{aligned}$$

La fonction  $\exp(\cdot)$  est analytique sur tout le plan complexe  $\mathbb{C}$ , par conséquent, l'erreur de quadrature unidimensionnelle d'ordre  $m_n$  suivant une direction  $n$  est donnée par  $e^{-g_n m_n}$ . Numériquement, une approche a posteriori décrite au chapitre 3 donne les valeurs des poids  $\{\alpha_n\}_{n=1}^N$  pour  $N = 10$ .

$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$
2.9	5.19	6.64	7.79	8.48	9.21	9.83	10.36	10.83	11.26

TABLE 4.1 – Les  $N = 10$  valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori.

Niveau $w$	$w=0$	$w = 1$	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
$\mathcal{H}_\alpha(w, N)$	1	3	7	21	57	137	325	759	1649

TABLE 4.2 – Le cardinal de la grille clairsemée anisotrope  $\#\mathcal{H}_\alpha(w, N)$ , les poids d'anisotropie sont donnés dans le tableau 4.1 .

Les tableaux 4.3 et 4.4 montrent une comparaison entre l'indicateur d'erreur classique ( $\text{Err}(\text{GCA})$ ) avec celui basé sur l'extrapolation de Richardson ( $\text{Err}(\text{ER})$ ) pour des niveaux  $w = 2, \dots, 8$ . On définit pour la moyenne les deux quantités suivantes :

$$\begin{aligned} \text{Err}(\text{GCA})(w) &:= \left| \mathbb{E}[\mathcal{A}_\alpha(N, w)Z] - \mathbb{E}[\mathcal{A}_\alpha(N, w+1)Z] \right| \\ \text{Err}(\text{ER})(w) &:= \left| \mathbb{E}[\mathcal{A}_\alpha(N, w)Z] - \mathcal{R}(\mathbb{E}[Z]) \right|, \end{aligned}$$

où  $\mathcal{R}(\mathbb{E}[Z])$  est calculé comme dans (4.43) (avec  $Q(Z) = Z$ ).

Pour la variance, les deux indicateurs sont donnés par :

$$\begin{aligned} \text{Err(GCA)}(w) &:= \left| \text{Var}[\mathcal{A}_\alpha(N, w)Z] - \text{Var}[\mathcal{A}_\alpha(N, w+1)Z] \right| \\ \text{Err(ER)}(w) &:= \left| \text{Var}[\mathcal{A}_\alpha(N, w)Z] - \mathcal{R}(\text{Var}[Z]) \right|, \end{aligned}$$

avec  $\mathcal{R}(\text{Var}[Z])$  est calculé comme dans (4.43) ( avec  $Q(Z) = Z^2 - \mathbb{E}[Z]$ ).

On donne aussi le nombre de points de la grille  $\mathcal{H}_\alpha(w, N)$ , donnant ainsi une idée sur l'effort qu'on évite sur chaque niveau de calcul . La comparaison montre que l'extrapolation

Niveau $w$	$w = 1$	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
Err(ER)	0.019	$2.92e - 3$	$8.62e - 04$	$2.8e - 04$	$5.9e - 05$	$1.05e - 05$	$1.59e - 06$	$2.29e - 07$
Err(GCA)	0.042	$2.9e - 2$	$6.3e - 03$	$9.4e - 04$	$1.6e - 05$	$2.5e - 05$	$2.59e - 06$	$2.97e - 07$

TABLE 4.3 – Indicateur sur l'approximation de la moyenne  $\mathbb{E}[Z]$  utilisant la grille clairsemée anisotrope (GCA) et l'extrapolation de Richardson (ER).

Niveau $w$	$w = 1$	$w = 2$	$w = 3$	$w = 4$	$w = 5$	$w = 6$	$w = 7$	$w = 8$
Err(ER)	0.97	0.09	0.029	0.013	$4.85e - 03$	$2.18e - 03$	$6.45e - 04$	$1.26e - 04$
Err(GCA)	1.29	0.68	0.33	0.077	0.033	$9.3e - 03$	$1.41e - 03$	$2.53e - 04$

TABLE 4.4 – Indicateur sur l'approximation de la variance  $\text{Var}[Z]$  utilisant la grille clairsemée anisotrope (GCA) et l'extrapolation de Richardson (ER).

de Richardson fonctionne très bien et offre une bonne approximation de l'intégrale au niveau suivant. Le tableau 4.2 donne une idée sur l'effort de calcul à éviter sur le niveau où nous avons utilisé l'extrapolation.

### 4.3 Application à un problème elliptique stochastique

Cette section illustre la convergence et l'efficacité de l'approche adaptative proposée pour un problème stochastique linéaire. Pour tester la validité de cette approche comme un moyen d'accélération de la convergence de la méthode avec une grille clairsemée anisotrope et de

réduire le coût des calculs, nous considérons un problème bidimensionnel, à deux dimensions spatiales, avec un paramètre de diffusion  $k$  exponentiel et paramétrisé par des variables aléatoires non bornées. Les résultats des calculs présentés ici sont réalisés par le logiciel Matlab.

$$\begin{cases} -\operatorname{div}(k\nabla u) = f & \text{dans } \Omega \times D := \Omega \times ]0; 1[^2 \\ u|_{\partial D} = 0 \end{cases} \quad (4.47)$$

Le paramètre de diffusion  $k$  est un champ aléatoire log-normal de la forme  $k = e^G$ , où  $G(\cdot, \cdot)$  est un champ gaussien.

Pour simplifier la présentation, on suppose que le champ  $G$  est un champ aléatoire qui dépend seulement de  $z \in [0, 1]$ , et qu'il admet une fonction de covariance exponentielle :

$$\operatorname{cov}[x_1, x_2] = \sigma^2 \exp\left\{-\frac{|x_1 - x_2|}{l_c}\right\},$$

où  $\sigma^2$  est la variance de  $G$  et  $l_c$  la longueur de corrélation. Le champ  $G$  est approché par sa troncature de Karhunen-Loève :

$$G_N(\omega, x) \approx \sum_{n=1}^N \sqrt{\lambda_n} \varphi_n(x) Y_n(\omega). \quad (4.48)$$

Les valeurs propres  $\lambda_n$  et les vecteurs propres  $\varphi_n$  de l'opérateur de covariance sont donnés explicitement (voir annexe) comme suit :

$$\lambda_n = \frac{2l_c\sigma^2}{1 + l_c^2\varpi_n^2},$$

$$\varphi_n = \beta_n \left( \sin(\varpi_n x) + l_c \varpi_n \cos(\varpi_n x) \right),$$

où  $(\varpi_n)_{n \geq 1}$  est la suite des racines positives de l'équation  $(l_c^2\varpi^2 - 1)\tan(\varpi) - 2l_c\varpi = 0$  (ordonnées dans l'ordre croissant), et  $\beta_n = \frac{1}{\sqrt{\frac{(1+l_c^2\varpi_n^2)}{2} + l_c}}$ .

Les variables aléatoires  $Y_1, \dots, Y_N$  sont gaussiennes centrées réduites et indépendantes.

La donnée  $f$  est une variable aléatoire,  $f(\omega, x_1, x_2) = 10 + \sum_{n=1}^N Y_n(\omega)$ .

On rappelle que nous avons (théorème de trace) :  $\sum_{n=1}^{\infty} \lambda_n = \sigma^2$ , ainsi selon une valeur de tolérance  $\operatorname{tol}_N$  désirable sur la moyenne et la variance de  $u$ , on peut choisir le nombre de termes  $N$  dans la série (4.48) satisfaisant le critère :

$$\left| \sum_{n=1}^N \lambda_n - \sigma^2 \right| \approx \operatorname{tol}_N.$$

Nous allons approcher  $u$  par la formule  $\mathcal{A}_\alpha(w, N)u$  et calculer à la fois  $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$  et

$l_c$	$\text{tol}_N=0.1$	$\text{tol}_N=0.05$	$\text{tol}_N=0.025$	$\text{tol}_N=0.01$
1	3	5	9	21
0.5	5	9	17	41
0.1	21	41	82	204
0.05	41	81	163	408

TABLE 4.5 – L'ordre de troncature  $N$  dans la série (4.48) par rapport à la longueur de corrélation  $l_c$  et la tolérance souhaitée  $\text{tol}_N$ .

$\text{var}[\mathcal{A}(w, N)u]$  comme des approximations de l'espérance et la variance de la fonction  $u$  respectivement. On prend un écart-type  $\sigma = 1$ , le tableau 4.5 montre l'ordre de troncature  $N$  dans la série (4.48) en fonction de la tolérance  $\text{tol}_N$  et la longueur de corrélation  $l_c$ .

Nous considérons deux cas de coefficients de corrélation  $l_c = 1$  et  $l_c = 0.5$ , et on prend l'ordre de troncature dans la série (4.48) comme  $N = 10$  pour le premier cas et  $N = 21$  pour l'autre.

Les poids d'anisotropie  $\alpha_1, \dots, \alpha_N$  sont estimés en utilisant l'approche décrite dans le chapitre 3, où nous avons utilisé l'indicateur  $\lambda = k^{-1}f$  et l'erreur de l'approximation unidimensionnelle  $O(e^{-\alpha_n m_n})$ . Le tableau 4.6 montre les valeurs des poids  $(\alpha_n)_{n=1}^N$  pour deux valeurs de  $l_c$ .

$l_c$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$
1	3.02	4.12	4.94	5.67	6.07	6.50	6.86	7.17	7.44	7.67
$l_c$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$
0.5	3.08	3.84	4.56	5.07	5.60	6.01	6.21	6.51	6.78	6.99
$\alpha_{11}$	$\alpha_{12}$	$\alpha_{13}$	$\alpha_{14}$	$\alpha_{15}$	$\alpha_{16}$	$\alpha_{17}$	$\alpha_{18}$	$\alpha_{19}$	$\alpha_{20}$	$\alpha_{21}$
7.20	7.39	7.57	7.72	7.87	8.01	8.14	8.08	8.20	8.31	8.41

TABLE 4.6 – Les  $N = 21$  valeurs des poids d'anisotropie obtenues en utilisant une information a posteriori,  $N = 10$  pour le cas  $l_c = 1$  et  $N = 21$  pour le cas  $l_c = 1/2$ .

La méthode des grilles clairsemées serait plus compétitive que celle de Monte Carlo si le taux de décroissance est plus grand que  $\frac{1}{2}$ . L'erreur de convergence unidimensionnelle dans ce cas de variables non bornées est de la forme  $O(e^{-g_n m_n})$ , on applique alors l'estimation du théorème 4.1. Elle donne une estimation de l'erreur, optimale et plus consistante. Pour cet exemple, nous avons un taux similaire donné par :

$$\mu := \frac{\log(2)eg}{2 \log(2) + \log\left(1 + \sum_{n=1}^N \frac{g}{g_n}\right)}. \quad (4.49)$$

Dans le tableau 4.7, on compare le taux de décroissance  $\mu$  avec celui montré comme dans le

théorème 1.4  $\tilde{\mu}$  (cf [50]),

$$\tilde{\mu} := \frac{\log(2)eg}{2\log(2) + \sum_{n=1}^N \frac{g}{g_n}}. \quad (4.50)$$

Rappelons que les  $g_n$  sont les taux de convergence théoriques suivant chaque direction  $y_n$  de la fonction  $u$ , nous les calculons numériquement en utilisant l'approche détaillée dans le chapitre 3, qui utilise l'indicateur  $k^{-1}f$ .

On considère une longueur de corrélation  $l_c = 0.1$ , pour plusieurs dimensions stochastiques  $N$ , nous comparons les deux taux  $\mu$  et  $\tilde{\mu}$ .

$N$	10	20	30	40	50	60	70
$\mu$	1.92	1.68	1.57	1.499	1.44	1.40	1.37
$\tilde{\mu}$	0.74	0.46	0.34	0.28	0.23	0.20	0.18

TABLE 4.7 – Comparaison du taux de convergence amélioré  $\mu$  comme dans (4.49) avec celui  $\tilde{\mu}$  comme dans (4.50).

La méthode d'interpolation anisotrope des grilles clairsemées est plus compétitive que celle de Monte Carlo si  $\mu > \frac{1}{2}$ , et c'est ce qu'on trouve numériquement, Tableau tab424 . Au contraire, le taux de décroissance  $\tilde{\mu}$  ne reflète pas cette compétitivité et efficacité puisqu'on trouve numériquement que  $\tilde{\mu} < \frac{1}{2}$ , dès que  $N \geq 20$ .

Les deux figures 4.1 et 4.2 comparent l'indicateur classique avec celui basé sur l'extrapolation de Richardson, pour l'espérance  $\mathbb{E}[u]$  et la variance  $\text{Var}[u]$ . L'indicateur classique compare les deux approximations successives à deux niveaux  $w$  et  $w + 1$ , tandis que l'indicateur de Richardson est basé sur l'extrapolation de  $u$  au niveau  $w + 1$ .

## 4.4 Conclusion

Dans ce chapitre nous avons repris la démonstration de l'estimation de l'erreur a priori de la méthode de collocation sur des grilles clairsemées anisotropes [50]. Une convergence algébrique plus fine de la méthode a été redémontrée. Cette estimation montre une compatibilité avec celle de la méthode isotrope présentée dans [49], ce qui confirme encore l'efficacité de la méthode en comparaison avec l'approche de Monte Carlo. Des estimations de l'erreur de la méthode pour un ensemble stochastique non borné sont aussi établies.

Nous avons aussi appliqué au cas anisotrope une technique d'adaptation donnée dans [11, 48] pour le cas de la méthode isotrope. Elle est basée sur l'extrapolation de Richardson et consiste à utiliser l'information de l'erreur a priori et une combinaison linéaire des approximations calculées à des niveaux précédents pour prédire une amélioration des résultats. Le calcul

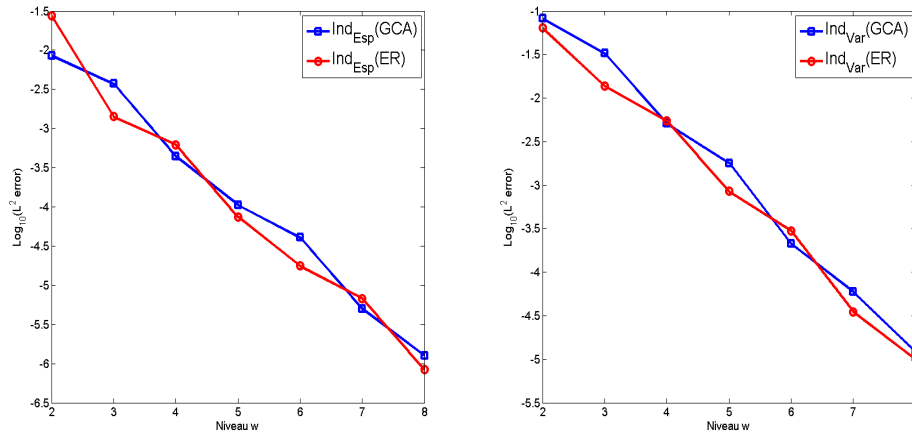


FIGURE 4.1 – Comparaison de l'indicateur classique avec l'indicateur basé sur l'extrapolation de Richardson dans le cas  $l_c = 1$  avec  $N = 10$ . À gauche l'indicateur de l'espérance et à droite celui de la variance .

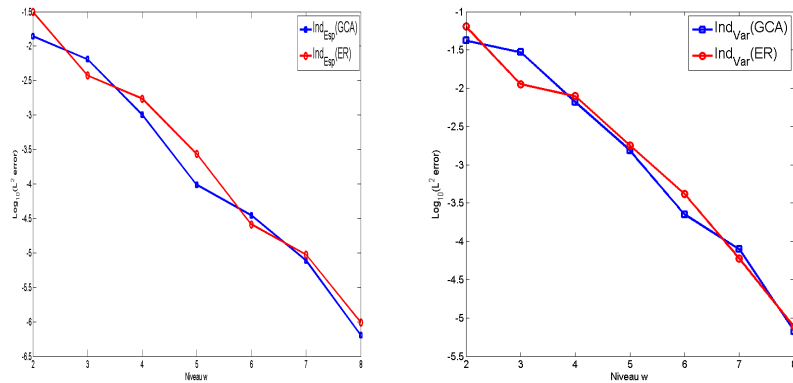


FIGURE 4.2 – Comparaison de l'indicateur classique avec celui basé sur l'extrapolation de Richardson dans le cas  $l_c = 0.5$  avec  $N = 21$ . À gauche l'indicateur de l'espérance et à droite celui de la variance.

des poids d'anisotropie est estimé en utilisant l'indicateur introduit dans le chapitre 3. Cette stratégie de résolution permet de réduire nettement le coût de calcul exigé par l'approche standard.

Dans le travail [28] les auteurs ont présenté une méthode de collocation multi-élément (ME-PCM). Lorsque la dimension stochastique est modérée, cette approche montre une accélération de la convergence de l'approximation, en comparaison avec la méthode de collocation de tensorisation pleine, comme c'est décrit dans le chapitre 1. On pense qu'il est intéressant d'étendre cette technique pour cette approche des grilles clairsemées avec une analyse de

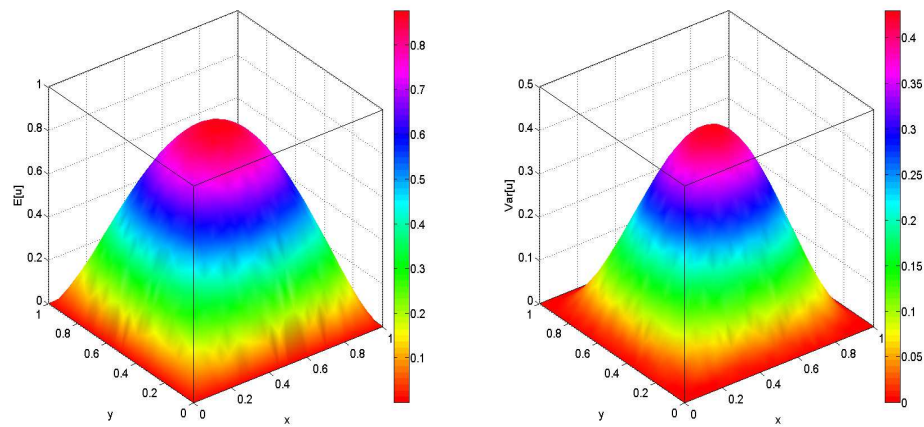


FIGURE 4.3 – L'espérance  $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$  et la variance  $\text{Var}[\mathcal{A}_\alpha(w, N)u]$ , le cas  $l_c = \frac{1}{2}$  avec  $N = 21$  et  $w = 4$ .

l'erreur a priori, pour augmenter la performance de l'adaptation.



## Première partie

# Analyse numérique du couplage de l'équation d'écoulement et de l'équation d'advection-diffusion

# Chapitre 5

## Analyse numérique de l'équation d'advection-diffusion dans un milieu aléatoire

On s'intéresse dans ce chapitre à l'approximation numérique de la solution du couplage de l'équation d'écoulement dans un milieu poreux aléatoire avec l'équation d'advection-diffusion, plus précisément, à l'approximation de la valeur moyenne de l'extension d'un soluté introduit au temps initial ( $t = 0$ ), et la valeur moyenne de la macro-dispersion. On propose une approche présentée dans [15, 53, 9, 23] basée sur la méthode de Monte-Carlo pour tenir compte des incertitudes. L'équation de l'écoulement peut être résolue en utilisant des éléments finis mixtes. L'équation d'advection-diffusion est vue comme une équation de Fokker-Planck, ainsi sa solution est approchée par une méthode particulière probabiliste. L'extension peut en effet être exprimée comme l'espérance d'une fonction de la solution de l'équation différentielle stochastique correspondante, et est calculée par un schéma d'Euler et une méthode de Monte-Carlo.

Dans [15, 53, 9, 23], le calcul de la dispersion comme dérivée de l'extension moyenne est donnée en utilisant une approximation par différences finies. Ceci exige un contrôle efficace et compatible entre le pas de discrétisation du temps et celle des différences finies.

Nous proposons ici de calculer la dispersion explicitement en utilisant la formule d'Itô. En se basant sur l'erreur faible du schéma d'Euler, et sur l'erreur spatiale, nous établissons des estimations a priori sur l'extension et la dispersion moyenne.

## 5.1 Problème physique

### 5.1.1 Équation de l'écoulement

Nous considérons un milieu poreux isotrope de porosité  $\epsilon = 1$ . Le domaine  $D$  est un pavé borné de  $\mathbb{R}^d$ , avec  $d = 1, 2, 3$ . Soit  $(\Omega, \mathfrak{A}, \mathbb{P})$  un espace de probabilité complet, pour prendre en compte les hétérogénéités des formations géologiques et le manque d'informations sur les données, nous considérons un modèle d'écoulement stochastique, où les paramètres et les données sont considérés comme des fonctions aléatoires. Un exemple classique fréquemment rencontré dans la pratique, est une perméabilité  $k$  donnée par un champ log-normal donné par :

$$k(\omega, x) = e^{G(\omega, x)}, \quad x \in D, \text{ et } \omega \in \Omega, \quad (5.1)$$

où  $G$  est un champ gaussien dont la fonction de covariance est exponentielle ou gaussienne, donnée sous la forme suivante :

$$\text{cov}(x, y) = \sigma^2 \exp\left(-\frac{\|x - y\|_2^\delta}{l_c}\right), \quad (x, y) \in D^2,$$

avec  $\delta = 1$  pour une covariance exponentielle, ou  $\delta = 2$  pour une covariance gaussienne,  $l_c$  est la longueur de corrélation,  $\sigma^2$  la variance de  $G$ .

On considère un écoulement stationnaire dans un milieu poreux incompressible donné par la loi de la conservation de la masse,  $\text{div} v = 0$  et l'équation de Darcy,  $v = k\nabla u$ , où  $v$  est la vitesse de Darcy et  $u$  la charge. Ces deux équations sont complétées par des conditions aux limites, par exemple une condition de Dirichlet sur une partie de la frontière  $\partial_1 D$ , et une condition de Neumann sur l'autre partie  $\partial_2 D$  ;

$$\begin{cases} -\text{div}(v) = 0 & \text{dans } \Omega \times D \\ v = -k\nabla u, \\ u|_{\partial_1 D} = g_1, \quad \frac{\partial u}{\partial \nu}|_{\partial_2 D} = g_2. \end{cases} \quad (5.2)$$

En pratique, le coefficient de perméabilité est donné par une série de  $N_s$  termes, comme dans le cas de la troncature de K-L dans l'Exemple 1.2. Ainsi, on suppose que le champ aléatoire  $k(\cdot, \cdot)$  est donné par une paramétrisation de  $N_s$  variables aléatoires indépendantes identiquement distribuées.

### 5.1.2 Équation d'advection-diffusion

On suppose qu'un soluté inerte est injecté dans le milieu poreux. La migration de ce soluté est gouvernée par la convection, la diffusion moléculaire et la dispersion cinématique. Pour

simplifier l'étude, la dispersion cinématique est négligée. Ici, nous considérons une diffusion moléculaire, supposée homogène et isotrope. Ce type de migration du soluté est décrit par l'équation d'advection-diffusion :

$$\frac{\partial c(\omega, x, t)}{\partial t} + v(\omega, x)\nabla c(\omega, x, t) - D_m\Delta c(\omega, x, t) = 0 \quad (5.3)$$

où  $c$  représente la concentration du soluté dans le milieu poreux,  $v$  la vitesse de Darcy, et  $D_m$  le tenseur de la diffusion moléculaire. La condition initiale à  $t = 0$  est donnée par l'injection du soluté, à savoir que  $c(0, x) = \frac{\mathbb{I}_A}{|A|}$ , où  $A$  représente un sous-ensemble contenu dans le domaine physique  $D$ , et  $|A|$  représente son volume.

### 5.1.3 Quantités d'intérêts

Nous définissons maintenant les deux quantités [33] que nous voulons calculer. On note par  $|\cdot|$  et  $\langle \cdot, \cdot \rangle$  respectivement, la norme euclidienne et le produit scalaire dans  $\mathbb{R}^d$ . On introduit le centre de masse de la distribution de soluté :

$$C_m(\omega, t) = \int_D c(\omega, t, x)x dx. \quad (5.4)$$

On s'intéresse à l'extension  $S$ ,

$$S(\omega, t) = \int_D c(\omega, t, x)|x - C_m(\omega, t)|^2 dx, \quad (5.5)$$

et à la dispersion définie comme la dérivée de l'étalement  $S(\cdot, \cdot)$  :

$$D(\omega, t) = \frac{1}{2} \frac{d}{dt} \int_D c(\omega, t, x)|x - C_m(\omega, t)|^2 dx. \quad (5.6)$$

Notre objectif est l'approximation des quantités d'intérêt suivantes :

$$\text{L'extension moyenne : } S(t) = \mathbb{E}_\omega[S(\omega, t)]. \quad (5.7)$$

$$\text{La dispersion moyenne : } \mathcal{D}(t) = \mathbb{E}_\omega[D(\omega, t)]. \quad (5.8)$$

## 5.2 Description de l'approche numérique

### 5.2.1 Méthodes particulières stochastiques

Considérons l'équation d'advection-diffusion dans un milieu déterministe, qui décrit l'évolution spatiale et temporelle de la concentration  $c$  d'un soluté. Cette évolution dépend de la

vitesse de l'écoulement  $v$  et du tenseur de diffusion moléculaire  $D_m$ , et est donné par :

$$\begin{cases} \frac{\partial c}{\partial t} + v(x)\nabla c - D_m\Delta c = 0 & \text{dans } [0, T] \times \mathbb{R}^d \\ c(0, x) = c_0(x) & x \in \mathbb{R}^d \end{cases} \quad (5.9)$$

Puisque  $\text{div}(v) = 0$  et  $D_m$  est une diffusion constante, l'équation (5.9) est une équation de Fokker-Planck.

La méthode particulière stochastique ou "random-walk" est une méthode qui consiste à simuler un nuage de particules sur le domaine  $D$ . Le mouvement des particules est caractérisé par un processus aléatoire, dont la densité de probabilité est donnée par la concentration  $c(\cdot, \cdot)$  [36, 33]. Puisqu'on s'intéresse à calculer des quantités données par des intégrales de la concentration  $c$ , la méthode de "random-walk" est plus avantageuse que les méthodes d'approximation directes, car nous n'avons pas besoin de calculer une approximation de  $c$  sur chaque point de l'espace.

Pour approcher la solution de l'équation de Fokker-Planck (5.9) avec une méthode probabiliste particulière, nous définissons l'équation différentielle stochastique associée :

$$\begin{cases} dX_t = v(X_t)dt + \sqrt{2D_m}dW_t \\ X_0 = \zeta. \end{cases} \quad (5.10)$$

où  $(W_t)$  est un mouvement brownien défini sur un autre espace de probabilité  $(\Theta, \mathcal{F}_t, \mathbf{P})$ ,  $\zeta$  est une variable aléatoire  $\mathcal{F}_t$ -indépendante qui admet  $c_0$  comme densité de probabilité. Pour notre exemple, où  $c_0 = \frac{1_A}{|A|}$ ,  $\zeta$  est une variable aléatoire de loi uniforme sur l'ensemble  $A$ .

**Remarque 5.1.**

L'équation d'advection-diffusion (5.9) n'est en réalité définie que sur le domaine physique  $D$ . Le lien entre cette équation et l'EDS (5.10) est bien établi seulement quand l'équation (5.9) est considérée sur tout l'ensemble  $\mathbb{R}^d$ . Pour l'étendre, on suppose que le domaine  $D$  est considéré tel que la concentration est presque nulle sur le bord, une extension continue du champ de la vitesse  $v$  par zéro est alors considérée en dehors de voisinage du domaine  $D$ .

$$v = v|_D\psi, \text{ où } \psi \in \mathcal{C}^\infty(\mathbb{R}^d) \begin{cases} 1 & \text{sur } D \\ 0 & \text{en dehors de } D_\epsilon. \end{cases}$$

où  $D_\epsilon$  est un petit ouvert contenant  $D$ .

Le processus  $X_t$  admet la concentration  $c(\cdot, \cdot)$  comme densité de probabilité. Le centre de masse de la distribution est donné par l'espérance de  $X_t$  :

$$C_m(t) = \int_D c(t, x)xdx = \mathbb{E}[X_t],$$

et l'extension  $S$ , qui est la propagation des particules autour de la masse  $C_m$ , est donnée par la variance de  $X_t$

$$S(t) = \int_D c(t, x)|x - C_m|^2dx = \mathbb{E}[|X_t|^2] - |\mathbb{E}[X_t]|^2. \quad (5.11)$$

La macro-dispersion temporelle est définie par  $D(t) = \frac{1}{2} \frac{d}{dt} S(t)$ , et exprime la vitesse avec laquelle le soluté s'étend.

Différentes approximations ont été proposées dans [15, 53, 9, 23, 22, 57] pour le calcul de la dispersion  $D(t)$ , cependant, il est difficile de quantifier l'erreur commise. Dans [15, 9, 23] les auteurs utilisent une différence finie, pour approcher cette dérivée en temps, et une estimation de l'erreur commise par cette approximation est donnée dans [15]. Elle montre une dépendance des pas de différences avec le nombre des simulations de Monte Carlo, qui limite l'efficacité de l'approximation. Ici on propose d'utiliser la formule d'Itô pour donner une expression explicite de cette dispersion comme fonction du processus  $X_t$ , permettant ainsi d'établir une estimation de l'erreur d'approximation temporelle et spatiale.

**Proposition 5.1.**

Soit  $X_t$  solution de l'équation différentielle (5.10),  $S(\cdot)$  l'extension donnée par (5.11), et notons par  $D(t)$  la dérivée  $\frac{1}{2} \frac{d}{dt} S(t)$ , alors  $D(t)$  est donnée explicitement comme suit :

$$D(t) = \mathbb{E}[\langle X_t, v(X_t) \rangle] - \langle \mathbb{E}[X_t], \mathbb{E}[v(X_t)] \rangle + \text{trace}(D_m). \quad (5.12)$$

**Preuve**

Par la formule d'Itô, le processus  $Y_t := |X_t|^2$  satisfait l'EDS suivante :

$$dY_t = (2 \langle X_t, v(X_t) \rangle + \text{trace}(2D_m))dt + 2\sqrt{2D_m}X_s dW_s$$

Une intégration de cette dernière EDS entre 0 et  $t$ , donne :

$$Y_t - Y_0 = \int_0^t (2 \langle X_s, v(X_s) \rangle + \text{trace}(2D_m))ds + 2 \int_0^t \sqrt{2D_m}X_s dW_s.$$

On en déduit que :

$$\mathbb{E}[Y_t] = \mathbb{E}[|X_0|^2] + \int_0^t (2\mathbb{E}[\langle X_s, v(X_s) \rangle] + \text{trace}(2D_m))ds.$$

Sachant que les trajectoires de  $X_t$  sont continues, on obtient

$$\frac{d}{dt} \mathbb{E}[|X_t|^2] = 2\mathbb{E}[\langle X_t, v(X_t) \rangle] + \text{trace}(2D_m) \quad (5.13)$$

De même, sachant que  $\frac{d}{dt} \mathbb{E}[X_t] = \mathbb{E}[v(X_t)]$ , alors :

$$\frac{d}{dt} |\mathbb{E}[X_t]|^2 = 2 \langle \mathbb{E}[X_t], \mathbb{E}[v(X_t)] \rangle, \quad (5.14)$$

et en combinant (5.13) et (5.14), on conclut que la fonction  $S(t)$  est dérivable et une expression explicite de  $D(t)$  est donnée par :

$$\begin{aligned} D(t) &= \frac{1}{2} \frac{d}{dt} (\mathbb{E}[|X_t|^2] - |\mathbb{E}[X_t]|^2) \\ &= \mathbb{E}[\langle X_t, v(X_t) \rangle] - \langle \mathbb{E}[X_t], \mathbb{E}[v(X_t)] \rangle + \text{trace}(D_m). \end{aligned}$$

■

Le tenseur de diffusion  $D_m$  a des composantes constantes, les schémas de discrétisation d'ordres élevés n'ont pas de grand intérêt dans ce cas. On utilise ainsi le schéma d'Euler pour approcher le processus  $X$ , donné par la suite des variables aléatoires  $\{X_n(t_l)\}_{l=0}^n$ ,

$$\begin{cases} X_n(t_{l+1}) = X_n(t_l) + v(X_n(t_l))\delta t + \sqrt{2D_m}\delta t\xi_l, & l = 0, \dots, n-1 \\ X_n(0) = \zeta \end{cases}$$

où  $t_0 = 0 < t_1 < \dots < t_n = T$  une partition de  $[0; T]$ ,  $\delta t = \max_{0 \leq l \leq n-1} (t_{l+1} - t_l)$  et  $\xi := (\xi_0, \dots, \xi_{n-1})$  un vecteur aléatoire gaussien de composantes indépendantes.

On simule  $M$  réalisations indépendantes  $\{\xi_j\}_{j=1}^M$  du vecteur aléatoire  $\xi$ , on obtient des simulations du processus  $\{X_n^j\}_{j=1}^M$ . On peut ainsi donner des approximations des quantités d'intérêts  $S(t)$  et  $D(t)$  par des moyennes empiriques  $S_M(t)$  et  $D_M(t)$  données par :

$$S_M(t) = \frac{1}{M} \sum_{j=1}^M |X_n^j(t)|^2 - \frac{1}{M^2} \left| \sum_{j=1}^M X_n^j(t) \right|^2. \quad (5.15)$$

$$D_M(t) = \frac{1}{M} \sum_{j=1}^M \langle X_n^j(t), v(X_n^j(t)) \rangle - \frac{1}{M^2} \sum_{j,k=1}^M \langle X_n^j(t), v(X_n^k(t)) \rangle + \text{trace}(D_m). \quad (5.16)$$

## 5.2.2 Approximation dans un milieu aléatoire

La perméabilité  $k$  est un champ aléatoire permettant de modéliser les incertitudes dans le milieu. Le champ  $k$  est paramétrisé par une série de variables aléatoires supposées indépendantes. On suppose que les incertitudes sont importantes, ainsi le nombre des variables dans  $k$  est important. On utilise la méthode de Monte Carlo pour traiter ces incertitudes.

Plus précisément, nous considérons  $N$  réalisations indépendantes du champ de la perméabilité  $k(\omega_1, \dots), k(\omega_N, \dots)$ . Pour chaque  $i \in [1, N]$ , on calcule la réalisation du champ de la vitesse  $v(\omega_i, \dots)$ , ainsi en utilisant l'approche ci-dessus (du cas déterministe), on calcule, l'extension  $S^i$  et la dispersion  $D^i$ , correspondant à chaque réalisation  $v(\omega_i, \dots)$ . On approche ainsi l'extension moyenne par  $\mathcal{S}(t) \approx \frac{1}{N} \sum_{i=1}^N S^i(t)$  et la dispersion moyenne par  $\mathcal{D}(t) \approx \frac{1}{N} \sum_{i=1}^N D^i(t)$ .

Pour la suite, l'indice  $i$  va désigner une simulation du processus par rapport à  $\omega$  et l'indice  $j$  une réalisation par rapport à  $\theta$ .

L'extension moyenne peut être approchée par

$$\mathcal{S}(t) \approx \frac{1}{N} \left( \sum_{i=1}^N \frac{1}{M} \sum_{j=1}^M |X_n^{i,j}(t)|^2 - \left| \frac{1}{M} \sum_{j=1}^M X_n^{i,j}(t) \right|^2 \right), \quad (5.17)$$

et la dispersion moyenne par,

$$\mathcal{D}(t) \approx \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{M} \sum_{j=1}^M \langle X_n^{i,j}(t), v(X_n^{i,j}(t)) \rangle - \frac{1}{M^2} \sum_{j,k=1}^M \langle X_n^{i,j}(t), v(X_n^{i,k}(t)) \rangle \right) + \text{trace}(D_m). \quad (5.18)$$

### 5.3 Analyse de l'approche numérique

Dans cette section, on s'intéresse à l'estimation de l'erreur commise en approchant des quantités d'intérêt par les approximations (5.17) et (5.18). Naturellement, cette erreur est donnée en fonction du pas de temps  $\delta t$ , du pas de la discrétisation spatiale  $h$  et du nombre de réalisations  $M$  et  $N$  qui contrôlent l'erreur statistique.

L'analyse numérique de l'approche décrite ci-dessus nécessite que la vitesse  $v$  soit suffisamment régulière par rapport à la variable  $x \in D$ . Cette régularité est naturellement associée à celle du champ de la perméabilité  $k$ . Nous considérons alors l'hypothèse suivante :

**Hypothèse 5.1.**

Le champ de perméabilité  $k \in L^2(\Omega, \mathcal{C}^{r,\alpha}(D))$  avec  $0 < \alpha \leq 1$  et  $r \geq 0$ , est tel que pour tout  $x \in D$ ,

$$0 < k_{min}(\omega) \leq k(\omega, x) \leq k_{max}(\omega), \quad \forall p.s \omega \in \Omega.$$

où les deux variables aléatoires  $k_{min}$ ,  $k_{max}$  et l'inverse  $k_{min}^{-1}$  sont dans  $L^2(\Omega)$ .

L'espace  $\mathcal{C}^{r,\alpha}(D)$  désigne l'espace des fonctions  $r$  fois dérivables, dont la dérivée  $r^{ième}$  est  $\alpha$ -höldérienne.

Notons aussi que l'approche numérique proposée suppose que l'équation d'advection-diffusion est posée dans tout  $\mathbb{R}^d$ . Pour surmonter cette difficulté, nous supposons que les conditions au bords de l'équation (5.3) sont de type Dirichlet, puisque dans la pratique le domaine  $D$  est choisi très grand par rapport à l'ensemble  $A$  et une très petite quantité de soluté atteint le bord. On peut ainsi étendre continûment par zéro la vitesse sur tout l'ensemble  $\mathbb{R}^d$ .

#### 5.3.1 Approximation de l'équation de l'écoulement

On souhaite construire une approximation précise de la vitesse discrète  $v_h$ . Ceci permet de minimiser l'erreur globale commise sur la simulation de l'équation du transport. On considère



ainsi, une approximation par la méthode des éléments finis-mixtes. Cette méthode permet de fournir simultanément une approximation de la charge hydraulique  $u$  et la vitesse  $v$  avec une même précision. Considérons le problème elliptique suivant :

$$\begin{cases} -\operatorname{div}(k\nabla u) = f & \text{dans } \Omega \times D \\ u|_{\partial D} = 0. \end{cases} \quad (5.19)$$

où le second membre  $f$  prend en compte les conditions aux bords de (5.2) tel que  $f \in \mathcal{C}^\alpha(D)$ .

**Proposition 5.2.** *Sous l'Hypothèse 5.1, le problème (5.19) admet une solution unique  $u \in L^2(\Omega; \mathcal{C}^{r+1, \alpha}(D))$ , et  $v \in L^2(\Omega; \mathcal{C}^{r, \alpha}(D))$ .*

Le problème (5.19) admet pour chaque  $\omega \in \Omega$  une solution unique  $u(\omega, \cdot)$  dans  $H_0^1(D)$ , pourvu que le bord  $\partial D$  et le second membre  $f$  soient réguliers,

$$\|u(\omega, \cdot)\|_{H_0^1(D)} \leq C_D \frac{\|f\|_{L^2(D)}}{k_{\min}(\omega)}$$

puisque  $k_{\min}^{-1} \in L^2(\omega)$  nous avons  $u \in L^2(\Omega; H_0^1(D))$ .

D'autre part, une intégration par parties montre que nous avons :

$$\begin{aligned} \|\sqrt{k^{-1}(\omega, \cdot)}v(\omega, \cdot)\|_{L^2(D)}^2 &= - \int_D u(\omega, \cdot) \operatorname{div}(v(\omega, \cdot)) dx \\ &\leq \|u(\omega, \cdot)\|_{L^2(D)} \|f\|_{L^2(D)} \end{aligned}$$

on en déduit que

$$\|v(\omega, \cdot)\|_{L^2(D)}^2 \leq k_{\max}(\omega) \|u(\omega, \cdot)\|_{L^2(D)} \|f\|_{L^2(D)}$$

Par Cauchy-shwartz, on conclut que  $v \in L^2(\Omega; L^2(D))$ .

Si de plus les trajectoires de  $k$  sont dans  $\mathcal{C}^{r, \alpha}(D)$ , alors un théorème de régularité elliptique [20, 40] montre que les trajectoires de la charge hydraulique sont de classe  $\mathcal{C}^{r+1, \alpha}(D)$ . Ainsi, les trajectoires de  $v$  sont dans  $\mathcal{C}^{r, \alpha}(D)$ . ■

Une formulation mixte du problème (5.19), pour chaque réalisation  $\omega \in \Omega$  est donnée par la formulation faible suivante :

$$\begin{cases} \text{Chercher } (u, v)(\omega) \in L^2(D) \times H(\operatorname{div}, D) \text{ telle que} \\ \int_D k^{-1} v w dx + \int_D u \operatorname{div}(w) dx = 0 & \forall w \in H(\operatorname{div}, D) \\ \int_D \operatorname{div}(v) \mu dx = \int_D -f \mu dx & \forall \mu \in L^2(D) \end{cases} \quad (5.20)$$

On considère une triangulation régulière  $\{\mathcal{T}_h\}_{h>0}$  du domaine  $D$ . On approche la vitesse  $v_h$  dans l'espace d'éléments finis de Raviart-Thomas  $RT_0(\mathcal{T}_h)$  d'ordre 0 comme un sous espace de  $H(\operatorname{div}, D)$ , et la charge hydraulique  $u$  dans l'espace des constantes par morceaux  $\mathcal{M}_h$ . On

cherche ainsi, une approximation discrète  $(u_h, v_h)$  solution du problème discret suivant, pour presque sûrement tout  $\omega \in \Omega$  :

$$\begin{cases} \int_D k^{-1} v_h w_h dx + \int_D u_h \operatorname{div}(w_h) = 0 & \forall w_h \in RT_0(\mathcal{T}_h) \\ \int_D \operatorname{div}(v_h) \mu_h = \int_D -f \mu_h dx, & \forall \mu_h \in \mathcal{M}_h, \end{cases} \quad (5.21)$$

Le problème (5.21) est bien posé [13] et, pourvu que la solution exacte  $(u, v)$  soit suffisamment régulière, nous pouvons appliquer l'estimation d'erreur déterministe de l'approximation mixte [13]. Ainsi, nous avons l'estimation suivante pour la norme  $L^2(\Omega; L^2(D)) \times L^2(\Omega; H(\operatorname{div}, D))$  :

**Proposition 5.3.**

Soit  $(u, v)$  solution de (5.20) et  $(u_h, v_h)$  solution du problème discret (5.21). Alors nous avons l'estimation suivante :

$$\|v - v_h\|_{L^2(\Omega; H(\operatorname{div}))} + \|u - u_h\|_{L^2(\Omega; L^2(D))} \leq Ch \left( \|v\|_{L^2(\Omega; H^1(D))} + \|u\|_{L^2(\Omega; H^1(D))} \right), \quad (5.22)$$

où la constante  $C$  est indépendante de  $h$ .

Nous rappelons le résultat classique du calcul stochastique [37, 2] qui relie la solution de l'équation (5.9) avec celle de l'EDS (5.10).

**Problème 5.1.**

Supposons que  $v \in C_b^{r, \alpha}(\mathbb{R}^d)$ , avec  $r \geq 0$  et  $0 < \alpha < 1$ ,  $c_0 \in C_0^0(\mathbb{R}^d)$ . Alors l'équation (5.9) admet une solution unique  $c \in C^{1, 2+r}([0, T], \mathbb{R}^d)$ , et admet  $X_t$  solution de l'EDS (5.10) comme densité de probabilité.

### 5.3.2 Erreur de discrétisation en temps

On présente dans la suite un résultat préliminaire, utile pour l'estimation de l'erreur en temps. Considérons  $(\Theta, \mathcal{F}, \mathbb{P})$  un espace de probabilité complet, et notons par  $C_b^{r, \alpha}(\mathbb{R}^d)$  l'espace des fonctions de classe  $C^r$ , qui sont bornées, ainsi que leurs dérivées, et la dérivées  $r^{\text{ième}}$  est  $\alpha$ -Höldérienne. Notons aussi par  $C_{pol}^r(\mathbb{R}^d)$  l'espace des fonctions de classe  $C^r$ , qui sont à croissance polynomiale, ainsi que leurs dérivées. Soit  $X_t$  un processus sur  $\Theta$  solution de l'EDS suivante :

$$\begin{cases} dX_t = v(X_t)dt + \sigma dW_t \\ X_0 = \zeta. \end{cases} \quad (5.23)$$

où  $\sigma > 0$ , le drift  $v \in C_b^{r, \alpha}(\mathbb{R}^d)$ , avec  $0 < \alpha < 1$  et  $r \geq 0$ . La variable aléatoire  $\zeta$  est  $\mathcal{F}_t$ -indépendante, où  $\mathcal{F}_t$  et la filtration canonique de  $W_t$ . On considère une partition quasi-uniforme de  $[0, T]$ ,  $t_0 = 0 < t_1 < \dots < t_n = T$  avec un pas de temps  $\delta t = \max_{0 \leq l \leq n-1} (t_{l+1} - t_l)$ .

Soit  $X_n(t)$  une approximation continue de  $X_t$  obtenue par le schéma d'Euler [52, 61] :

$$\begin{cases} dX_n(t) = v(X_n(t_l))dt + \sigma dW_t & t_l < t \leq t_{l+1} \\ X_n(0) = \zeta. \end{cases}$$

où  $l = 0, \dots, n-1$ .

suppose aussi que pour tout  $p > 0$ , la variable aléatoire  $\zeta$  vérifie  $\mathbb{E}[|\zeta|^{2p}] < \infty$ , ce qui permet d'assurer pour tout  $q \leq p$ , l'estimation suivante (cf [30]) :

$$\mathbb{E}[|X_t|^{2q}] \leq (1 + \mathbb{E}[|\zeta|^{2p}])e^{C_q t}; \quad \forall t \in [0, T],$$

où la constante  $C_q > 0$  ne dépend que de  $T$ ,  $q$  et  $\|v\|_{C_b^{1,\alpha}(\mathbb{R}^d)}$ . On en déduit ainsi que pour tout  $g \in C_{pol}^r(\mathbb{R}^d)$ , il existe  $m_p > 0$  telle que

$$\mathbb{E}[|g(X_t)|^{2p}] \leq m_p < \infty. \quad (5.24)$$

On s'intéresse à l'erreur faible du schéma d'Euler lorsque le drift n'est pas  $\mathcal{C}^2$ , cette erreur avait été étudiée dans [47] où on démontre une convergence d'ordre  $\frac{1}{2-\alpha}$  pour un drift de classe  $C_b^{1,\alpha}(\mathbb{R}^d)$ . Une estimation d'ordre  $\frac{1+\alpha}{2}$  était présentée dans [15] pour un bruit additif. Nous donnons ici une autre preuve du même ordre de convergence pour des raisons de complétude. L'estimation peut être étendue facilement pour un bruit multiplicatif.

**Proposition 5.4.**

Supposons que  $\mathbb{E}[|X_0|^p] < \infty$  pour tout  $p > 1$ , la fonction  $v$  est de classe  $C_b^{1,\alpha}(\mathbb{R}^d)$ , alors le schéma d'Euler est d'ordre  $\frac{\alpha+1}{2}$  pour des fonctions tests  $g$  de classe  $C_{pol}^{r,\alpha}(\mathbb{R}^d)$ , avec  $r \geq 1$  et  $0 < \alpha < 1$ .

$$\exists C(T, g) > 0, \quad |\mathbb{E}[g(X_T)] - \mathbb{E}[g(X_n(T))]| \leq C(T, g)(\delta t)^{\frac{\alpha+1}{2}}.$$

**Preuve :**

Considérons l'équation de Kolmogorov associée à l'EDS (5.23) :

$$\begin{cases} \frac{\partial u(t, x)}{\partial t} + v \cdot \nabla u + \frac{\sigma \sigma^t}{2} \Delta u = 0 & 0 \leq t < T. \\ u(T, x) = g(x). \end{cases} \quad (5.25)$$

Alors, la solution  $u \in \mathcal{C}^{1,3}([0, T[, \mathbb{R}^d)$  est donnée par [37, 2] :

$$u(t, x) = \mathbb{E}[g(X_T) | X_t = x],$$

en particulier  $u(0, X_0) = \mathbb{E}[g(X_T)]$ , l'erreur faible est donnée de la façon suivante :

$$\begin{aligned} e_T &:= \mathbb{E}[g(X_n(T))] - \mathbb{E}[g(X_T)] \\ &= \mathbb{E}[u(T, X_n(T))] - \mathbb{E}[u(0, X_0)] = \sum_{k=0}^{n-1} e_{k+1}, \end{aligned}$$

où on a posé  $e_{k+1} = \mathbb{E}[u(t_{k+1}, X_n(t_{k+1}))] - \mathbb{E}[u(t_k, X_n(t_k))]$ . Nous avons par la formule d'Itô

$$\begin{aligned} u(t_{k+1}, X_n(t_{k+1})) - u(t_k, X_n(t_k)) &= \int_{t_k}^{t_{k+1}} \left( \frac{\partial u}{\partial t} + v(X_n(t_k)) \cdot \nabla u + \frac{\sigma \sigma'}{2} \Delta u \right) (s, X_n(s)) ds \\ &+ \int_{t_k}^{t_{k+1}} \sigma \nabla u(s, X_n(s)) dW_s, \end{aligned}$$

par conséquent, l'erreur partielle  $e_{k+1}$  satisfait

$$e_{k+1} = \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \left( \frac{\partial u}{\partial t} + v(X_n(t_k)) \cdot \nabla u + \frac{\sigma \sigma'}{2} \Delta u \right) (s, X_n(s)) \right] ds.$$

En utilisant l'équation de Kolmogorov (5.25) au point  $X_n(s)$ , nous obtenons :

$$\begin{aligned} e_{k+1} &= \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \left( v(X_n(t_k)) - v(X_n(s)) \right) \cdot \nabla u(s, X_n(s)) \right] ds \\ &= J_1 + J_2, \end{aligned} \tag{5.26}$$

où nous avons posé  $J_2 := \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \left( v(X_n(t_k)) - v(X_n(s)) \right) \cdot \nabla u(s, X_n(t_k)) \right] ds$ , et le premier terme  $J_1$  par :

$$\begin{aligned} J_1 &:= \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \left( v(X_n(t_k)) - v(X_n(s)) \right) \cdot \nabla \left( u(s, X_n(s)) - u(s, X_n(t_k)) \right) \right] ds \\ &\leq \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \left( \|v\|_{C^{1,\alpha}} \|\mathcal{D}^2 u(s, Z)\| |X_n(t_k) - X_n(s)|^2 \right) \right] ds \\ &\leq \int_{t_k}^{t_{k+1}} \|v\|_{C^{1,\alpha}} \|\mathcal{D}^2 u(s, Z)\|_{L^2(\Theta)} \mathbb{E} \left[ |X_n(t_k) - X_n(s)|^4 \right]^{\frac{1}{2}}, \end{aligned}$$

où on a utilisé la formule de Taylor, avec  $Z = aX_n(t_k) + (1-a)X_n(s)$  et  $a \in [0, 1]$ , ainsi que l'inégalité de Cauchy-Schwartz.

En utilisant l'inégalité (5.24) et le fait que  $\mathbb{E}[|X_n(t_k) - X_n(s)|^4] = O(\delta t)^2$  pour tout  $s \in [t_k, t_{k+1}]$ , on conclut qu'il existe  $C_1$  indépendante de  $\delta t$  :

$$J_1 \leq C_1 (\delta t)^2. \tag{5.27}$$

Posons  $\varepsilon_n = X_n(s) - X_n(t_k)$ , par la formule de Taylor :

$$v(X_n(s)) - v(X_n(t_k)) = \mathcal{D}v(X_n(t_k)) \cdot \varepsilon_k + \int_0^1 (\mathcal{D}v(r\varepsilon_k + X_n(t_k)) - \mathcal{D}v(X_n(t_k))) \cdot \varepsilon_k dr \tag{5.28}$$

Nous avons  $\varepsilon_k = v(X_n(t_k))(s - t_k) + \sigma(W_s - W_{t_k})$ , les deux variables  $X_n(t_k)$  et  $W_s - W_{t_k}$  sont indépendantes, et  $\mathbb{E}[W_s - W_{t_k}] = 0$ , alors on en déduit que :

$$\begin{aligned} \mathbb{E} \left[ \mathcal{D}v(X_n) \cdot \varepsilon_k \nabla u(s, X_n(t_k)) \right] &= \mathbb{E} \left[ \mathcal{D}v(X_n(t_k)) \cdot v(X_n(t_k)) \nabla u(s, X_n(t_k)) \right] (s - t_k) \\ &\leq m_q \|v\|_{C^{1,\alpha}(\mathbb{R}^d)}^2 \delta t \end{aligned} \tag{5.29}$$

D'autre part, puisque  $v \in \mathcal{C}^{1,\alpha}(\mathbb{R}^d)$  on a :

$$\begin{aligned} \mathbb{E} \left[ \int_0^1 (\mathcal{D}v(r\varepsilon_k + X_n(t_k)) - \mathcal{D}v(X_n(t_k))) \cdot \varepsilon_k \nabla u dr \right] &\leq \|v\|_{\mathcal{C}^{1,\alpha}} \mathbb{E} \left[ |\nabla u(s, X_n(t_k))| |\varepsilon_k|^{\alpha+1} \right] \\ &\leq \|v\|_{\mathcal{C}^{1,\alpha}} m_p \left\| |\varepsilon_k|^{\alpha+1} \right\|_{L^2(\Theta)} \\ (\text{Inégalité de Hölder}) &\leq \|v\|_{\mathcal{C}^{1,\alpha}} m_p (\delta t)^{\frac{1+\alpha}{2}} \end{aligned} \quad (5.30)$$

En combinant les deux estimations (5.30), (5.29) avec l'égalité (5.28), on déduit que pour tout  $s \in [t_k, t_{k+1}]$  :

$$\mathbb{E}[(v(X_n(s)) - v(X_n(t_k))) \nabla u(s, X_n(t_k))] \leq \tilde{C}(\delta t)^{\frac{\alpha+1}{2}}, \quad (5.31)$$

ainsi, on obtient une estimation de terme  $J_2$ ,

$$J_2 = \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ (v(X_n(s)) - v(X_n(t_k))) \nabla u(s, X_n(t_k)) \right] ds \leq \tilde{C}(\delta t)^{\frac{\alpha+1}{2}+1}. \quad (5.32)$$

En combinant les deux estimations (5.32) et (5.27), on déduit que

$$e_{k+1} \leq C(\delta t)^{\frac{\alpha+1}{2}+1},$$

où la constante  $C$  ne dépend que de  $v$ ,  $X_0$ ,  $g$  et  $T$ . Le résultat final de la proposition est obtenu par sommation sur  $k$ . ■

Lorsque un drift  $v$  est seulement dans  $C_b^{0,\alpha}(\mathbb{R}^d)$ , nous avons l'estimation de l'erreur faible du schéma de la forme suivante :

**Proposition 5.5.**

$$\exists C(T, g) > 0, \quad |\mathbb{E}[g(X_T)] - \mathbb{E}[g(X_n(T))]| \leq C(T, g)(\delta t)^{\frac{\alpha}{2}}. \quad (5.33)$$

**Preuve :**

Considérons  $u$  solution du problème (5.25), en reprenant le début de la preuve de la proposition 5.4, nous avons

$$e_T = \sum_{k=1}^{n-1} e_{k+1},$$

où chaque erreur partielle  $e_{k+1}$  est donnée comme dans (5.26),

$$\begin{aligned} e_{k+1} &= \int_{t_k}^{t_{k+1}} \mathbb{E} \left[ \left( v(X_n(t_k)) - v(X_n(s)) \right) \cdot \nabla u(s, X_n(s)) \right] ds \\ &\leq \int_{t_k}^{t_{k+1}} \left( \mathbb{E}[|\nabla u(s, X_n(s))|^2] \right)^{\frac{1}{2}} \left( \mathbb{E}[|v(X_n(t_k)) - v(X_n(s))|^2] \right)^{\frac{1}{2}} ds \\ &\leq m_p \int_{t_k}^{t_{k+1}} \left( \mathbb{E}[|X_n(s) - X_n(t_k)|^{2\alpha}] \right)^{\frac{1}{2}} ds. \end{aligned} \quad (5.34)$$

Nous avons par l'inégalité de Hölder,

$$\begin{aligned} \mathbb{E} \left[ |X_n(s) - X_n(t_k)|^{2\alpha} \right] &\leq \left( \mathbb{E} \left[ |X_n(s) - X_n(t_k)|^2 \right] \right)^\alpha \\ &\leq \left( \mathbb{E} \left[ |v(X_n(t_k))(s - t_k) + \sigma(W_s - W_{t_k})|^2 \right] \right)^\alpha \\ &\leq c(\delta t)^\alpha, \end{aligned}$$

en substituant dans (5.34), on obtient

$$e(t_k) \leq cm_p \int_{t_k}^{t_{k+1}} (\delta t)^{\frac{\alpha}{2}} ds,$$

ainsi, on conclut par sommation sur les  $k$ . ■

### 5.3.3 Erreur de la discrétisation spatiale

On considère le processus continu  $\tilde{X}_n(t)$  qui est une perturbation de  $X_n(t)$  donnée par le schéma suivant .

$$\begin{cases} \tilde{X}_n(t) = \tilde{v}(\tilde{X}_n(t_l))dt + \sigma dW_t & t_l < t \leq t_{l+1}, l = 0, \dots, n-1 \\ \tilde{X}_n(t_0) = \zeta \end{cases}$$

où le drift  $\tilde{v}$  est une extension par zéro en dehors d'un voisinage de  $D$  de la vitesse  $v_h$ , approximation de  $v$  par les éléments finis mixtes. La fonction  $v$  est aussi prolongée par zéro comme la fonction  $\tilde{v}$ . Les deux fonctions  $v$  et  $\tilde{v}$  sont dans  $L^2(\mathbb{R}^d)$ , le prolongement est considéré de sorte que,

$$\|v - \tilde{v}\|_{L^2(\mathbb{R}^d)} = O(\|v|_D - v_h\|_{L^2(D)}).$$

La proposition suivante concerne l'erreur dans  $L^2(\Theta)$  commise en approchant  $X_n$  par  $\tilde{X}_n$ .

**Proposition 5.6.**

*Soit  $v$  et  $\tilde{v}$  dans  $L^2(\mathbb{R}^d)$ , avec  $v$   $L$ -lipschitzienne. Alors, pour tout  $T > 0$ , il existe une constante  $C > 0$ , telle que l'estimation suivante soit vérifiée :*

$$\left( \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_n(t) - \tilde{X}_n(t)|^2 \right] \right)^{\frac{1}{2}} \leq C \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)}. \quad (5.35)$$

**Preuve**

Posons pour tout  $0 \leq l \leq n-1$ ,  $t_l \leq t \leq t_{l+1}$ ,  $e_t = \tilde{X}_n(t) - X_n(t)$ , et  $\mu_{l+1} = \sup_{t_l < t \leq t_{l+1}} \mathbb{E}[|e_t|^2]$ .

Nous avons :

$$\begin{aligned}
|e_t| &= \left| e_{t_l} + (\tilde{v}(\tilde{X}_n(t_l)) - v(X_n(t_l)))(t - t_l) \right| \\
&\leq |e_{t_l}| + \left( \left| v(X_n(t_l)) - v(\tilde{X}_n(t_l)) \right| + \left| \tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l)) \right| \right) \delta t \\
&\leq |e_{t_l}|(1 + L\delta t) + \left| \tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l)) \right| \delta t,
\end{aligned}$$

On pose  $\nu := (1 + L\delta t)$ , on déduit de ce qui précède que :

$$\begin{aligned}
|e_t|^2 &\leq |e_{t_l}|^2 \nu^2 + |\tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l))|^2 (\delta t)^2 + 2\langle |e_t| \nu \sqrt{\delta t}, |\tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l))| \sqrt{\delta t} \rangle \\
&\leq |e_{t_l}|^2 \nu^2 (1 + \delta t) + |\tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l))|^2 (\delta t + (\delta t)^2).
\end{aligned} \tag{5.36}$$

Alors d'après la borne (5.36) on obtient :

$$\mu_{l+1} \leq \nu^2 (1 + \delta t) \mu_l + 2\delta t \mathbb{E}[|\tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l))|^2]. \tag{5.37}$$

D'autre part, considérons  $\tilde{c}(t, \cdot)$  la densité de probabilité de  $\tilde{X}_n(t)$ , alors nous avons :

$$\begin{aligned}
\mathbb{E}[|\tilde{v}(\tilde{X}_n(t_l)) - v(\tilde{X}_n(t_l))|^2] &= \int_{\mathbb{R}^d} |v(y) - \tilde{v}(y)|^2 \tilde{c}(t_l, y) dy \\
&\leq \sup_{0 \leq t \leq T} \sup_{y \in \mathbb{R}^d} \tilde{c}(t, y) \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)}^2 \\
&\leq \tilde{C} \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)}^2,
\end{aligned}$$

la constante  $\tilde{C} < \infty$  dépend de  $\sigma$ ,  $c_0$  et  $\|v\|_\infty$ . Par conséquent, en substituant dans l'estimation (5.37), et par une récurrence on obtient,

$$\begin{aligned}
\mu_{l+1} &\leq \nu^2 (1 + \delta t) \mu_l + 2\delta t \tilde{C} \|v - \tilde{v}\|_{\mathbb{R}^d}^2 \\
&\leq \left( \nu^2 (1 + \delta t) \right)^l \mu_1 + 2\delta t \tilde{C} \|v - \tilde{v}\|_{\mathbb{R}^d}^2 \sum_{j=0}^{l-1} \left( \nu^2 (1 + \delta t) \right)^j
\end{aligned} \tag{5.38}$$

D'autre part nous avons :

$$\begin{aligned}
\mu_1 &= \sup_{t_0 \leq t \leq t_1} \mathbb{E}[|v(X_0) - \tilde{v}(X_0)|(t - t_0)|^2] \\
&= \|c_0\|_\infty \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)} \delta t
\end{aligned}$$

En substituant dans (5.38), on obtient pour tout  $0 \leq k \leq n$

$$\begin{aligned}
\mu_{l+1} &\leq \hat{C} \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)}^2 \sum_{j=0}^l \left( \nu^2 (1 + \delta t) \right)^j \delta t \\
&\leq \hat{C} \|v - \tilde{v}\|_{\mathbb{R}^d}^2 \sum_{j=0}^l e^{(2L+1)T} \delta t \\
&\leq C \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)}^2.
\end{aligned}$$

La constante  $C = \hat{C}Te^{(2L+1)T}$  dépend de  $\|c_0\|_\infty$ ,  $\|\mathcal{D}v\|_\infty$  et  $\|v\|_\infty$ . ■

Notons que la proposition 5.6 donne une erreur forte, c'est-à-dire dans la norme de  $L^2(\Theta)$ . Pour la suite de l'analyse, nous n'avons besoin que d'une erreur faible. Le corollaire qui suit montre qu'une borne de cette erreur est obtenue à partir de la norme  $L^2(\Theta)$ .

### Corollaire 5.1.

Considérons une fonction test  $g$  dans  $\mathcal{C}_{pol}^r(\mathbb{R}^d)$ , alors il existe une constante  $C > 0$  telle que pour tout  $T > 0$ , l'estimation suivante est vérifiée :

$$\sup_{0 \leq t \leq T} \left| \mathbb{E} \left[ g(X_n(t)) - g(\tilde{X}_n(t)) \right] \right| \leq C \|v - \tilde{v}\|_{L^2(\mathbb{R}^d)}. \quad (5.39)$$

### Preuve

Par la formule de Taylor, nous avons :

$$g(X_n(t)) - g(\tilde{X}_n(t)) = \int_0^1 \mathcal{D}g \left( rX_n(t) + (1-r)\tilde{X}_n(t) \right) \left( X_n(t) - \tilde{X}_n(t) \right) dr$$

On en déduit que :

$$\mathbb{E} \left[ g(X_n(t)) - g(\tilde{X}_n(t)) \right] \leq \mathbb{E} \left[ |X_n(t) - \tilde{X}_n(t)|^2 \right]^{\frac{1}{2}} \int_0^1 \mathbb{E} \left[ \left| \mathcal{D}g \left( rX_n(t) + (1-r)\tilde{X}_n(t) \right) \right|^2 \right]^{\frac{1}{2}} dr \quad (5.40)$$

La fonction  $\mathcal{D}g$  est à croissance polynomiale, on en déduit d'après la borne (5.24) qu'il existe  $m_p$  finie tel que  $\mathbb{E} \left[ \mathcal{D}g \left( rX_n(t) + (1-r)\tilde{X}_n(t) \right) \right]^2 < m_p$ . On en déduit,

$$\mathbb{E} \left[ g(X_n(t)) - g(\tilde{X}_n(t)) \right] \leq \tilde{m}_p \mathbb{E} \left[ |X_n(t) - \tilde{X}_n(t)|^2 \right]^{\frac{1}{2}}.$$

Ainsi, par la borne (5.35) on conclut l'estimation du corollaire. ■

### Remarque 5.2.

On note que l'estimation (5.39) du corollaire précédent reste aussi vraie dans le cas où la fonction test  $g$  est seulement lipschitzienne.

## 5.3.4 Erreur sur l'extension moyenne

Nous rappelons que  $(\Omega, \mathfrak{A}, P)$  et  $(\Theta, \mathcal{F}, \mathbb{P})$  sont deux espaces de probabilités, et  $v$  un champ aléatoire sur  $(\Omega, \mathfrak{A}, P)$ ,  $W$  un mouvement brownien d-dimensionnel sur  $(\Theta, \mathcal{F}, \mathbb{P})$ , et  $D_m$  est



le tenseur de diffusion. Pour presque tout  $\omega \in \Omega$ ,  $X_t(\omega, \cdot)$  solution de l'EDS suivante :

$$\begin{cases} dX_t(\omega, \theta) = v(\omega, X_t(\omega, \theta))dt + \sqrt{2D_m}dW_t(\theta) & t \in [0, T], \\ X_0(\omega, \theta) = X_0(\theta), \end{cases} \quad (5.41)$$

où la variable aléatoire  $X_0$  est  $W_t$ -indépendante et admet  $c_0$  comme densité.

On considère une suite de réalisations  $\{v_h^i\}_{i=1}^N$  de la vitesse  $v_h$ , on construit ainsi la suite des processus  $\{\tilde{X}_n(\omega_i, \cdot)\}_{i=1}^N$ , où chaque processus est solution du schéma d'Euler de l'équation (5.41) qui correspond à la réalisation  $v_h^i$  :

$$\begin{cases} d\tilde{X}_n(\omega_i, \theta, t) = v_h^i(X_n(\omega_i, \theta, t))dt + \sqrt{2D_m}dW_t(\theta) & t_l < t \leq t_{l+1} \\ \tilde{X}_n(\omega_i, \theta, 0) = X_0(\theta), \end{cases} \quad (5.42)$$

Pour chaque  $1 \leq i \leq N$ , on considère  $M$  réalisations  $\{(\tilde{X}_n(\omega_i, \theta_j, t))\}_{j=1}^M$ , et on note pour des raisons de simplification chaque réalisation  $\tilde{X}_n(\omega_i, \theta_j, t)$  par  $\tilde{X}_n^{i,j}(t)$ . On définit une approximation  $\tilde{\mathcal{S}}_{MN}(\cdot)$  de l'extension moyenne  $\mathcal{S}(\cdot)$  par :

$$\tilde{\mathcal{S}}_{MN}(t) = \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{M} \sum_{j=1}^M |\tilde{X}_n^{i,j}(t)|^2 - \left| \frac{1}{M} \sum_{i=1}^M \tilde{X}_n^{i,j}(t) \right|^2 \right). \quad (5.43)$$

La proposition suivante montre l'erreur commise en approchant  $\mathcal{S}(\cdot)$  par l'approximation  $\tilde{\mathcal{S}}_{MN}(\cdot)$ , en particulier la précision en temps est d'ordre 1 sous l'hypothèse que  $v$  est de classe  $\mathcal{C}^1$  à dérivée lipschitzienne.

**Proposition 5.7.**

Soit  $v \in \mathcal{C}_b^{1,\alpha}(\mathbb{R}^d)$ , et soit  $\tilde{\mathcal{S}}_{MN}$  l'approximation de l'extension moyenne  $\mathbb{E}_\omega[S_t]$  comme définie par (5.43), alors nous avons l'estimation suivante :

$$\|\mathcal{S} - \tilde{\mathcal{S}}_{NM}\|_\infty \leq C \left( (\delta t)^{\frac{1+\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}} \right),$$

où la constante  $C$  est indépendante de  $\delta t$ ,  $h$ ,  $M$  et  $N$ .

**Preuve :**

La preuve suit l'approche développée dans [15, 16]. Naturellement, l'erreur peut être divisée en quatre quantités :

$$\mathcal{S} - \tilde{\mathcal{S}}_{NM} = \text{Er1} + \text{Er2} + \text{Er3} + \text{Er4},$$

où Er1, Er2, Er3 et Er4 sont définis par :

$$\begin{aligned} \text{Er1} &= \mathcal{S} - \mathbb{E}_\omega \left[ \mathbb{E}_\theta[|X_n(t)|^2] - |\mathbb{E}_\theta[X_n(t)]|^2 \right] \\ \text{Er2} &= \mathbb{E}_\omega \left[ \mathbb{E}_\theta[|X_n(t)|^2] - |\mathbb{E}_\theta[X_n(t)]|^2 \right] - \mathbb{E}_\omega \left[ \mathbb{E}_\theta[|\tilde{X}_n(t)|^2] - |\mathbb{E}_\theta[\tilde{X}_n(t)]|^2 \right] \\ \text{Er3} &= \mathbb{E}_\omega \left[ \mathbb{E}_\theta[|\tilde{X}_n(t)|^2] - |\mathbb{E}_\theta[\tilde{X}_n(t)]|^2 \right] - \frac{1}{N} \sum_{i=1}^N \left( \mathbb{E}_\theta[|\tilde{X}_n^i(t)|^2] - |\mathbb{E}_\theta[\tilde{X}_n^i(t)]|^2 \right) \\ \text{Er4} &= \frac{1}{N} \sum_{i=1}^N \left( \mathbb{E}_\theta[|\tilde{X}_n^i(t)|^2] - |\mathbb{E}_\theta[\tilde{X}_n^i(t)]|^2 \right) - \tilde{\mathcal{S}}_{MN}. \end{aligned}$$

Le premier terme Er1 tient compte de la discrétisation en temps, ainsi en appliquant la Proposition 5.4 avec  $g = |x|^2$  puis  $g = x$ , on obtient :

$$\mathbb{E}_\omega \left[ \mathbb{E}_\theta[|X_t|^2] - \mathbb{E}_\theta[|X_n(t)|^2] \right] \leq c_1 \delta t \quad (5.44)$$

$$\begin{aligned} \mathbb{E}_\omega \left[ |\mathbb{E}_\theta[X_t]|^2 - |\mathbb{E}_\theta[X_n(t)]|^2 \right] &= \mathbb{E}_\omega \left[ \left( \mathbb{E}_\theta[X_t] + \mathbb{E}_\theta[X_n(t)] \right) \left( \mathbb{E}_\theta[X_t] - \mathbb{E}_\theta[X_n(t)] \right) \right] \\ &\leq \mathbb{E}_\omega \left[ \left( \|X_t\|_{2,\theta} + \|X_n(t)\|_{2,\theta} \right) \left| \mathbb{E}_\theta[X_t] - \mathbb{E}_\theta[X_n(t)] \right| \right] \\ &\leq c_2 \delta t. \end{aligned}$$

On combine ainsi cette dernière estimation avec celle (5.44), on déduit :

$$\text{Er1} \leq C_1 \delta t. \quad (5.45)$$

En utilisant le corollaire 5.1, l'erreur de la discrétisation spatiale Er2 peut être majorée de la façon suivante :

$$\text{Er2} \leq C_2 \mathbb{E}_\omega [\|v - v_h\|_{L^2(\mathbb{R}^d)}],$$

et puisque  $\|v - v_h\|_{L^2(\mathbb{R}^d)} = O(\|v - v_h\|_{L^2(D)})$ , on déduit en utilisant la borne (5.22),

$$\text{Er2} \leq \tilde{C}_2 \mathbb{E}_\omega [\|v - v_h\|_{L^2(D)}] \leq h. \quad (5.46)$$

Psons  $Y_{1,i}(\omega) := \mathbb{E}_\theta[|\tilde{X}_n^i(t)|^2]$  et  $Y_{2,i}(\omega) := |\mathbb{E}_\theta[\tilde{X}_n^i(t)]|^2$ , pour tout  $i = 1, \dots, N$ , les deux systèmes de variables aléatoires  $(Y_{1,i})_{i=1}^N$  et  $(Y_{2,i})_{i=1}^N$  sont respectivement indépendantes et identiquement distribuées, par conséquent :

$$\begin{aligned} \mathbb{E}_\omega [\mathbb{E}_\theta[|\tilde{X}_n^i(t)|^2]] - \frac{1}{N} \sum_{i=1}^N \mathbb{E}_\theta[|\tilde{X}_n^i(t)|^2] &\leq \frac{m_1}{\sqrt{N}} \\ \mathbb{E}_\omega [|\mathbb{E}_\theta[\tilde{X}_n^i(t)]|^2] - \frac{1}{N} \sum_{i=1}^N |\mathbb{E}_\theta[\tilde{X}_n^i(t)]|^2 &\leq \frac{m_2}{\sqrt{N}} \end{aligned}$$

On en déduit que le terme Er3 vérifie :

$$\text{Er3} \leq \frac{m_1 + m_2}{\sqrt{N}}, \quad (5.47)$$

De manière analogue, pour chaque  $i \in [1, N]$ , les variables aléatoires  $(\tilde{X}_n^{i,j}(t))_{j=1}^M$  comme fonction de  $\theta$  sont indépendants et identiquement distribués, par conséquent les deux estimations sont vérifiées :

$$\begin{aligned} \left| \mathbb{E}_\theta[|\tilde{X}_n^i(t)|^2] - \frac{1}{M} \sum_{j=1}^M |\tilde{X}_n^{i,j}(t)|^2 \right| &\leq \frac{m_1}{\sqrt{M}} \\ \left| \mathbb{E}_\theta[\tilde{X}_n^i(t)] \right|^2 - \left| \frac{1}{M} \sum_{j=1}^M \tilde{X}_n^{i,j}(t) \right|^2 &\leq 2 \|\tilde{X}_n^i\|_{2,\theta} \left| \mathbb{E}_\theta[\tilde{X}_n^i(t)] - \frac{1}{M} \sum_{j=1}^M \tilde{X}_n^{i,j}(t) \right| \\ &\leq \frac{2m_2}{\sqrt{M}}, \end{aligned}$$

par conséquence, le terme Er4 satisfait l'estimation suivante :

$$|\text{Er4}(\omega, \theta)| \leq \frac{1}{N} \sum_{i=1}^N \frac{m_1 + 2m_2}{\sqrt{M}} \leq \frac{m_1 + 2m_2}{\sqrt{M}}. \quad (5.48)$$

On déduit ainsi, l'estimation de la proposition en rassemblant les bornes (5.45), (5.46), (5.47) et (5.48). ■

En utilisant l'estimation de la proposition 5.5, et en adoptant les mêmes démarches de la preuve précédente nous avons l'estimation suivante pour une vitesse moins régulière  $v \in \mathcal{C}^{0,\alpha}(\mathbb{R}^d) \cap H^1(\mathbb{R}^d)$  comme le montre la proposition qui suit.

**Proposition 5.8.**

Si  $v \in \mathcal{C}^{0,\alpha}(\mathbb{R}^d) \cap H^1(\mathbb{R}^d)$ , et soit  $\tilde{\mathcal{S}}_{MN}$  l'approximation de l'extention moyenne  $\mathbb{E}_\alpha[S_t]$  comme définie par (5.43), alors nous avons l'estimation suivante :

$$\|\mathcal{S} - \tilde{\mathcal{S}}_{NM}\|_\infty \leq \tilde{C} \left( (\delta t)^{\frac{\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}} \right),$$

où la constante  $\tilde{C}$  est indépendante de  $\delta t$ ,  $h$ ,  $M$  et  $N$ .

**Remarque 5.3.**

Soit  $\psi \in \mathcal{C}_{pol}^1(\mathbb{R}^d)$  et  $\phi \in \mathcal{C}_{pol}^2(\mathbb{R}^d)$ , on peut définir d'une manière générale l'erreur faible suivante :

$$Er := \mathbb{E}_\omega \left[ \psi \left( \mathbb{E}_\theta[\phi(X_T)] \right) \right] - \frac{1}{N} \sum_{i=1}^N \psi \left( \frac{1}{M} \sum_{j=1}^M \phi \left( \tilde{X}_n^{i,j}(T) \right) \right).$$

Alors en utilisant la formule de Taylor et en suivant les mêmes étapes de la preuve de la proposition 5.7 on montre l'estimation suivante :

$$Er = O \left( (\delta t)^{\frac{l+1}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}} \right),$$

$v \in L^2(\Omega, \mathcal{C}^{l,\alpha}(D) \cap H^1(D))$  avec  $0 < \alpha \leq 1$  et  $l = 0, 1$ .

### 5.3.5 Erreur sur la dispersion moyenne

La formule d'Itô nous permet de calculer une expression explicite ( Proposition 5.1) de la dispersion  $D(\omega_i, t)$ , pour chaque  $\omega_i \in \Omega$ , et  $i \in [1, N]$  donnée comme suit :

$$D(\omega_i, t) = \mathbb{E}_\theta[\langle X_t^i, v(\omega_i, X_t^i) \rangle] - \langle \mathbb{E}_\theta[X_t^i], \mathbb{E}_\theta[v(\omega_i, X_t^i)] \rangle + \text{trace}(D_m). \quad (5.49)$$

Considérons pour chaque  $i \in [1, N]$ , une suite de réalisations  $\{\tilde{X}_n^{i,j}\}_{j=1}^M$  du processus  $\tilde{X}_n^i$  défini par le schéma comme dans (5.41). Alors en utilisant (5.49), on construit une approximation

de la dispersion moyenne  $\mathcal{D}_t = \mathbb{E}_\omega[D(\cdot, t)]$  par la formule suivante :

$$\tilde{\mathcal{D}}_{MN}(t) = \frac{1}{MN} \sum_{i=1}^N \left( \sum_{j=1}^M \langle \tilde{X}_n^{i,j}, v_h^i(\tilde{X}_n^{i,j}) \rangle - \frac{1}{M} \left\langle \sum_{j=1}^M \tilde{X}_n^{i,j}, \sum_{j=1}^M v_h^i(\tilde{X}_n^{i,j}) \right\rangle \right) + \text{trace}(D_m). \quad (5.50)$$

La proposition suivante montre l'erreur commise en approchant  $\mathbb{E}_\omega[\mathcal{D}(\cdot, t)]$  par l'approximation  $\tilde{\mathcal{D}}_{MN}(t)$ . En particulier, la précision en temps est d'ordre 1 si la vitesse est de classe  $\mathcal{C}_b^2(\mathbb{R}^d)$  et une précision d'ordre  $\frac{1}{2}$  si  $v$  est lipschitzienne.

**Proposition 5.9.**

Considérons l'approximation de la dispersion moyenne donnée par (5.50), et supposons que  $v$  est de classe  $\mathcal{C}_b^{1,\alpha}(\mathbb{R}^d)$ , alors nous avons l'estimation suivante :

$$\|\mathcal{D}_t - \tilde{\mathcal{D}}_{NM}\|_\infty \leq C_1 \left( (\delta t)^{\frac{1+\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}} \right).$$

La constante  $C_1$  est positive et indépendante des deux pas  $\delta t$  et  $h$ , et des nombres de simulations  $N$  et  $M$ .

**Preuve :**

Supposons que  $v \in \mathcal{C}_b^{1,\alpha}(\mathbb{R}^d)$ , et posons

$$\begin{aligned} E_I &= \mathbb{E}_\omega \left[ \mathbb{E}_\theta[\langle X_t, v(X_t) \rangle] \right] - \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M \langle \tilde{X}_n(t)^{i,j}, v_h^i(\tilde{X}_n(t)^{i,j}) \rangle \\ E_{II} &= \mathbb{E}_\omega \left[ \langle \mathbb{E}_\theta[X_t], \mathbb{E}_\theta[v(X_t)] \rangle \right] - \frac{1}{MN} \sum_{i=1}^N \left\langle \sum_{j=1}^M \tilde{X}_n(t)^{i,j}, \sum_{j=1}^M v_h^i(\tilde{X}_n(t)^{i,j}) \right\rangle \end{aligned}$$

Nous allons montrer que  $E_I$  et  $E_{II}$  sont d'ordre  $O\left((\delta t)^{\frac{1+\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}}\right)$ . Considérons pour  $x \in \mathbb{R}^d$  la fonction  $g(x) = \langle x, v(x) \rangle$ , nous avons  $g \in \mathcal{C}_{pol}^{1,\alpha}(\mathbb{R}^d)$ . La quantité  $E_I$  peut être divisée en quatre termes :

$$E_I = \text{Er1} + \text{Er2} + \text{Er3} + \text{Er4},$$

où Er1, Er2, Er3 et Er4 sont définis comme suit :

$$\begin{aligned} \text{Er1} &= \mathbb{E}_\omega[\mathbb{E}_\theta[g(X_t)]] - \mathbb{E}_\omega[\mathbb{E}_\theta[g(X_n(t))]] \\ \text{Er2} &= \mathbb{E}_\omega[\mathbb{E}_\theta[g(X_n(t))]] - \mathbb{E}_\omega[\mathbb{E}_\theta[g(\tilde{X}_n(t))]] \\ \text{Er3} &= \mathbb{E}_\omega[\mathbb{E}_\theta[g(\tilde{X}_n(t))]] - \frac{1}{N} \sum_{i=1}^N \mathbb{E}_\theta[g(\tilde{X}_n^i(t))] \\ \text{Er4} &= \frac{1}{N} \sum_{i=1}^N \mathbb{E}_\theta[g(\tilde{X}_n^i(t))] - \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M g(\tilde{X}_n^{i,j}(t)). \end{aligned}$$

Le premier terme Er1 tient compte de la discrétisation en temps, puisque  $g$  est de classe  $C_{pol}^{1,\alpha}(\mathbb{R}^d)$ , en appliquant la proposition 5.4, on en déduit qu'il existe une constante  $C_1 > 0$  telle que

$$\text{Er1} \leq C_1(\delta t)^{\frac{1+\alpha}{2}}. \quad (5.51)$$

En utilisant le corrolaire 5.1, l'erreur de la discrétisation spatiale Er2 peut être majorée de la façon suivante :

$$\text{Er2} \leq C_2 \mathbb{E}_\omega[\|v - v_h\|_{L^2(\mathbb{R}^d)}],$$

et puisque  $\|v - v_h\|_{L^2(\mathbb{R}^d)} = O(\|v - v_h\|_{L^2(D)})$ , on déduit en utilisant la borne (5.22),

$$\text{Er2} \leq \tilde{C}_2 \mathbb{E}_\omega[\|v - v_h\|_{L^2(D)}] \leq \tilde{C}_3 h. \quad (5.52)$$

Le système des variables aléatoires  $Y_i = \mathbb{E}_\theta[g(\tilde{X}_n^i(t))]$ , pour tout  $i \in [1, N]$  comme fonction de  $\omega$  sont indépendantes et identiquement distribuées, par conséquent :

$$|\text{Er3}| \leq \frac{m_1}{\sqrt{N}} \quad (5.53)$$

De la même façon, pour chaque  $i \in [1, N]$ , les variables aléatoires  $(g(\tilde{X}_n^{i,j}(t)))_{j=1}^M$  comme fonction de  $\theta$  sont indépendantes et identiquement distribuées, alors on obtient :

$$\left| \mathbb{E}_\theta[g(\tilde{X}_n^i(t))] - \frac{1}{M} \sum_{j=1}^M g(\tilde{X}_n^{i,j}(t)) \right| \leq \frac{m_2}{\sqrt{M}}$$

On en déduit que :

$$\begin{aligned} |\text{Er4}| &\leq \frac{1}{N} \sum_{i=1}^N \left| \mathbb{E}_\theta[g(\tilde{X}_n^i(t))] - \frac{1}{M} \sum_{j=1}^M g(\tilde{X}_n^{i,j}(t)) \right| \\ &\leq \frac{1}{N} \sum_{i=1}^N \frac{m_2}{\sqrt{M}} \leq \frac{m_2}{\sqrt{M}}. \end{aligned} \quad (5.54)$$

En combinant les estimations (5.51),(5.52),(5.53) et (5.54), nous avons :

$$E_I = O\left((\delta t)^{\frac{1+\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}}\right).$$

D'une manière analogue on montre que  $E_{II} = O\left((\delta t)^{\frac{1+\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}}\right)$ , en remarquant que le terme Er qui prend en compte la discrétisation en temps satisfait pour chaque  $\omega \in \Omega$  :

$$\begin{aligned} Er &:= \langle \mathbb{E}_\theta[X_t], \mathbb{E}_\theta[v(\omega, X_t)] \rangle - \langle \mathbb{E}_\theta[X_n(t)], \mathbb{E}_\theta[v(\omega, X_n(t))] \rangle \\ &= \langle \mathbb{E}_\theta[X_t - X_n(t)], \mathbb{E}_\theta[v(\omega, X_t)] \rangle + \langle \mathbb{E}_\theta[X_n(t)], \mathbb{E}_\theta[v(\omega, X_t) - v(\omega, X_n(t))] \rangle \\ &\leq \|v(\omega, \cdot)\|_\infty |\mathbb{E}_\theta[X_t - X_n(t)]| + m_1 |\mathbb{E}_\theta[X_t - X_n(t)]| = O(\delta t) + O((\delta t)^{\frac{1+\alpha}{2}}) \\ &\leq C_1(\delta t)^{\frac{1+\alpha}{2}}. \end{aligned}$$

Par conséquent, on conclut l'estimation de la proposition, en combinant les deux estimations de  $E_I$  et  $E_{II}$ . ■

**Remarque 5.4.**

Si  $v$  est seulement de classe  $C_b^{0,\alpha}(\mathbb{R}^d)$ , nous pensons que nous avons une estimation sur la dispersion moyenne donnée par :

$$\|\mathcal{D} - \tilde{\mathcal{D}}_{NM}\|_\infty \leq \tilde{C}_1 \left( (\delta t)^{\frac{\alpha}{2}} + h + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M}} \right),$$

où la constante  $\tilde{C}_1 > 0$  est indépendante de  $\delta t$ ,  $h$ ,  $M$  et  $N$ .

**Remarque 5.5.**

Dans [15, 53, 9, 23] les auteurs ont proposé d'approcher la dispersion  $\mathcal{D}(t)$  par différences finies, introduisant un pas de discrétisation  $\delta s$  pour approcher la dérivée  $\frac{d}{dt}\mathcal{S}$ ,

$$\mathcal{D}(t) := \frac{1}{2} \frac{d\mathcal{S}}{dt} \approx \frac{1}{2} \frac{\mathcal{S}_{t+\delta s} - \mathcal{S}_t}{\delta s}.$$

On trouve une estimation de l'erreur commise par cette approximation dans [15, 16] d'ordre suivant,

$$O \left( (\delta t)^{\frac{1+\alpha}{2}} + \delta s + h |\ln(h)| + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{M\delta s}} \right).$$

Cette estimation montre que le pas de la discrétisation de différence finie  $\delta s$  peut affecter l'approximation de la dispersion à cause du terme  $\frac{1}{\sqrt{M\delta s}}$ . Un grand nombre de particules serait nécessaire pour maintenir le terme  $\frac{1}{\sqrt{M\delta s}}$  proche de zéro.

## 5.4 Erreur de la troncature

Dans cette section, nous considérons le cas où le coefficient de perméabilité  $k$  est donné par une troncature de Karhunen-Loève (1.4) et convenablement tronqué à un ordre  $N_s$ . Par conséquent, l'erreur de troncature  $k - k_{N_s}$  est non nulle et contribue à l'erreur sur la vitesse  $v$ , ainsi à l'erreur totale sur l'extension et la dispersion moyenne. Cette contribution doit être considérée aussi bien dans l'analyse de l'erreur.

L'erreur de troncature sur la vitesse  $v - v_{N_s}$  est équivalente à l'erreur commise sur la perméabilité  $k - k_{N_s}$  et la charge  $u - u_{N_s}$ . Cette dernière erreur est aussi liée à celle commise sur la perméabilité, et été quantifiée dans le cas de la troncature de K-L dans le Chapitre 1. On montre que pour tout  $\omega \in \Omega$ , il existe une constante  $\tilde{C}_\omega > 0$  telle que l'estimation suivante est vérifiée,

$$\|v - v_{N_s}\|_{L^2(D)} \leq \tilde{C}_\omega \|k - k_{N_s}\|_{L^2(D)}, \quad (5.55)$$

En effet, considérons  $u, v$  solutions du problème (5.19) et  $u_{N_s}, v_{N_s}$  solutions avec un paramètre tronqué  $k_{N_s}$ , nous avons pour tout  $\omega \in \Omega$  et tout  $\mu \in H_0^1(D)$ ,

$$\begin{aligned} \int_D k \nabla(u - u_{N_s}) \nabla \mu dx &= \int_D (k_{N_s} - k) \nabla u_{N_s} \nabla \mu dx + \int_D k \nabla u \nabla \mu dx - \int_D k_{N_s} \nabla u_{N_s} \nabla \mu dx \\ &\leq \|u_{N_s}\|_{C^1(\bar{D})} \|k - k_{N_s}\|_{L^2(D)} \|\nabla \mu\|_{L^2(D)}, \end{aligned}$$

en prenant  $\mu = u - u_{N_s}$ , on en déduit pour tout  $\omega \in \Omega$

$$\|u - u_{N_s}\|_{L^2(D)} \leq \frac{\|u_{N_s}\|_{C^1(\bar{D})}}{k_{min}} \|k - k_{N_s}\|_{L^2(D)}. \quad (5.56)$$

D'autre part, nous avons pour tout  $\omega \in \Omega$ ,

$$\begin{aligned} \|v - v_{N_s}\|_{L^2(D)} &= \|(k - k_{N_s})\nabla u + k_{N_s}\nabla(u - u_{N_s})\|_{L^2(D)} \\ &\leq \|k - k_{N_s}\|_{L^2(D)}\|u\|_{C^1(\bar{D})} + \|u - u_{N_s}\|_{H_0^1(D)}\|k_{N_s}\|_\infty \end{aligned}$$

par la borne (5.56), on en déduit l'estimation (5.56), où nous avons posé

$$\tilde{C} = \|u\|_{C^1(\bar{D})} + \frac{\|k_{N_s}\|_\infty \|u_{N_s}\|_{C^1(\bar{D})}}{k_{min}}. \quad (5.57)$$

■

Considérons  $Y_t$  le processus stochastique sur  $\Omega \times \Theta$  solution de l'EDS suivante :

$$\begin{cases} dY_t = v(\omega, Y_t)dt + \sqrt{D_m}dW_t \\ Y_0 = \zeta \end{cases} \quad (5.58)$$

où le drift  $v$  est la vitesse de Darcy avec un champ de perméabilité non tronqué .

Soit  $X_t$  le processus stochastique sur  $\Omega \times \Theta$  solution de l'EDS suivante :

$$\begin{cases} dX_t = v_{N_s}(\omega, X_t)dt + \sqrt{D_m}dW_t \\ X_0 = \zeta \end{cases} \quad (5.59)$$

le drift  $v_{N_s}$  est la vitesse de Darcy avec un champ de perméabilité tronqué,

La proposition qui suit fournit une estimation de l'erreur de la troncature sur des quantités d'intérêt données comme une moyenne d'une fonctionnelle du processus stochastique  $X_t$ .

**Proposition 5.10.**

Supposons que les trajectoires de  $v$  sont dans  $C_b^{0,\alpha}(\mathbb{R}^d)$  avec  $0 < \alpha \leq 1$ , et que ceux de  $v_{N_s}$  sont  $L$ -lipschitzienne, soit  $g$  une fonction lipschitzienne où  $g \in C_{pol}^l(\mathbb{R}^d)$ ,  $l \geq 1$ . Alors, nous avons l'estimation suivante :

$$\mathbb{E}_\omega \left[ \sup_{0 \leq t \leq T} |\mathbb{E}_\theta[g(Y_t) - g(X_t)]| \right] \leq C \|k - k_{N_s}\|_{L^2(\Omega; L^2(D))}, \quad (5.60)$$

où la constante  $C > 0$  est indépendante de  $N_s$ .

**Preuve**

Remarquons d'abord que pour chaque  $\omega \in \Omega$ , il existe  $C(v, g) > 0$  telle que l'estimation suivante est vérifiée :

$$|\mathbb{E}_\theta[g(Y_t) - g(X_t)]| \leq C(v, g) \|Y_t - X_t\|_{L^2(\Theta)}. \quad (5.61)$$

Posons pour chaque  $\omega$ ,  $e_t = Y_t - X_t$ , alors d'après la formule d'Itô,  $|e_t|^2$  satisfait l'équation :

$$\begin{aligned} |e_t|^2 &= \int_0^t 2 \langle v(Y_r) - v_{N_s}(X_r), e_r \rangle dr \\ &= \int_0^t 2 \langle v_{N_s}(Y_r) - v_{N_s}(X_r), e_r \rangle dr + \int_0^t 2 \langle v(Y_r) - v_{N_s}(Y_r), e_r \rangle dr \\ &\leq (2L+1) \int_0^t |e_r|^2 dr + \int_0^t |v(Y_r) - v_{N_s}(Y_r)|^2 dr, \end{aligned}$$

où nous avons utilisé le fait que  $v_{N_s}$  est L-lipschitzienne et l'inégalité  $2xy \leq x^2 + y^2$ . Ainsi, on en déduit que :

$$\mathbb{E}_\theta[|e_t|^2] \leq (2L+1) \int_0^t \mathbb{E}_\theta[|e_r|^2] dr + \int_0^t \mathbb{E}_\theta[|v(Y_r) - v_{N_s}(Y_r)|^2] dr \quad (5.62)$$

D'autre part, notons par  $c_Y(t, \omega, x)$  la densité de probabilité de  $Y_t$ , alors :

$$\begin{aligned} \mathbb{E}_\theta[|v(Y_s) - v_{N_s}(Y_s)|^2] &= \int_{\mathbb{R}^d} |v(x) - v_{N_s}(x)|^2 c_Y(t, \omega, x) dx \\ &\leq \|c_Y\|_\infty \|v - v_{N_s}\|_{L^2(\mathbb{R}^d)}^2 \end{aligned}$$

En substituant cette dernière estimation dans (5.62), on obtient pour chaque  $\omega \in \Omega$  :

$$\mathbb{E}_\theta[|e_t|^2] \leq (2L+1) \int_0^t \mathbb{E}_\theta[|e_s|^2] ds + \|c_Y\|_\infty T \|v - v_{N_s}\|_{L^2(\mathbb{R}^d)}^2,$$

On en déduit, par le lemme de Gronwall :

$$\mathbb{E}_\theta[|e_t|^2] \leq \|c_Y\|_\infty T e^{(2L+1)T} \|v - v_{N_s}\|_{L^2(\mathbb{R}^d)}^2. \quad (5.63)$$

En combinant la borne (5.61) avec (5.63), on obtient :

$$|\mathbb{E}_\theta[g(Y_t) - g(X_t)]| \leq C(v, g) \sqrt{\|c_Y\|_\infty T e^{(2L+1)T}} \|v - v_{N_s}\|_{L^2(\mathbb{R}^d)},$$

En prenant la moyenne des deux membres et par la borne (5.55), nous avons :

$$\mathbb{E}_\omega \left[ \sup_{0 \leq t \leq T} |\mathbb{E}_\theta[g(Y_t) - g(X_t)]| \right] \leq \mathbb{E}_\omega \left[ C(v, g) \sqrt{\|c_Y\|_\infty T e^{(2L+1)T}} \tilde{C} \|k - k_{N_s}\|_{L^2(D)} \right], \quad (5.64)$$

ainsi, par l'inégalité de Cauchy-Schwarz on conclut le résultat final où la constante

$$C = \|C(v, g) \sqrt{\|c_Y\|_\infty T e^{(2L+1)T}} \tilde{C}\|_{L^2(\Omega)} \text{ est indépendante de } N_s. \blacksquare$$

En utilisant le résultat de l'estimation de la Proposition 5.10, on peut en déduire une estimation de l'erreur a priori commise sur l'extension moyenne et la dispersion moyenne. Ces estimations données dans les deux Propositions 5.7 et 5.9 seront augmentées par l'erreur  $\|k - k_{N_s}\|_{L^2(\Omega; L^\infty(D))}$ , permettant ainsi de tenir compte l'erreur de la troncature sur les quantités d'intérêt.

### Remarque 5.6.

*Les hypothèses de régularité de la Proposition sur les deux vitesses  $v$  et  $v_{N_s}$  sont largement remplies. Par exemple, dans le cas où le champ  $k$  a une covariance analytique par morceaux. Aussi dans le cas d'un champ décrit dans le début du chapitre.*



## 5.5 Exemples numériques

On considère le problème elliptique unidimensionnel suivant :

$$\begin{cases} v' = 0 & \text{dans } \Omega \times ]-1, 1[ \\ v = -ku', \\ u(-1) = 0, \quad (ku')(1) = 1. \end{cases} \quad (5.65)$$

où  $k$  est un champ aléatoire de la forme  $e^G$ , où  $G$  est un champ gaussien.

La solution exacte de (5.65) est donnée par  $u(\omega, x) = \int_{-1}^x \frac{1}{k(\omega, z)} dz$ , ainsi nous avons une vitesse déterministe  $v(\omega, x) = -1$ .

On considère l'équation d'advection-diffusion suivante :

$$\begin{cases} \frac{\partial c(\omega, x, t)}{\partial t} + v(\omega, x)c'(\omega, x, t) - c''(\omega, x, t) = 0 & \text{dans } \Omega \times ]-1, 1[ \\ c(t, -1) = c(t, 1) = 0 & 0 \leq t \leq 1 \\ c(0, x) = \mathbb{I}_{[-0.5, 0.5]} & -1 \leq x \leq 1. \end{cases} \quad (5.66)$$

On prolonge continûment la vitesse par zéro au voisinage des deux points  $x = -1$  et  $x = 1$ . On approche l'extension  $\mathcal{S}$ , et la dispersion  $\mathcal{D}_t$  en utilisant l'approche détaillée précédemment. On prend le pas de discrétisation du schéma d'Euler  $\delta = \frac{1}{50}$ , et le nombre  $M = 10^6$  pour la simulation du mouvement brownien.

La figure 5.2 montre une approximation de la dispersion en utilisant une différence finie avec un pas de discrétisation  $\delta s$  comme proposées dans [16]. On remarque que même si l'approximation en temps est optimale par rapport à celle explicite donnée dans (5.1), l'approximation présente une instabilité par rapport au pas de différence  $\delta s$ . Plus le pas  $\delta s$  est petit, plus l'approximation diverge.

## 5.6 Conclusion

Dans ce chapitre, nous avons présenté une analyse numérique d'une méthode probabiliste basée sur la marche aléatoire et permettant de quantifier la migration d'un contaminant dans

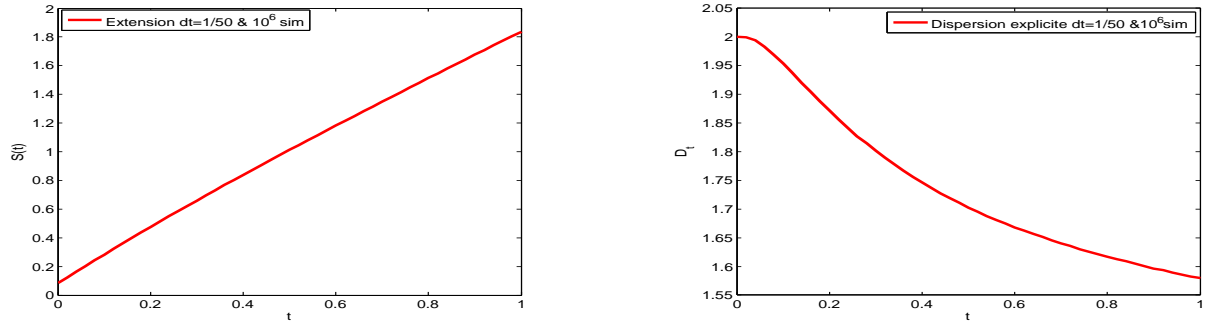


FIGURE 5.1 – L’extension (à gauche) et la dispersion explicite (à droite) avec un pas de temps  $\delta t = 1/50$  et  $M = 10^6$  de simulations

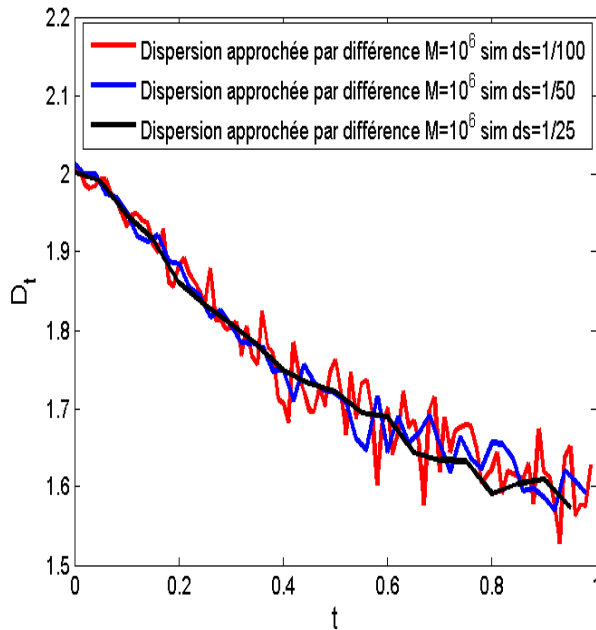


FIGURE 5.2 – La dispersion approchée par une différence finie.

un milieu aléatoire. Plus précisément, nous avons considéré le problème d’écoulement avec le couplage de l’équation d’advection-diffusion. Nous nous sommes intéressés à l’approximation

de l'extension moyenne et à la dispersion moyenne de soluté.

Le modèle d'écoulement est discrétisé par une méthode d'éléments finis mixtes, garantissant la conservation locale de la masse et permettant ainsi de calculer à la fois la charge hydraulique et la vitesse avec une bonne précision. La concentration du soluté est vue comme une fonction de densité d'un processus stochastique solution d'une équation différentielle stochastique donnée. Celle-ci est discrétisée par un schéma d'Euler. Une expression explicite de la dispersion moyenne est démontrée en utilisant la formule d'Itô comme fonction du processus stochastique, ainsi les deux moyennes des deux quantités d'intérêt sont estimées à l'aide d'échantillonnage de Monte Carlo. Nous avons aussi démontré des estimations de l'erreur optimales de la méthode, à savoir l'erreur commise sur la moyenne de l'extension et de la dispersion.

Ces estimations de l'erreur montrent qu'on peut exploiter d'autres orientations futures, pour optimiser le temps de calcul de la méthode, à savoir l'utilisation des lois de probabilités jointes. On pense aussi à introduire des méthodes de collocation stochastiques, telle que la méthode anisotrope des grilles clairsemées pour l'approximation du problème de l'écoulement.

# Conclusions et perspectives

La quantification d'incertitude dans les modèles constitue une étape importante dans la modélisation des problèmes d'écoulement et de transport souterrains. Elle conduit généralement à considérer une "EDP" ou une "EDO" stochastique, où les paramètres de ces équations sont considérés comme des champs aléatoires. Cependant, les modèles numériques de tels problèmes sont particulièrement coûteux en temps et en calculs, à cause de la nature complexe et de la dépendance de l'effort de simulation vis-à-vis de la dimension stochastique des paramètres du modèle.

L'objectif de cette thèse était de développer et d'analyser des approches numériques permettant de réduire la complexité de ces problèmes. Ainsi, dans ce manuscrit, nous avons présenté quelques résultats concernant des méthodes numériques déterministes et probabilistes pour des équations aux dérivées partielles à coefficients aléatoires, dont les applications peuvent concerner le domaine de l'hydrogéologie.

La première partie a concerné l'équation de l'écoulement dans un milieu poreux en régime stationnaire avec un coefficient de perméabilité stochastique. Nous avons présenté une méthode de projection pour le cas unidimensionnel, permettant de calculer efficacement la moyenne de la charge hydraulique sur une base de variable aléatoire de nombre assez réduit. La dimension de cette base augmente linéairement avec la dimension stochastique, alors que la projection classique de type Galerkin construit une projection de la solution sur une base dont la dimension augmente exponentiellement vis-à-vis de la dimension stochastique.

Les deux chapitres 3 et 4 ont concerné la méthode de collocation stochastique pour l'approximation du problème de l'écoulement dans l'espace stochastique. Elle est basée sur l'interpolation anisotrope sur des grilles clairsemées. D'abord, nous avons introduit un indicateur de l'erreur satisfaisant une borne supérieure de l'erreur qui nous a permis de calculer les poids d'anisotropie de la méthode avec moins d'effort. Ensuite, nous avons démontré une amélioration de l'erreur a priori de la méthode anisotrope des grilles clairsemées, et ceci pour un ensemble stochastique borné ou non borné. Cette estimation confirme l'efficacité de la méthode en comparaison avec l'approche de Monte Carlo, elle montre aussi une compatibilité et une consistance avec le résultat de la méthode des grilles clairsemées isotropes.

Nous avons aussi appliqué une technique d'adaptation basée sur l'extrapolation de Richardson, qui a consisté à utiliser l'information de l'erreur a priori précédente et une combinaison linéaire des résultats calculés auparavant pour prédire une amélioration des résultats. Ceci présente un grand avantage, puisqu'on évite un effort de calcul considérable à un niveau supplémentaire.

La deuxième partie a été consacrée au couplage de l'équation d'écoulement avec une équation d'advection-diffusion. Nous avons présenté une analyse numérique de la méthode permettant de quantifier l'extension et la vitesse avec laquelle un contaminant s'étend dans un milieu poreux aléatoire. Il s'agit d'une méthode de Monte-Carlo combinant une méthode des éléments finis mixtes pour l'équation d'écoulement et un schéma d'Euler explicite pour l'équation différentielle stochastique associée à l'équation d'advection-diffusion. Nous avons donné une expression explicite de la dispersion moyenne comme fonction du processus stochastique solution de l'EDS associée à l'équation d'advection-diffusion. Ainsi les deux moyennes des deux quantités d'intérêt sont estimées à l'aide d'échantillonnage de Monte Carlo. Cette forme explicite nous a permis de quantifier l'estimation de l'erreur commise par la discrétisation temporelle. Des estimations de l'erreur optimales de la méthode ont été présentées, à savoir des estimations optimales de l'erreur commise sur la moyenne de l'extension et de la dispersion.

Plusieurs orientations futures concernant le travail de cette thèse peuvent être envisagées. D'abord une vérification ou une extension de la projection donnée dans le chapitre 2 pour des dimensions quelconques (2 ou 3), ou au moins la revoir numériquement. Ensuite pour la méthode anisotrope des grilles clairsemées, on pense qu'il est intéressant d'étudier une alternative de cette méthode à savoir une méthode multi-éléments anisotropes des grilles clairsemées, de faire une analyse numérique et une comparaison avec celle de multi-niveaux de Monte Carlo [16]. Enfin pour la deuxième partie de cette thèse, on pense d'abord à établir une extension de la méthode présentée dans le chapitre 5 utilisant des lois jointes pour le tirage de Monte Carlo. Ensuite on pense aussi à combiner la méthode de collocation anisotrope des grilles clairsemées pour le problème d'écoulement et la méthode de Monte Carlo pour la marche aléatoire. Finalement, une autre direction pour ces problèmes stochastiques qui peut être aussi envisagée, elle concerne le contrôle de la discrétisation spatiale par élément finis.

# Annexe

## 5.7 Développement de Karhunen-Loève pour les champs aléatoires

En général, les champs aléatoires ne satisfont pas l'hypothèse de dimension finie ( chapitre 1, section 1.1). Cette hypothèse est absolument cruciale pour la résolution des équations aux dérivées partielles stochastiques.

Afin de travailler sous cette hypothèse de bruits finis, on utilise une approximation des champs en séries de fonctions déterministes et aléatoires [3, 34]. La décomposition de Karhunen-Loève est l'une des discrétisations cruciales utilisées dans les méthodes comme dans le cas de la méthode de Galerkin stochastiques. Le développement consiste essentiellement à représenter un champ aléatoire sur la base des fonctions propres de sa fonction de covariance.

### 5.7.1 Formulation du problème

Considérons un champ aléatoire  $a$  de second ordre,  $a \in L^2(\Omega) \otimes L^2(D)$ , qui est continue pour la norme quadratique, i.e pour tout  $x \in D$ , nous avons :

$$\lim_{z \rightarrow x} \mathbb{E} \left[ \|a(\cdot, z) - a(\cdot, x)\|_{L^2(D)} \right] = 0.$$

La continuité de  $a$  permet d'assurer que la fonction de covariance  $\text{cov}[a]$  définie par :

$$\text{cov}[a](x, z) := \mathbb{E} \left[ a(\cdot, x) a(\cdot, z) \right] - \mathbb{E} \left[ a(\cdot, x) \right] \mathbb{E} \left[ a(\cdot, z) \right],$$

est continue dans  $D \times D$ . On définit ainsi,  $T$  l'opérateur linéaire sur  $L^2(D)$ , donnée par le noyau de covariance  $\text{cov}[a]$  comme suit :

$$T(\phi)(x) := \int_D \text{cov}[a](x, z) \phi(z) dz.$$

La fonction de covariance est borné, symétrique et définie positive. Par conséquent,  $T$  est symétrique définie positif (opérateur de Hilbert-Schmidt) sur  $L^2(D)$ , Nous avons ainsi les propriétés pertinentes suivantes[7] .

- le problème aux valeurs propres suivant :

$$\int_D \text{cov}[a](x, z)\phi(z)dz = \lambda\phi(x), \quad (5.67)$$

admet un système dénombrable de valeurs et de fonctions propres  $\{\lambda_n, \phi_n\}_{n=1}^\infty$ .

- La suite des fonctions propres  $\{\lambda_n, \phi_n\}_{n=1}^\infty$ . forme une base orthonormée de l'espace  $L^2(D)$ .
- Les valeurs propres  $\{\lambda_n\}_{n=1}^\infty$  sont positives et vérifiant :

$$\begin{aligned} \lambda_1 &< \lambda_2 < \dots < \lambda_n \longrightarrow_{n \rightarrow \infty} 0, \\ \sum_{n=1}^{\infty} \lambda_n &= \int_D \text{cov}[a](x, x)dx < \infty. \end{aligned}$$

- Le noyau  $\text{cov}[a]$  possède alors la décomposition spectrale suivante :

$$\text{cov}[a](x, z) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \sqrt{\lambda_m} \phi_n(x) \phi_m(z).$$

Une conséquence de ces propriétés et que le champ aléatoire  $a$  peut être exprimé en séries de Fourier généralisées sur la base des fonctions propres de sa fonction de covariance, cette décomposition est donnée comme suit :

$$a(\omega, x) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \phi_n(x) Y_n(\omega), \quad (5.68)$$

où la suite  $\{Y_j\}_{j=1}^\infty$  sont des variables aléatoires données par :

$$Y_n(\omega) = \frac{1}{\sqrt{\lambda_n}} \int_D (a(\omega, x) - \mathbb{E}[a(x)]) \phi_n(x) dx. \quad (5.69)$$

et satisfaisant  $\mathbb{E}[Y_n] = 0$  et  $\mathbb{E}[Y_n Y_m] = \delta_{n,m}$ .

On définit ainsi, la troncature de Karhunen-Loève  $a_N$  comme la troncature de la série (5.68) à l'ordre  $N$  comme suit :

$$a_N(\omega, x) = \mathbb{E}[a](x) + \sum_{n=1}^N \sqrt{\lambda_n} \phi_n(x) Y_n(\omega).$$

Cette troncature est la décomposition optimale du processus  $a$  au sens de  $L^2(\Omega) \otimes L^2(D)$  sur l'ensemble de toutes les décompositions de type  $\sum_{n=1}^N v_n \zeta_n$ , telles que  $(\zeta_n, v_n) \in L^2(\Omega) \times L^2(D)$  [44, 45].

## 5.8 Champ Gaussian de covariance exponentielle

Un champ aléatoire Gaussian  $G$  est parfaitement déterminé par son espérance et sa fonction de covariance, nous considérons comme exemple de fonction de covariance de  $G$  la fonction exponentielle donnée par :

$$C(x_1, x_2) = \sigma^2 \exp\left(-\frac{|x_1 - x_2|}{l_c}\right), \quad \forall (x_1, x_2) \in [0, 1]^2.$$

où  $\sigma^2$  est la variance de  $G$  et  $l_c$  le coefficient de corrélation. L'avantage de cette fonction de covariance est que l'équation intégrale de Fredholm (5.67) peut être résolue analytiquement [34]. En dérivant deux fois (5.67), nous obtenons l'équation différentielle du second ordre suivante :

$$\phi'' + \varpi^2 \phi = 0, \quad (5.70)$$

où le coefficient  $\varpi^2 = \frac{2\sigma^2 - \lambda}{l_c \lambda}$ , et les deux conditions aux limites suivantes sont satisfaites :

$$l_c \phi'(0) - \phi(0) = 0 \quad (5.71)$$

$$l_c \phi'(1) + \phi(1) = 0. \quad (5.72)$$

Une solution non triviale de l'équation (5.70) existe si  $\varpi \geq 0$ , elle est donnée par :

$$\phi(x) = a \cos(\varpi x) + b \sin(\varpi x), \quad (5.73)$$

L'équation (5.71) montre que les deux constantes  $a$  et  $b$  vérifient :

$$\frac{a}{b} = l_c \varpi, \quad (5.74)$$

en combinant (5.74) avec la condition (5.72), on déduit que le réel  $\varpi$  satisfait l'équation :

$$\tan(\varpi) = \frac{2\varpi l_c}{l_c^2 \varpi^2 - 1}. \quad (5.75)$$

Cette dernière équation admet une infinité de racines  $(\varpi_n)_n^\infty$ , où chaque  $\varpi_n$  lui correspond la valeur propre :

$$\lambda_n = \frac{2\sigma^2 l_c}{l_c^2 \varpi_n^2 + 1}, \quad (5.76)$$

et la fonction propre  $\phi_n$  :

$$\phi_n(x) = \beta_n (l_c \varpi_n \cos(\varpi_n x) + \sin(\varpi_n x)), \quad (5.77)$$

la constante  $\beta_n$  est calculée de sorte que  $\|\phi_n\|_{L^2(D)} = 1$ , i.e  $\beta_n = \frac{1}{\sqrt{(1 + l_c^2 \varpi_n^2)/2 + l_c}}$ .

Figure 5.3 montre les 10 premières valeurs propres de cette covariance exponentielle pour différentes valeurs de la longueur de corrélation  $l_c$ . On peut remarquer que les valeurs propres



décroissent de façon monotone lorsque  $n$  augmente. Plus la longueur de corrélation est grande, plus la décroissance des valeurs propres vers zéro est plus grande.

Puisque  $\tan(n\pi) = 0$ , on constate de l'équation (5.75) que ses racine  $\varpi_n \approx n\pi$  au voisinage de  $n \rightarrow \infty$ , ainsi de l'expression (5.76), nous déduisons que les valeurs propres décroissent algébriquement vers zéros comme suit

$$\lambda_n \approx_{n \rightarrow \infty} \frac{2\sigma^2}{nl_c}.$$

Cette décroissance est due du fait que la fonction de covariance n'est pas analytique au voisinage de la première bissectrice. Lorsque la fonction de covariance est analytique sur  $D^2$  comme le cas d'une fonction Gaussienne, les valeurs propres décroissent rapidement vers zéros avec une décroissance exponentielle (cf,[56]).

Figure 5.4 montre les 6 premières vecteurs propres. Lorsque  $n$  augmente, ces fonctions deviennent de plus en plus oscillatoires.

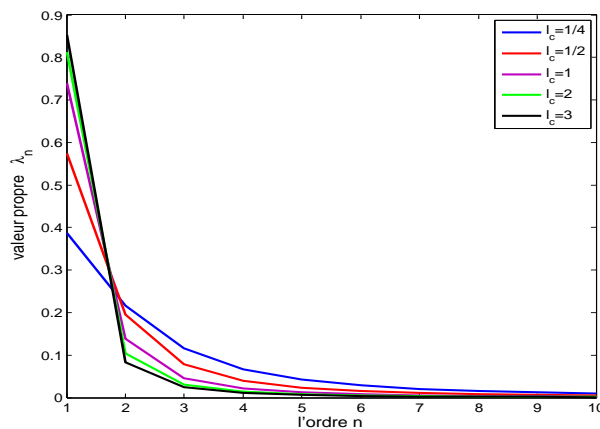


FIGURE 5.3 – Valeurs propres de la covariance exponentielle,  $l_c = 1$ ,  $\sigma = 1$

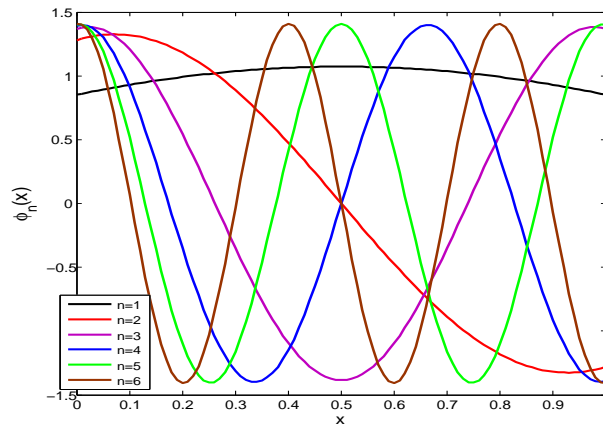


FIGURE 5.4 – Vecteurs propres de la covariance exponentielle,  $l_c = 1$ ,  $\sigma = 1$ .

# Bibliographie

- [1] R. AGARWAL AND P. WONG "*Error inequalities in polynomial interpolation and their applications*", vol. 262, Kluwer, 1993.
- [2] ALESSANDRA LUNARDI. "*Analytic semigroups and optimal regularity in parabolic problems. Progress in Nonlinear Differential Equations and their Applications*", 16. Birkhäuser Verlag, Basel, 1995
- [3] ANTHONY NOUY, "*Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations*". Archives of Computational Methods in Engineering (2009) doi :10.1007/s11831-009-9034-5.
- [4] I.BABUSKA AND P.CHATZIPANTELIDIS : "*On solving linear elliptic stochastic partial differential equations*", Comput.Methods Appl.Mech.Engrg., 191 :4093-4122, 2002.
- [5] I. M. BABUŠKA, R. TEMPONE, AND G. E. ZOURARIS "*Solving elliptic boundary value problems with uncertain coefficients by the finite element method : the stochastic formulation*", Comput. Methods Appl. Mech. Engrg., 194(12- 16) :1251-1294, 2005.
- [6] I.M. BABUŠKA, F. NOBILE, AND R. TEMPONE : "*A stochastic collocation method for elliptic partial differential equations with random input data*", SIAM J. Numer. Anal, 43(3) :1005-1034, 2007
- [7] BACHMAN, G., NARICI, L. "*Functional Analysis*". Dover, New York (2000)
- [8] BARTHELMANN, V., E. NOVAK, AND K. RITTER, 2000A "*High dimensional polynomial interpolation on sparse grids*". Adv. Comput. Math., 12(4), 273-288, multivariate polynomial interpolation. 9, 34, 45
- [9] BEAUDOIN, A., DREUZY, J.R., ERHEL, J., "*An efficient parallel particle tracker for advection-diffusion simulations in heterogeneous porous media*", Europar 2007.
- [10] J. P. BOYD. "*Chebyshev and Fourier Spectral Methods*". Dover Publication, New York, 2000.
- [11] H. BUNGARTZ, M. GRIEBEL, U. RÜDE, "*Extrapolation, combination, and sparse grid techniques for elliptic boundary value problems*", Comput. Methods Appl. Mech. Engrg. 116 (1994) 243-252.
- [12] J. BURKARDT : "*SGMGA, Sparse Grid Mixed Growth Anisotropic Rules*", [http ://people.sc.fsu.edu/~jburkardt/cpp-src/sgmga/sgmga.html](http://people.sc.fsu.edu/~jburkardt/cpp-src/sgmga/sgmga.html)
- [13] BREZZI, F. FORTIN, M., "*Mixed and Hybrid Finite Element Methods*". Springer-Verlag, New York, NY, 1991

- [14] R. E. CAFLISCH. " *Monte carlo and quasi-monte carlo methods*". Acta. Numer., 7 :1-49, 1998.
- [15] CHARRIER J. " *Numerical analysis of the advection-diffusion of a solute in random media*", Rapport d'INRIA, Num RR-7585 (2011).
- [16] CHARRIER J. " *Analyse numérique d'équations aux dérivées partielles à coefficients aléatoires, applications à l'hydrogéologie*".
- [17] CHARRIER J. " *Strong and weak error estimates for the solutions of elliptic partial differential equations with random coefficients*" , SIAM Journal on Numerical Analysis 2012, Vol. 50, No. 1, pp. 216-246
- [18] P. G. CIARLET. : " *The Finite Element Method for Elliptic Problems*". North-Holland, 1978.
- [19] P.J. DAVIS, P. RABINOWITZ : " *Methods of Numerical Integration*", Academic Press, New York, 1975.
- [20] DAVID GILBARG AND NEIL S. TRUDINGER. " *Elliptic partial differential equations of second order*". Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition
- [21] M. DEB, I. BABUŠKA, AND J. T. ODEN. " *Solution of stochastic partial differential equations using galerkin finite element techniques*". Computer Methods in Applied Mechanics and Engineering, 190 :6359-6372, 2001.
- [22] DENTZ, M. KINZELBACH, H. ATTINGER, S. AND W. KINZELBACH. " *Temporal behavior of a solute cloud in a heterogeneous porous medium*" : 1. Point-like injection, Water Resources Research, 36 (12), pp. 3591-3604
- [23] DREUZY, J.R. AND BEAUDOIN, A. AND ERHEL, J., " *Asymptotic dispersion in 2D heterogeneous porous media determined by parallel numerical simulations*", Water Resources Research, 2007, 43
- [24] I. ELISHAKO. " *Whys and Hows in Uncertainty Modelling-Probability, Fuzziness, and Anti-Optimization*". Springer-Verlag, Berlin, 1999
- [25] J. ERHEL, Z. MGHAZLI, M. OUMOUNI : " *Calcul de l'espérance de la solution d'une EDP stochastique unidimensionnelle à l'aide d'une base réduite*", C.R.Acad.Sci.Paris, Ser. I, 349 (2011), 861-865.
- [26] J. ERHEL, J.-R. DE DREUZY, A. BEAUDOIN, E. BRESCIANI, AND D. TROMEUR-DERVOU. " *A parallel scientific software for heterogeneous hydrogeology*. In Ismail H. Tuncer, Ulgen Gulcat, David R. Emerson, and Kenichi Matsuno, editors, Parallel Computational Fluid Dynamics 2007, volume 67 of Lecture Notes in Computational Science and Engineering, pages 39-48. Springer, 2009. invited plenary talk
- [27] G. FISHMAN " *Monte Carlo : Concepts, Algorithms, and Applications*", Springer-Verlag, New York, 1996.
- [28] J. FOO, X. WAN, G. E. KARNIADAKIS. " *The multi-element probabilistic collocation method (ME-PCM) : Error analysis and applications*", J. Comput. Phys. 227 (2008) 9572-9595

- [29] P. FRAUENFELDER, C. SCHWAB AND R.A. TODOR, "*Finite Elements for Elliptic Problems with Stochastic Coefficients*". Computer Methods in Applied Mechanics and Engineering, Volume 194, Issues 2-5, 4 Feb 2005, Pages 205-228.
- [30] FRIEDMAN, A. "*Stochastic Differential Equations and Applications*", Academic Press, New-York, 1975
- [31] GARCKE, J. AND M. GRIEBEL, 2002 "*Classification with anisotropic sparse grids using simplicial basis functions*". Intelligent Data Analysis, 6, 483-502. 47
- [32] T. GERSTNER AND M. GRIEBEL. "*Dimension-adaptive tensor-product quadrature*". Computing, 71 :65-87, 2003.
- [33] GELHAR, L. "*Stochastic Subsurface Hydrology*", Engelwood Cliffs, New Jersey (1993).
- [34] R. G. GHANEM AND P. D. SPANOS : "*Stochastic finite elements : a spectral approach*", Springer-Verlag, New York, 1991.
- [35] T. GERSTNER, M. GRIEBEL "*Numerical integration using sparse grid*", Numer. Algorithms 18 (1998) 209-232.
- [36] IOANNIS KARATZAS AND STEVEN E. SHREVE. "*Brownian motion and stochastic calculus, volume 113 of Graduate Texts in Mathematics*". Springer-Verlag, New York, second edition, 1991
- [37] KUNITA H. "*Stochastic Differential Equations and Stochastic Flows of Diffeomorphisms*", Cours de l'école d'été de Probabilité de Saint-Flour, 1982, Lect. Notes in Math., Vol 1097, Springer-Verlag, 1982.
- [38] M. KLEIBER AND T.D. HIEN. "*The Stochastic Finite Element Method. Basic Perturbation Technique and Computer Implementation*". John Wiley & sons, Chichester, 1992.
- [39] A. KLIMKE, B. WOHLMUTH, "*Algorithm 847 : spinterp : Piecewise Multilinear Hierarchical Sparse Grid Interpolation in MATLAB*", ACM Transactions on Mathematical Software 31 (2005).
- [40] LADYZHENSKAYA, O. AND URALTSEVA, N., "*Linear and quasilinear elliptic equations*", Academic Press, 1968
- [41] O.P. LE MAÎTRE, O.M. KNIO "*Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*", Computational Science 2010.
- [42] O.P. LE MAÎTRE, O. M. KNIO, B. J. DEBUSSCHERE, H. N. NAJM, AND R. G. GHANEM "*A multigrid solver for two-dimensional stochastic diffusion equations*". Comput. Methods Appl. Mech. Engrg., 192(41-42), 4723-4744. 42
- [43] O.P. LE MAÎTRE, O. M. KNIO, H. N. NAJM, AND R. G. GHANEM : "*Uncertainty propagation using wiener-Haar expansions*", Journal of Computational Physics, 197(1) :28-57, 2004.
- [44] M. LOÈVE : "*Probability Theory I*", fourth edition, in : Graduate Texts in Mathematics, vol. 45, Springer-Verlag, New York, 1977.
- [45] M. LOÈVE : "*Probability Theory II*", fourth edition, Graduate Texts in Mathematics, Vol. 46. Springer-Verlag, New York, 1978

- [46] X. MA, N. ZABARAS, " *An adaptive hierarchical sparse grid collocation method for the solution of stochastic differential equations*, J. Comput. Phys. 228 (2009) 3084-3113.
- [47] REMIGIUS MIKULEVIČIUS AND ECKHARD PLATEN. " Rate of convergence of the Euler approximation for diffusion processes". Math. Nachr., 151 :233-239, 1991
- [48] MEILIN LIU , ZHEN GAOB, JAN S. HESTHAVEN " *Adaptive sparse grid algorithms with applications to electromagnetic scattering under uncertainty*" Applied Numerical Mathematics 61 (2011) 24-7
- [49] F. NOBILE, R. TEMPONE, AND C.WEBSTER : " *A sparse grid stochastic collocation method for elliptic partial differential equations with random input data*", SIAM Journal on Numerical Analysis. Volume 46 Issue 5, May 2008, pp. 2309–2345.
- [50] F. NOBILE, R. TEMPONE, AND C.WEBSTER : " *An anisotropic sparse grid stochastic collocation method for elliptic partial differential equations with random input data*", SIAM J. Numer. Anal., 2008, vol. 46/5, pp. 2411–2442.
- [51] J. T. ODEN AND S. PRUDHOMME. " *Estimation of modeling error in computational mechanics*". Journal of Computational Physics, 182 :496515, 2002
- [52] PETER E. KLOEDEN ECKHARD PLATEN. " *Numerical Solution of Stochastic Differential Equations*" Series : Stochastic Modelling and Applied Probability, Vol. 23, Springer 1992, XXXVI, 636 p.
- [53] SALANDIN P. BELLIN, A. AND RINALDO A. " *Simulation of dispersion in heterogeneous porous formations : Statistics, first-order theories, convergence of computations*". Water Resour. Res., 28 :2211-2227, 1992
- [54] A. SCHREIBER. " *Smolyak's method for multivariate interpolation*". PhD thesis, Georg-August-Universität Göttingen, Germany, 2000
- [55] G. SCHAFFER. " *A Mathematical Theory of Evidence*". Princeton University Press, Princeton, 1976
- [56] CH. SCHWAB, R.-A. TODOR, " *Karhunen-Loève approximation of random fields by generalized fast multipole methods*", J. Comput. Phys. 217 (1) (2006) 100-122.
- [57] H. SCHWARZE, U. JAEKEL AND H. VEREECKEN. " *Estimation of Macrodispersion by Different Approximation Methods for Flow and Transport in Randomly Heterogeneous Media*. Transport in Porous Media 43 : 265287, 2001
- [58] S. SMOLYAK : " *Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions*, Doklady Akademii Nauk SSSR, Volume 4, 1963, pages 240-243.
- [59] I.M. SOBOL. " *On quasi-monte carlo integrations*". Math. Comput. Simulat., 47 :103-112, 1998.
- [60] G. SZEGÖ. " *Orthogonal Polynomials*". Amer. Math. Soc., fourth edition, 1975.
- [61] D. TALAY, " *Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution*", RAIRO Modél. Math. Anal. Numér. 20 (1986), no. 1, 141179
- [62] G. W. WASILKOWSKI AND H. WOŹNIAKOWSKI : " *Explicit cost bounds of algorithms for multivariate tensor product problems*", Journal of Complexity, 11 :1-56, 1995.

- [63] R. VERFÜRTH. *"A review of a posteriori error estimation and adaptive mesh refinement techniques"*. Wiley-Teubner, 1996.
- [64] D. ZHANG, Z. LU : *"An efficient, high-order perturbation approach for flow in random porous media via Karhunen-Loève and polynomial expansions"*, J. Comput. Phys. 194 (2004) 773-794.

# Résumé

Ce travail présente un développement et une analyse des approches numériques déterministes et probabilistes efficaces pour les équations aux dérivées partielles avec des coefficients et données aléatoires.

On s'intéresse au problème d'écoulement stationnaire avec des données aléatoires. Une méthode de projection dans le cas unidimensionnel est présentée, permettant de calculer efficacement la moyenne de la solution.

Nous utilisons la méthode de collocation anisotrope des grilles clairsemées. D'abord, un indicateur de l'erreur satisfaisant une borne supérieure de l'erreur est introduit, il permet de calculer les poids d'anisotropie de la méthode. Ensuite, nous démontrons une amélioration de l'erreur a priori de la méthode. Elle confirme l'efficacité de la méthode en comparaison avec celle de Monte Carlo et elle sera utilisée pour accélérer la méthode par l'extrapolation de Richardson.

Nous présentons aussi une analyse numérique d'une méthode probabiliste pour quantifier la migration d'un contaminant dans un milieu aléatoire. Nous considérons le problème d'écoulement couplé avec l'équation d'advection-diffusion, où on s'intéresse à la moyenne de l'extension et de la dispersion du soluté. Le modèle d'écoulement est discrétisé par une méthode des éléments finis mixtes, la concentration du soluté est une densité d'une solution d'une équation différentielle stochastique, qui sera discrétisée par un schéma d'Euler. Enfin, nous présentons une formule explicite de la dispersion et des estimations de l'erreur a priori optimales.

**Mots clés :** Quantification des incertitudes, EDP à coefficients aléatoires, Méthode de collocation des grilles clairsemées, Méthode de Monte-Carlo, Équation d'advection-diffusion, Extension et dispersion, marche aléatoire, schéma d'Euler pour les EDS.



## Abstract

This work presents a development and an analysis of an effective deterministic and probabilistic approaches for partial differential equation with random coefficients and data.

We are interesting in the steady flow equation with stochastic input data. A projection method in the one-dimensional case is presented to compute efficiently the average of the solution.

An anisotropic sparse grid collocation method is also used to solve the flow problem. First, we introduce an indicator of the error satisfying an upper bound of the error, it allows us to compute the anisotropy weights of the method. We demonstrate an improvement of the error estimation of the method which confirms the efficiency of the method compared with Monte Carlo and will be used to accelerate this method by the Richardson extrapolation technique.

We also present a numerical analysis of one probabilistic method to quantify the migration of a contaminant in random media. We consider the previous flow problem coupled with the advection-diffusion equation, where we are interested in the computation of the mean extension and the mean dispersion of the solute. The flow model is discretized by a mixed finite elements method and the concentration of the solute is the density of the solution of the stochastic differential equation, this latter will be discretized by an Euler scheme. We also present an explicit formula of the dispersion and optimal a priori error estimates.

**Keywords** : Uncertainty quantification, PDEs with stochastic coefficients, Stochastic collocation method, Anisotropic sparse grids, Monte-Carlo method, Monte-Carlo method, Advection-diffusion equation, Spread and dispersion, Random walk, Euler scheme for SDE.