



**HAL**  
open science

# Analyse de structures à dimension stochastique élevée : application aux toitures bois sous sollicitation sismique

Hassen Riahi

## ► To cite this version:

Hassen Riahi. Analyse de structures à dimension stochastique élevée: application aux toitures bois sous sollicitation sismique. Architecture, aménagement de l'espace. Université Blaise Pascal - Clermont-Ferrand II, 2013. Français. NNT: 2013CLF22348 . tel-00881187

**HAL Id: tel-00881187**

**<https://theses.hal.science/tel-00881187>**

Submitted on 7 Nov 2013

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

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

---

# UNIVERSITE BLAISE PASCAL – CLERMONT II

---

*ECOLE DOCTORALE*

*SCIENCES POUR L'INGENIEUR DE CLERMONT-FERRAND*

## Thèse

Présentée et soutenue publiquement par

**Hassen RIAHI**

**Docteur en Génie Mécanique de l'ENIM**

le 8 avril 2013

pour obtenir le grade de

**DOCTEUR D'UNIVERSITÉ**

**SPÉCIALITÉ : GÉNIE CIVIL**

**Analyse de structures à dimension stochastique  
élevée : application aux toitures bois sous  
sollicitation sismique**

**-Jury-**

M. Franck SCHOEFS  
M. Frédéric DUPRAT  
M. Laurent DAUDEVILLE  
M<sup>me</sup> . Carole FAYE  
M. Julien BAROTH  
M. Eric FOURNELY  
M. Philippe BRESSOLETTE  
M. Alaa CHATEAUNEUF

Professeur à l'Université de Nantes  
Maître de conférences-HDR à l'INSA de Toulouse  
Professeur à l'Université Joseph Fourier  
Docteur Ingénieur au FCBA  
Maître de conférences à l'Université Joseph Fourier  
Maître de conférences à l'Université Blaise Pascal  
Maître de conférences à l'Université Blaise Pascal  
Professeur à l'Université Blaise Pascal

Rapporteur  
Rapporteur  
Examineur  
Examinatrice  
Invité  
Invité  
Co-directeur de thèse  
Directeur de thèse



# Remerciements

Les travaux de recherche présentés dans ce mémoire ont été réalisés à l'Institut Pascal de l'Université Blaise Pascal-Clermont II.

Ce mémoire est le fruit d'un travail d'équipe dirigé par Monsieur Alaa CHATEAUNEUF, Professeur à Polytech'Clermont-Ferrand. Son aide, de diverses façons, fut pour moi déterminante et inestimable. Ses qualités professionnelles et humaines m'ont profondément marqués. Qu'il veuille bien trouver ici l'expression de ma très sincère gratitude.

Ma profonde reconnaissance va également à Monsieur Philippe BRESSOLETTE, Maître de Conférences à Polytech'Clermont-Ferrand, qui m'a suivi de près tout au long de ce travail et qui a été pour moi un modèle de rigueur, de disponibilité et de persévérance. Qu'il trouve ici l'expression de ma sincère gratitude.

Je tiens à remercier Messieurs Franck SCHOEFS et Frédéric DUPRAT qui ont accepté de rapporter ce travail. Leurs appréciations positives constituent pour moi un tremplin pour la suite de mes travaux de recherche.

Je remercie également Messieurs Laurent DAUDEVILLE, Eric FOURNELY et Julien BAROTH, ainsi que Madame Carole FAYE qui ont accepté d'examiner ce travail et de faire partie du jury.

Mes pensées vont également vers tous les doctorants et stagiaires du département génie civil de Polytech'Clermont-Ferrand avec lesquels j'ai passé de bons moments.

Cette thèse a été financée dans le cadre du projet ANR SISBAT. Je tiens ainsi à remercier ce soutien financier qui m'a permis d'avancer dans les meilleures conditions.

Pour finir, je remercie tout mon entourage, en particulier ma famille, qui malgré la distance, était toujours présente dans les moments difficiles.

# Résumé

Le problème de la dimension stochastique élevée est récurrent dans les analyses probabilistes des structures. Il correspond à l'augmentation exponentielle du nombre d'évaluations du modèle mécanique lorsque le nombre de paramètres incertains est élevé. Afin de pallier cette difficulté, nous avons proposé dans cette thèse, une approche à deux étapes. La première consiste à déterminer la dimension stochastique efficace, en se basant sur une hiérarchisation des paramètres incertains en utilisant les méthodes de criblage. Une fois les paramètres prépondérants sur la variabilité de la réponse du modèle identifiés, ils sont modélisés par des variables aléatoires et le reste des paramètres est fixé à leurs valeurs moyennes respectives, dans le calcul stochastique proprement dit. Cette tâche fut la deuxième étape de l'approche proposée, dans laquelle la méthode de décomposition de la dimension est utilisée pour caractériser l'aléa de la réponse du modèle, par l'estimation des moments statistiques et la construction de la densité de probabilité. Cette approche permet d'économiser jusqu'à 90% du temps de calcul demandé par les méthodes de calcul stochastique classiques. Elle est ensuite utilisée dans l'évaluation de l'intégrité d'une toiture à ossature bois d'une habitation individuelle installée sur un site d'aléa sismique fort. Dans ce contexte, l'analyse du comportement de la structure est basée sur un modèle éléments finis, dans lequel les assemblages en bois sont modélisés par une loi anisotrope avec hystérésis et l'action sismique est représentée par huit accélérogrammes naturels fournis par le BRGM. Ces accélérogrammes permettent de représenter différents types de sols selon en se référant à la classification de l'Eurocode 8. La défaillance de la toiture est définie par l'atteinte de l'endommagement, enregistré dans les assemblages situés sur les éléments de contreventement et les éléments d'anti-flambement, d'un niveau critique fixé à l'aide des résultats des essais. Des analyses déterministes du modèle du modèle éléments finis ont montré que la toiture résiste à l'aléa sismique de la ville du Moule en Guadeloupe. Les analyses probabilistes ont montré que parmi les 134 variables aléatoires représentant l'aléa dans le comportement non linéaire des assemblages, 15 seulement contribuent effectivement à la variabilité de la réponse mécanique ce qui a permis de réduire la dimension stochastique dans le calcul des moments statistiques. En s'appuyant sur les estimations de la moyenne et de l'écart-type on a montré que la variabilité de l'endommagement dans les assemblages situés dans les éléments de contreventement est plus importante que celle de l'endommagement sur les assemblages situés sur les éléments d'anti-flambement. De plus, elle est plus significative pour les signaux les plus nocifs sur la structure.

# Abstract

The problem of the curse of dimensionality is frequently encountered in practical applications. It can be defined as the significant increase of the number of mechanical model calls with the number of uncertain parameters. To overcome this difficulty, a two-steps stochastic approach has been developed in this work. The first step of this approach consists in calculating the stochastic effective dimension by the means of Morris screening method. Once the most significant uncertain parameters on the variability of the mechanical responses are identified, they are modeled as random variables and the remaining parameters are fixed to their respective mean values. This allows us to reduce significantly the stochastic dimension of the problem in the second step of the approach where the decomposition method is used to estimate the statistical characteristics of the mechanical responses. The efficiency and the accuracy of this approach are evaluated through an academic problem dealing with the assessment of the integrity of a three-span five-story frame structure subjected to horizontal loads. We have demonstrate that we can reduce about 90% of the computation time required by the classical stochastic methods. Then, the proposed approach is used to the analysis of the integrity of timber roofs under seismic loading. The behaviour of this structure is described through a finite element model where the timber joints are modeled by anisotropic hysteresis law, and the seismic action is represented by eight real earthquake ground motion records. These accelerograms provided by the French institution involved in geosciences BRGM allow us to take into account different soil types according to the classification provided by the european design code dealing with seismic events Eurocode 8. The failure of timber roofs is reached when the damage levels in the timbers joints localized on the buckling and bracing members reach the critical value. It is shown, through a deterministic analysis, that the structure resists the seismic hazard representing the city of Le Moule in Guadeloupe. The stochastic analysis has shown that, among the 134 random variables representing the uncertainty in the nonlinear behaviour of the timber joints, only 15 have a significant effect on the variability of the structural response, which allow us to reduce the stochastic dimension in the computation of the statistical moments. According to the estimates of the mean and the standard deviation, we have shown that the variability of bracing members damage is greater than the variability of buckling members damage. Moreover, the variability of the bracing members damage is more significant for the earthquake ground motion records having the lowest collapse PGA.



# Table des matières

Remerciements .....	i
Résumé .....	ii
Abstract .....	iii
Introduction générale.....	1
I Chapitre I : méthodes de calcul stochastique .....	7
I.1 Introduction .....	8
I.2 Formulation générale.....	8
I.3 Analyse de tendance centrale .....	10
I.3.1 Présentation du problème.....	10
I.3.2 Méthodes d'estimation des moments statistiques .....	10
I.4 Analyse de fiabilité.....	16
I.4.1 Présentation du problème.....	16
I.4.2 Méthodes d'estimation de la probabilité de défaillance.....	17
I.5 Analyse de sensibilité.....	38
I.5.1 Présentation du problème.....	38
I.5.2 Méthodes d'estimation des indices de sensibilité .....	38
I.5.3 Etude comparative des méthodes d'analyse de sensibilité.....	62
I.6 Conclusion.....	71
II Chapitre II : problèmes à dimension stochastique élevée .....	73
II.1 Introduction .....	74
II.2 Vaincre le problème de dimension stochastique élevée.....	75
II.2.1 Dimension stochastique efficace .....	75
II.2.2 Stratégie de résolution .....	77
II.3 Développement en chaos polynômial .....	79
II.3.1 Chaos polynômial de Weiner .....	79
II.3.2 Chaos polynômial généralisé .....	81
II.3.3 Représentation de la réponse d'un modèle.....	82
II.3.4 Détermination des coefficients du DCP .....	83
II.3.5 Post-traitements basés sur le DCP.....	102
II.4 Méthode de décomposition de la dimension .....	108
II.4.1 Représentation d'une fonction $N$ -dimensionnelle .....	108

## Table des matières

II.4.2	Intégration par réduction de la dimension.....	112
II.4.3	Résolution de problèmes d'analyse de tendance centrale.....	117
II.4.4	Méta-modèle basé sur les polynômes de Lagrange.....	120
II.4.5	Estimation de la fiabilité.....	124
II.4.6	Méthode de décomposition polynômiale de la dimension.....	126
II.5	Validation de la stratégie de résolution.....	133
II.6	Conclusion.....	137
III	Chapitre III : application aux structures bois sous sollicitation sismique.....	140
III.1	Introduction.....	141
III.2	Modèle éléments finis d'un assemblage bois.....	142
III.2.1	Comportement des assemblages bois par plaques métalliques.....	142
III.2.2	Loi anisotrope avec hystérésis.....	143
III.2.3	Analyse du comportement d'un assemblage bois.....	147
III.3	Analyse stochastique d'un assemblage bois.....	148
III.3.1	Etude de sensibilité.....	148
III.3.2	Evaluation de l'aléa de la réponse de l'assemblage.....	151
III.4	Analyse de fiabilité d'une fermette en bois.....	153
III.4.1	Comportement des fermettes en bois.....	153
III.4.2	Modèle éléments finis de la fermette en bois.....	154
III.4.3	Evaluation de l'intégrité de la fermette en bois.....	158
III.5	Evaluation de l'endommagement d'une toiture à ossature bois.....	160
III.5.1	Modèle éléments finis de la toiture.....	160
III.5.2	Critère d'endommagement dans la toiture.....	163
III.5.3	Modélisation de l'action sismique.....	164
III.5.4	Sélection des signaux sismiques.....	168
III.6	Analyses stochastiques d'une toiture à ossature bois.....	170
III.6.1	Stratégie d'évaluation de l'effet de l'aléa.....	170
III.6.2	Stratégie de mise en œuvre du calcul stochastique.....	171
III.6.3	Analyse de l'effet de l'aléa des propriétés du matériau bois.....	172
III.6.4	Analyse de l'effet de l'aléa des propriétés des assemblages.....	174
III.7	Conclusion.....	185
	Conclusion et perspectives.....	189
	Bibliographie.....	195

# Liste des figures

Figure A : Schéma organisationnel du projet SISBAT.....	3
Figure I. 1 : Géométrie de la sphère sous pression interne.....	15
Figure I. 2 : Concepts de domaine de défaillance, de domaine de sûreté et de surface d'état limite.....	17
Figure I. 3 : Illustration du principe des approches FORM et SORM.....	19
Figure I. 4 : Illustration de la technique des tirages d'importance dans l'espace aléatoire standard.....	23
Figure I. 5 : Illustration de la technique des tirages conditionnés dans l'espace aléatoire standard.....	24
Figure I. 6 : Illustration de la technique des tirages adaptatifs asymptotiques dans l'espace standard.....	27
Figure I. 7 : Etapes de la méthode des tirages adaptatifs asymptotiques .....	29
Figure I. 8 : Evolution du coefficient de variation $\delta$ de l'estimation de la probabilité de défaillance en fonction de la probabilité seuil $p_0$ .....	33
Figure I. 9 : Etapes de la méthode Subset simulations.....	34
Figure I. 10 : Evolution de la probabilité de défaillance au cours des étapes de la méthode Subset simulations .....	35
Figure I. 11 : Différents plans d'expériences utilisés par la méthode SRQ .....	37
Figure I. 12 : Evolution de la fonction de transformation définie par l'équation (I.95) et histogramme de la distribution empirique des points $x_i$ .....	46
Figure I. 13 : Evolution de la fonction de transformation définie par l'équation (I.96) et histogramme de la distribution empirique des points $x_i$ .....	46
Figure I. 14 : Evolution de la fonction de transformation définie par l'équation (I.98) et histogramme de la distribution empirique des points $x_i$ .....	47
Figure I. 15 : Distribution des points générés dans le cas bidimensionnel en utilisant (a) équation (I.95), (b), équation (I.96), (c) équation (I.98) et (d) équation (I.99) .....	48
Figure I. 16 : Distribution des points générés dans les cas bidimensionnel en utilisant l'équation (I.99), (a) $N_r = 1$ et (b) $N_r = 2$ .....	49
Figure I. 17 : Illustration de la région recommandée pour le choix des paramètres $\omega_i$ et $N_r$ Saltelli et al. [52] .....	51
Figure I. 18 : Evolution du modèle d'Ishigami en fonction des paramètres $x_1$ , $x_2$ et $x_3$ .....	51
Figure I. 19 : Hiérarchisation des paramètres d'entrée d'un modèle par la méthode de Morris .....	53
Figure I. 20 : Illustration de l'importance du choix de la perturbation sur le recouvrement du domaine de définition des paramètres d'entrée [65] .....	55
Figure I. 21 : Organigramme de la méthode de Morris .....	56

## Liste des figures

Figure I. 22 : Hiérarchisation des différents paramètres du modèle basée sur les mesures (a) $\mu_i$ et $\sigma_i$ , (b) $\mu_i^*$ et $\sigma_i$ pour le modèle de Morris .....	59
Figure I. 23 : Hiérarchisation des groupes de paramètres $G_1 = \{x_1, x_2, x_4\}$ , $G_2 = \{x_3, x_5\}$ , $G_3 = \{x_6, x_7, x_8, x_9, x_{10}\}$ et $G_4 = \{x_{11}, \dots, x_{20}\}$ du modèle basée sur les mesures (a) $\mu_i$ et $\sigma_i$ , (b) $\mu_i^*$ et $\sigma_i$ .....	60
Figure I. 24 : Comparaison entre les différentes estimations des indices de sensibilité de premier ordre et les indices de sensibilité totaux pour la fonction A .....	64
Figure I. 25 : Comparaison entre les différentes estimations des mesures de sensibilité $\mu_i^*$ et $\sigma_i^*$ pour la fonction A.....	65
Figure I. 26 : Comparaison entre les différentes estimations des indices de sensibilité de premier ordre et les indices de sensibilité totaux pour la fonction B .....	66
Figure I. 27 : Comparaison entre les différentes estimations des mesures de sensibilité $\mu_i^*$ et $\sigma_i^*$ pour la fonction B.....	67
Figure I. 28 : Evolutions de l'erreur sur les estimations de $\mu_1^*$ données par la méthode de Morris et la méthode MSDP dans le cas du modèle B.....	68
Figure I. 29 : Comparaison entre les différentes estimations des indices de sensibilité de premier ordre et les indices de sensibilité totaux pour la fonction C .....	68
Figure I. 30 : Comparaison entre les différentes estimations des mesures de sensibilité $\mu_i^*$ et $\sigma_i^*$ pour la fonction C .....	69
Figure I. 31 : Evolutions de l'erreur sur les estimations de $\mu_1^*$ données par la méthode de Morris et la méthode MSDP dans le cas du modèle C.....	69
Figure II. 1 : (a) hiérarchisation des paramètres incertains en se référant à la mesure de sensibilité $G_i^*$ , (b) identification des paramètres les plus importants pour un quantile seuil $p = 95\%$ .....	78
Figure II. 2 : Stratégie de résolution du problème de dimension stochastique élevée .....	79
Figure II. 3 : Illustration de la procédure des tirages en hyper-cube latin pour la génération de cinq réalisations dans le cas de deux paramètres incertains .....	86
Figure II. 4 : Illustration de l'effet de la réduction de corrélation sur la répartition des réalisations obtenues par tirages en hyper-cube latin .....	87
Figure II. 5 : Répartition des points obtenus par un générateur de nombres pseudo-aléatoires et par tirages en hyper-cube latin dans le cas bidimensionnel .....	88
Figure II. 6 : Répartition des points obtenus par les séquences de Halton et Hammersley dans le cas bidimensionnel .....	90
Figure II. 7 : Grilles de calcul associées respectivement aux points d'intégration de Gauss-Hermite et Gauss-Legendre, utilisées dans les formules de quadrature de Smolyak d'ordre 3, 5 et 6, dans le cas de deux variables aléatoires .....	93
Figure II. 8 : Comparaison dans le cas bidimensionnel entre les grilles basées sur les points d'intégration de Gauss-Hermite utilisées respectivement par une formule de quadrature d'ordre $M = 4$ construite à partir d'un produit tensoriel complet et par une formule de quadrature de Smolyak d'ordre $l = 3$ .....	94

## Liste des figures

Figure II. 9 : <i>Points d'intégration utilisés par les formules de quadrature de Stroud d'ordre 2 et d'ordre 3 dans le cas bidimensionnel</i> .....	97
Figure II. 10 : <i>Comparaison pour différents échantillons de l'évolution de l'erreur <math>\epsilon_{LOF}</math> en fonction du nombre de simulations</i> .....	105
Figure II. 11 : <i>Comparaison de la densité de probabilité <math>f_Y(y)</math> obtenue par les simulations de Monte-Carlo sur le développement en chaos polynômial de la réponse du modèle</i> .....	106
Figure II. 12 : <i>Evolutions du rapport du nombre d'appels au modèle demandé par la méthode d'intégration par réduction de la dimension et par la méthode d'intégration directe en fonction des paramètres <math>s</math> et <math>N</math> : (a) <math>n = 3</math> et (b) <math>n = 4</math></i> .....	118
Figure II. 13 : <i>Comparaison de la densité de probabilité <math>f_Y(y)</math> de la réponse du modèle obtenue par la méthode des moments et les simulations de Monte-Carlo</i> .....	120
Figure II. 14 : <i>Plan d'expérimentation : (a) <math>s = 1, N = 1</math>, (b) <math>s = 1, N = 2</math>, (c) <math>s = 2, N = 2</math></i>	122
Figure II. 15 : <i>Evolution de l'approximation de la réponse du modèle par décomposition de la dimension en fonction des différents paramètres incertains</i> .....	123
Figure II. 16 : <i>Comparaison de la densité de probabilité <math>f_Y(y)</math> de la réponse du modèle obtenue par la méthode indirecte (simulations de Monte-Carlo sur la surface de réponse) et la méthode directe (simulations de Monte-Carlo sur le modèle)</i> .....	124
Figure II. 17 : <i>Comparaison des résultats obtenus par développement en chaos polynômial et par décomposition polynômiale de la dimension pour (a) les indices de Sobol de premier ordre et (b) les indices de Sobol totaux</i> .....	131
Figure II. 18 : <i>Structure à cinq niveaux sollicitée par des charges latérales (a) géométrie (b) déformée</i> .....	133
Figure II. 19 : <i>Hiérarchisation des paramètres incertains pour le problème de la structure à 5 étages, basée sur les caractéristiques statistiques des effets élémentaires, (a) non normalisées, (b) normalisées</i> .....	135
Figure II. 20 : (a) <i>hiérarchisation des paramètres incertains en se référant à la mesure de sensibilité <math>G_i^*</math></i> (b) <i>identification des paramètres les plus importants pour un quantile seuil <math>p = 95\%</math></i> .....	135
Figure II. 21 : <i>Comparaison des densités de probabilité obtenues par la méthode de Pearson et les simulations de Monte-Carlo</i> .....	136
Figure III. 1 : <i>Modes de défaillance des assemblages bois par plaques métalliques embouties : (a) arrachement des dents de la plaque du bois, (b) plastification de la plaque, (c) rupture d'un élément bois [140]</i> .....	143
Figure III. 2 : <i>Loi d'hystérésis unidimensionnelle [139] développée dans le cadre du projet SISBAT</i> .....	144
Figure III. 3 : <i>Définition du déplacement relatif <math>u</math>, de la force résultante <math>F</math> et du plan d'anisotropie <math>(x, y)</math></i> .....	145
Figure III. 4 : <i>Modèle ressort utilisé dans la modélisation du comportement anisotrope des assemblages bois</i> .....	146
Figure III. 5 : <i>Evolution du déplacement au pic <math>d_1</math> en fonction (a) de l'angle de chargement et (b) des déplacements au pic équivalent dans le plan d'anisotropie <math>(x, y)</math></i> .....	146

## Liste des figures

Figure III. 6 : (a) Configuration géométrique de l'assemblage testé, (b) évolution du chargement .....	147
Figure III. 7 : Comparaison entre les prévisions données par le modèle éléments finis et les résultats expérimentaux : la courbe force-déplacement et évolution du cumul de l'énergie dissipée .....	148
Figure III. 8 : Evolution de l'approximation de la capacité résistante de l'assemblage bois par un développement en chaos polynômial en fonction des différents paramètres incertains ...	151
Figure III. 9 : Comparaison des évolutions de (a) la densité de probabilité et (b) la répartition de probabilité .....	153
Figure III. 10 : Accélérogramme enregistré lors du séisme de Kobe au Japon [148] .....	155
Figure III. 11 : Géométrie de la fermette en bois .....	155
Figure III. 12 : Géométrie des assemblages bois utilisés pour (a) les essais d'ancrage et (b) les essais de cisaillements .....	156
Figure III. 13 : Modes de défaillances observés au niveau des assemblages de la fermette en bois [139] .....	156
Figure III. 14 : Evolutions du déplacement global et de l'effort normal dans l'assemblage A1 .....	157
Figure III. 15 : Evolutions du déplacement global et de l'effort normal dans l'assemblage B1 .....	157
Figure III. 16 : Illustration schématique du critère de défaillance de la fermette en bois .....	158
Figure III. 17 : Géométrie et répartition des fermettes de la toiture à combles perdus étudiée dans le cadre du projet SISBAT [7] .....	160
Figure III. 18 : Géométrie de la maquette de toiture à combles perdus testée dans le cadre du projet SISBAT [7] .....	161
Figure III. 19 : Géométrie et dimensions des poutres au vent [7] .....	162
Figure III. 20 : Illustration schématique du critère de défaillance adopté pour l'évaluation de la résistance de la toiture .....	164
Figure III. 21 : Evolution du spectre de réponse élastique selon l'Eurocode 8 .....	165
Figure III. 22 : Spectre de réponse élastique pour différents type de sol selon l'Eurocode 8, (a) zones de sismicité faible et (b) zones de sismicité élevée .....	165
Figure III. 23 : Quatre signaux sismiques fournis par le BRGM respectivement pour une zone d'aléa modéré, une zone aléa moyen, une zone d'aléa fort-proche et une zone d'aléa fort-lointain .....	167
Figure III. 24 : Définition de la durée efficace du signal sismique .....	168
Figure III. 25 : Illustration de la détermination de l'amplification correspondant au critère de défaillance de la structure pour les signaux sismiques 7 et 8 (cf. tableau. III.8) .....	169
Figure III. 26 : Stratégie d'évaluation de l'aléa associé à l'endommagement cumulé dans une toiture à ossature bois sous l'action sismique .....	171
Figure III. 27 : Organigramme de la mise en œuvre de la stratégie du calcul stochastique .	172
Figure III. 28 : Hiérarchisation des paramètres incertains vis-à-vis de l'endommagement dans les assemblages situés dans les éléments d'anti-flambement et de contreventement .....	174

## Liste des figures

Figure III. 29 : Hiérarchisation des paramètres incertains du modèle basée sur les mesures $\mu_i^*$ et $\sigma_i$ pour l'endommagement enregistré au niveau des éléments (a) d'anti-flambement AFA, (b) de contreventement CVS.....	176
Figure III. 30 : Localisation des assemblages les plus importants pour la résistance de la toiture à ossature bois.....	177
Figure III. 31 : Densité de probabilité du deuxième indicateur d'endommagement pour les signaux sismiques de la ville du Moule – scénario lointain.....	180
Figure III. 32 : Evolutions de la densité de probabilité du déplacement maximal de la structure pour les signaux sismiques 1 et 8.....	181
Figure III. 33 : Illustration schématique de la procédure de construction des courbes probabilistes d'évolution de l'endommagement.....	182
Figure III. 34 : Courbes probabilistes d'endommagement associés aux deux critères de défaillance de la structure – signal sismique 1.....	183
Figure III. 35 : Distributions du PGA du signal sismique pour un niveau d'endommagement seuil $D = 0,5$ – signal sismique 1.....	184
Figure III. 36 : Courbes probabilistes d'endommagement associés aux deux critères de défaillance de la structure – signal sismique 8.....	184
Figure III. 37 : Distributions du PGA du signal sismique pour un niveau d'endommagement seuil $D = 0,5$ – signal sismique 8.....	185

# Liste des tableaux

Tableau I. 1 : <i>Caractéristiques statistiques des paramètres incertains du problème de la sphère sous pression</i> .....	16
Tableau I. 2 : <i>Comparaison des estimations des moments statistiques obtenues par les simulations de Monte-Carlo, la méthode de quadrature et la méthode de collocation stochastique</i> .....	16
Tableau I. 3 : <i>Fréquences nécessaires pour le calcul des indices de sensibilité par la méthode FAST</i> .....	45
Tableau I. 4 : <i>Fréquences générées par l’algorithme proposé par Saltelli et al. [52]</i> .....	50
Tableau I. 5 : <i>Indices de sensibilité obtenus par les méthodes FAST et EFAST</i> .....	52
Tableau I. 6 : <i>Statistiques <math>\mu_i^*</math> et <math>\sigma_i</math> des paramètres d’entrée pour le modèle de Morris</i> .....	59
Tableau I. 7 : <i>Indices de sensibilité de premier ordre <math>S_i</math> et totaux <math>S_{T_i}</math> obtenus par la méthode de Sobol-LHS et la méthode EFAST pour la fonction A</i> .....	64
Tableau I. 8 : <i>Mesures de sensibilité <math>\mu_i^*</math> et <math>\sigma_i^*</math> et <math>G_i^*</math> obtenus par la méthode MSDP-MC et la méthode de Morris pour la fonction A</i> .....	65
Tableau II. 1 : <i>Correspondance (schéma d’Askey) entre les distributions de variables aléatoires et les polynômes orthogonaux</i> .....	82
Tableau II. 2 : <i>Nombre de points de la grille utilisée dans la quadrature de Smolyak pour les points d’intégration de Clenshaw-Curtis, Gauss-Pattersson et Gauss-Legendre</i> .....	95
Tableau II. 3 : <i>Comparaison des estimations des moments statistiques de la réponse du modèle défini par l’équation (II.67)</i> .....	105
Tableau II. 4 : <i>Erreur d’intégration associée à la méthode quasi-Monte Carlo et à la méthode de réduction de la dimension pour différentes valeurs de la dimension N de l’intégrale</i> .....	116
Tableau II. 5 : <i>Moments statistiques de la réponse du modèle pour <math>\sigma = 0,4</math></i> .....	119
Tableau II. 6 : <i>Moments statistiques de la réponse du modèle obtenus par simulations de Monte-Carlo sur la surface de réponse construite par la méthode de décomposition de la dimension</i> .....	123
Tableau II. 7 : <i>Résultats d’analyse de fiabilité par les méthodes d’approximation FORM et SORM</i> .....	125
Tableau II. 8 : <i>Estimations de la probabilité de défaillance par couplage de la méthode de décomposition de la dimension et les simulations de Monte-Carlo</i> .....	126
Tableau II. 9 : <i>Caractéristiques statistiques des paramètres incertains du problème de la mécanique de la rupture</i> .....	131
Tableau II. 10 : <i>Caractéristiques statistiques des paramètres incertains du problème de d’oscillateur à deux degrés de liberté</i> .....	132
Tableau II. 11 : <i>Indice de Sobol totaux obtenus par le développement en chaos polynômial et la décomposition polynômiale de la dimension pour le problème de l’oscillateur à deux degrés de liberté</i> .....	133

## Liste des tableaux

Tableau II. 12 : <i>Propriétés géométriques et matérielles des éléments de la structure à cinq niveaux</i> .....	134
Tableau II. 13 : <i>Caractéristiques statistiques des paramètres incertains pour le problème de la structure à cinq niveaux</i> .....	134
Tableau II. 14 : <i>Estimations des moments statistiques – problème de la structure à cinq niveaux</i> .....	136
Tableau III. 1 : <i>Caractéristiques statistiques des paramètres incertains [139]</i> .....	148
Tableau III. 2 : <i>Indices de Sobol de premier ordre <math>S</math> et totaux <math>S_T</math> associés aux paramètres incertains de l'assemblage bois pour un déplacement appliqué inférieur à <math>d_y</math></i> .....	149
Tableau III. 3 : <i>Indices de Sobol de premier ordre <math>S</math> et totaux <math>S_T</math> associés aux paramètres incertains de l'assemblage bois pour un déplacement appliqué supérieur à <math>d_y</math></i> .....	150
Tableau III. 4 : <i>Estimations des moments statistiques de la capacité résistante de l'assemblage</i> .....	152
Tableau III. 5 : <i>Caractéristiques statistiques des paramètres incertains pour l'analyse de fiabilité d'une fermette en bois sous sollicitation sismique-rupture des assemblages [149]</i> . 159	
Tableau III. 6 : <i>Résultats de l'analyse de la fiabilité d'une fermette en bois sollicitée par une action sismique</i> .....	160
Tableau III. 7 : <i>Coordonnées géographiques et caractéristiques sismiques des sites choisis pour l'aléa modéré, moyen et fort [153]</i> .....	167
Tableau III. 8 : <i>Amplifications correspondant à l'atteinte du critère de ruine pour les signaux de la ville du Moule en Guadeloupe – scénario lointain</i> .....	169
Tableau III. 9 : <i>Désignation et localisation des paramètres incertains dans la toiture à ossature bois-étude de l'effet de l'aléa des propriétés du matériau bois</i> .....	173
Tableau III. 10 : <i>Caractéristiques statistiques des paramètres incertains- étude de l'effet de l'aléa des propriétés du matériau bois</i> .....	173
Tableau III. 11 : <i>Affectation de l'aléa dans les assemblages de la toiture à ossature bois</i> ...	175
Tableau III. 12 : <i>Identification des paramètres importants sur la résistance de la toiture</i> ...	176
Tableau III. 13 : <i>Estimations des deux premiers moments statistiques des deux indicateurs d'endommagement de la toiture pour les signaux de la ville du Moule en Guadeloupe – scénario lointain</i> .....	178

---

# **Introduction générale**

---

## Contexte

Le bois est un matériau dont l'utilisation dans la construction des habitations individuelles est abondante dans des pays comme le Canada, les Etats Unis, ou la Nouvelle Zélande. En France, l'utilisation du bois comme matériau de construction est encouragée. En effet, il s'agit d'une ressource naturelle renouvelable. Il ne nécessite pas beaucoup de transformations pour être utilisé, ainsi il est relativement peu coûteux comparé à d'autres matériaux de construction tels que l'acier ou le béton. Cependant, malgré ses différents avantages, d'un point de vue mécanique, le bois est considéré comme un matériau fragile, dont la rupture s'effectue d'une façon soudaine sans signe précurseur. De plus, il ne permet pas de dissiper l'énergie induite par les charges appliquées sur la structure. Par conséquent, de grandes précautions sont à prendre lorsqu'on conçoit des structures bois, en particulier lorsque ces dernières sont installées dans des territoires à sismicité élevée.

Généralement, ces structures sont constituées d'éléments en bois assemblés ou fixés à d'autres éléments structuraux par des organes métalliques qui peuvent être de simples clous, des pointes, des agrafes, des boulons, des équerres ou des plaques métalliques à dents. La plupart du temps, les éléments en bois sont conçus pour résister à des charges axiales. Malheureusement, l'utilisation de connecteurs induit localement des contraintes multiaxiales. Par conséquent, les assemblages constituent le plus souvent les points faibles de la structure. La singularité au niveau des assemblages est principalement due à la présence de différents types de matériaux (i.e. les éléments en bois et les connecteurs métalliques) et au comportement orthotrope du matériau bois. Les structures bois doivent en majeure partie leur ductilité aux connecteurs métalliques, qui sont souvent utilisés dans les fermettes des toitures de maisons individuelles.

D'un point de vue conceptuel, les assemblages bois utilisant des connecteurs métalliques à dents sont constitués de deux ou plusieurs éléments en bois connectés de part et d'autre par deux plaques métalliques à dents de mêmes dimensions et de même orientation. L'expérience a montré, en particulier les essais expérimentaux sous chargements monotones et cycliques réalisés par Gupta et al [1], que ce type d'assemblages permet d'absorber l'énergie induite par les charges appliquées. Ainsi, il est fortement suggéré d'évaluer et de quantifier l'effet de cette dissipation d'énergie dans le cas de structures situées dans un contexte sismique. En effet, cette dissipation d'énergie, sous-estimée par les codes de conception, peut contribuer à l'amélioration de la résistance des structures sollicitées par des actions sismiques, et ainsi répondre au mieux aux exigences des réglementations de constructions parasismiques. De plus, le comportement de ce type d'assemblages est hautement non linéaire en raison de la plastification des connecteurs métalliques et du contact entre les éléments en bois. Par conséquent, l'analyse du comportement de ces structures implique implicitement l'étude de phénomènes complexes tels que la non linéarité, l'orthotropie ou encore la dissipation d'énergie. Cela justifie en grande partie les insuffisances dans la modélisation de la réponse mécanique réelle des structures bois vis-à-vis de conditions exceptionnelles de chargement telles que les événements sismiques. Ce constat est justifié par le nombre de défaillances observées dans les structures bois lors des événements sismiques majeurs qu'a connus l'humanité. En effet, Ayoub [2] rapporte qu'après le séisme de Northridge qui s'est produit le 17 janvier 1994, le service du logement de la ville de Los Angeles a recensé 330000 habitations individuelles à ossatures bois endommagées, ce qui représentent 90% des logements situés dans la zone sinistrée. Parmi ceux-ci, 65000 logements ont été lourdement endommagés, et les pertes matérielles globales ont été estimées à 20 millions de dollars américains. Ce constat était alarmant pour les autorités publiques du fait que, dans l'état de Californie, 99% des habitations individuelles sont des structures à ossatures bois. De plus, sur

## Introduction générale

l'ensemble des Etats-Unis, les structures à ossatures bois représentent de 80% à 90% de l'ensemble des habitations individuelles.

L'analyse de défaillances dans les structures bois a montré que plusieurs modes de défaillances peuvent se manifester voire même coexister. Dans ce contexte, Frühwald-Hansson [3] a effectué récemment des analyses statistiques sur 127 structures à ossatures bois endommagées en Suède. Après avoir réparti les sources de défaillances en dix catégories, elle a montré que 50% des défaillances sont dues à des mauvaises conceptions, et près du quart à des erreurs humaines commises pendant la phase de construction. La qualité du matériau bois et la fabrication n'interviennent ensemble que sur 11% des défaillances totales recensées. En conjonction, elle a montré que les éléments structuraux les plus touchés par l'endommagement sont les éléments bois des fermettes et les connecteurs métalliques assurant les assemblages puisqu'ils représentent respectivement 34% et 23% des sites défaillants dans les structures à ossatures bois étudiées. De plus, parmi les 127 structures à considérer, 34% d'entre elles atteignent la ruine durant les trois premières années de leur utilisation, 48% sont défaillantes après cette durée d'exploitation, et les 19% restant sont déjà défaillantes dans la phase de construction. Cette étude était motivée par le débat lancé dans certains pays nordiques tels que la Suède après les défaillances catastrophiques observées sur les structures à ossatures bois. En effet, dans ces pays, le dimensionnement des structures bois s'effectue selon les recommandations de l'Eurocode 5 [4] mais en adaptant le niveau de fiabilité cible au contexte et exigences du pays. Ainsi, la question se posait de savoir si ce niveau de fiabilité cible était approprié.

Ces difficultés à cerner le comportement réel des structures à ossatures bois vis-à-vis de certaines conditions de chargement particulières telles que les événements sismiques sont la cause d'inquiétudes et d'interrogations sur la compétitivité du matériau bois sur le marché des matériaux de construction. Ces interrogations se posent aussi dans le contexte de la France pour laquelle la réévaluation de la carte du zonage sismique en 2004 a montré que 60% du territoire présente une forte probabilité d'occurrence d'évènements sismiques, tandis qu'il était évalué à 20% dans l'ancienne réglementation française PS92 [6]. De plus, 75% des structures à porteurs verticaux maçonnés, abondamment utilisés en France, possèdent une charpente en bois. Par conséquent, des réponses claires doivent être apportées à ces interrogations, d'autant que le risque de défaillance a de graves conséquences sur les vies humaines et engendre des pertes économiques considérables.

Dans cette perspective, le projet SISBAT [7] a été lancé en 2009. L'objectif était de clarifier certains aspects du comportement des structures à ossatures bois vis-à-vis d'évènements sismiques, afin de concevoir des structures à la fois robustes et économiques. Ce projet à grande échelle fût le fruit d'une collaboration entre plusieurs partenaires de référence dans leurs domaines de compétences respectifs tels que la modélisation du matériau bois, les constructions parasismiques, l'expérimentation sur les matériaux et les structures, ainsi que le calcul stochastique. La figure A donne le schéma organisationnel du projet SISBAT, ainsi que les rôles des différents partenaires. L'équipe chargée du calcul stochastique est représentée par des membres de la thématique de recherche Mécanique Probabiliste des Matériaux et des Structures (MPMS) de l'Institut Pascal (ex. LaMI). Elle intervient principalement, comme le montre la figure A, dans les tâches 5-1, 6 et 11 du projet. Le travail engagé dans la présente thèse s'articule principalement autour de la tâche 6 qui a pour objectif l'intégration des différentes sources d'aléa dans le calcul mécanique d'une toiture à ossature bois et la quantification de leurs effets sur les indicateurs d'endommagement de cette structure en adoptant une stratégie de couplage mécano-probabiliste. Dans ce contexte, les trois finalités du calcul stochastique seront abordées, à savoir l'analyse de tendance centrale, l'analyse de sensibilité basée sur la décomposition de la variance, et l'analyse de fiabilité.

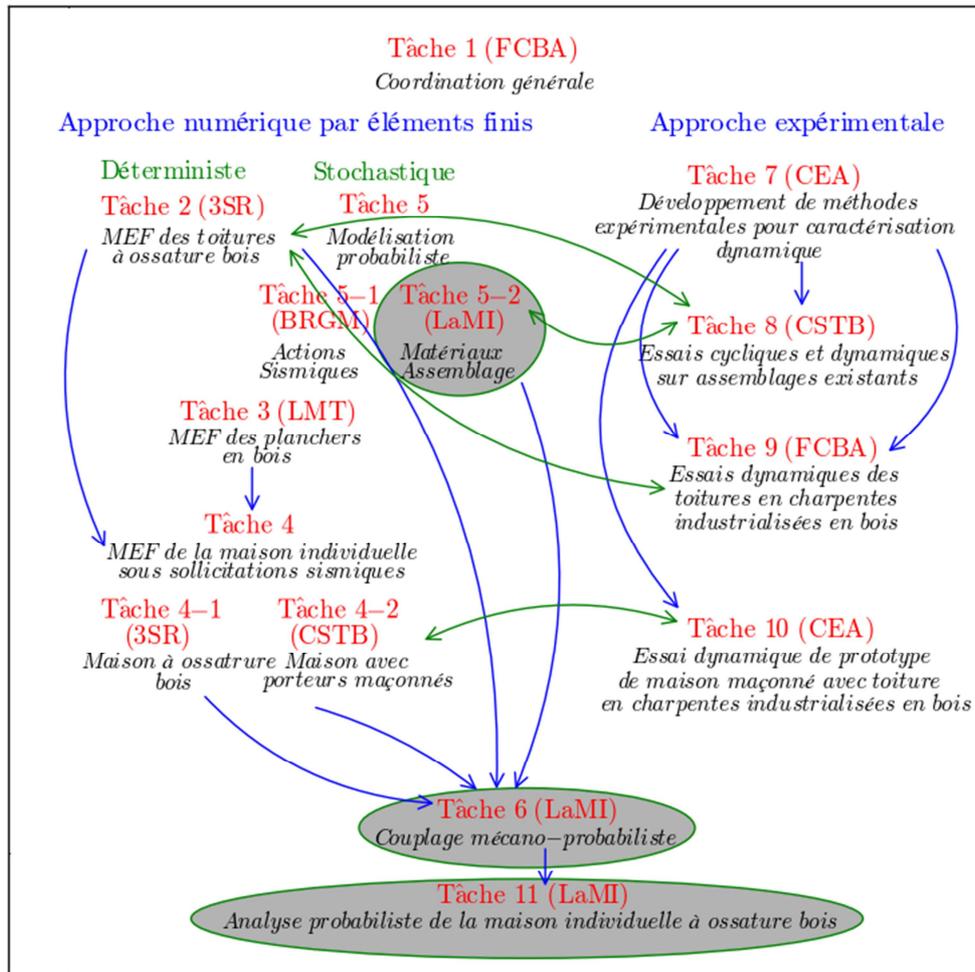


Figure A : Schéma organisationnel du projet SISBAT

## Problématique

Généralement le dimensionnement des structures vis-à-vis d'événements sismiques s'effectue en se référant aux réglementations parasismiques, représentées par l'Eurocode 8 [5] dans le contexte européen. Afin de simplifier la vérification parasismique, ces réglementations préconisent le plus souvent une analyse linéaire élastique équivalente au lieu d'une analyse chronologique complexe. Les effets de dissipation ou de plastification sont alors pris en compte par l'intermédiaire d'un coefficient de comportement. Ce dernier joue le même rôle des coefficients de sécurité rencontrés dans les procédures de dimensionnement classiques dans le contexte d'ingénieries mécanique et civile. Le coefficient de comportement a pour rôle de pondérer le résultat final pour pallier les lacunes d'un calcul simplifié ou notre méconnaissance du modèle réel de certains paramètres. Par conséquent, sa signification physique est difficile à justifier. De plus, dans la plupart, si ce n'est dans toutes, les réglementations parasismiques, la prise en compte de l'aléa qui peut entacher certains paramètres représentant la géométrie de la structure, les propriétés des matériaux ou le chargement, est effectuée d'une manière implicite, soit en fixant ces paramètres incertains à leurs valeurs moyennes (au risque d'omettre certains scénarios de défaillance), soit à des valeurs extrêmes qui peuvent induire un surdimensionnement. Elles s'avèrent non adéquates pour être appliquées pour les structures à ossatures bois installées dans des sites de forte sismicité.

Dans ce contexte, les sources d'incertitudes sont variées. En effet, à cause de son origine naturelle, des défauts sous forme d'irrégularités ou de nœuds sont inhérents au matériau bois, qui introduisent un aléa sur ces propriétés mécaniques. De plus, les enregistrements des mouvements sismiques montrent que l'action sismique est entachée d'une forte dispersion intrinsèque. Cette dispersion est amplifiée si en lui ajoute l'aléa sur les paramètres du sol. Dans ce contexte, Sommerville [8] a montré que la dispersion de l'action sismique enregistrée dans différents sites correspondant à différents types de sols est plus significative que celle de plusieurs enregistrements effectués sur le même site. Cette notion de dispersion a été mise en évidence dans le cadre du projet SISBAT. En effet, les essais expérimentaux ont montré que certains paramètres de la loi avec hystérésis utilisée dans la modélisation du comportement non linéaire des assemblages bois sont entachés d'incertitude. Ainsi, il est nécessaire de prendre en compte ces sources d'incertitude dans la modélisation, d'une part pour quantifier leurs effets sur la capacité résistante de l'assemblage et sa dissipation d'énergie et d'autre part pour identifier les paramètres incertains prépondérants sur la variabilité de ces deux réponses mécaniques. Cela n'est possible que par l'utilisation des approches de calcul stochastique.

D'un point de vue de la formulation, ces approches sont partagées en deux classes : les méthodes intrusives et les méthodes non-intrusives. Les méthodes intrusives exigent la modification de la formulation éléments finis du problème mécanique. En effet, la prise en compte de l'aléa sur le matériau induit une modification de la matrice de rigidité dans l'équation d'équilibre, qui se traduit par un développement en série de cette matrice ainsi que des vecteurs représentant les déplacements nodaux et le chargement extérieur. Cette manipulation de l'équation d'équilibre n'est possible que pour les problèmes linéaires et donc inapplicable dans notre situation. Les méthodes non-intrusives permettent de résoudre ce problème en découplant le calcul stochastique du calcul éléments finis. Cette classe de méthodes de calcul stochastique offre la possibilité d'utiliser les codes éléments finis commerciaux et par conséquent une grande variété de problèmes mécaniques peut être abordée. En effet, les grandeurs (i.e. moments statistiques ou probabilité de défaillance) à appréhender sont déterminées à partir de quelques réalisations du modèle éléments finis déterministe en un ensemble de points de l'espace aléatoire. Cependant, à l'échelle d'une toiture à ossature bois, prendre en compte l'aléa qui peut se manifester dans le comportement des différents assemblages qui la constitue peut induire un nombre exorbitant de variables aléatoires même si on ne garde que les paramètres incertains potentiels à l'échelle d'un assemblage élémentaire. A l'exception des simulations de Monte-Carlo, dont la convergence est lente et s'avèrent ainsi inapplicables lorsque le modèle éléments finis est exigeant en terme de temps de calcul, la convergence de la plupart des méthodes de calcul stochastique est affectée par la dimension stochastique du problème (i.e. nombre de paramètres incertains). Ce problème est connu dans le domaine du calcul stochastique par le problème de dimension stochastique élevée. Heureusement, dans la plupart des problèmes d'ingénierie mécanique et génie civil, on constate que, parmi tous les paramètres incertains pris en compte dans la modélisation, peu d'entre eux contribuent réellement à la variabilité de la réponse du modèle. A première vue, si on arrive à distinguer entre la dimension stochastique nominale du problème et la dimension stochastique efficace, le problème de dimension stochastique élevée peut être facilement résolu en réduisant le nombre de paramètres incertains dans le calcul stochastique proprement dit. Malheureusement, l'identification de la dimension stochastique efficace d'un problème est en soit non triviale.

### **Objectifs du travail**

L'objectif de ce travail est de développer une procédure de calcul stochastique permettant d'analyser, à coût raisonnable de calcul, l'effet des incertitudes qui entachent certains

paramètres entrant dans la modélisation du comportement non linéaire des assemblages bois réalisés par connecteurs métalliques, sur les indicateurs d'endommagement dans une toiture à ossature bois d'une habitation individuelle. Pour ce faire, le cheminement est composé de deux axes : d'une part développer une stratégie pour vaincre le problème de dimension stochastique élevée. Cette stratégie est articulée de deux étapes : la première a pour but d'identifier la dimension stochastique efficace du problème, pour que, dans la deuxième étape, dédiée au calcul stochastique proprement dit, le nombre de paramètres incertains soit réduits en ne retenant que les plus importants sur la variabilité de la réponse du modèle. D'autre part développer une procédure de post-traitement des résultats du calcul stochastique afin de fournir des informations supplémentaires sur l'endommagement accumulé dans la structure. Afin de répondre à ces objectifs, le travail est scindé en trois chapitres.

Après une brève présentation de quelques notions de base sur la théorie des probabilités, le **premier chapitre** du mémoire permet de faire un état des lieux sur les méthodes de calcul stochastique. Dans ce contexte, nous nous intéressons aux trois finalités du calcul stochastique, à savoir l'analyse de tendance centrale, l'analyse de sensibilité et l'analyse de fiabilité. Pour chacune de ces finalités, nous présentons les méthodes existantes en nous focalisant sur la présentation de leurs formulations mathématiques ainsi que leurs avantages et inconvénients en se basant sur des exemples d'illustrations. A la fin de ce chapitre, nous nous focalisons sur la présentation d'une technique de criblage, en insistant sur la consistance des mesures de sensibilités qualitatives qu'elle fournit pour la hiérarchisation des paramètres incertains d'un modèle et sur son efficacité lorsque le nombre de ces paramètres est élevé.

Dans le **deuxième chapitre**, et après avoir présenté la notion de la dimension stochastique efficace, nous développons une stratégie de résolution du problème de dimension stochastique élevée, basée sur la méthode de criblage de Morris. Ensuite, nous présentons deux méthodes de construction de méta-modèles désignées respectivement par développement en chaos polynômial et décomposition de la dimension. Nous nous focalisons d'une part sur la présentation de leurs formulations mathématiques soutenue par des exemples d'illustrations. D'autre part, des alternatives sont proposées pour améliorer l'efficacité du développement en chaos polynômial. Enfin, la stratégie de résolution du problème de dimension stochastique élevée est validée sur l'étude d'une structure à cinq étages sollicitée par des charges horizontales.

Après avoir présenté, analysé et validé la loi anisotrope avec hystérésis en traitant un assemblage bois élémentaire réalisé par plaques métalliques à dents, le **troisième chapitre** présente dans un premier temps une analyse de sensibilité basée sur la décomposition de la variance et le calcul des indices de Sobol pour déterminer la contribution de l'aléa associé à chaque paramètre de la loi de comportement avec hystérésis sur la variabilité de la capacité résistante d'un assemblage bois et sur la variabilité de l'énergie dissipée au cours d'un chargement cyclique. Ensuite, nous présentons une analyse de fiabilité d'une fermette bois sollicitée par une action sismique dont l'intégrité structurale est évaluée par rapport à plusieurs modes de défaillance. Enfin, après avoir présenté d'une façon rapide le modèle éléments finis représentant une toiture à ossature bois d'une habitation individuelle. Nous utilisons, d'une part, la stratégie de résolution du problème de dimension stochastique élevée, développée dans le deuxième chapitre, pour quantifier l'effet de l'aléa associé au comportement des assemblages sur la variabilité des indicateurs d'endommagement de la toiture à ossature bois. D'autre part, nous présentons une méthodologie de post-traitement des résultats du calcul stochastique ayant pour but de concrétiser leur utilisation dans la pratique par les ingénieurs concepteurs et de fournir des informations supplémentaires sur l'endommagement accumulé dans la structure lors d'un évènement sismique.

Le manuscrit se termine par une conclusion générale et une proposition de perspectives.



---

# **Chapitre I : méthodes de calcul stochastique**

---

## I.1 Introduction

Dans le contexte classique des approches déterministes utilisées dans la conception en génie civil et mécanique, les incertitudes observées sur les chargements, les propriétés des matériaux ainsi que les paramètres définissant la géométrie de la structure, sont prises en compte en représentant ces paramètres soit par leurs valeurs extrêmes, soit par leurs valeurs moyennes respectives. De cette manière, les incertitudes ne sont pas introduites de façon explicite dans la simulation du comportement. En plus, pour s'assurer d'être à l'abri de toute mauvaise surprise, on emploie des coefficients de sécurité qui viennent pondérer le résultat d'un dimensionnement. Cette façon de gérer le risque de défaillance n'a aucune consistance physique et elle est le fruit d'une prudence parfois excessive. Ce pessimisme s'exprime en grande partie dans les exigences normatives des codes de conception, en particulier celles qui s'adressent aux enjeux humains liés à un scénario de défaillance tels que les séismes. En effet, l'ingénieur a tendance à concevoir des structures fiables en terme de sécurité, même si leur réalisation engendre un surcoût. Cela réduit l'espace d'innovation chez l'ingénieur et l'éloigne de son rôle principal qui consiste à trouver une conception optimale qui garantit le meilleur compromis entre exigences de sûreté et exigences économiques. Cette optimalité n'est possible que par un traitement rationnel des incertitudes. Dans ce contexte, les méthodes de calcul stochastique ont été proposées pour les applications de génie civil et mécanique dans les années 1970. Ces méthodes consistent, dans un premier temps, à modéliser les sources d'incertitudes par des variables, champs ou processus aléatoires, puis, dans un deuxième temps, à les introduire dans l'analyse du comportement des structures en employant une technique de couplage mécano-probabiliste, et enfin à quantifier l'effet de ces incertitudes sur les réponses du modèle.

D'un point de vue conceptuel, les méthodes de calcul stochastique peuvent être réparties en deux catégories : les méthodes intrusives pour lesquelles on est contraint de manipuler les équations gouvernant le comportement mécanique pour introduire l'aléa associé aux paramètres incertains, et les méthodes non intrusives qui permettent de prendre en compte cet aléa seulement en simulant le modèle mécanique en un certain nombre de points de l'espace aléatoire. Une autre classification des méthodes de calcul stochastique peut être faite en prenant comme critère la finalité visée. Dans ce contexte, on distingue trois catégories : les méthodes d'analyse de tendance centrale dont le but est de déterminer les caractéristiques statistiques (i.e. moments statistiques et densité de probabilité) de la réponse mécanique, les méthodes d'analyse de sensibilité qui permettent de quantifier la part de l'aléa associé à chaque paramètre d'entrée sur la variance de la réponse, et les méthodes fiabilistes qui visent l'évaluation de la probabilité de défaillance vis-à-vis d'un scénario donné. Dans ce chapitre, nous nous basons sur cette dernière classification pour présenter les méthodes de calcul stochastique classiques, en développant pour chaque classe la formulation du problème à résoudre. Une analyse critique basée sur l'évaluation de la robustesse, de l'efficacité et de la précision des différentes méthodes est présentée pour identifier les méthodes susceptibles d'être améliorées. Mais avant cela, nous commençons par rappeler quelques notions de base de la théorie des probabilités nécessaires pour la compréhension des différentes méthodes de calcul stochastique présentées par la suite.

## I.2 Formulation générale

Considérons un modèle quelconque ayant un nombre  $N$  de paramètres incertains groupés dans le vecteur  $\mathbf{x} = (x_1, \dots, x_N)^T$  et dont l'aléa associé est représenté par le vecteur aléatoire  $\mathbf{X} = (X_1, \dots, X_N)^T$ . Mathématiquement, ce modèle peut être représenté par la fonction générique suivante :

$$y = f(\mathbf{x}) \tag{I.1}$$

où  $y$  est la réponse du modèle, obtenue soit par une représentation analytique de la fonction  $f$ , soit par l'intermédiaire d'une procédure numérique (par exemple par une modélisation éléments finis dans le cas de problèmes d'ingénierie). Dans notre travail, nous nous plaçons dans le cas le plus général où la fonction générique  $f$  n'est pas définie d'une façon explicite. Les informations associées à la fonction  $f$  ne sont disponibles que d'une façon ponctuelle, par la simulation d'un modèle explicite (modèle éléments finis) en un certain nombre de points appartenant à son domaine de définition. Soit  $Y = f(\mathbf{X})$  le modèle probabiliste (i.e. modèle à variables aléatoires) associé au modèle défini par l'équation (I.1), où le vecteur  $\mathbf{X}$  est une variable aléatoire  $N$ -dimensionnelle dont les composantes sont supposées indépendantes, et  $Y$  une variable aléatoire scalaire. Notons, à ce niveau, que le fait de représenter la réponse du modèle par une grandeur scalaire ne doit pas être considéré comme une contrainte d'ordre mathématique mais juste comme une façon de simplifier l'écriture des formulations évoquées dans ce qui suit. Ainsi, le traitement du cas où la réponse du modèle est une grandeur vectorielle n'ajoute rien à la complexité du problème à résoudre.

La modélisation des incertitudes et leur prise en compte dans le calcul mécanique implique l'emploi des notions de la théorie des probabilités. Nous nous proposons donc de passer en revue les principales d'entre elles afin de simplifier la compréhension des développements que l'on souhaite effectuer dans ce chapitre. Dans ce contexte, soit  $\Omega$  l'espace des événements aléatoires et  $\Theta$  l'espace des fonctions qui permettent d'associer à un événement  $\omega \in \Omega$  une valeur dans  $\mathbb{R}$ . Cette valeur est couramment désignée dans la théorie des probabilités par la *mesure de probabilité* qui n'est autre que la mesure de la *probabilité d'occurrence* de l'évènement  $\omega$ . Ainsi, une variable aléatoire réelle  $\theta$  est une fonction  $\theta: \omega \in \Omega \mapsto \mathbb{R}$ . Tout au long de ce manuscrit,  $\mathbf{X} = (X_1, \dots, X_N)^T$  désigne une variable aléatoire vectorielle définie dans l'espace des événements aléatoires  $\Omega$ , qui représente l'aléa associé aux paramètres incertains qui peuvent être liés aux conditions de chargement, aux propriétés des matériaux et aux caractéristiques géométriques. Notons ici que le  $i^{\text{ème}}$  composant  $X_i$ ,  $i = 1, \dots, N$ , du vecteur  $\mathbf{X}$  est une variable aléatoire continue définie sur l'espace de probabilité  $\Omega_i \subset \Omega$ , dont la densité de probabilité  $p_{X_i}(x_i)$  est supposée connue. Dans notre travail, nous nous proposons de nous placer dans le cas général, où la modélisation de l'aléa des paramètres incertains  $\mathbf{x} = (x_1, \dots, x_N)$  implique différents types de distributions de probabilité. Autrement dit, les composantes du vecteur aléatoire  $\mathbf{X}$  sont des variables aléatoires de distribution quelconque, et corrélées. Dans ce cas, la formulation du calcul stochastique n'est pas triviale, puisqu'on est souvent contraint à manipuler des équations complexes dont la résolution implique la connaissance de techniques mathématiques élaborées. En guise de simplification, il est souvent conseillé de trouver une formulation standard du problème : l'idée est d'écrire la variable aléatoire  $\mathbf{X}$  en fonction d'une variable aléatoire Gaussienne standard  $\mathbf{U}$  (i.e. les composantes de  $\mathbf{U}$  sont des variables aléatoires Gaussiennes indépendantes de moyenne nulle et d'écart-type unitaire). Cette tâche généralement désignée par le passage de l'espace aléatoire physique (i.e. espace formé par les distributions réelles des paramètres incertains) à l'espace aléatoire standard (i.e. espace formé par des variables aléatoires Gaussiennes standard indépendantes), peut être effectuée au moyen d'une transformations iso-probabiliste  $T$ . Nous pouvons écrire alors :

$$\mathbf{X} = T(\mathbf{U}) \tag{I.2}$$

Par conséquent, l'équation (I.1) s'écrit dans l'espace aléatoire standard sous la forme composée suivant :

$$y = f \circ T(\mathbf{u}) \equiv h(\mathbf{u}) \tag{I.3}$$

où  $h \equiv f \circ T$  est la fonction générique représentant le modèle physique dans l'espace aléatoire standard. Tout au long de ce manuscrit, les différentes formulations seront écrites dans l'espace aléatoire standard.

### I.3 Analyse de tendance centrale

#### I.3.1 Présentation du problème

Comme nous l'avons évoqué au paragraphe précédent, les paramètres d'entrée  $\mathbf{x} = (x_1, \dots, x_N)^T$  du modèle défini par la fonction générique  $f$  sont entachés d'incertitudes représentées par le vecteur aléatoire  $\mathbf{X} = (X_1, \dots, X_N)^T$ . Suite à la propagation de ces incertitudes dans le modèle, sa réponse  $y$  est elle-même entachée d'incertitude. En se référant à l'hypothèse indiquée dans le paragraphe précédent, postulant que la réponse  $y$  est scalaire, il s'ensuit que l'aléa qui lui est associé peut être représenté par une variable aléatoire réelle  $Y$ . Cet aléa est caractérisé par ses moments statistiques ainsi que par sa densité de probabilité  $p_Y(y)$ . Ce problème est généralement connu sous l'appellation de problème de propagation d'incertitude. Par définition, le moment statistique  $m_Y^l$  d'ordre  $l$  de la variable aléatoire  $Y$  est donné par :

$$m_Y^l = \mathbb{E}[Y^l(\mathbf{X})] = \int_{\mathbb{R}^N} f^l(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^N} y^l p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (I.4)$$

où  $p_{\mathbf{X}}(\mathbf{x})$  est la densité de probabilité jointe de la variable aléatoire  $\mathbf{X}$ .

Compte tenue de la standardisation du problème, l'équation (I.4) s'écrit dans l'espace aléatoire standard :

$$m_Y^l = \int_{\mathbb{R}^N} f^l(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^N} (f \circ T)^l(\mathbf{u}) \varphi_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} = \int_{\mathbb{R}^N} h^l(\mathbf{u}) \varphi_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \quad (I.5)$$

où  $\varphi_{\mathbf{U}}(\mathbf{u})$  est la densité de probabilité jointe d'une variable aléatoire Gaussienne standard  $N$ -dimensionnelle.

Il est clair à partir des équations (I.4) et (I.5) que l'évaluation d'un moment statistique revient à résoudre un problème mathématique complexe, défini par le calcul d'une intégrale  $N$ -dimensionnelle. En effet, en pratique, les modèles représentant le phénomène physique à analyser ne sont disponibles que sous forme implicite, ce qui rend impossible le calcul analytique de cette intégrale. Pour pallier ce problème, des méthodes numériques [9] sont souvent employées, telles que le développement en série et les méthodes de simulations. Dans ce qui suit, nous nous restreignons à la présentation des méthodes les plus classiques.

#### I.3.2 Méthodes d'estimation des moments statistiques

##### I.3.2.1 Simulations de Monte-Carlo

Les simulations de Monte-Carlo figurent parmi les méthodes les plus simples pour résoudre un problème de tendance centrale. Leur idée fondatrice consiste dans un premier temps à simuler la réponse du modèle en un certain nombre de points de l'espace aléatoire (i.e. espace défini par l'aléa associé aux paramètres incertains), et, dans un deuxième temps, à effectuer des statistiques (moments statistiques,...) à partir de l'échantillon obtenu. Dans le but de simplifier la présentation de cette approche, nous nous focalisons sur le calcul des deux premiers moments statistiques, à savoir la moyenne  $\mu_Y$  et l'écart-type  $\sigma_Y$ . On considère  $M$  réalisations  $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M\}$  de la variable aléatoire  $N$ -dimensionnelle  $\mathbf{U}$  représentant l'aléa associé aux paramètres incertains dans l'espace aléatoire standard, et générées selon la densité

de probabilité jointe  $\varphi_{\mathbf{U}}(\mathbf{u})$ . Les estimations  $\hat{\mu}_Y$  et  $\hat{\sigma}_Y$  obtenues grâce aux simulations de Monte-Carlo, pour la moyenne et l'écart-type de la variable aléatoire  $Y$ , sont respectivement définies par :

$$\hat{\mu}_Y = \frac{1}{M} \sum_{i=1}^M f \circ T(\mathbf{u}^i) = \frac{1}{M} \sum_{i=1}^M h(\mathbf{u}^i) \quad (I.6)$$

$$\hat{\sigma}_Y = \frac{1}{M-1} \sum_{i=1}^M (f \circ T(\mathbf{u}^i) - \hat{\mu}_Y)^2 = \frac{1}{M-1} \sum_{i=1}^M (h(\mathbf{u}^i) - \hat{\mu}_Y)^2 \quad (I.7)$$

La précision de ces estimations dépend essentiellement de l'ordre  $l$  du moment statistique et de la taille  $M$  de l'échantillon. D'un point de vue statistique, ces estimations sont des variables aléatoires. En se référant au théorème central limite,  $\hat{\mu}_Y$  est l'estimation asymptotique d'une variable aléatoire Gaussienne de moyenne  $\mu_Y$  et d'écart-type  $\sigma_Y/\sqrt{M}$ . Aussi est-il conseillé de fournir un intervalle de confiance pour les estimations issues des simulations de Monte-Carlo. En plus des moments statistiques, l'échantillon  $\{y^1, y^2, \dots, y^M\}$  obtenu par simulations de Monte-Carlo peut être utilisé pour la construction de la densité de probabilité de la variable aléatoire  $Y$ . Dans la pratique, un nombre important de simulations est nécessaire pour obtenir une bonne approximation de la densité de probabilité. De plus, une technique de lissage [10] peut être utilisée pour obtenir une meilleure représentation graphique.

L'avantage des simulations de Monte-Carlo réside dans leur robustesse, puisque la convergence est pratiquement indépendante de la complexité du phénomène physique à analyser (i.e. non linéarité, dynamique,...), ainsi que dans la simplicité de son implémentation et son indépendance de la dimension de l'espace aléatoire (i.e. du nombre de paramètres incertains). Cependant, sa limite majeure réside dans son faible taux de convergence, qui est de l'ordre de  $M^{-1/2}$ . Beaucoup de travaux ont montré que la convergence des simulations de Monte-Carlo peut être améliorée si la répartition des points de simulation dans l'espace aléatoire est plus uniforme. Dans ce contexte, on suggère souvent d'utiliser les tirages dans un hyper-cube latin [11] et les séquences quasi-Monte Carlo [12] pour la génération des points de simulation et ainsi d'améliorer l'efficacité des simulations de Monte-Carlo. Cependant, il est important de noter que, dans le cas d'utilisation des séquences quasi-Monte Carlo, la précision des estimations peut être affectée lorsque la dimension de l'espace aléatoire est élevée.

### I.3.2.2 Méthode de quadrature

La méthode de quadrature a été utilisée par Baldewck [13] pour calculer les moments statistiques de la réponse mécanique d'un système défini implicitement, et dont certains des paramètres d'entrée sont incertains. Son idée fondatrice consiste à calculer les intégrales présentes dans les expressions de ces moments statistiques par quadrature.

- *Cas unidimensionnel*

Afin de simplifier la présentation de la méthode de quadrature, nous nous proposons de traiter le cas unidimensionnel, pour lequel un seul paramètre d'entrée est entaché d'incertitude. Ainsi, le vecteur aléatoire  $\mathbf{X}$  est réduit à une variable aléatoire scalaire  $X$ , et respectivement le vecteur aléatoire  $\mathbf{U}$  est réduit à une variable aléatoire Gaussienne standard scalaire  $U$ . Dans ce cas, l'évaluation du moment statistique  $m_Y^l$  consiste à calculer une intégrale unidimensionnelle. En se référant à la méthode de quadrature associée à la fonction de pondération  $\varphi_U(u)$  (i.e. la densité de probabilité de la variable aléatoire  $U$ ), l'intégrale (I.5)

peut être approximée par une somme pondérée de valeurs de l'intégrande. Ainsi, le moment statistique d'ordre  $l$  de la variable aléatoire  $Y$  est estimé par :

$$m_Y^l = E[h^l(U)] \approx \sum_{k=1}^{n_p} w_k [h(u_k)]^l \quad (I.8)$$

où  $\{w_k\}_{k=1}^{n_p}$  et  $\{u_k\}_{k=1}^{n_p}$  sont respectivement les pondérations et les points d'intégration associés à la densité de probabilité  $\varphi_U$ , et  $n_p$  est l'ordre du schéma de quadrature. On note que puisque  $U$  est une variable aléatoire Gaussienne standard, les points d'intégration correspondent aux points de Gauss-Hermite. D'une manière générale, les pondérations et les points d'intégration associés à d'autres types de densité de probabilité peuvent être facilement calculés ou fournis dans des tables. Baldeweck [13] les a fournis pour les cas des densités de probabilité Gaussienne et lognormal jusqu'à l'ordre 10.

- **Cas multidimensionnel**

Supposons que les paramètres incertains d'entrée d'un système soient décrits dans l'espace aléatoire standard par la variable aléatoire  $N$ -dimensionnelle  $\mathbf{U} = (U_1, \dots, U_N)^T$ , où les composantes  $\{U_i\}_{i=1}^N$  sont des variables aléatoires Gaussiennes standard indépendantes. Dans ce cas, la densité jointe de  $\mathbf{U}$  n'est autre que le produit des densités de probabilité de chacune des composantes. Ainsi, le moment statistique d'ordre  $l$  de la variable aléatoire  $Y = f \circ T(\mathbf{U}) = h(\mathbf{U})$  s'écrit :

$$m_Y^l = E[h^l(\mathbf{U})] = \int_{\mathbb{R}^N} [h(u_1, u_2, \dots, u_N)]^l \varphi_{U_1}(u_1) \dots \varphi_{U_N}(u_N) du_1 \dots du_N \quad (I.9)$$

où  $\varphi_{U_i}$ ,  $i = 1, \dots, N$ , est la densité de probabilité de la variable aléatoire Gaussienne standard  $U_i$  à valeur dans  $\mathbb{R}$  et  $h(u_1, u_2, \dots, u_N)$  est la réponse du modèle au point  $\mathbf{u} = (u_1, \dots, u_N)^T$  (i.e.  $\mathbf{x} = T(\mathbf{u})$ ).

En utilisant la méthode de quadrature associée à la mesure de probabilité Gaussienne, l'intégrale définie ci-dessus est calculée par :

$$m_Y^l \approx \sum_{k_1=1}^{n_p} \sum_{k_2=1}^{n_p} \dots \sum_{k_N=1}^{n_p} w_{k_1} \dots w_{k_N} [h(u_{k_1}, u_{k_2}, \dots, u_{k_N})]^l \quad (I.10)$$

La méthode de quadrature a été utilisée principalement pour calculer les quatre premiers moments statistiques (i.e. la moyenne, l'écart-type, le coefficient d'asymétrie et le coefficient d'aplatissement). A partir de ces moments statistiques, la densité de probabilité peut être obtenue par les approches dites de moments telles que la méthode de Pearson [14] et la méthode de Johnson [15]. On note cependant que ces techniques de construction de la densité de probabilité ne sont fiables que si les estimations des moments statistiques sont précises. D'un point de vue théorique, un schéma d'intégration par quadrature à  $n_p$  points de Gauss fournit la valeur exacte de l'intégrale d'un polynôme de degré inférieur ou égal à  $2n_p - 1$ . On constate que  $n_p^N$  évaluations (i.e. appels au modèle éléments finis) du système mécanique sont nécessaires pour calculer le moment statistique d'ordre  $l$ . Le nombre d'appels au modèle éléments finis augmente exponentiellement avec le nombre de paramètres incertains. Néanmoins, les termes dont les poids sont faibles par rapport à d'autres peuvent être négligés dans l'évaluation du moment statistique. On note aussi que l'efficacité de la méthode de quadrature peut être améliorée dans le cas où la dimension de l'espace aléatoire est élevée par l'utilisation de schémas de quadrature plus économiques en terme d'appels au modèle, tels

que la quadrature de Stroud [16] ou la quadrature basée sur les grilles de Smolayk [17], qui seront présentées dans le deuxième chapitre de ce manuscrit.

### I.3.2.3 Méthode de collocation stochastique

Initialement, la méthode de collocation stochastique a été proposée pour résoudre des équations différentielles stochastiques [18-19]. Ensuite, elle a été le sujet de plusieurs applications [20-23] portant sur le problème de propagation d'incertitude. La méthode de collocation stochastique peut être classée parmi les méthodes non intrusives. A l'image des simulations de Monte-Carlo, elle ne nécessite que l'exécution du modèle représentant le phénomène physique, en des points particuliers dans l'espace aléatoire. D'un point de vue mathématique, la méthode de collocation stochastique est basée sur des techniques simples connues en analyse numérique : l'interpolation par les polynômes de Lagrange et l'intégration par quadrature. Son principal avantage réside dans la construction d'une surface de réponse stochastique précise, qui peut être ainsi utilisée en conjonction avec les simulations de Monte-Carlo pour effectuer différents types de calcul stochastique (i.e. analyse de tendance centrale, analyse de fiabilité ou analyse de sensibilité). Cependant, nous nous proposons dans ce qui suit de traiter le problème d'analyse de tendance centrale, et nous n'évoquons que le calcul des deux premiers moments statistiques de la réponse  $Y$ . Dans le but d'évaluer l'intégrale (I.5) par la méthode de collocation stochastique, nous devons dans un premier temps construire une approximation analytique  $\tilde{h}(\mathbf{u})$  du modèle  $h(\mathbf{u})$  définie éventuellement sous forme implicite.

Soit  $\Pi_N^p$  une base polynômiale construite à partir des polynômes de Lagrange  $N$ -dimensionnels et de degré  $p$  et  $\{\mathbf{u}_i\}_{i=1}^{n_p}$  un ensemble de points de collocation définis dans l'espace aléatoire standard. La surface de réponse stochastique  $\tilde{h}(\mathbf{u})$  est définie par la projection du modèle  $h(\mathbf{u})$  sur la base des polynômes de Lagrange  $N$ -dimensionnels de degré  $p = n_p + 1$  :

$$\tilde{h}(\mathbf{u}) = \sum_{i_1=1}^{n_p} \dots \sum_{i_N=1}^{n_p} h_{i_1 \dots i_N} [L_{i_1} \otimes \dots \otimes L_{i_N}](\mathbf{u}) = \sum_{i_1=1}^{n_p} \dots \sum_{i_N=1}^{n_p} h_{i_1 \dots i_N} [L_{i_1} \times \dots \times L_{i_N}](\mathbf{u}) \quad (I.11)$$

où  $\forall k \in \{1, \dots, N\}$ ,  $\{L_{i_k}\}_{i_k=1}^{n_p}$  sont les polynômes de Lagrange construits autour des points de collocation associés à la mesure de probabilité Gaussienne,  $h_{i_1 \dots i_N}$  sont des coefficients réels à déterminer,  $\otimes$  désigne le produit tensoriel dans l'espace des polynômes unidimensionnels de degré  $p$  et  $\times$  désigne la multiplication dans  $\mathbb{R}$ .

En se référant aux propriétés des polynômes de Lagrange à savoir que  $L_{i_k}(u_{k,i_l}) = \delta_{i_k i_l}$  (i.e.  $\delta_{i_k i_l}$  est l'opérateur de Kronecker qui vaut 1 si  $i_k = i_l$  et 0 sinon), les coefficients réels  $h_{i_1 \dots i_N}$  inconnus peuvent être obtenus facilement en remplaçant  $\mathbf{u}$  dans l'équation (I.11) par le point de collocation  $\mathbf{u}_i = (u_{1,i_1}, \dots, u_{N,i_N})^T$ . Ainsi, les coefficients réels inconnus  $h_{i_1 \dots i_N}$  sont directement obtenus en simulant le modèle implicite aux points de collocation. Nous avons donc l'égalité suivante :

$$h_{i_1 \dots i_N} = h(u_{1,i_1}, \dots, u_{N,i_N}) \quad (I.12)$$

Le nombre de points de collocation (i.e. le plan d'expérience) doit être choisi au moins égal au nombre des coefficients inconnus. Par conséquent, l'approximation  $\tilde{h}(\mathbf{u})$  est construite par le calcul de tous les coefficients  $h_{i_1 \dots i_N}$  pour  $1 \leq i_1, \dots, i_N \leq n_p$ , ce qui nécessite  $n_p^N$  appels au modèle implicite. On note que les points de collocation sont choisis pour garantir une bonne précision dans l'estimation des moments statistiques de la réponse  $Y$ . Dans le but de

clarifier cette idée et de simplifier la présentation du calcul des moments statistiques par la méthode de collocation stochastique, nous nous intéressons dans un premier temps au cas unidimensionnel pour lequel un seul paramètre d'entrée est entaché d'incertitude (i.e.  $N = 1$ ). Dans ce cas, l'approximation polynômiale donnée par la méthode de collocation stochastique est définie par :

$$\tilde{h}(u) = \sum_{i=1}^{n_p} h_i L_i(u) \quad (I.13)$$

En substituant  $h(u)$  dans l'équation (I.5) par l'approximation  $\tilde{h}(u)$  donnée par la méthode de collocation stochastique, la moyenne et l'écart-type de la variable aléatoire  $Y$  sont définis par :

$$\mu_Y \cong \sum_{i=1}^{n_p} h_i \int_{\mathbb{R}} L_i(u) \varphi_U(u) du \quad (I.14)$$

$$\sigma_Y^2 \cong \sum_{i=1}^{n_p} \sum_{j=1}^{n_p} h_i h_j \int_{\mathbb{R}} L_i(u) L_j(u) \varphi_U(u) du - \mu_Y^2 \quad (I.15)$$

Les intégrales dans les équations (I.14) et (I.15) peuvent être facilement calculées en choisissant convenablement les points de collocation  $\{u_i\}_{i=1}^{n_p}$ , de telle manière que la quadrature suivante :

$$I = \int_{\mathbb{R}} L_i(u) \varphi_U(u) du = \sum_{k=1}^{n_p} w_k L_i(u_k) \quad (I.16)$$

soit exacte  $\forall L_i \in \Pi^{2n_p+1}$  (i.e.  $\Pi^{2n_p+1}$  est l'ensemble des polynômes de Lagrange de degré inférieur ou égal à  $2n_p + 1$ ). Les grandeurs  $\{w_i\}_{i=1}^{n_p}$  et  $\{u_i\}_{i=1}^{n_p}$  sont les poids et les points d'intégration associés à la mesure de probabilité Gaussienne puisque l'approximation  $\tilde{h}(u)$  est exprimée dans l'espace aléatoire standard. Ainsi, les estimations de la moyenne et de l'écart-type sont données par :

$$\hat{\mu}_Y \cong \sum_{i=1}^{n_p} h_i \sum_{k=1}^{n_p} w_k L_i(u_k) = \sum_{k=1}^{n_p} w_k h_k \quad (I.17)$$

$$\hat{\sigma}_Y^2 \cong \sum_{i=1}^{n_p} \sum_{j=1}^{n_p} \sum_{k=1}^{n_p} w_k h_i h_j L_i(u_k) L_j(u_k) - \hat{\mu}_Y^2 = \sum_{k=1}^{n_p} w_k h_k^2 - \hat{\mu}_Y^2 \quad (I.18)$$

Dans le cas multidimensionnel (i.e.  $N > 1$ ), la procédure décrite ci-dessus est appliquée pour construire l'approximation polynômiale  $\tilde{h}(\mathbf{u})$ , et les moments statistiques de la variable aléatoire  $Y$  sont estimés par :

$$\hat{\mu}_Y \cong \sum_{k_1=1}^{n_p} \dots \sum_{k_N=1}^{n_p} w_{k_1} \dots w_{k_N} h_{k_1 \dots k_N} \quad (I.19)$$

$$\hat{\sigma}_Y^2 \cong \sum_{k_1=1}^{n_p} \dots \sum_{k_N=1}^{n_p} w_{k_1} \dots w_{k_N} h_{k_1 \dots k_N}^2 - \hat{\mu}_Y^2 \quad (I.20)$$

Une fois les estimations des moments statistiques calculées, la densité de probabilité peut être construite comme nous l'avons vu pour la méthode de quadrature, par l'utilisation d'une méthode de moments. Néanmoins, la densité de probabilité peut être construite d'une façon plus précise en effectuant des simulations de Monte-Carlo sur la surface de réponse  $\tilde{h}(\mathbf{u})$ . Dans ce contexte, le temps de calcul du modèle implicite ne sera pas une contrainte pour la convergence des simulations de Monte-Carlo.

La méthode de collocation a été utilisée pour le traitement du problème de propagation d'incertitude, portant sur des applications telles que le comportement non linéaire des structures [20, 21], la dynamique [22] ou la propagation de fissure par fatigue [23]. Elle est jugée précise, mais son efficacité est affectée dans le cas où la dimension de l'espace aléatoire est élevée. Afin de pallier ce problème, il a été suggéré, dans le cas où une grande précision sur le calcul des moments statistiques n'est pas exigée, d'utiliser la quadrature de Stroud [18], dans l'évaluation des intégrales  $N$ -dimensionnelles. Bien que cette proposition résolve le problème de dimension stochastique élevée, elle fait perdre à la méthode de collocation stochastique l'un de ses avantages majeurs, qui consiste à construire une approximation polynômiale du modèle implicite. Ce problème peut être surmonté en utilisant des schémas de quadrature plus adaptés, tel que celui basé sur les grilles de Smolyak [17].

### I.3.2.4 Exemple d'illustration

Cette application est tirée de la thèse de Baldeweck [13], dans laquelle les grandeurs à appréhender (i.e. les moments statistiques et la densité de probabilité) ont été estimées par la méthode de quadrature. On considère une sphère métallique creuse de rayon interne  $R_0$  et de rayon extérieur  $R_1$ , sollicitée par une pression interne  $P_i$  (figure I.1).

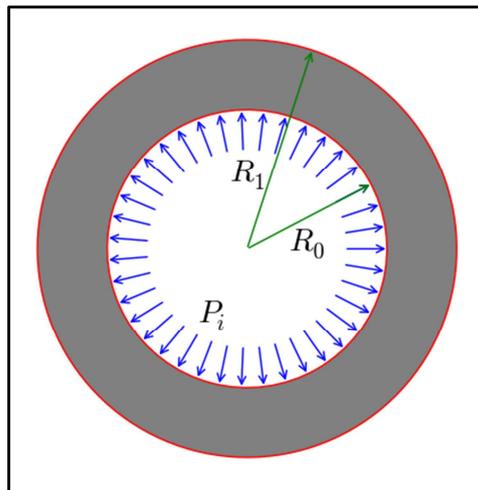


Figure I. 1 : Géométrie de la sphère sous pression interne

On s'intéresse à l'évaluation de la contrainte équivalente de Von-Mises au niveau de la paroi interne (i.e.  $r = R_0$ ) sur laquelle apparaît la première plastification. Etant donnée que le problème est simple, cette contrainte équivalente peut être obtenue analytiquement :

$$\sigma_{eq\ VM}(r = R_0) = \frac{3P_i}{2} \frac{R_1^3}{R_1^3 - R_0^3} \quad (I. 21)$$

Les paramètres incertains du problème sont le rayon interne  $R_0$ , le rayon externe  $R_1$  et la pression appliquée  $P_i$ . Ils sont modélisés par des variables aléatoires non corrélées du type lognormal, dont les caractéristiques statistiques sont données dans le tableau I.1.

**Tableau I. 1 :** *Caractéristiques statistiques des paramètres incertains du problème de la sphère sous pression*

<i>Paramètre</i>	<i>Nature</i>	$\mu$	$\sigma$
$R_0$	lognormal	50 mm	2,5 mm
$R_1$	lognormal	100 mm	5 mm
$P_i$	lognormal	130 MPa	8 MPa

Nous nous proposons d'étudier l'effet de l'aléa associé aux différents paramètres incertains sur la variabilité de la contrainte équivalente de Von-Mises  $\sigma_{eq\ VM}$ . Le tableau I.2 présente les estimations des quatre premiers moments statistiques (i.e. la moyenne, l'écart-type, le coefficient d'asymétrie  $\delta_Y = \frac{1}{\sigma_Y^3}E[(Y-\mu_Y)^3]$  et le coefficient d'aplatissement  $\kappa_Y = \frac{1}{\sigma_Y^4}E[(Y-\mu_Y)^4]$ ) obtenus par les simulations de Monte-Carlo (SMC), par la méthode de quadrature (Q-i) et par la méthode de collocation stochastique basée sur les polynômes de Lagrange (L-i) ;  $i$  désigne le nombre de points d'intégration dans le cas de la méthode de quadrature et le nombre de points de collocation pour la méthode de collocation stochastique.

**Tableau I. 2 :** *Comparaison des estimations des moments statistiques obtenues par les simulations de Monte-Carlo, la méthode de quadrature et la méthode de collocation stochastique*

<i>Méthode</i>	<i>Moments statistiques</i>				<i>Nombre d'appels au modèle</i>
	$\hat{\mu}_Y$	$\hat{\sigma}_Y$	$\hat{\delta}_Y$	$\hat{\kappa}_Y$	
<b>Q - 4</b>	222,506	15,617	0,293	3,205	4 <sup>3</sup>
<b>L - 4</b>	223,197	15,583	0,296	3,256	4 <sup>3</sup>
<b>SMC</b>	223,100	15,580	0,301	3,308	100000

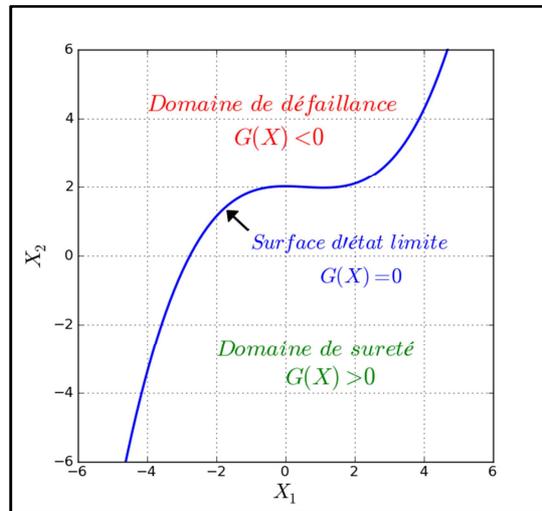
En prenant les résultats donnés par les simulations de Monte-Carlo comme solution de référence, on constate que la méthode de collocation stochastique fournit les meilleures estimations pour les quatre premiers moments statistiques. De plus, sa convergence est rapide puisqu'elle ne nécessite que 64 appels au modèle analytique défini par l'équation (I.21). On note tout de même que le nombre d'appels au modèle associé à la méthode de quadrature et la méthode de collocation stochastique augmente exponentiellement dans le cas où la dimension de l'espace aléatoire est élevée.

## I.4 Analyse de fiabilité

### I.4.1 Présentation du problème

L'objectif principal de l'analyse de fiabilité est de calculer la probabilité de défaillance, notée  $P_f$ , d'un système mécanique vis-à-vis d'un critère de défaillance qui prend en considération les incertitudes qui peuvent se manifester dans la géométrie, les propriétés du matériau et les paramètres du chargement. Le concept de critère de défaillance peut être défini comme la confrontation de deux quantités fondamentales désignées respectivement par la « Capacité » et la « Demande » du système. Dans le contexte de génie mécanique ou civil, la « Demande » est définie comme les effets des actions extérieures qui induisent une réalisation d'un état mécanique tel que l'état de distribution de contraintes ou de déplacements dans une structure. La « Capacité » est représentée par la résistance du matériau, qui définit un état limite qu'il faut respecter, tel qu'une contrainte ultime ou un déplacement critique, pour éviter la défaillance de la structure. Mathématiquement, le critère de défaillance est représenté par une fonction d'état limite  $G$  définie en fonction du vecteur des paramètres incertains  $\mathbf{X} = \{X_1, X_2, \dots, X_n\}^T$ , tel que  $\Omega_S = \{x|G(\mathbf{X}) > 0\}$  est le domaine de sûreté et  $\Omega_F = \{x|G(\mathbf{X}) \leq 0\}$

est le domaine de défaillance. La surface d'état limite est formée par les points vérifiant la condition  $G(\mathbf{X}) = 0$  ; autrement dit, elle définit la frontière délimitant le domaine de défaillance et le domaine de sûreté. La figure I.2 illustre les concepts de domaine de défaillance, de domaine de sûreté et de surface d'état limite dans le cas bidimensionnel.



**Figure I. 2 :** Concepts de domaine de défaillance, de domaine de sûreté et de surface d'état limite

Par définition, la fiabilité est le complémentaire de la probabilité de défaillance  $P_f$  qui est définie par :

$$P_f = \text{Prob}(G(\mathbf{X}) \leq 0) = \int_{G(\mathbf{X}) \leq 0} p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^N} 1_{\Omega_F}(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (I.22)$$

où  $1_{\Omega_F}$  désigne la fonction indicatrice dans  $\Omega_F$  et  $f_{\mathbf{X}}$  est la densité conjointe de probabilité du vecteur aléatoire  $\mathbf{X}$ . Ainsi, l'évaluation de la probabilité de défaillance n'est autre que le calcul d'une intégrale  $N$ -dimensionnelle. Il est évident que la surface d'état limite  $G(\mathbf{X}) = 0$  joue un rôle essentiel dans l'évaluation de cette intégrale. En effet, l'efficacité de la méthode d'intégration dépend de la complexité de la surface d'état limite, mais aussi de la dimension  $N$  de l'espace aléatoire et de la densité  $p_{\mathbf{X}}$  qui est souvent inconnue. En particulier, pour les problèmes pratiques où le nombre de paramètres incertains peut être élevé et la surface d'état limite est disponible sous forme implicite, l'évaluation de la probabilité de défaillance constitue un véritable défi. Ainsi, une grande variété d'approches [24] a été proposée dans la littérature pour la résolution des problèmes de fiabilité. Dans ce qui suit, nous présentons les approches fiabilistes classiques tout en précisant leurs avantages et leurs limites.

## I.4.2 Méthodes d'estimation de la probabilité de défaillance

### I.4.2.1 Simulations de Monte-Carlo

Parmi les approches les plus populaires pour l'estimation de la probabilité de défaillance, on distingue les méthodes de simulations. Ces techniques sont basées sur la traditionnelle méthode de simulations de Monte-Carlo déjà présentée au I.3.2.1 [25]. Dans ce contexte, son principe consiste à interpréter l'intégrale (I.22) comme l'espérance mathématique de la variable aléatoire  $1_{\Omega_F}(\mathbf{X})$  :

$$P_f = \int_{\mathbb{R}^N} 1_{\Omega_F}(\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \mathbb{E}[1_{\Omega_F}(\mathbf{x})] \quad (I.23)$$

En pratique, la méthode des simulations de Monte-Carlo n'est autre qu'une évaluation répétitive de la fonction d'état limite pour un nombre donné  $N_s$  de réalisations du vecteur  $\mathbf{X}$  dans l'espace aléatoire. Une meilleure appréciation de l'étendue du domaine de défaillance, et par conséquent une bonne estimation de la probabilité de défaillance, implique forcément l'utilisation d'un nombre élevé de simulations. Un estimateur non biaisé  $\hat{P}_f$  de la probabilité de défaillance est donné par :

$$\hat{P}_f = \frac{1}{M} \sum_{k=1}^M 1_{\Omega_F}(\mathbf{x}^k) \quad (I.24)$$

Le taux de convergence de la méthode de Monte-Carlo, mesuré au sens des moindres carrés par le coefficient de variation  $\Delta_{MC} = \sqrt{V[\hat{P}_f]}/P_f = \sqrt{(1 - P_f)/M P_f}$  de l'estimateur  $\hat{P}_f$ , est indépendant de la dimension du vecteur aléatoire  $\mathbf{X}$ . La principale limite des simulations de Monte-Carlo réside dans son inefficacité dans l'estimation des probabilités de défaillance d'ordre faible à cause du nombre élevé de tirages  $M$  (i.e. proportionnel à  $1/P_f$ ). Dans le but de pallier cette difficulté, un grand nombre de techniques dites de réduction de la variance ont été développées [26]. Citons la méthode des tirages d'importance utilisant le concept du point de conception, la méthode des tirages adaptatifs et la méthode des tirages directionnels.

Malgré ces efforts dans le développement des méthodes de simulations, le temps de calcul, représenté par les appels au modèle mécanique, peut être rédhibitoire lorsqu'on traite des problèmes pratiques. En effet, pour les structures et installations réelles pour lesquelles les géométries sont complexes et les conditions d'exploitation multiples, le comportement mécanique est rarement décrit d'une façon explicite et il en découle une modélisation par éléments finis conduisant à des calculs coûteux en terme de temps. De plus, les probabilités de défaillance sont très faibles. L'utilisation des méthodes de simulation n'est donc suggérée que pour des problèmes académiques ou lorsqu'il est nécessaire de disposer d'une solution de référence pour valider d'autres types d'approches. Pour ces raisons, des méthodes d'approximations ont été développées lors des trois dernières décennies. Une revue exhaustive de l'ensemble de ces méthodes est fournie par Ditlevsen et Madsen [24]. Dans ce qui suit, nous nous restreignons à la présentation des principaux concepts.

#### I.4.2.2 Méthodes de fiabilité du premier et du second ordre : FORM/SORM

Les méthodes de fiabilité du premier et du second ordre, respectivement FORM et SORM, sont basées sur le concept de l'indice de fiabilité de Hasofer et Lind  $\beta_{HL}$  [27]. Afin d'assurer l'invariance de ce paramètre vis-à-vis d'un éventuel changement de la formulation mathématique de la fonction d'état limite, les auteurs (i.e. Hasofer et Lind) ont suggéré de formuler le problème de fiabilité dans l'espace standard formé par des variables aléatoires Gaussiennes standard et indépendantes. Soit  $T$  la transformation probabiliste permettant le passage de l'espace physique  $\mathbf{X}$  défini par les variables aléatoires de base, à l'espace standard  $\mathbf{U}$ . La fonction d'état limite  $H(\mathbf{u})$  exprimée dans l'espace standard est obtenue par l'égalité :

$$G(\mathbf{X}) = G(T(\mathbf{U})) \equiv H(\mathbf{U}) \quad (I.25)$$

Ainsi, la probabilité de défaillance peut être réécrite sous la forme suivante :

$$P_f = \int_{G(\mathbf{X}) \leq 0} p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{H(\mathbf{U}) \leq 0} \varphi_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \quad (I.26)$$

avec  $\varphi_U$  la densité de probabilité de la variable Gaussienne standard  $N$ -dimensionnelle  $\mathbf{U}$ , donnée par :

$$\varphi_U = \frac{1}{(2\pi)^{n/2}} \exp \left[ -\frac{1}{2} \|\mathbf{u}\|^2 \right] \quad (I.27)$$

où  $\|\cdot\|$  est la norme Euclidienne dans  $\mathbb{R}^n$ . La particularité de la densité de probabilité  $\varphi_U$  est qu'elle possède une symétrie de révolution, centrée à l'origine de l'espace standard et décroît exponentiellement par rapport au carré de la norme  $\|\mathbf{u}\|$ . Par conséquent, les points qui ont une contribution significative dans l'intégrale (I.26), sont les points de la surface d'état limite les plus proches de l'origine. Ce résultat nous amène à la définition de l'indice de fiabilité  $\beta$  comme solution du problème d'optimisation suivant :

$$\begin{cases} \beta = \boldsymbol{\alpha}^T \cdot \mathbf{u}^* \\ \mathbf{u}^* = \operatorname{argmin} \{ \|\mathbf{u}\| \mid H(\mathbf{U}) \leq 0 \} \end{cases} \quad (I.28)$$

où  $\boldsymbol{\alpha} = -\frac{\nabla H(\mathbf{u})}{\|\nabla H(\mathbf{u})\|} \Big|_{\mathbf{u}=\mathbf{u}^*}$  est le vecteur gradient normé de la fonction d'état limite  $H(\mathbf{u})$ , évalué au point  $\mathbf{u}^*$ , et désigné en fiabilité par le vecteur des cosinus directeurs.

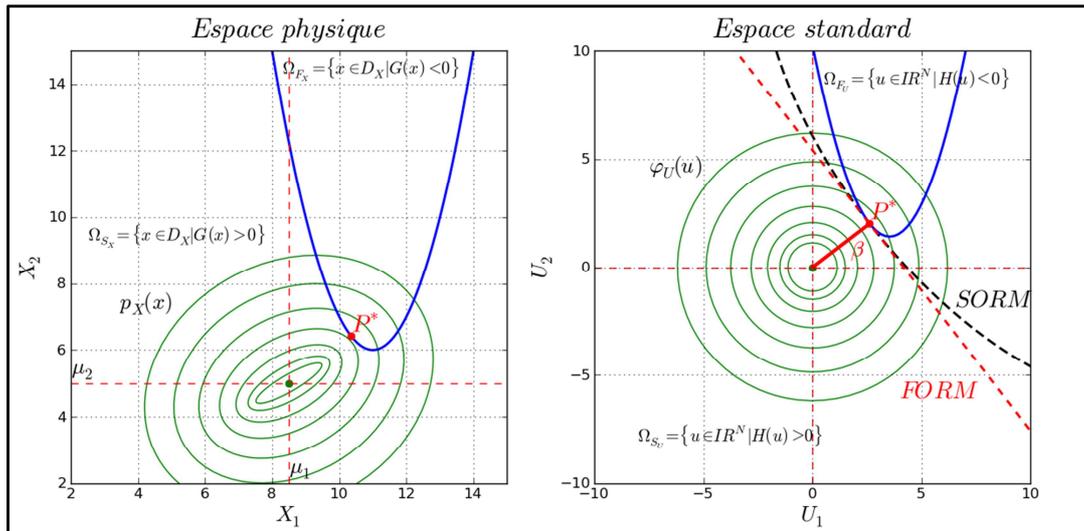
La solution  $\mathbf{u}^*$  du problème d'optimisation avec contrainte (Eq. (I.28)) est appelée point de conception ou encore point de défaillance le plus probable. Il est évident que si la fonction d'état limite  $H(\mathbf{u})$  est linéaire vis-à-vis de  $\mathbf{u}$ , la probabilité de défaillance, définie par l'intégrale (I.26), est obtenue par :

$$P_f = \Phi(-\beta) \quad (I.29)$$

où  $\Phi$  est la fonction de répartition d'une variable Gaussienne standard.

Quand la fonction d'état limite  $H(\mathbf{u})$  est non-linéaire, ce qui est presque toujours le cas dans la pratique, la méthode de fiabilité du premier ordre (FORM) suggère après détermination de l'indice de fiabilité  $\beta$ , d'approximer la fonction d'état limite par un état limite linéaire  $H_L(\mathbf{u})$  défini par :

$$H_L(\mathbf{u}) = \boldsymbol{\alpha}^T \cdot (\mathbf{u} - \mathbf{u}^*) = \beta - \boldsymbol{\alpha}^T \mathbf{u} \quad (I.30)$$



**Figure I.3 :** Illustration du principe des approches FORM et SORM

La figure I.3 illustre le principe des méthodes de fiabilité du premier et du second ordre (FORM/SORM). Du point de vue mathématique,  $H_L(\mathbf{u})$  n'est autre que le développement en série de Taylor du premier ordre. Du point de vue géométrique, ceci est équivalent à

remplacer le domaine de défaillance par un hyperplan tangent à la surface d'état limite au point de conception  $\mathbf{u}^*$ . L'estimation au premier ordre  $P_{f_{FORM}}$  de la probabilité de défaillance  $P_f$  est obtenue par :

$$P_f \approx P_{f_{FORM}} = \Phi(-\beta) \quad (I.31)$$

Généralement, l'approche FORM fournit une bonne estimation pour des valeurs élevées de l'indice de fiabilité  $\beta$ , et elle est exacte si la fonction d'état limite est linéaire dans l'espace standard. Néanmoins, l'approche FORM n'est pas valable quand la surface d'état limite est irrégulière ou présente plusieurs minima locaux. De plus, elle manque de sélectivité du fait qu'elle ne tient pas compte de la courbure du domaine de défaillance au point de conception. En effet, pour deux domaines de défaillance à géométries différentes et à  $\beta$  identiques, la méthode FORM donne une même estimation de la probabilité de défaillance.

Afin d'améliorer la précision, l'approche SORM suggère de remplacer la surface d'état limite par une surface quadratique. Parmi les méthodes proposées dans la littérature, deux types d'approximations sont intéressantes. Le premier type, développé par Breitung [28], est basé sur l'évaluation des courbures principales de la surface d'état limite au point de conception, qui nécessite le calcul de la matrice Hessienne de la fonction d'état limite  $H(\mathbf{u})$ . La deuxième approximation est proposée par Der Kiureghian et al. [29], et suggère d'interpoler la surface d'état limite en un ensemble de points situés autour du point de conception par une semi-paraboloïde. Dans cette étude, nous n'évoquons que la première approche.

Soit  $\{\kappa_i\}_{i=1}^{n-1}$  les courbures principales de la surface d'état limite au point de conception  $\mathbf{u}^*$ , l'approximation SORM de la probabilité de défaillance, fournie par Breitung, s'écrit :

$$P_f \approx P_{f_{SORM,Breitung}} = \Phi(-\beta) \cdot \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - \beta\kappa_i}} = P_{f_{FORM}} \cdot \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - \beta\kappa_i}} \quad (I.32)$$

On note que les courbures principales  $\{\kappa_i\}_{i=1}^{n-1}$  sont prises positives dans le cas où le domaine de défaillance est de forme convexe. Breitung a montré que l'estimation  $P_{f_{SORM,Breitung}}$  tend vers la valeur exacte de la probabilité de défaillance  $P_f$  si l'indice de fiabilité  $\beta$  tend vers l'infini et la quantité  $\beta\kappa_i$  est constante.

Il est clair, d'après l'équation (I.32), que l'approximation SORM consiste en une pondération de l'estimation du premier ordre fournie par l'approche FORM par un facteur de correction qui contient compte des courbures principales.

Pour les petites et moyennes valeurs de  $\beta$  ( $\leq 3$ ), Tvedt [30] a proposé une formule asymptotique meilleure en terme de précision. Cette formule est basée sur un développement de Taylor au second ordre de la fonction d'état limite  $H(\mathbf{u})$ . Elle est désignée souvent par la formule TT (i.e. Three Term) ou formule à trois termes, et est définie par :

$$P_f \approx P_{f_{SORM,Tvedt}} = T_1 + T_2 + T_3 \quad (I.33)$$

avec :

$$T_1 = \Phi(-\beta) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - \beta\kappa_i}}$$

$$T_2 = [\beta\Phi(-\beta) - \varphi(-\beta)] \left\{ \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - \beta\kappa_i}} - \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - (1 + \beta)\kappa_i}} \right\}$$

$$T_3 = (1 + \beta)[\beta\Phi(-\beta) - \varphi(-\beta)] \left\{ \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - \beta\kappa_i}} - \operatorname{Re} \left\{ \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - (j + \beta)\kappa_i}} \right\} \right\}$$

où  $\operatorname{Re}\{\cdot\}$  désigne la partie réelle d'un nombre complexe et  $j = \sqrt{-1}$ . On constate que le premier terme de la formule de Tvedt correspond à l'estimation de la probabilité de défaillance fournie par Breitung. Dans ce même contexte (i.e. approximation du second ordre de la probabilité de défaillance), Tvedt a développé d'autres formules. On peut citer ainsi la formule de l'intégrale unique SI (i.e. Single Integral) et la formule de l'intégrale double DI (i.e. Double Integral) qui sont explicitées dans les travaux de Der Kiureghian et al. [29].

Comme nous l'avons évoqué auparavant, l'idée principale de l'approche de fiabilité du second ordre est d'ajuster à la surface d'état limite réelle une fonction quadratique ayant les mêmes courbures principales au point de conception  $\mathbf{u}^*$ . Ces courbures sont définies par les valeurs propres de la matrice des dérivées secondes de la fonction d'état limite. Cette dernière doit être continue et deux fois dérivable au voisinage du point de conception. Le calcul de la matrice des dérivées secondes est souvent coûteux, particulièrement quand la dimension du vecteur  $\mathbf{X}$  représentant les variables aléatoires de base est élevée et quand l'évaluation de la fonction d'état limite implique l'emploi de procédures numériques compliquées qui introduisent des irrégularités (i.e. bruit numérique) dans la surface d'état limite. Ainsi, les courbures calculées peuvent donner des résultats erronés, sauf si une procédure de lissage est utilisée. De plus, dans certaines situations, les courbures principales peuvent ne pas donner une image réaliste de la forme de la surface d'état limite au voisinage du point de conception  $\mathbf{u}^*$ . A titre d'exemple, quand  $\mathbf{u}^*$  est un point d'inflexion, les courbures principales sont nulles et la surface quadratique d'approximation est réduite à un hyperplan tangent et par conséquent l'estimation de la probabilité de défaillance n'est autre qu'une approximation FORM. Ainsi, l'utilisation des approximations du second ordre est liée au rapport entre précision et temps de calcul. En effet, l'approche SORM est abordable si l'amélioration en précision dans l'estimation de la probabilité de défaillance par rapport à l'approche FORM n'est pas coûteuse en temps de calcul.

### I.4.2.3 Méthode des tirages d'importance

Les méthodes de fiabilité du premier et du second ordre (FORM/SORM) permettent de fournir une estimation de la probabilité de défaillance à faible coût, par comparaison aux simulations de Monte-Carlo. Cependant, elles donnent des résultats erronés dans le cas où la fonction d'état limite est complexe, en particulier si la surface d'état limite présente de fortes courbures ou contient un bruit de forme. Dans ce cas, l'imprécision des méthodes FORM/SORM est due principalement à la difficulté de la recherche du point de défaillance le plus probable et de bien cerner le domaine de défaillance par une approximation du premier ou du second ordre. De plus, aucune estimation de l'erreur ne peut être fournie, contrairement aux simulations de Monte-Carlo pour lesquelles l'estimation de la probabilité de défaillance est fournie en conjonction avec un intervalle de confiance. Malgré leur robustesse et leur indépendance de la dimension de l'espace aléatoire, les simulations de Monte-Carlo perdent de leur efficacité dans le cas où la probabilité de défaillance est faible (i.e.  $P_f < 10^{-3}$ ). Dans ce cas, la seule alternative possible pour améliorer la précision, est d'effectuer des simulations supplémentaires, ce qui conduit à un temps de calcul inabordable, qui s'amplifie lorsque le modèle implicite lui-même est coûteux en terme de temps de simulation. Ainsi, pour pallier l'imprécision des méthodes FORM/SORM dans le cas où la surface d'état limite est complexe, et l'inefficacité des simulations de Monte-Carlo classiques dans le cas où la probabilité de défaillance est faible, les méthodes de simulations basées sur les tirages

d'importance sont proposées. Cette technique compte parmi les méthodes de simulations les plus utilisées dans le contexte d'analyse de fiabilité. En raisonnant dans l'espace standard, l'idée de base consiste à construire une densité de probabilité dite instrumentale ou auxiliaire  $\varphi_U^{TI}(\mathbf{u})$  qui permet de générer des tirages  $\mathbf{u} = (u_1, \dots, u_N)^T$  des paramètres incertains, situés dans la *région d'importance* de l'espace aléatoire, définie par le domaine de défaillance (cf. figure I.4). Par conséquent, la probabilité de défaillance peut être réécrite en fonction de la densité de probabilité d'importance  $\varphi_U^{TI}(\mathbf{u})$  :

$$\begin{aligned} P_f &= \int_{\mathbb{R}^N} 1_{\Omega_F}(T(\mathbf{u})) \varphi_U(\mathbf{u}) d\mathbf{u} = \int_{\mathbb{R}^N} \frac{1_{\Omega_F}(T(\mathbf{u})) \varphi_U(\mathbf{u})}{\varphi_U^{TI}(\mathbf{u})} \varphi_U^{TI}(\mathbf{u}) d\mathbf{u} \\ &= \mathbb{E} \left[ \frac{1_{\Omega_F}(T(\mathbf{u})) \varphi_U(\mathbf{u})}{\varphi_U^{TI}(\mathbf{u})} \right] \end{aligned} \quad (I.34)$$

En utilisant un estimateur semblable à celui adopté pour les simulations de Monte-Carlo classiques, l'estimation de la probabilité de défaillance  $\hat{P}_{f,TI}$  associée aux tirages d'importance est donnée par :

$$\begin{aligned} \hat{P}_{f,TI} &= \frac{1}{M} \sum_{k=1}^M \frac{1_{\Omega_F}(T(\mathbf{u}^k)) \varphi_U(\mathbf{u}^k)}{\varphi_U^{TI}(\mathbf{u}^k)} \end{aligned} \quad (I.35)$$

Dans cette expression,  $\mathbf{u}^k$ ,  $k = 1, \dots, M$ , sont des tirages dans l'espace aléatoire générés selon la densité de probabilité d'importance  $\varphi_U^{TI}$ . Puisque les tirages  $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M\}$  générés dans l'espace standard sont indépendants et identiquement distribués, le coefficient de variation  $\Delta_{TI}$  de l'estimation  $\hat{P}_{f,TI}$  est donné par :

$$\begin{aligned} \Delta_{TI} &= \frac{\sqrt{V[\hat{P}_{f,TI}]}}{P_f} = \frac{\sqrt{\frac{1}{M} V \left[ \frac{1_{\Omega_F}(T(\mathbf{u}^k)) \varphi_U(\mathbf{u}^k)}{\varphi_U^{TI}(\mathbf{u}^k)} \right]}}{P_f} \\ &= \frac{\sqrt{\frac{1}{M} \int_{\mathbb{R}^N} \frac{1_{\Omega_F}(T(\mathbf{u}^k)) [\varphi_U(\mathbf{u}^k)]^2}{[\varphi_U^{TI}(\mathbf{u}^k)]^2} \varphi_U^{TI}(\mathbf{u}^k) d\mathbf{u} - P_f^2}}{P_f} \end{aligned} \quad (I.36)$$

A partir de l'équation ci-dessus, on vérifie que le coefficient de variation  $\Delta_{TI}$  de l'estimation  $\hat{P}_{f,TI}$  est nul, si la densité de probabilité d'importance  $\varphi_U^{TI}(\mathbf{u})$  est définie par :

$$\varphi_U^{TI}(\mathbf{u}^k) = \frac{1_{\Omega_F}(T(\mathbf{u}^k)) \varphi_U(\mathbf{u})}{P_f} \quad (I.37)$$

Ainsi, il est clair que la construction d'une densité de probabilité d'importance  $\varphi_U^{TI}(\mathbf{u})$  optimale est infaisable, puisqu'elle nécessite la connaissance préalable de la valeur de la probabilité de défaillance. Afin de contourner ce problème, plusieurs techniques [26] ont été proposées pour la construction d'une approximation de la densité de probabilité d'importance optimale afin que la variance de l'estimateur de la probabilité de défaillance soit minimum. Dans ce contexte, la manière la plus simple de la construire est de traduire, dans l'espace aléatoire, le centre de la densité jointe  $\varphi_U^{TI}(\mathbf{u})$  au point de défaillance le plus probable  $\mathbf{u}^*$ . Ce

dernier, défini comme le point situé sur la surface d'état limite le plus proche de l'origine, est aussi le point de maximum de vraisemblance. Par conséquent, des informations supplémentaires seront disponibles concernant la surface d'état limite réelle. La densité de probabilité d'importance  $\varphi_U^{TI}(\mathbf{u})$  est définie par :

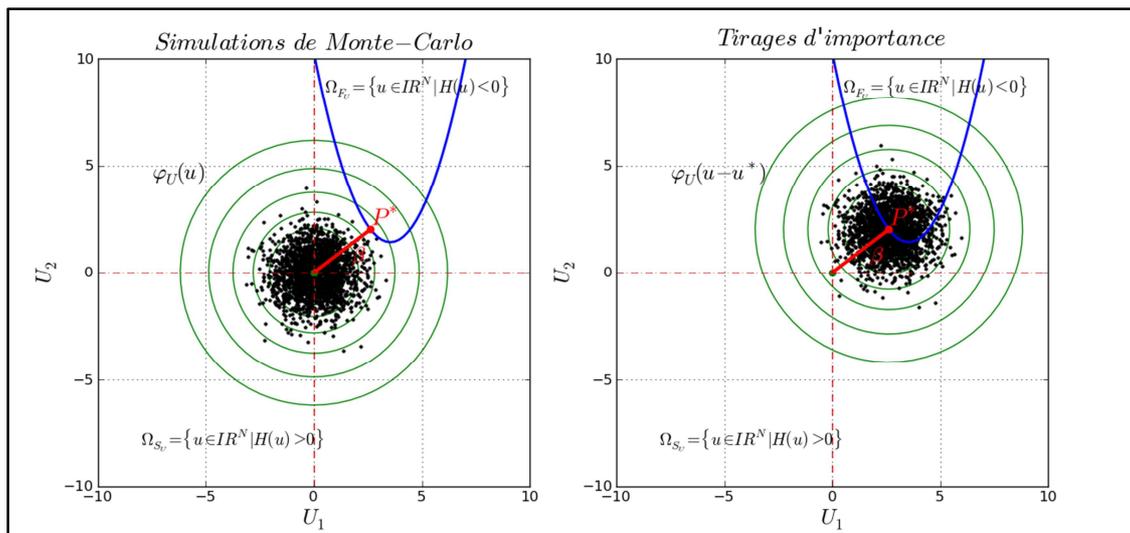
$$\varphi_U^{TI}(\mathbf{u}) = \varphi_U(\mathbf{u}-\mathbf{u}^*) \quad (I.38)$$

En se référant aux simulations de Monte-Carlo, l'estimation de la probabilité de défaillance est donnée par :

$$\hat{P}_{f, TI} = \frac{\exp[\beta^2/2]}{M} \sum_{k=1}^M 1_{\Omega_F}(T(\mathbf{u}^k)) \exp[-\mathbf{u}^k \cdot \mathbf{u}^*] \quad (I.39)$$

où  $\mathbf{u}^k$ ,  $k = 1, \dots, M$ , sont des tirages générés selon la densité de probabilité d'importance  $\varphi_U^{TI}$ .

La figure I.4 illustre le principe de construction de la densité de probabilité d'importance, qui consiste à centrer la densité de probabilité  $\varphi_U$  au point de défaillance le plus probable  $\mathbf{u}^*$ . On constate que cette technique permet une meilleure identification du domaine de défaillance. En effet, un grand pourcentage des points générés selon la densité de probabilité d'importance  $\varphi_U^{TI}(\mathbf{u}) = \varphi_U(\mathbf{u}-\mathbf{u}^*)$  est situé dans le domaine de défaillance, qui représente, du point de vue mathématique, le support de l'intégrale (I.22). Par conséquent, une meilleure estimation de la probabilité de défaillance est possible en se basant seulement sur un échantillon de points de taille réduite par rapport à celui utilisé par les simulations de Monte-Carlo classiques.



**Figure I. 4 :** Illustration de la technique des tirages d'importance dans l'espace aléatoire standard

Les simulations de Monte-Carlo basées sur les tirages d'importances convergent rapidement par rapport aux simulations de Monte-Carlo classiques, et comme pour toute autre méthode de simulations, elles permettent de fournir un intervalle de confiance pour l'estimation de la probabilité de défaillance. De plus, l'estimation de la probabilité de défaillance n'est pas très sensible à la position exacte du point de défaillance le plus probable. Par conséquent, il n'est pas impératif de le déterminer avec une grande précision. On note que, si le domaine de défaillance est convexe, le centre des tirages d'importance doit être déplacé dans la région de défaillance, ou dans le domaine de sûreté si ce dernier est convexe. Comme on l'a déjà évoqué dans le paragraphe précédent, la détermination du point de défaillance le plus probable nécessite la résolution d'un problème d'optimisation sous contrainte. Dans le cas où plusieurs solutions sont possibles, le point de défaillance le plus probable n'est pas unique.

Par conséquent, la densité de probabilité d'importance est représentée par une distribution multimodale construite sur l'ensemble des points de défaillance les plus probables. Leur détermination nécessite l'utilisation d'algorithmes d'optimisation sophistiqués tel que celui développé par Der Kiureghian et Dakessian [31]. D'une manière générale, le temps de calcul nécessaire à la détermination des ou du point de défaillance le plus probable occupe une proportion considérable du temps de calcul total alloué aux simulations de Monte-Carlo basées sur les tirages d'importance, surtout dans les situations où la surface d'état limite est fortement concave, fortement convexe ou présente un bruit de forme. Dans ces cas, l'application de la procédure de tirages d'importance sera probablement inefficace pour l'estimation de la probabilité de défaillance. Bien que la notion de point de défaillance le plus probable soit un bon moyen pour construire la densité de probabilité d'importance, elle peut être, dans les cas où la dimension de l'espace aléatoire est élevée, imprécise pour la caractérisation de la région d'importance. Afin de pallier ce problème, divers techniques ont été proposées ; nous nous restreignons à citer celle développée par Au et Beck [32], basée sur un pré-échantillonnage dans le domaine de défaillance et utilisant les chaînes de Markov. Nous ne la détaillons pas dans ce travail, mais le lecteur est invité à consulter la référence citée.

#### I.4.2.4 Méthode des tirages conditionnés

La notion de tirages conditionnés a été développée par Harbitz [33] dans le but d'améliorer l'efficacité des simulations de Monte-Carlo. L'idée de base est d'exclure les points situés dans une sphère de rayon  $r$  de l'ensemble des points utilisés par les simulations de Monte-Carlo. Etant donné que ces points sont probablement situés dans le domaine de sûreté, ils ne contribuent naturellement pas au calcul de la probabilité de défaillance. Le choix optimal pour le rayon  $r$  de la sphère est qu'il soit égal à l'indice de fiabilité  $\beta$ , qui est par définition la distance minimale entre la surface d'état limite et l'origine de l'espace standard. La figure I.5 illustre la technique des tirages conditionnés dans l'espace standard. Pour deux échantillons de points de tailles égales, on constate que la technique des tirages conditionnés permet de générer plus de points dans le domaine de défaillance et par conséquent de gagner en efficacité par rapport aux simulations de Monte-Carlo classiques.

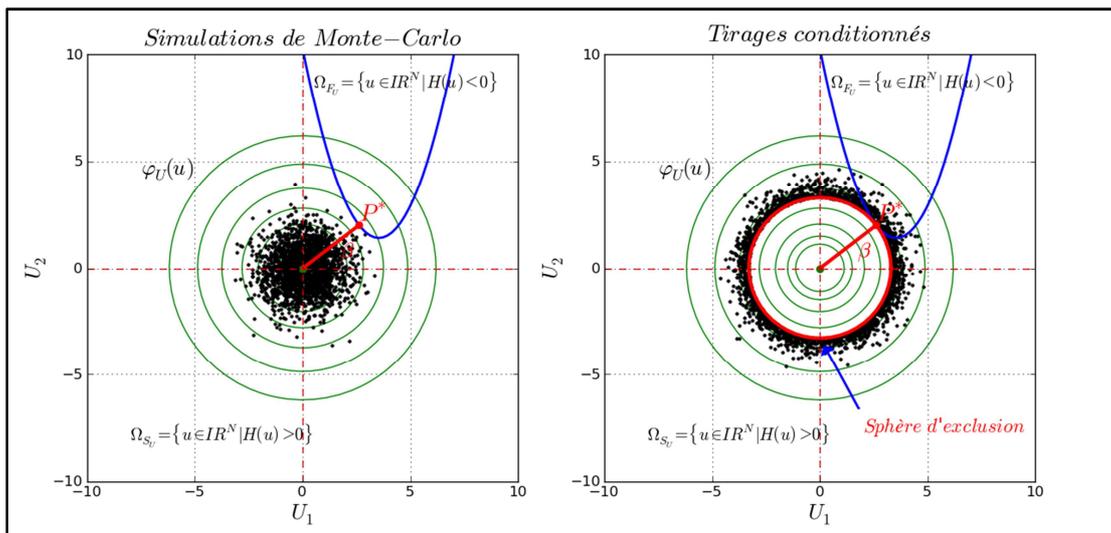


Figure I. 5 : Illustration de la technique des tirages conditionnés dans l'espace aléatoire standard

Le contenu  $P_\beta$  en terme de probabilité des points situés dans la sphère d'exclusion ou de conditionnement de rayon  $\beta$ , peut être facilement calculé en se basant sur la distribution du  $\chi_N^2$  à  $N$  degrés de liberté :

$$P_\beta = \text{Prob}\{\|\mathbf{u}\| \leq \beta\} = \text{Prob}\{\|\mathbf{u}\|^2 \leq \beta^2\} = \chi_N^2(\beta^2) \quad (I.40)$$

La probabilité de défaillance définie par l'équation (I.22) peut être exprimée en terme de la probabilité conditionnelle :

$$P_f = \text{Prob}\{H(\mathbf{u}) \leq 0 \mid \|\mathbf{u}\| > \beta\} \cdot \text{Prob}\{\|\mathbf{u}\| > \beta\} \quad (I.41)$$

Le premier terme  $\text{Prob}\{H(\mathbf{u}) \leq 0 \mid \|\mathbf{u}\| > \beta\}$  est calculé par simulations de Monte-Carlo basées sur les points situés à l'extérieur de la sphère de rayon  $\beta$ . Le terme  $\text{Prob}\{\|\mathbf{u}\| > \beta\}$  est facilement obtenu comme étant le complémentaire de la probabilité  $P_\beta$  définie par l'équation (I.40). Ainsi, l'estimation  $\hat{P}_f$  par les simulations de Monte-Carlo conditionnées est donnée par :

$$\hat{P}_f = \frac{1}{M} \sum_{k=1}^M 1_{\Omega_{F_U} / \|\mathbf{u}\| > \beta}(\mathbf{u}^k) \cdot [1 - \chi_N^2(\beta^2)] \quad (I.42)$$

où  $\Omega_{F_U}$  désigne le domaine de défaillance dans l'espace standard.

Dans le cas des simulations de Monte-Carlo classiques, les points de simulation sont générés dans tout l'espace aléatoire. En excluant les points situés dans la sphère de rayon  $\beta$ , nous gagnons en efficacité dans le calcul des probabilités de défaillance d'ordre faible, puisque le modèle implicite ne sera pas simulé en ces points, tout en gardant la robustesse et la précision. La méthode des simulations conditionnées proposée par Harbitz [33] admet comme hypothèse la connaissance préalable du point de défaillance le plus probable. Normalement, quand on démarre la procédure d'analyse de fiabilité, on n'a aucune information sur la fonction d'état limite. La manière la plus simple pour avoir des informations sur la localisation du point de défaillance le plus probable est d'effectuer une analyse de fiabilité par la méthode FORM. Ainsi, la réussite de la procédure de simulations conditionnées est étroitement liée à la réussite de l'identification du point de défaillance le plus probable par l'algorithme d'optimisation adopté. Cela peut être difficile à réaliser lorsque la surface d'état limite est de forme complexe. Afin de pallier ce problème, Grooteman [34, 35] a proposé une procédure adaptative qui permet de déterminer le rayon  $\beta$  optimal de la sphère d'exclusion.

#### I.4.2.5 Méthode des tirages adaptatifs asymptotiques

La méthode des tirages adaptatifs a été proposée par Bucher [36] comme alternative aux tirages d'importance, dans le but de construire la densité de probabilité d'importance par une procédure adaptative, sans avoir recours à la détermination du point de défaillance le plus probable. Malheureusement, cette technique d'amélioration des simulations de Monte-Carlo perd de son efficacité lorsque la dimension de l'espace aléatoire est élevée. En effet, la procédure d'adaptation de la densité de probabilité d'importance implique l'estimation d'une matrice de covariance de dimension  $N \times N$  dont la construction n'est pas triviale quand le nombre de paramètres incertains est élevé. Afin de pallier ce problème, la méthode des tirages adaptatifs a été améliorée [37] en se basant sur le fait que la probabilité de défaillance présente dans l'espace aléatoire standard un comportement asymptotique [38] en fonction de l'écart-type  $\sigma_U$  (i.e. ce paramètre commande la dispersion, autour de l'origine de l'espace aléatoire standard, des points générés selon la densité de probabilité d'échantillonnage), de la densité de probabilité d'échantillonnage  $\varphi_U$ . L'idée de base de la méthode des tirages

adaptatifs asymptotiques consiste, dans un premier temps, à construire une fonctionnelle  $\beta(\sigma_U)$ , qui décrit l'évolution de l'indice de fiabilité  $\beta$  en fonction de l'écart-type d'échantillonnage  $\sigma_U$ , en partant d'un ensemble d'indices de fiabilité  $\{\beta_i\}_{i=1}^{n_B}$  calculés par simulations de Monte-Carlo basées sur des échantillons obtenus pour différents écarts-types  $\sigma_U$  supérieurs à l'unité. Une fois l'équation de la fonctionnelle  $\beta(\sigma_U)$  déterminée par régression, l'indice de fiabilité cible est obtenu par extrapolation pour  $\sigma_U = 1$ .

Prenons l'exemple d'un cas d'étude académique, pour lequel le scénario de défaillance est représenté par une fonction d'état limite linéaire. En utilisant une transformation appropriée des coordonnées de l'espace aléatoire, ce problème peut être réduit à une fonction d'une seule variable aléatoire. Ainsi, l'indice de fiabilité  $\beta(\sigma_U)$  est donné par :

$$\beta(\sigma_U) = \frac{\beta_1}{\sigma_U} \quad (I.43)$$

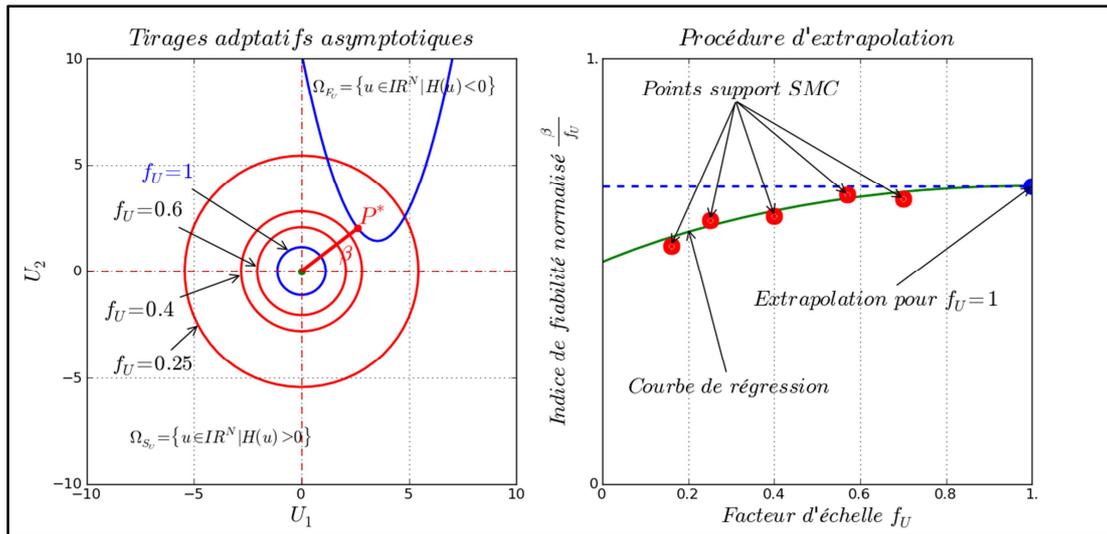
où  $\beta_1$  est l'indice de fiabilité obtenu pour  $\sigma_U = 1$ . En introduisant la variable d'échelle  $f_U = \frac{1}{\sigma_U}$ , l'équation (I.43) peut être réécrite sous la forme linéaire suivante :

$$\beta(f_U) = f_U \cdot \beta_1 \quad (I.44)$$

Afin d'obtenir une estimation précise de l'indice de fiabilité cible  $\beta_1$ , on calcule dans un premier temps l'indice de fiabilité  $\beta(f_U)$  pour des valeurs faibles du paramètre  $f_U$  (réciproquement pour des valeurs élevées de l'écart-type d'échantillonnage  $\sigma_U$ ) en utilisant les simulations de Monte-Carlo classiques. Ensuite, à partir de ces points de référence, nous effectuons une extrapolation pour un facteur d'échelle  $f_U$  égal à l'unité.

Le concept des tirages adaptatifs consiste à déterminer la probabilité de défaillance cible, à partir de probabilités de défaillance dont les estimations sont basées sur des points d'échantillonnage plus dispersés dans l'espace aléatoire standard. En effet, la dispersion des points d'échantillonnage augmente d'une manière adaptative en augmentant l'écart-type  $\sigma_U$  de la densité d'échantillonnage  $\varphi_U$ . Ainsi, la relation entre l'indice de fiabilité  $\beta$  et le facteur d'échelle  $f_U$  tend asymptotiquement vers une relation linéaire, surtout lorsque la dimension  $N$  de l'espace aléatoire est élevée. Cette relation est construite à partir des estimations de l'indice de fiabilité pour  $\sigma_U > 1$  ; elle est ensuite extrapolée pour  $\sigma_U = 1$  afin de déterminer l'indice de fiabilité cible. Dans ce contexte, Bucher [37] a proposé la forme linéaire suivante pour décrire la relation entre l'indice de fiabilité  $\beta$  et le facteur d'échelle  $f_U$  :

$$\beta(f_U) = Af_U + \frac{B}{f_U} \quad (I.45)$$



**Figure I. 6 :** Illustration de la technique des tirages adaptatifs asymptotiques dans l'espace standard

Ce choix est motivé par le fait qu'il garantit un comportement linéaire asymptotique de l'indice de fiabilité  $\beta$  en fonction du facteur d'échelle  $f_U$ , et n'exclut pas l'emploi de modèles alternatifs qui permettent de représenter ce comportement. Les coefficients  $A$  et  $B$  peuvent être obtenus par une procédure de régression basée sur des estimations obtenues par simulations de Monte-Carlo de l'indice de fiabilité  $\beta$  pour différentes valeurs du facteur d'échelle  $f_U < 1$ . La figure I.6 illustre le concept des tirages adaptatifs asymptotiques.

Afin que les points supports utilisés dans la procédure de régression puissent avoir le même poids, l'équation (I.45) est réécrite sous la forme suivante :

$$\frac{\beta(f_U)}{f_U} = A + \frac{B}{f_U^2} \quad (I.46)$$

L'avantage majeur de cette procédure d'estimation de la probabilité de défaillance réside dans l'indépendance vis-à-vis de la dimension  $N$  de l'espace aléatoire. La précision est gouvernée seulement par la relation entre le nombre de points supports utilisés dans la procédure de régression, la précision des probabilités de défaillance intermédiaires associées à ces points supports et la complexité de la forme de la surface d'état limite. L'efficacité de la procédure des tirages adaptatifs asymptotiques peut être améliorée en utilisant des points de simulations générés par hyper-cube latin ou par séquences quasi-Monte Carlo. La fiabilité de la procédure des tirages adaptatifs asymptotiques est étroitement liée au choix de la valeur de départ du paramètre d'échelle  $f_U$  et au nombre de simulations utilisées pour l'estimation des probabilités de défaillance associées aux points supports. Ce couple de paramètres doit être choisi de telle façon qu'à chaque itération de la procédure des tirages d'importance asymptotiques, on obtienne un point support. La validité de chaque itération est évaluée par un critère d'arrêt stipulant que le nombre de points dans le domaine de défaillance intermédiaire actuel soit supérieur ou égal à un nombre de points seuil, fixé préalablement par l'utilisateur. En absence d'informations a priori, la valeur de départ du paramètre d'échelle doit être égale à l'unité. Au cours de la première itération, on s'attend naturellement à ce que le nombre de points dans le domaine de défaillance initial soit limité. Par conséquent, on soustrait de la valeur actuelle du paramètre d'échelle la quantité 0,9 et on répète l'itération jusqu'à ce que le critère de validation soit vérifié. La procédure se poursuit en diminuant la valeur du paramètre d'échelle d'un incrément  $\Delta f_U$ , jusqu'à ce que le nombre de points supports nécessaire dans l'analyse de régression soit atteint.

L'efficacité et la précision des simulations de Monte-Carlo améliorées par les tirages adaptatifs asymptotiques ont été évaluées par Bucher [37] dans un contexte d'analyse de fiabilité impliquant un grand nombre de paramètres incertains. Il a jugé que cette technique est efficace pour l'estimation de la probabilité de défaillance dans la plage  $[10^{-7}, 10^{-3}]$ , et que sa précision est indépendante de la dimension de l'espace aléatoire. Cette technique a été améliorée par Sichani et al. [39], en intégrant une procédure d'optimisation qui permet de déterminer les bornes inférieure et supérieure optimales du facteur d'échelle  $f_U$ . Bien que cette procédure implique des appels supplémentaires au modèle implicite, l'efficacité des simulations de Monte-Carlo basées sur les tirages adaptatifs asymptotiques n'est pas pour autant affectée.

Dans le but d'illustrer l'implémentation de la méthode des tirages adaptatifs asymptotiques, et d'en analyser l'efficacité et la précision, nous considérons une surface d'état limite parabolique définie par :

$$G(\mathbf{x}) = (x_1 - 11)^2 - x_2 + 6 \quad (I.47)$$

où  $x_1$  et  $x_2$  sont deux paramètres incertains représentés par deux variables aléatoires Gaussiennes indépendantes de moyennes respectives  $\mu_1 = 8,5$  et  $\mu_2 = 5$ , et d'écart-type  $\sigma_1 = \sigma_2 = 0,5$ .

Les simulations de Monte-Carlo basées sur un échantillon de  $10^6$  points obtenus par un générateur classique de nombres pseudo-aléatoires, ont donné une estimation de la probabilité de défaillance  $P_f^{MC} = 3,26 \cdot 10^{-4}$ , ce qui correspond à un indice de fiabilité  $\beta = 3,409$ . Ces estimations sont prises comme solution de référence.

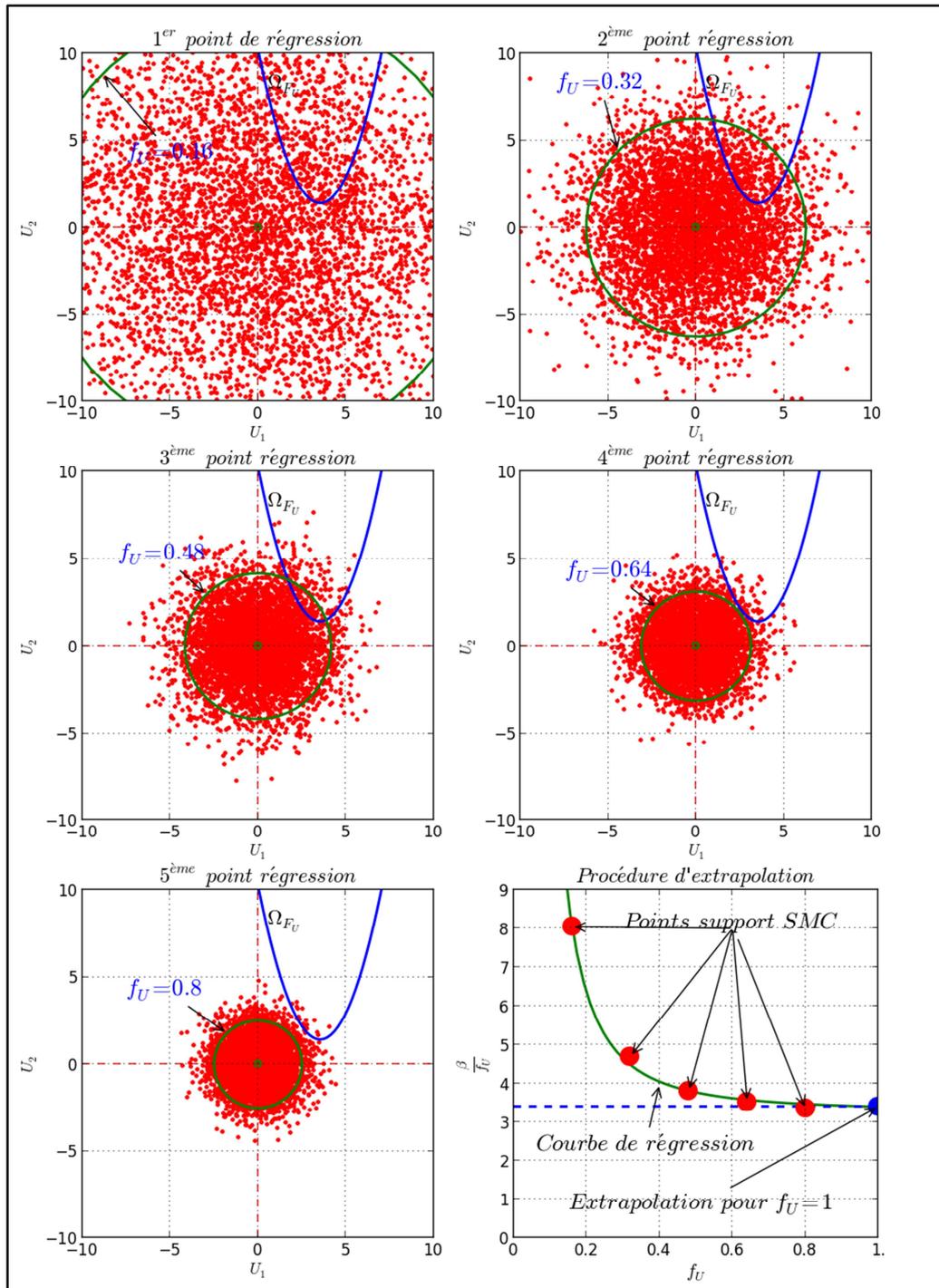


Figure I. 7 : Etapes de la méthode des tirages adaptatifs asymptotiques

Nous avons appliqué la méthode des tirages adaptatifs asymptotiques basée sur l'algorithme décrit auparavant. L'estimation de la probabilité de défaillance est obtenue par une analyse de régression basée sur cinq points représentant les estimations de la probabilité de défaillance obtenues pour différentes valeurs de l'écart-type de la densité de probabilité d'échantillonnage. On note que chacun des points de régression est obtenu par simulations de Monte-Carlo basées sur un échantillon de 5000 points. La figure I.7 illustre les différentes étapes de la méthode ainsi que la procédure de régression. L'estimation de la probabilité de défaillance  $P_f^{TA} = 3,253 \cdot 10^{-4}$ , correspondant à un indice de fiabilité  $\beta = 3,409$ , est en bon accord avec la solution de référence fournie par les simulations de Monte-Carlo classiques,

puisque l'écart entre les deux estimations est de l'ordre de 0,02%. Cette bonne précision est garantie grâce à 25000 appels à la fonction d'état limite. La méthode des tirages adaptatifs asymptotiques permet d'améliorer l'efficacité des simulations de Monte-Carlo puisque, pour une même précision, on économise environ 97% des évaluations de la fonction d'état limite.

#### I.4.2.6 Subset simulations

La méthode de Subset simulations a été proposée par Au et Beck [40] pour l'analyse de fiabilité des structures sollicitées en dynamique. Elle permet de surmonter l'inefficacité des simulations de Monte-Carlo classiques pour l'estimation d'une probabilité de défaillance d'ordre faible, en l'exprimant comme un produit de probabilités conditionnelles d'ordre élevé dont l'évaluation n'est pas coûteuse en temps de calcul. Le concept de base de cette méthode consiste à définir un ensemble d'évènements ou domaines de défaillance intermédiaires  $\{\Omega_{F_i}\}_{i=1}^m$  tel que  $\Omega_{F_m} = \Omega_F$  et  $\Omega_{F_1} \subset \Omega_{F_2} \cdots \subset \Omega_{F_m} = \Omega_F$ . Du fait qu'on peut écrire  $\cap_{i=1}^k \Omega_{F_i} = \Omega_{F_k} \forall k \leq m$ , la probabilité de défaillance cible  $P_f$  (i.e. associée au domaine de défaillance cible) peut être définie par le produit de probabilités conditionnelles :

$$P_f = P[\Omega_{F_m}] = P[\cap_{i=1}^m \Omega_{F_i}] = P[\Omega_{F_1}] \prod_{i=2}^m P[\Omega_{F_i} | \Omega_{F_{i-1}}] \quad (I.48)$$

Afin de simplifier la notation dans la suite, soit  $P_1 = P[\Omega_{F_1}]$  et  $P_i = P[\Omega_{F_i} | \Omega_{F_{i-1}}]$ ,  $i = 2, \dots, m$ . En choisissant convenablement les domaines de défaillance intermédiaires  $\{\Omega_{F_i}\}_{i=1}^m$ , les probabilités  $P_1$  et  $P_i$ ,  $i = 2, \dots, m$ , peuvent être suffisamment élevées (i.e.  $\geq 10^{-3}$ ) pour qu'elles puissent être évaluées efficacement par simulations de Monte-Carlo. Ainsi, le problème initial, qui consiste à estimer une probabilité de défaillance d'ordre faible, est remplacé par un ensemble de  $m$  problèmes intermédiaires, dont la résolution de chacun est non coûteuse en terme de nombre de simulations. Le domaine de défaillance cible  $\Omega_F = \{\mathbf{u} \in \mathbb{R}^N | H(\mathbf{u}) < 0\}$  est défini par les points de l'espace aléatoire vérifiant la condition  $H(\mathbf{u}) < 0$  ( $H(\mathbf{u})$  désigne la fonction d'état limite définie dans l'espace aléatoire standard). En suivant le même raisonnement, les domaines de défaillance intermédiaires  $\{\Omega_{F_i}\}_{i=1}^m$  peuvent être choisis tel que  $\Omega_{F_i} = \{\mathbf{u} \in \mathbb{R}^N | H(\mathbf{u}) < b_i\}$ , où  $b_1 > b_2 > \cdots > b_m = 0$  sont des seuils intermédiaires. Il est clair que la formulation des fonctions d'état limite associées aux domaines de défaillance intermédiaires nécessite la détermination des seuils intermédiaires respectifs  $\{b_i\}_{i=1}^{m-1}$ . En effet, ces paramètres jouent un rôle clé dans l'efficacité de la méthode, puisqu'ils commandent le nombre  $m$  de domaines de défaillance intermédiaires  $\{\Omega_{F_i}\}_{i=1}^m$ . Plus le nombre de domaines intermédiaires est élevé (i.e. les seuils intermédiaires  $\{b_i\}_{i=1}^{m-1}$  décroissent lentement), plus les probabilités conditionnelles qui lui sont associées sont élevées et par conséquent le nombre de simulations  $M$  nécessaire pour l'estimation de chacune d'elles est limité. Cependant, le nombre total de simulations  $mM$  sera très élevé et l'efficacité de la méthode affectée. Inversement, si les seuils intermédiaires  $\{b_i\}_{i=1}^{m-1}$  décroissent rapidement de telle façon que les évènements de défaillance conditionnelles deviennent rares, nous aurons besoin d'un nombre de simulations plus conséquent pour l'estimation des probabilités de défaillance conditionnelles qui leurs sont associés. Par conséquent, le nombre total de simulations sera très élevé. Ainsi, les seuils intermédiaires  $\{b_i\}_{i=1}^{m-1}$  sont choisis de telle manière qu'ils garantissent le meilleur compromis entre le nombre  $m$  de domaines de défaillance intermédiaires et le nombre de simulations  $M$  nécessaire pour l'estimation des probabilités conditionnelles. Dans ce contexte, Au et Beck [40] ont suggéré que ces derniers soient déterminés d'une manière adaptative au cours des simulations de telle façon que les

probabilités de défaillance intermédiaires soient toutes égales à une valeur préalablement définie  $p_0$  (par exemple  $p_0 = 0,1$ ). La valeur de la probabilité conditionnelle  $p_0$  doit être choisie dans l'intervalle  $[0,1, 0,3]$ , bien que, dans la plupart des applications de la méthode, la valeur  $p_0 = 0,1$  soit adoptée. Le domaine de défaillance intermédiaire  $\Omega_{F_i}$  associé à l'étape  $i$  des simulations est défini de telle façon qu'il contienne  $Mp_0$  des points appartenant au domaine de défaillance intermédiaire  $\Omega_{F_{i-1}}$  associé à l'étape précédente. Les  $M(1 - p_0)$  points restant sont générés selon la densité de probabilité conditionnelle  $p_c(\mathbf{u}|\Omega_{F_{i-1}}) = 1_{\Omega_{F_{i-1}}} \varphi_U(\mathbf{u})/P_i$ .

Bien que les simulations de Monte-Carlo puissent être utilisées pour obtenir des points de simulations distribués selon la densité de probabilité conditionnelle  $p_c(\mathbf{u}|\Omega_{F_{i-1}})$ , elles sont jugées inefficaces puisque, en moyenne, il faut générer un échantillon de  $1/P_{i-1}$  points pour qu'on puisse avoir un point distribué selon la densité de probabilité conditionnelle  $p_c(\mathbf{u}|\Omega_{F_{i-1}})$ . Afin de pallier le problème d'inefficacité des simulations de Monte-Carlo, on utilise l'algorithme de Metropolis [41] basé sur la théorie des chaînes de Markov. L'idée fondatrice de cet algorithme est de construire une *trajectoire* ou une chaîne de Markov  $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M\}$  dont la distribution stationnaire corresponde à la densité de probabilité conditionnelle  $p_c$ , en partant d'un point initial appelé souvent *germe* ou *semence*  $\mathbf{u}^*$ . Le passage d'un point  $\mathbf{u}^j$  à un point  $\mathbf{u}^{j+1}$  de la chaîne est géré par une densité de probabilité  $N$ -dimensionnelle  $p^*$ , dite densité de probabilité de proposition, choisie de telle façon qu'elle soit symétrique  $p^*(\mathbf{u}|\mathbf{u}^*) = p^*(\mathbf{u}^*|\mathbf{u})$  et dont la formulation soit facile à manipuler. Dans ce contexte, la densité de probabilité Gaussienne ou la densité de probabilité uniforme peuvent être utilisées. Une revue plus détaillée sur ce choix a été fournie par Au et Beck [43] en fonction du type de problème de fiabilité à résoudre. Cependant, dans la plupart des applications, la densité de probabilité uniforme est utilisée.

Nous nous proposons de décrire brièvement l'algorithme de Metropolis. Supposons que la chaîne de Markov est à l'état  $j$  et qu'on souhaite construire l'état  $j + 1$ . Soit  $\mathbf{u}^*$  un point germe, on génère dans un premier temps un point candidat  $\hat{\mathbf{u}}$  selon la densité de probabilité de proposition  $p^*(\hat{\mathbf{u}}|\mathbf{u}^*)$  centrée au point germe  $\mathbf{u}^*$  et on calcule le rapport d'acceptation  $r = p^*(\hat{\mathbf{u}}|\mathbf{u}^*)\varphi_U(\hat{\mathbf{u}})/p^*(\mathbf{u}^*|\hat{\mathbf{u}})\varphi_U(\mathbf{u}^j) = \varphi_U(\hat{\mathbf{u}})/\varphi_U(\mathbf{u}^j)$ . Ainsi, la probabilité d'acceptation du point candidat  $\hat{\mathbf{u}}$  est  $\min\{1, r\}$  et la probabilité que la chaîne de Markov reste dans son état actuel, autrement dit  $\mathbf{u}^{j+1} = \mathbf{u}^j$ , est  $1 - \min\{1, r\}$ . Admettons que la probabilité d'acceptation soit supérieure à la probabilité de rejet : si le point candidat  $\hat{\mathbf{u}}$  appartient au domaine de défaillance  $\Omega_{F_{i-1}}$  alors on l'accepte et on a  $\mathbf{u}^{j+1} = \hat{\mathbf{u}}$ , sinon on le rejette et on a  $\mathbf{u}^{j+1} = \mathbf{u}^j$ . Ce processus est répété jusqu'à ce qu'on obtienne l'échantillon de la taille  $M$  voulue. Au et Beck [40] ont montré que l'algorithme de Metropolis est inefficace dans le cas où la dimension de l'espace aléatoire est élevée, à cause de la forte corrélation entre les points générés, principalement due à la répétition de certains états de la chaîne de Markov. Afin de pallier ce problème, l'algorithme de Metropolis a été le sujet d'une amélioration qui stipule de construire la chaîne de Markov dans chaque direction de l'espace aléatoire. Ainsi, les composantes  $u_i^j$ ,  $i = 1, \dots, N$  d'un état  $\mathbf{u}^j$  de la chaîne de Markov sont obtenues une par une en utilisant l'algorithme de Metropolis et en partant des composantes respectives de l'état  $\mathbf{u}^{j-1}$ . Le risque de répétition concerne ainsi une composante d'un état donné de la chaîne de Markov et pas l'état lui-même. Lorsque la dimension de l'espace aléatoire est élevée, il est peu probable que l'état candidat  $\hat{\mathbf{u}}$  soit identique au point actuel  $\mathbf{u}^j$ , puisque cela implique que toutes les composantes de l'état candidat sont rejetées, ce qui est presque impossible. Cette version de l'algorithme de Metropolis est souvent désignée par l'algorithme de Metropolis-Hasting [42]. On note que la stationnarité de la chaîne de Markov n'est pas

atteinte immédiatement, au démarrage de l'algorithme de Metropolis-Hasting. L'expérience a montré qu'une étape de préchauffe est nécessaire pour atteindre cette stationnarité et que les points obtenus par la chaîne de Markov soient distribués selon la densité de probabilité ciblée. Autrement dit, il faut que l'algorithme de Metropolis-Hasting tourne à vide jusqu'à un certain seuil à partir duquel les points générés sont comptabilisés. Il n'y a pas de règles précises à suivre pour définir ce seuil, mais généralement la comptabilisation commence après la génération de quelques centaines de points. On note à ce niveau que l'efficacité de la méthode n'est pas affectée, puisque, dans l'étape de préchauffe, il n'y a pas d'évaluation de la fonction d'état limite.

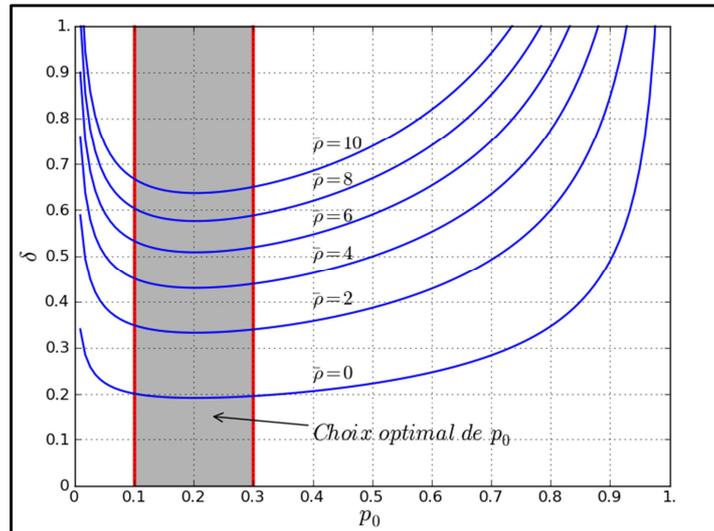
En résumé, la première étape de la méthode consiste à effectuer des simulations de Monte-Carlo sur un échantillon  $\{\mathbf{u}_0^1, \mathbf{u}_0^2, \dots, \mathbf{u}_0^M\}$  de  $M$  points générés dans l'espace aléatoire standard, pour obtenir  $M$  réalisations  $\{H_0^j = H(\mathbf{u}_0^j)\}_{j=1}^M$  de la fonction d'état limite qui sont arrangées d'une manière croissante,  $H_0^{j_1} \leq H_0^{j_2} \leq \dots \leq H_0^{j_M}$ . Le premier seuil intermédiaire  $b_1$  est choisi comme la  $((1 - p_0)M - 1)^{\text{ème}}$  valeur la plus grande des  $M$  réalisations  $\{H_0^j = H(\mathbf{u}_0^j)\}_{j=1}^M$ , de telle manière que la probabilité de défaillance  $P_1 = P[\Omega_{F_1}]$  associée au domaine de défaillance intermédiaire  $\Omega_{F_1}$  soit égale à la probabilité seuil  $p_0$ . Ainsi, le domaine de défaillance intermédiaire  $\Omega_{F_1}$  contient  $Mp_0$  points, parmi les points de l'échantillon  $\{\mathbf{u}_0^1, \mathbf{u}_0^2, \dots, \mathbf{u}_0^M\}$ , associés aux réalisations de la fonction d'état limite supérieures au seuil intermédiaire  $b_1$ . Ces points distribués naturellement selon la densité de probabilité conditionnelle  $p_c(\cdot | \Omega_{F_1})$ , sont utilisés comme des points germes dans l'algorithme de Metropolis-Hasting, pour générer les  $(1 - p_0)M$  points qui manquent. Le deuxième seuil intermédiaire  $b_2$  est choisi comme la  $((1 - p_0)M - 1)^{\text{ème}}$  valeur la plus grande des  $M$  nouvelles réalisations  $\{H_1^j = H(\mathbf{u}_1^j)\}_{j=1}^M$ . On note que la fonction d'état limite n'est évaluée qu'aux points germes récupérés de l'étape précédente de la méthode. Ainsi, le deuxième domaine de défaillance intermédiaire  $\Omega_{F_2}$  contient  $Mp_0$  points, parmi les points de l'échantillon  $\{\mathbf{u}_1^1, \mathbf{u}_1^2, \dots, \mathbf{u}_1^M\}$ , associés aux réalisations de la fonction d'état limite supérieures au seuil intermédiaire  $b_2$ , et la probabilité conditionnelle  $P_2 = P[\Omega_{F_2} | \Omega_{F_1}]$  est automatiquement égale à la probabilité seuil  $p_0$ . Ce processus est répété jusqu'à ce que la probabilité cible soit atteinte. On note que le critère d'arrêt est défini par l'atteinte du seuil intermédiaire  $b_i$  d'une valeur inférieure ou égale à 0. En supposant que toutes les probabilités conditionnelles sont égales à la probabilité seuil  $p_0$  et que le nombre  $M$  de points dans toutes les étapes de la méthode de Subset simulations est le même, le coefficient de variation  $\delta$  de l'estimation de la probabilité de défaillance est donné par :

$$\delta^2 \approx \frac{m(1 - p_0)}{Mp_0} (1 + \bar{\rho}) \quad (I.49)$$

où  $\bar{\rho}$  est la moyenne du coefficient de corrélation sur toutes les étapes des simulations, qui reflète la corrélation entre les points générés par l'algorithme de Metropolis-Hasting dans chaque étape, qui dépend à son tour de l'écart-type de la densité de probabilité de proposition  $p^*$ . Puisque le nombre total de simulations est  $M_T = mM$  et le nombre de domaines de défaillance intermédiaires est  $m = \log(P_f)/\log(p_0)$ , l'équation (I.49) peut être réécrite sous la forme suivante :

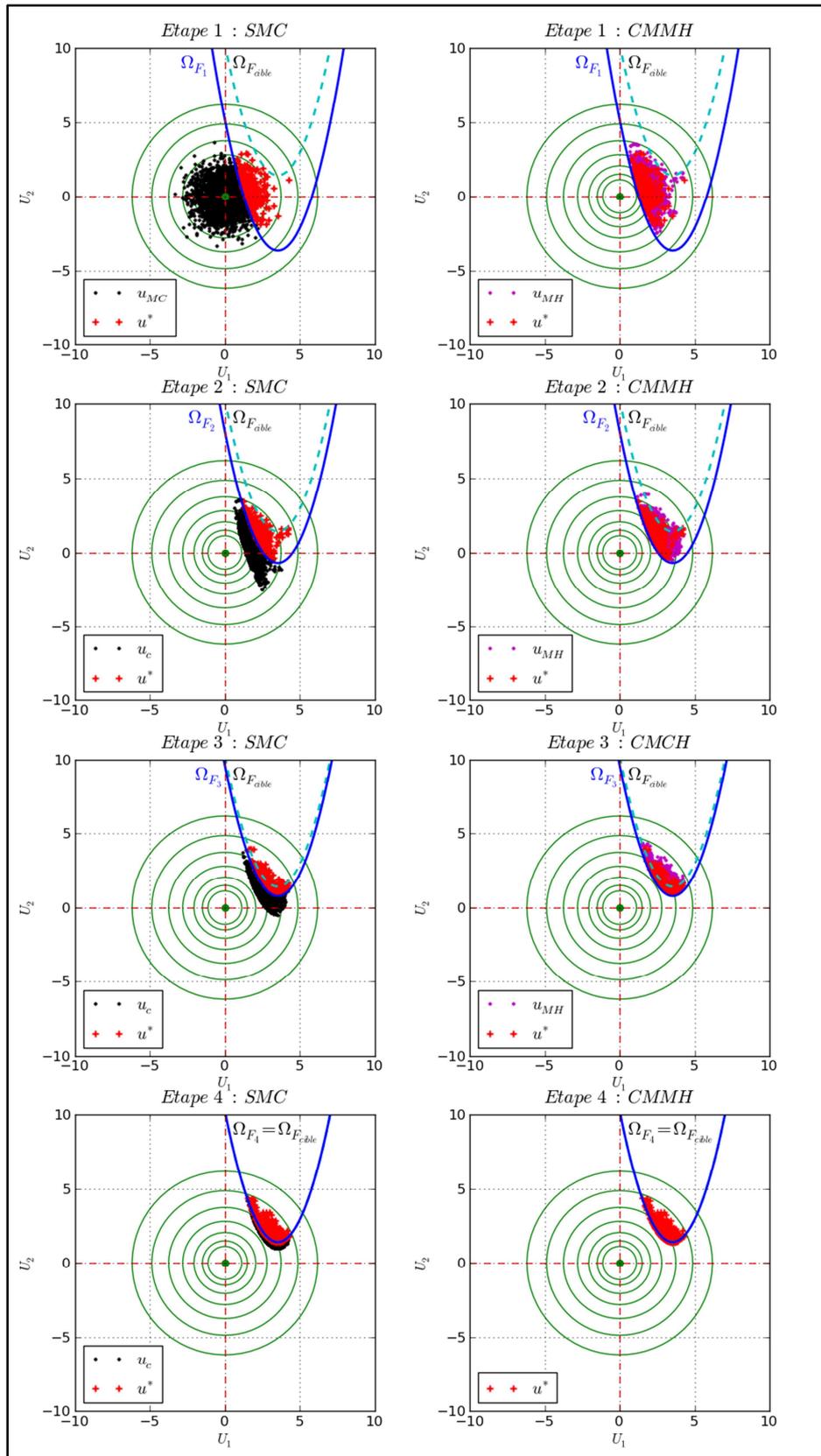
$$\delta^2 \approx \frac{(1 - p_0)}{p_0(\log(p_0))^2} \frac{(\log(P_f))^2}{M_T} (1 + \bar{\rho}) \quad (I.50)$$

La figure I.8 illustre l'évolution du coefficient de variation  $\delta$  en fonction de la probabilité seuil  $p_0$ , pour un nombre total de simulations  $M_T = 2000$  et une probabilité de défaillance cible  $P_f = 10^{-3}$ . On constate que la valeur optimale de la probabilité seuil  $p_0$  est égale à 0,2. De plus on constate que le coefficient de variation  $\delta$  est relativement peu sensible à la corrélation  $\bar{\rho}$  entre les points générés par l'algorithme de Metropolis-Hasting au voisinage de cette valeur optimale. Ainsi, l'efficacité de la méthode de simulation Subset est presque invariable si la probabilité seuil  $p_0$  reste dans l'intervalle  $[0,1, 0,3]$ .



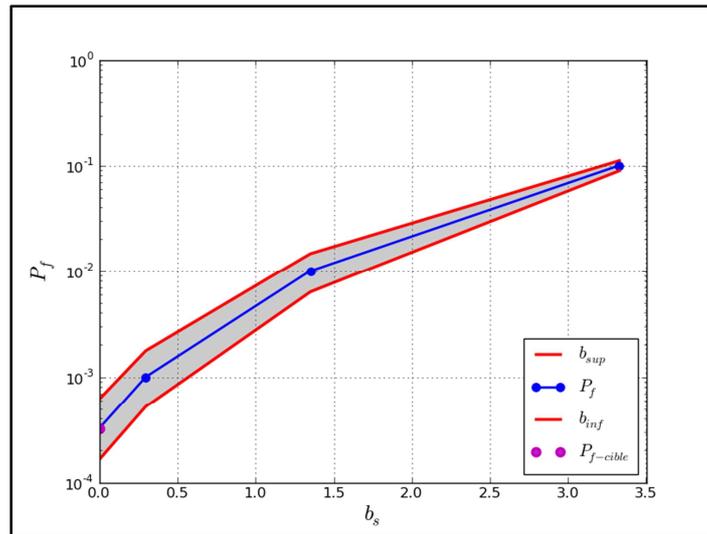
**Figure I. 8 :** Evolution du coefficient de variation  $\delta$  de l'estimation de la probabilité de défaillance en fonction de la probabilité seuil  $p_0$

Dans le but d'illustrer l'implémentation de la méthode, nous reprenons le problème de la surface d'état limite parabolique traité dans le paragraphe précédent. L'estimation de la probabilité de défaillance est obtenue par 3000 simulations par étape de la méthode Subset. La probabilité seuil  $p_0$  est égale à 0,1 et la densité de proposition  $p^*$  utilisée dans l'algorithme de Metropolis-Hasting est représentée par une densité de probabilité uniforme dont la largeur du support est égale à l'unité. La figure I.9 illustre les différentes étapes de la méthode Subset. On constate que le domaine de défaillance cible est identifié après quatre étapes ; au cours de chacune d'elles la probabilité conditionnelle est égale à 0,1. L'estimation de la probabilité de défaillance donnée par la méthode Subset, est  $P_f^{SS} = 3,267 \cdot 10^{-4}$ , ce qui est en bon accord avec la solution de référence. Pour atteindre une telle précision, on a besoin uniquement de 12000 simulations. Ainsi, cette méthode permet d'économiser plus de 98% des appels à la fonction d'état limite par rapport aux simulations de Monte-Carlo classiques, pour garantir la même précision sur les estimations de la probabilité de défaillance.



**Figure I. 9 :** Etapes de la méthode Subset simulations

La figure I.10 illustre l'évolution de la probabilité de défaillance au cours des étapes de la méthode, ainsi que l'intervalle de confiance des estimations.



**Figure I. 10 :** Evolution de la probabilité de défaillance au cours des étapes de la méthode Subset simulations

On constate que la probabilité de défaillance associée à une fonction d'état limite qui peut être définie en faisant varier la valeur du seuil intermédiaire  $b_s$  dans l'intervalle  $[0, 3,4]$ , peut être déterminée directement, sans aucun effort de calcul supplémentaire, à partir de la courbe  $P_f = f(b_s)$  donnée dans la figure I.10. Ainsi, on justifie une fois de plus l'efficacité de la méthode Subset.

#### I.4.2.7 Méthode de surface de réponse quadratique

Dans la pratique, le comportement d'un système est rarement décrit au moyen d'un modèle mécanique explicite. On a recours le plus souvent à une modélisation par éléments finis. Les différentes techniques présentées dans les paragraphes précédents constituent les ingrédients nécessaires pour effectuer le couplage dit « *direct* », entre l'analyse de la fiabilité et le calcul par éléments finis. Cependant, l'expérience a montré que ce couplage n'est abordable que si l'on dispose d'un algorithme robuste pour la détermination du point de conception dans l'espace standard et d'une méthode pratique pour l'évaluation du vecteur gradient de la fonction d'état limite. La première condition peut être remplie en employant la version améliorée de l'algorithme HLRF [44]. Pour le calcul du gradient de la fonction d'état limite, la méthode de différentiation directe est parfois suggérée. Cette méthode peut être appliquée pour des problèmes impliquant une non-linéarité d'origine géométrique ou matérielle et aussi pour des problèmes dynamiques. Cependant, elle nécessite des développements analytiques qui peuvent être lourds dans certains cas (par exemple problèmes non linéaires). De plus, ces développements et leurs implémentations ne sont pas universels et doivent être refaits à chaque fois qu'on aborde une nouvelle classe de problèmes. Pour cette raison, la méthode des différences finies est généralement retenue pour le calcul du gradient. L'application de cette méthode n'implique pas la modification de la formulation éléments finis (découplage du calcul par éléments finis et du calcul du gradient), mais nécessite plus d'évaluations du modèle mécanique. A titre illustratif, chaque calcul de gradient nécessite au moins  $N + 1$  évaluations de la fonction d'état limite ( $N$  est le nombre de variables aléatoires de base). Par conséquent, si la dimension de l'espace aléatoire de base est élevée et la méthode des différences finies utilisée pour le calcul du gradient, l'approche de couplage directe n'est pas abordable puisqu'elle peut engendrer un temps de calcul très important. Ainsi, les méthodes basées sur le concept de surface de réponse offrent une bonne alternative pour traiter une telle

situation puisque l'évolution des gradients n'est pas coûteuse. Parmi les plus populaires, nous pouvons citer la méthode de Surface de Réponse Quadratique (SRQ) présentée ci-dessous.

En raisonnant dans l'espace standard, l'idée de base de la méthode SRQ est de construire une représentation explicite  $\hat{H}(\mathbf{u})$  de la fonction d'état limite exacte  $H(\mathbf{u})$ , généralement définie d'une façon implicite. Dans la littérature, plusieurs auteurs [45-49] ont suggéré que cette représentation explicite  $\hat{H}(\mathbf{u})$  soit définie par une approximation polynômiale quadratique. Soit  $\mathbf{u} = \{u_1, u_2, \dots, u_n\}^T$  le vecteur désignant une réalisation des paramètres incertains dans l'espace standard, une approximation polynômiale quadratique de la fonction d'état limite s'écrit :

$$\hat{H}(\mathbf{u}) = a_0 + \sum_{i=1}^n a_i u_i + \sum_{i=1}^n a_{ii} u_i^2 + \sum_{i=1}^n \sum_{j=1, j \neq i}^n a_{ij} u_i u_j \quad (I.51)$$

où  $a_0$ ,  $a_i$ ,  $a_{ii}$  et  $a_{ij}$  sont des coefficients réels correspondant respectivement aux termes constants, linéaires (i.e. ordre 1), carrés (i.e. second ordre) et croisés de l'approximation polynômiale quadratique, qui sont des inconnues à déterminer. Une fois cette formulation analytique de la fonction d'état limite construite, elle se substitue à la fonction d'état limite réelle dans la procédure d'analyse de la fiabilité.

La détermination des coefficients  $\mathbf{A} = \{a_0, a_i, a_{ii}, a_{ij}\}$  peut être effectuée par moindres carrés impliquant la résolution de systèmes linéaires. Après avoir choisi une série de points d'expérimentation (i.e. un plan d'expérience)  $\{\mathbf{u}^k\}_{k=1}^{N_p}$ , pour lesquels la valeur exacte de la fonction d'état limite  $y^k = H(\mathbf{u}^k)$  est calculée, un système linéaire défini par  $N_p$  équations dont les inconnues sont les coefficients  $\mathbf{A}$  de l'approximation est obtenu. En considérant la méthode des moindres carrés, la résolution de ce système d'équations est effectuée en minimisant l'erreur suivante par rapport au paramètre  $\mathbf{A}$  :

$$err(\mathbf{A}) = \sum_{k=1}^{N_p} (y^k - \hat{H}(\mathbf{u}^k))^2 \quad (I.52)$$

Pour plus de commodité dans la représentation, on se propose d'écrire l'approximation polynômiale quadratique  $\hat{H}(\mathbf{u})$  de la fonction d'état limite définie par l'équation (I.51) sous la forme matricielle suivante :

$$\hat{H}(\mathbf{u}) = \{1, u_i, u_i^2, u_i u_j\}^T \cdot \{a_0, a_i, a_{ii}, a_{ij}\} \equiv \mathbf{B}^T(\mathbf{u}^k) \cdot \mathbf{A} \quad (I.53)$$

Ainsi le problème des moindres carrés s'écrit :

$$\mathbf{A} = \operatorname{argmin} \left\{ \sum_{k=1}^{N_p} (y^k - \mathbf{B}^T(\mathbf{u}^k) \cdot \mathbf{A})^2 \right\} \quad (I.54)$$

En effectuant quelques calculs algébriques, la solution du problème défini par l'équation (I.54) est donnée par [47] sous la forme suivante :

$$\mathbf{A} = [\mathbf{B}^T \cdot \mathbf{B}]^{-1} \cdot \mathbf{B}^T \cdot \mathbf{Y} \quad (I.55)$$

où  $\mathbf{B}$  est une matrice dont les lignes correspondent aux vecteurs  $\mathbf{B}(\mathbf{u}^k)$  et  $\mathbf{Y}$  est un vecteur dont les composantes sont  $y^k = H(\mathbf{u}^k)$ .

Une grande variété de méthodes SRQ a été présentée dans la littérature [48, 49]. Néanmoins, toutes ces approches sont basées sur le même concept général décrit ci-dessus. Elles diffèrent

soit par les termes retenus dans la formulation de l'approximation polynômiale quadratique (avec ou sans termes croisés), soit par le critère de convergence adopté, soit par le type de plan d'expérience (i.e. les points d'expérimentation utilisés dans la régression). On note que les points d'expérimentation doivent être choisis de telle manière que les équations formant le système à résoudre pour déterminer les coefficients  $\mathbf{A} = \{a_0, a_i, a_{ii}, a_{ij}\}$  soient indépendantes (i.e. il faut que la quantité  $\mathbf{B}^T \cdot \mathbf{B}$  soit inversible). De plus, il est nécessaire que  $N_p \geq N$  pour que l'équation (I.55) puisse être résolue.

Le rôle principal du plan d'expérience est de générer des réalisations du vecteur  $\mathbf{U}$  à partir de la combinaison de différents niveaux des variables aléatoires de base, qui serviront par la suite comme points d'évaluation pour la fonction d'état limite. Pour des raisons de simplicité, les points d'expérimentation sont générés autour du point moyen. Cependant, il faut définir les points d'expérimentation de façon à correspondre au mieux aux exigences du problème étudié. Il existe plusieurs types de plans d'expérience disponibles dans la littérature. Parmi eux, on peut citer le plan en étoile, le plan hyper-cube, le plan factoriel, le plan minimal et le plan composite.

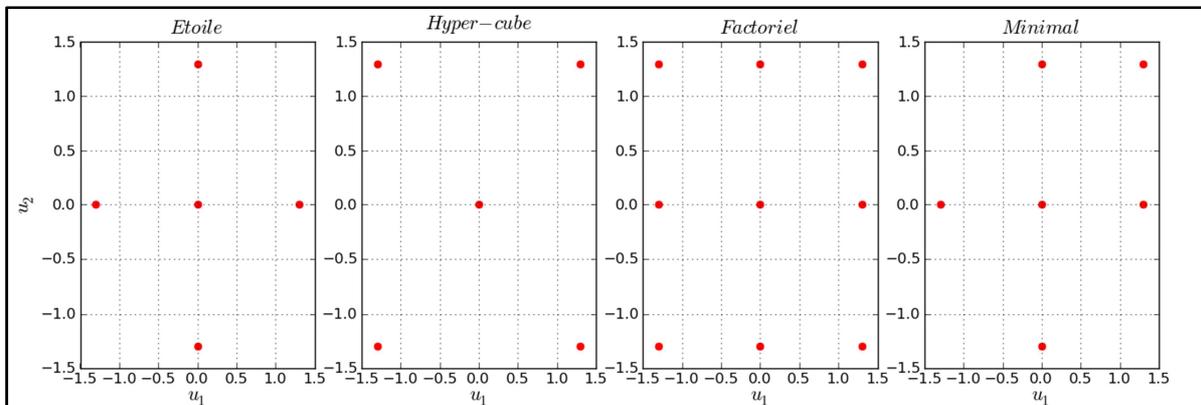


Figure I.11 : Différents plans d'expériences utilisés par la méthode SRQ

La figure I.11 illustre les différents plans d'expérience couramment utilisés par la méthode SRQ. Ils correspondent au cas bidimensionnel (i.e. deux variables aléatoires de base) et sont exprimés dans l'espace standard. Le choix du type de plan d'expérience est crucial pour l'efficacité de la méthode SRQ. A titre d'exemple, si on utilise un plan factoriel, le nombre total d'expérimentations  $2^N$  augmente d'une façon exponentielle avec le nombre de variables aléatoires de base. Par conséquent, le temps de calcul peut être significativement élevé. Dans le but de réduire le nombre d'expérimentations, Bucher et Bourgund [45] ont suggéré l'utilisation d'un plan d'expérience formé uniquement des points situés sur les axes (i.e. plan en étoile sur la figure I.11), pour lequel le nombre total d'expérimentations est  $2n + 1$ . Cette suggestion simplifie la forme de l'approximation polynômiale quadratique définie par l'équation (I.51) en éliminant les termes croisés.

Puisque la région avoisinant le point de conception possède la plus forte contribution dans la probabilité de défaillance, l'approximation polynômiale quadratique de la fonction d'état limite doit être seulement précise dans cette région. Cette précision est améliorée au fur et à mesure dans la procédure itérative de recherche du point de conception jusqu'à la vérification du critère de convergence. Autrement dit, à chaque itération, une nouvelle approximation est construite en utilisant un nouveau plan d'expérience centré au point de conception calculé à l'itération précédente.

## I.5 Analyse de sensibilité

### I.5.1 Présentation du problème

D'une manière générale, les analyses de sensibilité peuvent être réparties en trois catégories : les analyses de sensibilité locale, les analyses de criblage (screening en anglais) et les analyses de sensibilité globale. Dans ce qui suit, chacune de ces catégories est brièvement présentée. Pour commencer, on note que l'analyse de sensibilité locale est suggérée seulement dans les cas pratiques pour lesquels les analyses ne demandent pas un temps de calcul important. Ce type d'analyse fournit des informations sur le comportement du modèle seulement au voisinage d'un point de référence. Elle est basée sur l'estimation des dérivées partielles correspondant à chaque paramètre. Ces dernières peuvent être normalisées dans le but de leur donner plus de signification, soit par la valeur nominale du paramètre soit par son écart-type. Les méthodes de criblage permettent d'identifier les paramètres les moins influents sur le modèle. Du fait qu'elles sont moins exigeantes en terme de temps de calcul, elles sont conseillées dans le cas de modèles contenant un grand nombre de paramètres d'entrée. Dans ce contexte, la méthode proposée par Morris [50] est la plus populaire. Les méthodes de criblage fournissent des mesures de sensibilité qualitative. Autrement dit, elles permettent d'arranger les paramètres intervenant dans le modèle sur la base de la mesure de leurs importances sans quantification (i.e. elles ne permettent pas de quantifier de combien un paramètre donné est important par rapport à un autre paramètre). Une mesure quantitative peut être obtenue en appliquant des méthodes de régression, mais seulement dans le cas où le modèle est linéaire ou faiblement non linéaire. Enfin, les méthodes d'analyse de sensibilité globale s'intéressent à l'étude de l'effet des paramètres d'entrée sur le modèle tout en considérant la totalité de leur domaine de variation, et non pas une partie du domaine au voisinage d'un point de référence. L'idée fondatrice de cette catégorie de méthodes d'analyse de sensibilité est de quantifier la contribution de la variabilité relative à chaque paramètre d'entrée sur la variabilité de la réponse du modèle. Mathématiquement parlant, il s'agit d'estimer la contribution de chaque paramètre d'entrée sur la variance de la réponse du modèle. Pour cette raison, elles sont souvent désignées par méthodes d'analyse de la variance (ANOVA : ANalyse Of VAriance). Dans ce contexte, on peut citer la méthode proposée par Sobol [51], qui est basée sur la notion d'indices de sensibilité pour lesquels l'effet de chaque facteur est calculé par l'évaluation d'une intégrale multidimensionnelle en utilisant généralement les simulations de Monte-Carlo. Dans ce qui suit, nous nous proposons de présenter les principales méthodes d'analyse de sensibilité. Nous nous focalisons sur l'exposition de leur formulation mathématique ainsi qu'aux différents aspects liés à leur implémentation. Mais, auparavant, nous nous proposons d'introduire la notion de décomposition de la variance et les indices de sensibilité de Sobol.

### I.5.2 Méthodes d'estimation des indices de sensibilité

#### I.5.2.1 Notion de décomposition de la variance

Généralement, le but des analyses de sensibilité basées sur l'étude de la variance est de déterminer la contribution de chaque variable aléatoire d'entrée  $X_i$ , dans la variance  $V[Y]$  de la réponse  $Y$ . En se basant sur cette idée, Saltelli et al. [52] ont montré que cette variance peut être décomposée de la façon suivante :

$$V[Y] = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < k} V_{ijk} + \dots + V_{1,2,\dots,N} \quad (I.56)$$

où  $V_i, V_{ij}, \dots, V_{ij\dots N}$  sont appelées variances partielles, exprimées par :

$$V_i = V[E[Y|X_i = x_i^*]]$$

$$V_{ij} = V[E[Y|X_i = x_i^*, X_j = x_j^*]]$$

$$V_{ijk} = V[E[Y|X_i = x_i^*, X_j = x_j^*, X_k = x_k^*]]$$

avec  $E[Y|X_i = x_i^*]$  l'espérance de la variable aléatoire  $Y$  sachant que le paramètre  $X_i$  est déterministe (i.e. fixé à une valeur particulière  $x_i^*$ ). Elle est souvent désignée par espérance conditionnelle d'ordre 1. De même,  $E[Y|X_i = x_i^*, X_j = x_j^*, X_k = x_k^*]$  est l'espérance conditionnelle d'ordre  $k$ , évaluée en considérant  $k$  paramètres d'entrée comme déterministes. La variance partielle  $V_i$  mesure la contribution de l'aléa associé au paramètre d'entrée  $X_i$  dans la variance totale de la réponse  $Y$ , alors que les variances d'ordres supérieurs permettent de mesurer la contribution de l'interaction entre les différents paramètres d'entrée.

### I.5.2.2 Indices de sensibilité de Sobol

Afin de déterminer ces grandeurs, Sobol [51] a proposé la décomposition de la fonction  $f$  représentant le modèle en une somme de fonctions de dimensions croissantes. Autrement dit, elle est composée de termes constant  $f_0$ , unidimensionnel  $f(x_i)$ , bidimensionnel  $f(x_i, x_j)$  et ainsi de suite. Cette représentation s'écrit :

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^N f_i(x_i) + \sum_{i=1}^N \sum_{j=1}^N f_{ij}(x_i, x_j) + \dots + f_{1,2,\dots,N}(x_1, x_2, \dots, x_N) \quad (I.57)$$

Sobol [51] a montré, pour des variables uniformes sur  $[0, 1]$ , que cette représentation existe si et seulement si le terme  $f_0$  est constant et l'intégrale de n'importe quel terme d'ordre supérieur  $f_{1,2,\dots,s}$  par rapport à un seul paramètre  $x_{i_k}$  est nulle :

$$\int_0^1 f_{1,2,\dots,s}(\mathbf{x}) dx_{i_k} = 0, \quad \text{pour } 1 \leq k \leq s \quad (I.58)$$

Par conséquent, tous les termes  $f_{1,2,\dots,s}$  du développement défini par l'équation (I.57) sont orthogonaux. Mathématiquement, si  $(i_1, \dots, i_s) \neq (j_1, \dots, j_l)$  alors :

$$\int_{[0,1]^N} f_{i_1,\dots,i_s}(\mathbf{x}) f_{j_1,\dots,j_l}(\mathbf{x}) d\mathbf{x} = 0 \quad (I.59)$$

Ces termes (désignés souvent sous le nom de fonctions de Sobol) peuvent être calculés par l'évaluation d'intégrales multidimensionnelles, telles que :

$$f_0 = \int_{[0,1]^N} f(\mathbf{x}) d\mathbf{x} \quad (I.60)$$

$$f_i(x_i) = \int_{[0,1]^{N-1}} f(\mathbf{x}) d\mathbf{x}_{-i} - f_0 \quad (I.61)$$

$$f_{ij}(x_i, x_j) = \int_{[0,1]^{N-2}} f(\mathbf{x}) d\mathbf{x}_{-ij} - f_i(x_i) - f_j(x_j) - f_0 \quad (I.62)$$

où  $d\mathbf{x}_{-i}$  et  $d\mathbf{x}_{-ij}$  indiquent que l'intégration s'effectue par rapport à toutes les variables sauf par rapport à  $x_i$ , ou par rapport à  $x_i$  et  $x_j$  respectivement. Les termes d'ordres supérieurs peuvent être déterminés en suivant la même procédure.

Il est facile de vérifier que les grandeurs définies par les équations (I.60) à (I.62) correspondent respectivement aux espérances conditionnelles d'ordre 0, 1 et 2. On peut écrire alors :

$$E[Y] = f_0 \quad (I.63)$$

$$E[Y|X_i = x_i^*] = f_i(x_i) - f_0 \quad (I.64)$$

$$E[Y|X_i = x_i^*, X_j = x_j^*] = f_{ij}(x_i, x_j) - f_0 \quad (I.65)$$

On note que cette correspondance reste valable pour les ordres plus élevés.

Par définition, la variance de la variable aléatoire s'écrit :

$$V[Y] = \int_{[0,1]^N} f^2(\mathbf{x}) d\mathbf{x} - f_0^2 \quad (I.66)$$

En substituant  $f$  dans l'équation (I.66) par le développement défini par l'équation (I.57), on obtient une décomposition de la variance  $V[Y]$  telle qu'elle est définie par l'équation (I.56). Ainsi, par identification, les variances partielles  $V_i$  et  $V_{ij}$  peuvent être respectivement exprimées par :

$$V_i = V[f_i] = \int_{[0,1]} f_i^2(x_i) dx_i \quad (I.67)$$

$$V_{ij} = V[f_{ij}] = \int_{[0,1]^2} f_{ij}^2(x_i, x_j) dx_i dx_j \quad (I.68)$$

Généralement, la variance partielle  $V_{i_1, \dots, i_s}$  correspondant au terme générique  $f_{i_1, \dots, i_s}$ , tel que  $1 \leq i_1 < \dots < i_s \leq N$  et  $s = 1, \dots, N$ , est obtenue par :

$$V_{i_1, \dots, i_s} = V[f_{i_1, \dots, i_s}] = \int_{[0,1]^s} f_{i_1, \dots, i_s}^2(x_{i_1}, \dots, x_{i_s}) dx_{i_1} \dots dx_{i_s} \quad (I.69)$$

En normalisant ces variances partielles par rapport à la variance totale  $V[Y]$ , Sobol [51] a introduit un ensemble de mesures de sensibilité de différents ordres, désignées par indices de sensibilité de Sobol. Les indices de Sobol du premier ordre sont définis par :

$$S_i = \frac{V_i}{V[Y]} \quad (I.70)$$

et mesurent la contribution de l'aléa relatif au paramètre  $X_i$  dans la variance de la variable  $Y$ . Ils peuvent être interprétés comme des mesures des *effets principaux* qui sont souvent employés dans les approches de plans d'expériences. Ils sont désignés aussi par *mesures d'importance* qui peuvent servir à la classification des paramètres d'un modèle.

Les indices de Sobol du second ordre sont définis par :

$$S_{ij} = \frac{V_{ij}}{V[Y]} \quad (I.71)$$

Ils mesurent la contribution de l'interaction entre les paramètres  $X_i$  et  $X_j$  dans la variance de la variable aléatoire  $Y$ , c'est-à-dire la sensibilité de  $Y$  qui n'était pas prise en compte dans l'effet des paramètres  $X_i$  et  $X_j$  seuls.

Les indices de Sobol du troisième ordre sont définis par :

$$S_{ijk} = \frac{V_{ijk}}{V[Y]} \quad (I.72)$$

et mesurent la contribution de l'interaction entre les paramètres  $X_i$ ,  $X_j$  et  $X_k$  dans la variance du paramètre aléatoire  $Y$ , c'est-à-dire la sensibilité de  $Y$  qui n'était pas prise en compte ni dans l'effet des paramètres  $X_i$ ,  $X_j$  et  $X_k$  seuls, ni dans les interactions deux à deux.

Les indices de Sobol d'ordre plus élevé sont obtenus de la même manière. En général, les indices de Sobol  $S_{i_1, \dots, i_s}$  d'ordre  $s$ , sont définis par :

$$S_{i_1, \dots, i_s} = \frac{V_{i_1, \dots, i_s}}{V[Y]} \quad (I.73)$$

Etant donné que les indices de Sobol sont positifs et leur somme est égale à l'unité :

$$\sum_i S_i + \sum_{i < j} S_{ij} + \sum_{i < j < k} S_{ijk} + \dots + S_{1,2,\dots,N} = 1 \quad (I.74)$$

ils sont faciles à interpréter. Ainsi, plus la valeur de l'indice de Sobol est proche de l'unité, plus l'aléa relatif au paramètre auquel il est associé a un effet significatif sur la variabilité de la réponse du modèle.

Pour un modèle ayant  $N$  paramètres d'entrée, le nombre d'indices de Sobol de différents ordres qu'il faut évaluer est  $2^N - 1$ . Lorsque  $N$  est très grand, le nombre d'indices de Sobol augmente d'une manière exponentielle. Par conséquent, le coût de calcul en terme d'appels au modèle devient prohibitif. Dans le cas de modèles additifs (i.e. les interactions entre les différents paramètres n'ont pas d'effet), ce problème ne se pose pas puisque les indices de Sobol du premier ordre peuvent à eux seuls quantifier l'effet de chaque paramètre sur la réponse du modèle. Par contre, lorsque les interactions entre les paramètres ont une influence sur la réponse du modèle, ce problème se pose sérieusement.

Dans le but de surmonter cette difficulté, Homma et Saltelli [53] ont proposé de nouvelles mesures de sensibilité, désignées par indices de sensibilité totaux. L'indice de sensibilité total noté  $S_{T_i}$  est défini comme la somme de l'effet principal relatif au paramètre  $X_i$  et les effets des interactions dans lesquels intervient ce paramètre. Autrement dit, pour un paramètre  $X_i$ , l'indice de sensibilité total qui lui est associé correspond à la somme de tous les indices de Sobol faisant intervenir l'indice  $i$ . On peut écrire alors :

$$S_{T_i} = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots = 1 - \frac{V[E[Y|\mathbf{X}_{-i}]]}{V[Y]} \quad (I.75)$$

où  $\mathbf{X}_{-i}$  est un vecteur représentant tous les paramètres du modèle à l'exception du paramètre  $X_i$ . En se référant au théorème de la variance totale, la variance de la variable  $Y$  représentant la réponse du modèle est donnée par :

$$V[Y] = V[E[Y|\mathbf{X}_{-i}]] + E[V[Y|\mathbf{X}_{-i}]] \quad (I.76)$$

où  $V[E[Y|\mathbf{X}_{-i}]]$  et  $E[V[Y|\mathbf{X}_{-i}]]$  désignent, respectivement, l'effet principal et le résidu.

En remplaçant  $V[E[Y|\mathbf{X}_{-i}]]$  dans l'équation (I.75) par son expression issue de l'équation (I.68), l'indice total peut être réécrit sous la forme :

$$S_{T_i} = \frac{E[V[Y|\mathbf{X}_{-i}]]}{V[Y]} \quad (I.77)$$

La grandeur  $E[V[Y|\mathbf{X}_{-i}]]$  n'est autre que la moyenne de la variance conditionnelle  $V[Y|\mathbf{X}_{-i}]$ , et peut être estimée par simulations de Monte-Carlo. L'évaluation de l'indice de sensibilité total en se basant sur l'équation (I.77) est plus économique en terme d'appels au modèle mécanique, par rapport à l'approche classique qui implique le calcul des indices de Sobol de différents ordres. Le coût des estimations est égal à  $M(N + 1)$ , où  $M$  est le nombre de simulations utilisées pour le calcul d'un indice de Sobol du premier ordre. Si, en plus des indices totaux, on veut calculer les indices de Sobol du premier ordre, le coût total des estimations sera égal à  $M(N + 2)$ .

Dans la littérature, plusieurs méthodes [54, 55] ont été proposées pour calculer les indices de sensibilité de Sobol. Le choix le plus simple, et pas forcément le meilleur, est l'utilisation des simulations de Monte-Carlo. Dans ce contexte, les estimations de la moyenne et de la variance de la variable  $Y$ , ainsi que de la variance partielle relative à un paramètre  $X_i$  sont respectivement données par :

$$\hat{\mu}_Y = \frac{1}{N_s} \sum_{m=1}^{N_s} f(\mathbf{x}_m) \quad (I.78)$$

$$\hat{V}_Y = \frac{1}{N_s} \sum_{m=1}^{N_s} f(\mathbf{x}_m)^2 - \hat{\mu}_Y^2 \quad (I.79)$$

$$\hat{V}_i = \frac{1}{N_s} \sum_{m=1}^{N_s} f(\mathbf{x}_{-im}^{(1)}, x_{im}^{(1)}) f(\mathbf{x}_{-im}^{(2)}, x_{im}^{(1)}) - \hat{\mu}_Y^2 \quad (I.80)$$

où  $\mathbf{x}_{-im} = (x_{1m}, \dots, x_{(i-1)m}, \dots, x_{(i+1)m}, x_{N_s m})$ , et les exposants <sup>(1)</sup> et <sup>(2)</sup> indiquent qu'on utilise deux échantillons distincts pour l'estimation des variances partielles du premier ordre sauf pour la composante  $x_{im}$  pour laquelle on utilise le même point.

La variance conditionnelle qui intervient dans l'estimation des indices totaux est calculée par :

$$\hat{V}_{-i} = \frac{1}{N_s} \sum_{m=1}^{N_s} f(\mathbf{x}_{-im}^{(1)}, x_{im}^{(1)}) f(\mathbf{x}_{-im}^{(1)}, x_{im}^{(2)}) - \hat{\mu}_Y^2 \quad (I.81)$$

Les indices de sensibilité du premier ordre ainsi que les indices totaux sont directement obtenus en normalisant respectivement les variances partielles  $\hat{V}_i$  et les variances conditionnelles  $\hat{V}_{-i}$  par l'estimation de la variance totale  $\hat{V}_Y$ . On note que l'efficacité des simulations de Monte-Carlo peut être améliorée en utilisant le tirage en hyper-cube Latin ou des séquences quasi-aléatoires, telle que la séquence de Halton [56] ou la séquence de Hammersley [57]. Cependant, la réduction du nombre d'appels au modèle n'est pas significative.

### I.5.2.3 La méthode FAST

La méthode FAST a été développée par Cukier et al. [58-60] dans les années 70 dans le but d'effectuer des analyses de sensibilité pour des modèles de chimie. Depuis, elle a connu plusieurs améliorations, qui lui ont permis d'être considérée parmi les méthodes d'analyse de sensibilité les plus performantes. Elle permet de calculer les moments statistiques (i.e. la moyenne et l'écart-type) de la réponse d'un modèle ainsi que la quantification de la contribution de chaque paramètre d'entrée sur cette variance. Sa particularité est qu'elle peut être toujours appliquée, que le modèle soit monotone ou non. Dans ce qui suit, on se propose de présenter son idée fondatrice ainsi que sa formulation mathématique.

Cukier et al. [60] ont montré qu'en utilisant une transformation de Fourier  $N$ -dimensionnelle de  $f$ , il est possible de décomposer la variance de la variable aléatoire  $Y$  comme une somme de fonctionnelles de dimensions croissantes, dépendant des paramètres d'entrée. Ainsi, il est possible de déterminer les effets principaux (i.e. effet de la variabilité d'un paramètre d'entrée sur la variance du modèle) et les effets d'interaction (i.e. effet de l'interaction entre deux paramètres ou plus sur la variabilité du modèle). Cependant, la construction d'une telle transformation  $N$ -dimensionnelle n'est pas triviale dans la pratique et peut impliquer un calcul prohibitif. Afin de surmonter ce problème, l'idée fondatrice de la méthode FAST consiste à remplacer cette transformation  $N$ -dimensionnelle par un ensemble de transformations de Fourier unidimensionnelles le long d'une courbe parcourant l'espace  $[0,1]^N$ . Cette courbe est construite à partir d'un ensemble d'équations paramétrées définies par :

$$x_i = G_i(\sin \omega_i s) \quad (I.82)$$

où  $s$  est une variable scalaire réelle (i.e.  $s \in \mathbb{R}$ ),  $\{G_i\}_{i=1}^N$  sont des fonctions de transformation permettant un recouvrement uniforme de l'espace  $[0,1]^N$  et  $\{\omega_i\}_{i=1}^N$  est un ensemble de fréquences entières linéairement indépendantes. Le choix de ce domaine de variation (i.e.  $[0,1]^N$ ) est fait uniquement dans un but illustratif et pour standardiser la formulation de la méthode FAST. D'autres domaines de variation peuvent être introduits puisqu'on peut toujours trouver une transformation qui permet de passer au domaine de variation standard  $[0,1]^N$ .

Lorsque le paramètre  $s$  varie dans  $\mathbb{R}$ , l'ensemble des paramètres d'entrée  $\mathbf{x} = (x_1, \dots, x_N)$  varie simultanément le long d'une courbe de telle façon à explorer uniformément l'espace  $[0,1]^N$  et chaque paramètre  $x_i$  oscille périodiquement à la fréquence  $\omega_i$  qui lui est associée, quelle que soit la fonction de transformation  $G_i$ . De plus, la réponse du modèle  $y$  présente différentes périodicités relatives à l'ensemble des fréquences  $\{\omega_i\}_{i=1}^N$ , indépendamment du modèle  $f$ . Ainsi, si la réponse du modèle  $y$  enregistre une forte oscillation pour une fréquence  $\omega_i$ , le paramètre  $x_i$  qui lui correspond (i.e. à la fréquence  $\omega_i$ ) est susceptible d'avoir une forte influence. En effet, en se référant à la méthode FAST, la mesure de sensibilité relative au paramètre  $x_i$  est obtenue à partir des coefficients relatifs à la fréquence  $\omega_i$  et ses harmoniques.

Cukier et al. [58] ont montré que, lorsque les fréquences  $\{\omega_i\}_{i=1}^N$  sont linéairement indépendantes, c'est-à-dire qu'aucune fréquence  $\omega_i$  n'est obtenue par une combinaison linéaire des autres fréquences  $\{\omega_k\}_{k=1, k \neq i}^N$ , soit :

$$\sum_{i=1}^N c_i \omega_i \neq 0, c_i \in \mathbb{R} \quad (I.83)$$

le moment statistique d'ordre  $r$  peut être calculé en évaluant le modèle  $f$  le long de la courbe parcourue par le vecteur  $(x_1(s), \dots, x_N(s))$  sur  $[0,1]^N$ .

Par définition, le moment statistique d'ordre  $r$  de la variable aléatoire  $Y$  uniforme sur  $[0,1]$  s'écrit :

$$\langle Y \rangle^r = \int_{[0,1]^N} f^r(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \quad (I.84)$$

où  $p$  est la densité de probabilité jointe de la variable aléatoire  $N$ -dimensionnelle  $\mathbf{X}$  et  $[0,1]^N$  est le domaine de variation (i.e. domaine de définition) des paramètres d'entrée.

En suivant le raisonnement proposé par la méthode FAST, l'intégrale  $N$ -dimensionnelle correspondant au calcul du moment statistique d'ordre  $r$  est transformée en une intégrale unidimensionnelle définie par :

$$\langle Y \rangle^r = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f^r(\mathbf{x}(s)) ds \quad (I.85)$$

De plus, puisque les fréquences  $\{\omega_i\}_{i=1}^N$  sont entières et positives, et la courbe parcourue par le vecteur  $(x_1(s), \dots, x_N(s))$  ne remplit pas la totalité de l'espace  $[0,1]^N$  mais est périodique de période  $2\pi$ , Cukier et al. [58] ont montré que la moyenne et la variance de la variable aléatoire  $Y$  peuvent être respectivement exprimées par :

$$E[Y] = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) ds \quad (I.86)$$

$$V[Y] = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(s) ds - E[Y]^2 \quad (I.87)$$

En substituant  $f(s)$  par son développement en série de Fourier, Saltelli et al. [52] ont proposé une approximation de la variance de la variable  $Y$  par :

$$\begin{aligned} V[Y] &\approx \sum_{j=-\infty}^{+\infty} (A_j^2 + B_j^2) - (A_0^2 + B_0^2) \\ &\approx 2 \sum_{j=1}^{+\infty} (A_j^2 + B_j^2) \end{aligned} \quad (I.88)$$

où  $A_j$  et  $B_j$  sont les coefficients de Fourier définis par :

$$A_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \cos(js) ds \quad (I.89)$$

$$B_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \sin(js) ds \quad (I.90)$$

En évaluant les coefficients  $A_j$  et  $B_j$  à la fréquence  $\omega_i$  et ses harmoniques  $p\omega_i$ , on peut calculer la portion  $V_{\omega_i}[Y]$  de la variance totale  $V[Y]$  due à la variabilité du paramètre  $x_i$  :

$$V_{\omega_i}[Y] \approx 2 \sum_{j=1}^{+\infty} (A_{j\omega_i}^2 + B_{j\omega_i}^2) \quad (I.91)$$

De plus, puisque le spectre  $\Lambda_{j\omega_i} = A_{j\omega_i}^2 + B_{j\omega_i}^2$  du développement en série de Fourier évalué aux harmoniques  $p\omega_i$  décroît lorsque  $p$  croît, la quantité  $V_{\omega_i}[Y]$  peut être définie par :

$$V_{\omega_i}[Y] \approx 2 \sum_{j=1}^M (A_{j\omega_i}^2 + B_{j\omega_i}^2) \quad (I.92)$$

où  $M$  est le nombre maximal d'harmoniques, qui peut être pris égal à 4 ou 6 comme suggéré par Cukier et al. [59].

Le rapport  $V_{\omega_i}[Y]/V[Y]$  fournit alors une mesure de la contribution de la variabilité du paramètre  $x_i$  relatif à la fréquence  $\omega_i$  sur la variance totale de la réponse du modèle. Il peut aussi être identifié à la mesure de *l'effet principal* couramment employé dans la technique des plans d'expériences, et qui constitue un moyen de ranger les paramètres d'entrée du modèle dans un ordre d'importance (i.e. plus le rapport  $V_{\omega_i}[Y]/V[Y]$  est grand, plus la variabilité du paramètre auquel il est associé est influente sur la réponse du modèle). Ce rapport est souvent désigné dans le contexte d'analyse de sensibilité par indice de sensibilité. Il peut être exprimé en se basant sur les équations (I.85) et (I.92) par :

$$S_i = 1 - \frac{\sum_{j=1}^M (A_j^2 + B_j^2)}{\sum_{j=1}^M (A_{j\omega_i}^2 + B_{j\omega_i}^2)} \quad (I.93)$$

Saltelli et Bolado [61] ont montré que les indices de sensibilité  $\{S_i\}_{i=1}^N$  sont équivalents aux indices de Sobol du premier ordre.

La première étape pour le calcul des indices de sensibilité  $\{S_i\}_{i=1}^N$  est le choix de l'ensemble des fréquences entières  $\{\omega_i\}_{i=1}^N$  pour les paramètres du modèle, qui doivent être linéairement indépendantes. Pour assurer cette condition, Cukier et al. [59] ont proposé un algorithme pour générer cet ensemble de fréquences en fonction du nombre de paramètres d'entrée. De plus, ils ont montré, en se basant sur le critère de Nyquist, que le nombre optimal de points nécessaire pour le calcul des indices de sensibilité est :

$$N_s = 2M\omega_{max} + 1 \quad (I.94)$$

où  $M$  est le nombre maximal d'harmoniques et  $\omega_{max}$  est la fréquence maximale parmi l'ensemble des fréquences  $\{\omega_i\}_{i=1}^N$ . Le tableau I.3 fournit l'ensemble des fréquences  $\{\omega_i\}_{i=1}^N$  ainsi que le nombre de points  $N_s$  pour différents nombres de paramètres d'entrée  $N$ , et pour un nombre d'harmoniques  $M = 4$ .

**Tableau I. 3** : Fréquences nécessaires pour le calcul des indices de sensibilité par la méthode FAST

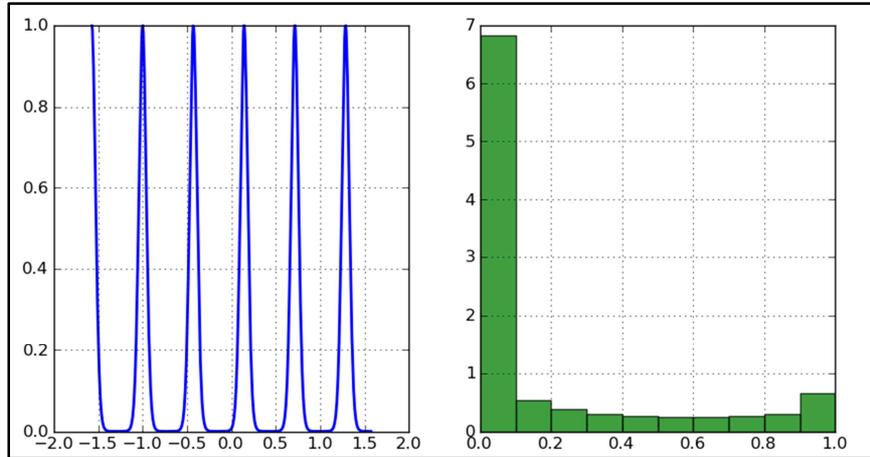
$N$	$N_s$	$\omega_{max}$	$\{\omega_i\}_{i=1}^N$
3	121	15	{1, 9, 15}
4	185	23	{5, 11, 19, 23}
5	625	39	{11, 21, 27, 35, 39}
6	393	49	{1, 21, 31, 37, 45, 49}
7	697	87	{17, 39, 59, 69, 75, 83, 87}
8	1001	125	{23, 55, 77, 97, 107, 113, 121, 125}

Outre l'ensemble des fréquences  $\{\omega_i\}_{i=1}^N$ , le choix des fonctions de transformation  $\{G_i\}_{i=1}^N$  qui permettent le passage de l'espace de variation des paramètres d'entrée à l'espace fréquentiel est primordial. En effet, une condition nécessaire que doivent vérifier ces fonctions de transformation est de générer un ensemble de points pour chaque paramètre d'entrée  $x_i$ ,  $\forall i = 1, \dots, N$  uniformément distribués. Dans ce contexte, différents types de fonctions ont été proposés dans la littérature [52]. Celle proposée par Cukier et al. [58] est donnée par :

$$x_i = \bar{x}_i e^{\bar{v}_i \sin(\omega_i s)} \quad (I.95)$$

où  $\bar{x}_i$  est la valeur nominale relative à  $x_i$  et  $\bar{v}_i$  un paramètre permettant de définir les bornes de définition du paramètre d'entrée  $x_i$  lorsque  $s$  varie entre  $-\pi/2$  et  $\pi/2$ .

La figure I.12 illustre l'évolution de la fonction de transformation définie par l'équation (I.95) pour  $\bar{x}_i = e^{-5}$ ,  $\bar{v}_i = 5$  et  $\omega_i = 11$ , ainsi que l'histogramme de la distribution empirique des points  $x_i$  (obtenu à partir d'un échantillon de 500 points).



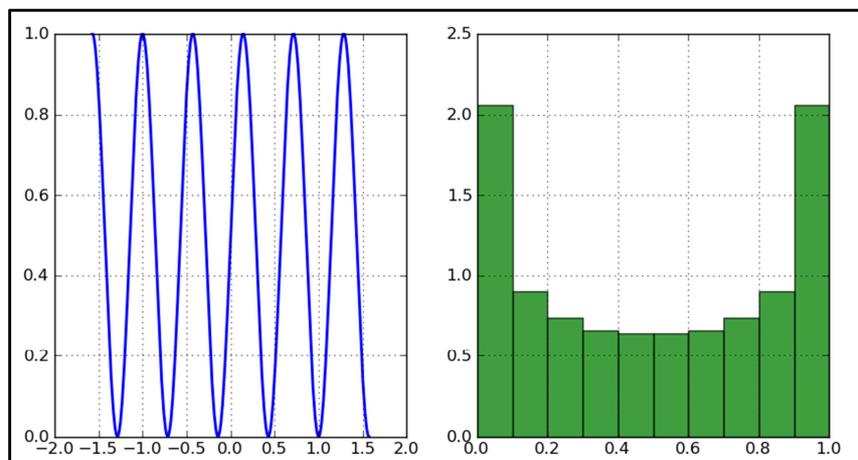
**Figure I. 12 :** Evolution de la fonction de transformation définie par l'équation (I.95) et histogramme de la distribution empirique des points  $x_i$

On constate que l'histogramme de la distribution des points  $x_i$  générées par la fonction de transformation définie par l'équation (I.95) est fortement asymétrique, puisque la majorité des points est située au niveau de la borne inférieure du domaine de variation du paramètre  $x_i$ . Ainsi, il est clair que ce type de transformation ne peut pas générer des points uniformément distribués et ne peut être employé que dans le cas où la densité de probabilité d'un paramètre d'entrée a une asymétrie positive et une longue queue.

D'autres auteurs tels que Koda et al. [62] ont proposé une fonction de transformation de la forme suivante :

$$x_i = \bar{x}_i [1 + \bar{v}_i \sin(\omega_i s)] \tag{I. 96}$$

La figure I.13 illustre l'évolution de cette fonction de transformation ainsi que l'histogramme de la distribution empirique d'un échantillon de 500 points.



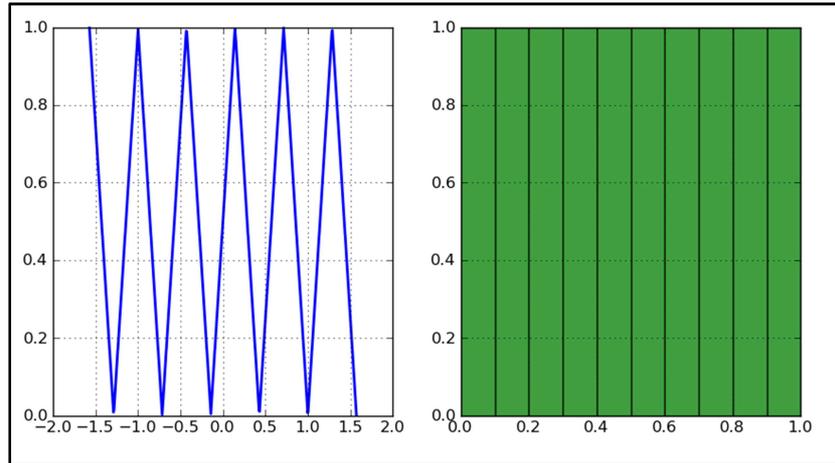
**Figure I. 13 :** Evolution de la fonction de transformation définie par l'équation (I.96) et histogramme de la distribution empirique des points  $x_i$

On constate que l'histogramme est symétrique mais ne permet pas non plus de générer des points uniformément distribués dans le domaine de variation des paramètres d'entrée.

Afin de surmonter le problème, Cukier et al. [60] ont proposé la fonction différentielle de forme générale suivante dont la résolution peut fournir des fonctions de transformation optimales :

$$(1 - x_i^2)^{1/2} p_i(G_i) \frac{dG_i(x_i)}{dx_i} = 1 \quad (I.97)$$

où  $p_i$  est la densité de probabilité du paramètre  $x_i$ .



**Figure I. 14 :** Evolution de la fonction de transformation définie par l'équation (I.98) et histogramme de la distribution empirique des points  $x_i$

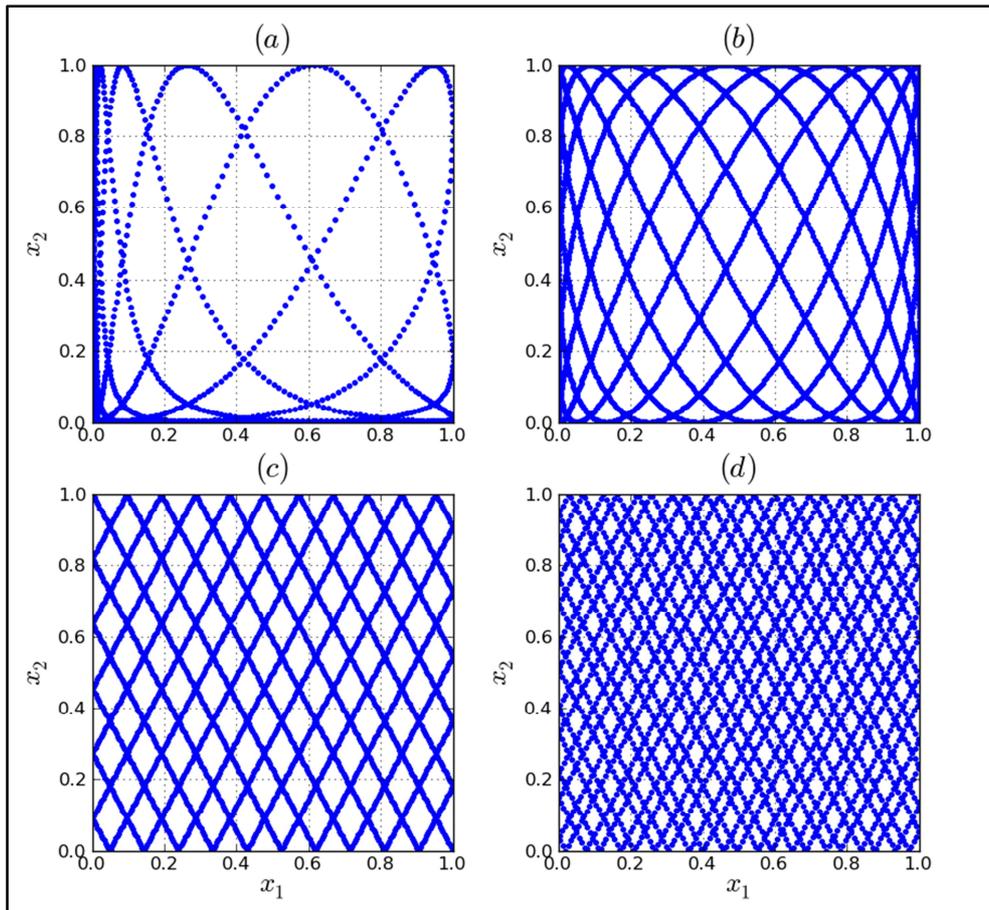
Dans ce contexte, Saltelli et al. [63] ont proposé une fonction de transformation de la forme suivante, obtenue par la résolution de l'équation (I.97) :

$$x_i = \frac{1}{2} + \frac{1}{\pi} \arcsin(\omega_i s) \quad (I.98)$$

L'évolution de cette fonction de transformation est définie comme un ensemble de lignes droites obliques illustré sur la figure I.14. On constate ce type de transformation permet de générer des points uniformément répartis puisque les différentes classes de l'histogramme de leur distribution empirique sont uniformément peuplées. Cependant, on note que le point de référence au voisinage duquel les points sont générés est toujours le même. Aussi pour donner plus de flexibilité au schéma de tirage, Saltelli et al. [63] ont suggéré la modification de l'équation (I.98) en intégrant le paramètre  $\varphi_i$  (angle de déphasage). La fonction de transformation est alors donnée par :

$$x_i = \frac{1}{2} + \frac{1}{\pi} \arcsin(\sin(\omega_i s + \varphi_i)) \quad (I.99)$$

où l'angle de déphasage  $\varphi_i$  est généré à partir d'une variable aléatoire uniforme ayant comme support l'intervalle  $[0, 2\pi[$ . Par conséquent, les tirages peuvent commencer d'un point quelconque du domaine de définition des paramètres d'entrée. En effet, en choisissant un ensemble d'angles de déphasage  $\{\varphi_i\}_{i=1}^N$  distincts, différentes fonctions de transformation peuvent être générées.



**Figure I. 15 :** *Distribution des points générés dans le cas bidimensionnel en utilisant (a) équation (I.95), (b), équation (I.96), (c) équation (I.98) et (d) équation (I.99)*

Dans le but de mieux illustrer les différences entre les fonctions de transformation présentées ci-dessus, on a tracé sur la figure I.15 les points générés par chacune d'elles dans le cas bidimensionnel (i.e. deux paramètres d'entrée) et ce pour les fréquences  $\omega_1 = 11$  et  $\omega_2 = 12$ . On constate que la répartition des points est plus uniforme dans le cas de la fonction de transformation définie par l'équation (I.99).

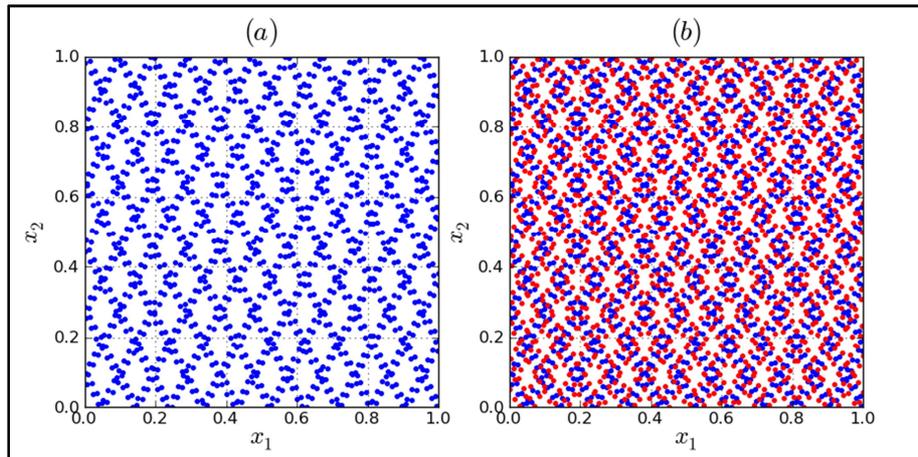
Saltelli et al. [63] ont proposé de calculer les indices de sensibilité en moyennant les résultats obtenus pour plusieurs réalisations des angles de déphasage  $\{\varphi_i\}_{i=1}^N$  (i.e. différentes fonctions de transformation). Si on désigne par  $N_r$  le nombre de réalisations utilisées, le nombre total de points utilisés dans le calcul des indices de sensibilité est donné par :

$$N_s = (2M\omega_{max} + 1)N_r \quad (I.100)$$

Les figures I.16a et I.16b illustrent les distributions des points d'expérimentation respectivement dans le cas d'une et de deux réalisations de la fonction de transformation définie par l'équation (I.99), soit pour  $N_r = 1$  et  $N_r = 2$ . On constate que la répartition des points est plus uniforme dans le cas où on effectue deux réalisations. De plus, on observe que l'ensemble des points générés dans une réalisation et dans une autre sont distincts.

Il est clair que la difficulté, avec la méthode FAST, est de choisir un ensemble de fréquences linéairement indépendantes. La fréquence maximale est proportionnelle au nombre  $N$  de paramètres d'entrée du modèle. Par conséquent, le coût de calcul en terme d'appels au modèle augmente d'une manière exponentielle avec  $N$ . Cukier et al. [59] ont montré que, pour un

nombre d'harmoniques  $M = 4$ , le nombre d'appels au modèle nécessaire pour l'estimation des indices de sensibilité est proportionnel à  $2,6N^{2,5}$ .



**Figure I. 16 :** Distribution des points générés dans les cas bidimensionnel en utilisant l'équation (I.99), (a)  $N_r = 1$  et (b)  $N_r = 2$

Dans le cas où le modèle est additif (i.e.  $\sum_{i=1}^N S_i = 1$ ), les indices de sensibilité du premier ordre fournissent une image complète de la sensibilité de la réponse du modèle vis-à-vis des paramètres d'entrée. Cependant, lorsque le modèle n'est pas additif, les interactions entre les différents paramètres d'entrée peuvent avoir une contribution significative dans la variance de la réponse du modèle. Ainsi, il est impératif de calculer les indices de sensibilité totaux. Malheureusement, ce calcul ne peut pas être effectué par la méthode FAST classique. Pour cette raison Saltelli et al. [63] ont proposé la méthode EFAST (Extended FAST), qui est une extension de la méthode FAST.

L'idée de base est de choisir des fréquences qui n'appartiennent pas à l'ensemble  $\{p_1\omega_1, p_2\omega_2, \dots, p_N\omega_N\}$  avec  $p_i = 1, 2, \dots, \infty$  et  $\forall i = 1, \dots, N$ . En effet, les fréquences autres que la fréquence  $\omega_i$  et ses harmoniques fournissent des informations concernant le résidu  $V[Y] - \sum_{i=1}^N V_{\omega_i}[Y]$  de la variance qui englobe les interactions entre les paramètres d'entrée. La procédure consiste dans un premier temps à affecter une fréquence  $\omega_i$  pour un paramètre  $x_i$ ; ensuite on attribue aux autres paramètres des fréquences  $\omega_{-i}$  (i.e.  $\omega_{-i} = \{\omega_k\}_{k=1, k \neq i}^N$ ) la plupart identiques mais différentes de  $\omega_i$ . Ainsi, en identifiant les coefficients de Fourier  $A_{p\omega_{-i}}$  et  $B_{p\omega_{-i}}$  attribués aux fréquences  $\omega_{-i}$  et leurs harmoniques  $p\omega_{-i}$ , la variance partielle  $V_{\omega_{-i}}$  peut être calculée par :

$$V_{\omega_{-i}}[Y] \approx 2 \sum_{j=1}^M (A_{j\omega_{-i}}^2 + B_{j\omega_{-i}}^2) \quad (I.101)$$

L'indice de sensibilité total  $S_{T_i} = V_{\omega_{-i}}[Y]/V[Y]$ , associé au paramètre  $x_i$ , est donné par :

$$S_{T_i} = 1 - \frac{\sum_{j=1}^M (A_{j\omega_{-i}}^2 + B_{j\omega_{-i}}^2)}{\sum_{j=1}^M (A_{j\omega_i}^2 + B_{j\omega_i}^2)} \quad (I.102)$$

Dans la pratique, on choisit la fréquence  $\omega_i$  très élevée par rapport aux fréquences  $\omega_{-i}$  attribuées aux autres paramètres. A titre d'exemple, on peut choisir  $\omega_i = 20$  et  $\omega_{-i} = 1$ . Les composantes du spectre correspondant aux harmoniques  $\{20, 40, 60, \dots\}$  contribuent à l'estimation de la variance  $V_{\omega_i}[Y]$  associée à la contribution de l'aléa relatif au paramètre  $x_i$ , et leurs amplitudes diminuent pour les harmoniques les plus élevées. La variance partielle

$V_{\omega_i}[Y]$  est estimée à partir seulement de quelques composantes du spectre correspondant aux harmoniques des fréquences  $\omega_{-i}$ . Les informations concernant la grandeur  $V_{i,\omega_i}[Y]$ , qui mesure la contribution de l'interaction entre le paramètre  $x_i$  et le reste des paramètres  $x_{-i}$ , peuvent être collectées à partir des composantes associées aux autres fréquences dans l'intervalle  $[1, M\omega_{max}]$ , où  $M$  est le nombre d'harmoniques (i.e. l'ordre de troncature) et  $\omega_{max}$  est la maximale des fréquences  $\{\omega_i\}_{i=1}^N$ . Saltelli et al. [63] ont montré que ce choix des fréquences peut mener à une surestimation des variances  $V_{\omega_i}[Y]$ , et ils ont proposé un algorithme automatique performant pour générer ces fréquences. Cet algorithme est défini par les étapes suivantes :

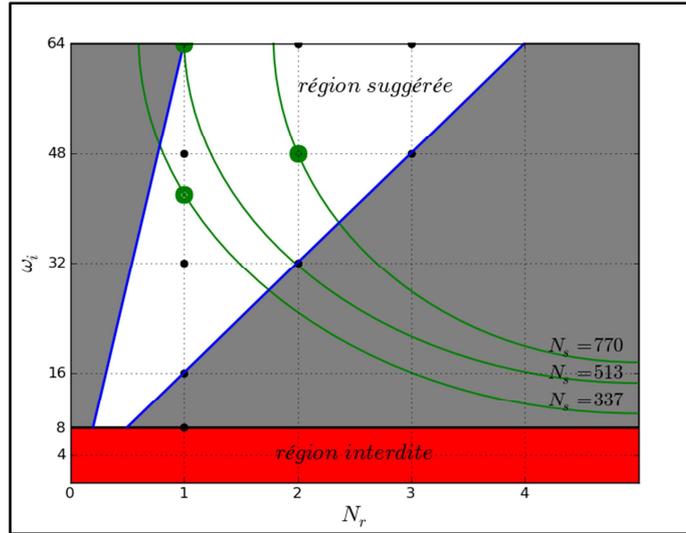
- a. Choisir une fréquence  $\omega_i$  pour le paramètre  $x_i$ , qui doit tenir compte du nombre optimal de simulations (équation (I.94)) nécessaire pour le calcul des indices de sensibilité.
- b. Définir une borne maximale pour les fréquences  $\omega_{-i}$  à attribuer pour les autres paramètres, telle que  $\max\{\omega_{-i}\} = \frac{\omega_i}{2M}$ .
- c. Générer les fréquences  $\omega_{-i}$  dans l'intervalle  $[1, M\omega_{max}]$  telles que le pas  $\Delta$  entre les fréquences soit aussi grand que possible et le nombre de paramètres  $x_i$  auquel on a attribué la même fréquence soit faible.

Le tableau I.4 donne les fréquences générées par cet algorithme dans le cas de 8 paramètres d'entrée, pour différents nombres de simulations  $N_s$  et en supposant qu'on s'intéresse au paramètre  $x_4$ . On constate que, pour un nombre  $N_s$  limité de simulations (les trois premières lignes du tableau I.4), la valeur de la fréquence maximale  $\max\{\omega_{-i}\}$  permise pour les paramètres complémentaires (i.e. autres que le paramètre  $x_4$ ) est inférieure au nombre de paramètres  $N = 8$ . Ainsi, plusieurs paramètres peuvent avoir la même fréquence. On constate aussi que le nombre de simulations  $N_s = 65$  peut être interprété comme une borne inférieure imposée par le choix d'une fréquence  $\omega_i$ .

**Tableau I. 4** : Fréquences générées par l'algorithme proposé par Saltelli et al. [52]

$N_s$	$\omega_i$	$\max\{\omega_{-i}\}$	$\Delta$	$\{\omega_i\}_{i=1}^N$
65	8	1	0	{1, 1, 1, 8, 1, 1, 1, 1}
129	16	2	1	{1, 2, 1, 16, 1, 2, 1, 2}
257	32	4	1	{1, 2, 3, 32, 1, 2, 3, 4}
513	64	8	1	{1, 2, 3, 64, 5, 6, 7, 8}
1025	128	16	2	{1, 3, 5, 128, 9, 11, 13, 15}

Dans le but d'obtenir de bonnes estimations, Saltelli et al. [52] ont proposé de calculer les indices de sensibilité non pas à partir d'un seul échantillon de points mais à partir de plusieurs échantillons de points générés à partir de la fonction de transformation définie par l'équation (I.99). Cette alternative peut être coûteuse en terme d'appels au modèle. Ainsi, pour obtenir le meilleur compromis entre la précision des estimations et le coût de calcul, Saltelli et al. [52] ont suggéré, pour un nombre de simulations  $N_s$  donné, de choisir la fréquence  $\omega_i$  et le nombre d'échantillons  $N_r$  de telle manière que le rapport  $\omega_i/N_r$  soit compris entre 16 et 64. Ce choix est justifié par le fait que, si  $\omega_i$  est faible et  $N_r$  élevé, il en résulte un échantillon de points très dispersés, alors que dans le cas contraire (i.e.  $\omega_i$  élevée et  $N_r$  faible), il en résulte des points très proches. La figure I.17 illustre l'étendue du domaine permettant un choix optimal du couple de paramètres  $\omega_i$  et  $N_r$ . On note que pour un nombre de simulations  $N_s = 65$ , il est impératif que la fréquence  $\omega_i$  soit égale à 8 et  $N_r = 1$  (cf. équation (I.100)).

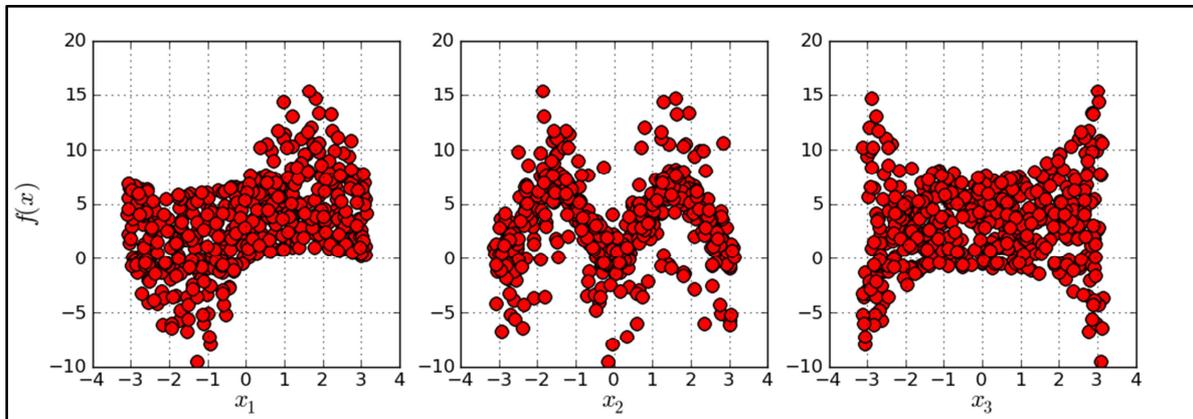


**Figure I. 17 :** Illustration de la région recommandée pour le choix des paramètres  $\omega_i$  et  $N_r$ , Saltelli et al. [52]

Dans ce qui suit, on se propose d'analyser l'efficacité de la méthode FAST et de sa version améliorée EFAST à travers un modèle simple. On considère la fonction d'Ishigami [64] suivante :

$$y = \sin(x_1) + a \sin(x_2)^2 + bx_3^4 \sin(x_1)$$

où  $a$  et  $b$  sont deux constantes réelles, et les paramètres  $x_i$ ,  $i = 1, 2, 3$  sont uniformément distribués sur l'intervalle  $[-\pi, \pi]$ . La figure I.18 représente la réponse  $y$  en fonction des paramètres  $x_1$ ,  $x_2$  et  $x_3$ .



**Figure I. 18 :** Evolution du modèle d'Ishigami en fonction des paramètres  $x_1$ ,  $x_2$  et  $x_3$

La particularité de la fonction d'Ishigami est qu'elle est non linéaire et non monotone. Cependant, la variance  $V$  de la réponse  $y$  du modèle, les variances partielles  $V_1$ ,  $V_2$  et  $V_3$  et les interactions  $V_{12}$ ,  $V_{13}$ ,  $V_{23}$  et  $V_{123}$  peuvent être calculées analytiquement. Elles sont définies par :

$$V = \frac{a^2}{8} + \frac{b\pi^4}{5} + \frac{b^2\pi^8}{18} + \frac{1}{2}$$

$$V_1 = \frac{b\pi^4}{5} + \frac{b^2\pi^8}{50} + \frac{1}{2}, V_2 = \frac{a^2}{8}, V_3 = 0$$

$$V_{12} = V_{23} = 0, V_{13} = \frac{8b^2\pi^8}{225}, V_{123} = 0$$

Les indices de sensibilité du premier ordre  $\{S_i = V_i/V\}_{i=1}^3$  et les indices de sensibilité totaux  $\{S_{T_i}\}_{i=1}^3$  sont calculés respectivement par les méthodes FAST et EFAST en choisissant  $a = 7$  et  $b = 0,1$ . On note que la méthode FAST est utilisée avec différents types de fonctions de transformation et que le calcul par la méthode EFAST est effectué pour différentes fréquences  $\omega_i$  et en utilisant la fonction de transformation I.99. Les résultats obtenus sont donnés dans le tableau I.5.

**Tableau I. 5 :** Indices de sensibilité obtenus par les méthodes FAST et EFAST

		<i>Indices de premier ordre</i>			<i>Indices totaux</i>			
		$S_1$	$S_2$	$S_3$	$S_{T_1}$	$S_{T_2}$	$S_{T_3}$	$N_s$
<i>Analytique</i>		0,3138	0,4424	0	0,5574	0,4424	0,2436	-
<i>FAST</i>	$x_i = \bar{x}_i e^{\bar{v}_i \sin(\omega_i s)}$	0,3775	0,0993	0,0140	-	-	-	121
	$x_i = \bar{x}_i [1 + \bar{v}_i \sin(\omega_i s)]$	0,4192	0,2917	0	-	-	-	121
	$x_i = \frac{1}{2} + \frac{1}{\pi} \arcsin(\omega_i s)$	0,3085	0,4725	0,0001	-	-	-	121
<i>EFAST</i>	$\omega_i = 16, N_r = 1$	0,3155	0,3046	0,0285	0,5435	0,3636	0,2407	129
	$\omega_i = 32, N_r = 1$	0,3091	0,4410	0,0259	0,5243	0,4871	0,2437	257
	$\omega_i = 48, N_r = 1$	0,3109	0,4438	0,0285	0,5372	0,4889	0,2383	385
	$\omega_i = 64, N_r = 1$	0,3124	0,4413	0,0161	0,5384	0,4869	0,2343	513

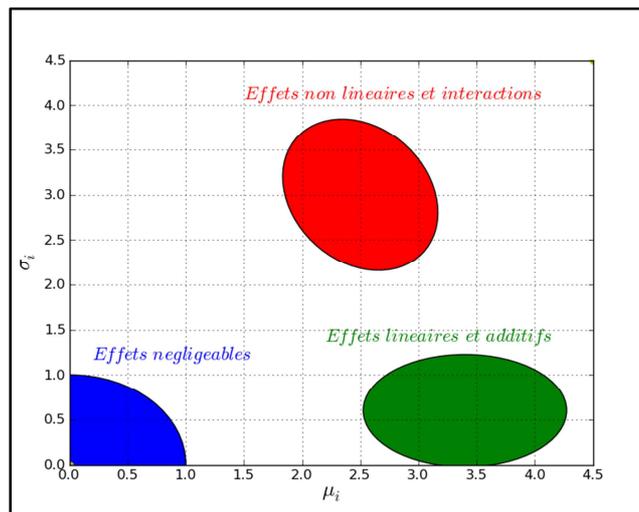
La méthode FAST, comme nous l’avons déjà évoqué, permet de calculer seulement les indices de sensibilité du premier ordre. On constate que les estimations sont en bon accord avec les résultats issus du calcul analytique, et ce seulement lorsqu’on emploie la fonction de transformation définie par l’équation (I.99). Ceci est justifié par le fait que l’emploi de ce type de fonction de transformation permet un meilleur recouvrement du domaine de définition des paramètres d’entrée. La méthode EFAST permet de calculer à la fois les indices de sensibilité du premier ordre et les indices de sensibilité totaux. On constate que, pour différentes fréquences  $\omega_i$ , les estimations correspondent aux résultats du calcul analytique. Cependant, on remarque une surestimation des indices de sensibilité du premier ordre pour les paramètres ayant une faible influence sur la réponse du modèle. Cela correspond au cas du paramètre  $x_3$  pour lequel l’indice de sensibilité du premier ordre doit être nul, tandis que, au meilleur des cas (i.e.  $\omega_i = 64$ ) la méthode EFAST donne une estimation de l’ordre de 1,5%.

L’application de la méthode FAST au modèle d’Ishigami, qui est non linéaire et non monotone, permet de conclure qu’elle peut être appliquée pour différents types de modèles. Grâce à sa version améliorée EFAST, elle permet, en plus des indices de sensibilité du premier ordre, d’estimer les indices de sensibilité totaux qui sont plus efficaces pour quantifier la contribution de l’aléa relatif à chaque paramètre d’entrée sur la variance totale de la réponse. Cette méthode est d’implémentation simple et assure un bon compromis entre le coût du calcul et la précision des estimations malgré la présence éventuelle d’un biais. Etant donné que la méthode n’est pas exacte, on souhaiterait avoir une mesure de l’intervalle de confiance sur les estimations, chose qui n’est pas possible puisque les estimations fournies par la méthode FAST sont déterministes pour un ensemble donné de fréquences.

#### I.5.2.4 Méthode de Morris

La méthode de sensibilité proposée par Morris [50] fait partie des méthodes dites de criblage telle que la méthode des plans d’expériences OAT (One-At-Time). Ces méthodes particulières d’analyse de sensibilité sont souvent désignées dans la littérature par le terme anglophone *screening methods*. Elles ne permettent qu’une évaluation qualitative de l’influence des

différents paramètres d'entrée d'un modèle sur sa réponse. Ainsi, elles sont la plupart du temps utilisées comme moyen de hiérarchisation des paramètres d'entrée selon leur ordre d'importance. Ceci s'avère très intéressant dans le cas de modèles faisant intervenir un nombre important de paramètres d'entrée, où l'application de ces méthodes peut être considérée comme une pré-étude de sensibilité permettant de déterminer les paramètres les plus influents, et ainsi de réduire le nombre de paramètres qui seront considérés par la suite dans l'analyse de sensibilité. Dans ce contexte, une des méthodes les plus évoluées est celle développée par Morris. Sa philosophie est de déterminer parmi tous les paramètres d'entrée ceux susceptibles d'avoir des effets : (a) négligeables, (b) linéaires et additifs ou (c) non linéaires ou en interaction avec d'autres paramètres. Cette hiérarchisation est basée sur la notion d'effet élémentaire, qui n'est autre, pour un paramètre donné, comme on le verra plus explicitement dans ce qui suit, que la mesure du gradient du modèle relatif à ce paramètre. Plus précisément, cette classification n'est pas effectuée directement au moyen des effets élémentaires mais à partir de leurs statistiques (i.e. la moyenne et l'écart-type) calculées à partir de plusieurs échantillons. En effet, pour un paramètre donné, une valeur élevée de la moyenne de l'effet élémentaire qui lui est associé indique l'importance de ce paramètre sur la réponse du modèle, alors qu'une valeur élevée de l'écart-type indique de fortes interactions avec d'autres paramètres. La figure I.19 illustre la procédure de hiérarchisation des paramètres d'entrée, sur laquelle est basée la méthode de Morris.



**Figure I. 19 :** Hiérarchisation des paramètres d'entrée d'un modèle par la méthode de Morris

Le calcul des effets élémentaires fait appel à un plan d'expérience de type OAT (One At Time) probabiliste. La particularité de ce plan d'expérience est qu'il n'est pas construit autour d'un point particulier telle que la valeur nominale, mais autour d'un point dont les composantes relatives à chaque paramètre sont tirées d'une manière aléatoire parmi une gamme de valeurs obtenues en discrétisant le domaine de définition de chaque paramètre en un nombre donné de niveaux, d'où la justification du terme probabiliste employé ci-dessus. Pour cette raison, certains auteurs [52] qualifient la méthode de Morris de méthode d'analyse de sensibilité globale bien qu'elle permette seulement une mesure qualitative de la sensibilité.

Dans le développement qui suit, en se référant aux pratiques usuelles en analyse de sensibilité, on considère que les paramètres d'entrée sont uniformément définis sur l'hyper-cube unitaire  $[0, 1]^N$ .

La première étape dans la méthode de Morris consiste à discrétiser chaque paramètre d'entrée sur son domaine de définition en un nombre donné de niveaux  $p$ . Compte tenu de l'hypothèse évoquée ci-dessus, une grille de valeurs  $\Omega = \{0, 1/(p - 1), 2/(p - 1), \dots, 1\}$  peut être

attribuée à chaque paramètre d'entrée. Soit  $\Delta$  une perturbation prédéfinie de manière qu'elle soit un multiple de  $1/(p - 1)$  et  $f$  la fonction générique représentant le modèle ; selon Morris [50], l'effet élémentaire  $EE_i$  associé à un paramètre  $x_i$  est défini par :

$$EE_i = \frac{f(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_N) - f(x)}{\Delta} \quad (I.103)$$

où  $\mathbf{x} = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_N)$  est un point de  $[0,1]$  choisi de telle manière que le point perturbé  $\mathbf{x} + \mathbb{1}_i \Delta = (x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_N)$  soit situé lui aussi dans l'intervalle  $[0,1]$ .

En effectuant plusieurs tirages du point  $\mathbf{x}$  parmi la grille de valeurs  $\Omega$ , on obtient, pour chaque paramètre d'entrée, une distribution finie notée  $F_i$  de l'effet élémentaire  $EE_i$ . Ainsi, chaque distribution  $F_i$  contient  $p^{N-1}[p - \Delta(p - 1)]$  éléments. La manière la plus simple de caractériser, dans un contexte statistique, la distribution  $F_i$  est de calculer sa moyenne et son écart-type. Ces grandeurs peuvent être estimées respectivement comme suit :

$$\mu_i = \frac{1}{N_d} \sum_{j=1}^{N_d} EE_i^j \quad (I.104)$$

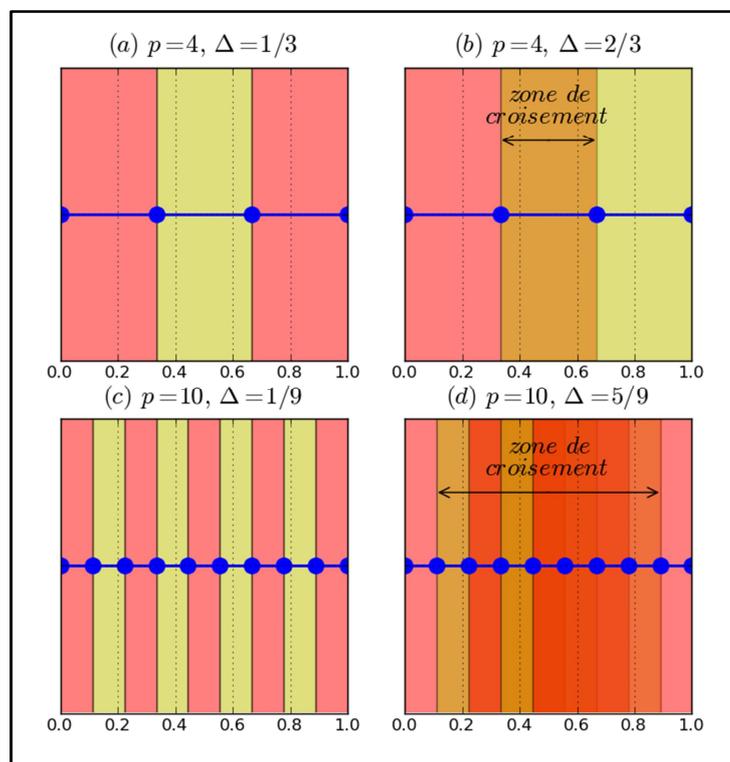
$$\sigma_i = \sqrt{\frac{1}{N_d} \sum_{j=1}^{N_d} (EE_i^j - \mu_i)^2} \quad (I.105)$$

où  $N_d$  est le nombre d'éléments dans la distribution  $F_i$ .

En suivant la procédure décrite ci-dessus, le nombre total d'appels au modèle est  $N_s = 2N_d N$  : chaque effet élémentaire  $EE_i$  nécessite l'évaluation du modèle deux fois. Afin d'évaluer l'efficacité de la procédure d'expérimentation utilisée, Morris [50] a introduit une mesure qu'il a nommé *économie*. Elle est définie comme le rapport entre le nombre d'effets élémentaires à déterminer et le nombre d'appels au modèle. Plus cette mesure est élevée, plus la procédure d'expérimentation est dite informative (i.e. plus elle fournit d'informations sur la sensibilité du modèle vis-à-vis des paramètres d'entrée). Ainsi, l'économie de la procédure d'expérimentation décrite auparavant est  $N_d N / 2N_d N = 1/2$ .

Dans le but de réduire le coût du calcul (i.e. le nombre d'appels au modèle), Morris [50] a proposé une procédure plus économique que celle décrite ci-dessus. Elle est basée sur la construction d'une matrice d'expérimentation notée  $\mathbf{B}^*$  constituée de  $N + 1$  lignes et  $N$  colonnes, où chaque ligne représente une réalisation du vecteur  $\mathbf{x}$ . La matrice d'expérimentation  $\mathbf{B}^*$ , désignée souvent par *matrice d'orientation* ou *trajectoire*, a la propriété intéressante suivante : pour toute colonne  $j = 1, 2, \dots, N$ , seules deux lignes de la matrice  $\mathbf{B}^*$  ont des composantes différentes. Grâce à cette propriété, les  $N + 1$  lignes de la matrice  $\mathbf{B}^*$ , représentant  $N + 1$  réalisations distinctes du vecteur  $\mathbf{x}$ , fournissent  $N + 1$  valeurs de la réponse du modèle qui permettent de calculer  $N$  effets élémentaires (i.e. un effet élémentaire pour chaque paramètre d'entrée). Ainsi, l'économie de cette procédure d'expérimentation est de  $N/N + 1$ , qui est meilleure que celle de la procédure d'expérimentation originale. Dans la construction d'une telle procédure d'expérimentation (i.e. la construction de la matrice d'orientation  $\mathbf{B}^*$ ), Morris [50] a postulé qu'il est plus convenable de choisir un nombre de niveaux  $p$  pair et une perturbation  $\Delta = p/2(p - 1)$ . En effet, ce choix permet de s'assurer que chacun des  $p^{N-1}[p - \Delta(p - 1)] = p^N/2$  effets élémentaires associés à un paramètre d'entrée a la même probabilité d'être sélectionné. Bien que ce choix de la perturbation  $\Delta$  soit adopté dans la plupart des études rencontrées dans la littérature, il a été critiqué par Van Houwelingen et al. [65]. En effet, en choisissant  $\Delta =$

$p/2(p - 1)$ , le domaine de définition de chaque paramètre d'entrée est divisé en  $p/2$  sous domaines qui se croisent entre eux. Le croisement entre deux domaines est d'autant plus marqué que le nombre de niveaux est élevé. Cependant, l'exploration du domaine de définition des paramètres d'entrée est optimale si ce croisement n'existe pas. Van Houwelingen et al. [65] ont montré, pour la fonction simple  $y = \cos(x)$  avec  $x \in [-\pi, 2\pi]$  et en choisissant un nombre de niveaux  $p = 4$ , que la perturbation  $\Delta = p/2(p - 1) = 2/3$  ne permet pas de détecter le caractère non monotone de la fonction ; or une perturbation  $\Delta = 1/(p - 1) = 1/3$  le permet. En effet, cette alternative permet d'éliminer le croisement entre les différentes subdivisions, comme on peut l'observer sur la figure I.20. En choisissant  $p = 4$  et  $\Delta = p/2(p - 1) = 2/3$ , l'intervalle  $[0, 1]$  est discrétisé en deux subdivisions  $[0, 2/3]$  et  $[1/3, 1]$  (cf. figure I.20b), qui se chevauchent sur  $[1/3, 2/3]$ . Par contre, si  $\Delta = 1/(p - 1) = 1/3$ , il en résulte trois subdivisions distinctes  $[0, 1/3]$ ,  $[1/3, 2/3]$  et  $[2/3, 1]$ , qui ne se croisent pas (cf. figure I.20a). De plus, on constate que plus le nombre de niveaux est élevé, plus la zone de croisement entre les subdivisions est large (cf. figure I.20d).



**Figure I. 20 :** Illustration de l'importance du choix de la perturbation sur le recouvrement du domaine de définition des paramètres d'entrée [65]

Afin de construire une matrice d'orientation  $\mathbf{B}^*$  vérifiant les propriétés décrites ci-dessus, la première étape consiste à construire une matrice triangulaire inférieure  $\mathbf{B}$  de dimension  $(N + 1) \times N$ . Ensuite, on construit une matrice  $\mathbf{B}'$  définie par :

$$\mathbf{B}' = \mathbf{J}_{N+1,N} \cdot \mathbf{x}^* + \Delta \cdot \mathbf{B} \quad (I. 106)$$

où  $\mathbf{J}_{N+1,N}$  est une matrice de dimension  $(N + 1) \times N$ , telle que  $\forall i = 1, \dots, N + 1$  et  $j = 1, \dots, N$ ,  $J_{i,j} = 1$ ,  $\Delta$  est une perturbation et  $\mathbf{x}^*$  est un point de référence dont les composantes  $x_i$ ,  $i = 1, \dots, N$ , sont tirées d'une manière aléatoire (i.e. en effectuant des permutations aléatoires) de l'ensemble  $\{0, 1/(p - 1), \dots, 1 - \Delta\}$ . La matrice  $\mathbf{B}'$  pourrait être utilisée comme matrice d'expérimentation, puisqu'elle permet de calculer un effet élémentaire pour chaque paramètre d'entrée. Cependant, ces effets élémentaires ne correspondent pas à un tirage

aléatoire des distributions  $\{F_i\}_{i=1}^N$ . Dans le but de surmonter ce problème, une version probabilisée de la matrice d'expérimentation peut être construite comme suit :

$$\mathbf{B}^* = (\mathbf{J}_{N+1,1} \cdot \mathbf{x}^* + (\Delta/2)[(2 \cdot \mathbf{B} - \mathbf{J}_{N+1,N}) \cdot \mathbf{D}^* + \mathbf{J}_{N+1,N}]) \cdot \mathbf{P}^* \quad (I.107)$$

où  $\mathbf{D}^*$  est une matrice carrée de dimension  $N$  telle que ses éléments sont des entiers choisis dans l'ensemble  $\{-1,1\}$  d'une manière équiprobable, et  $\mathbf{P}^*$  est une matrice carrée de permutation aléatoire de dimension  $N$ , obtenue par la permutation aléatoire de la matrice identité de dimension  $N$ . On peut constater que chaque matrice d'orientation permet de calculer seulement un effet élémentaire par paramètre d'entrée. Si  $N_d$  est le nombre de simulations nécessaire pour estimer les statistiques de la distribution  $F_i$  de l'effet élémentaire associé à un paramètre d'entrée, la procédure d'expérimentation nécessite la construction de  $N_d$  matrices d'orientation  $\mathbf{B}^*$ . Ainsi, le nombre total d'appels au modèle est  $N_s = N_d(N + 1)$ . Morris [50] préconise d'employer un nombre de niveaux  $p = 4$  et un nombre de trajectoires (i.e. matrices d'orientation)  $N_d = 10$  pour obtenir un bon compromis entre le coût de calcul et la précision des estimations des statistiques définies par les équations (I.104) et (I.105). L'organigramme de la figure I.21 résume la procédure d'expérimentation proposée par Morris pour hiérarchiser les paramètres d'entrée d'un modèle.

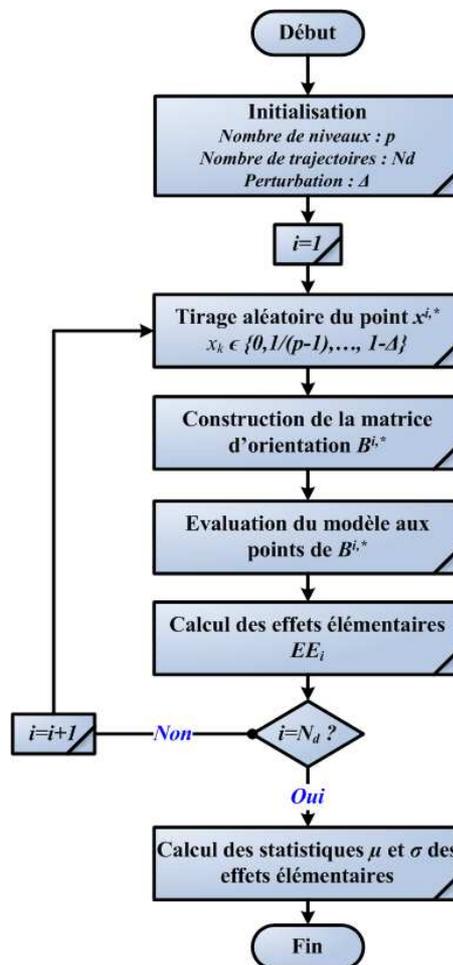


Figure I. 21 : Organigramme de la méthode de Morris

L'avantage principal de la méthode de Morris est qu'elle permet d'évaluer qualitativement la sensibilité de la réponse d'un modèle vis-à-vis des différents paramètres d'entrée tout en nécessitant un coût de calcul raisonnable. Cependant, son inconvénient majeur est qu'elle ne permet pas de dissocier les effets des interactions relatives à chacun des paramètres. En effet,

elle fournit une mesure globale de l'effet d'interaction d'un paramètre avec les autres paramètres, mais elle ne donne pas d'information spécifique sur l'identité de cette interaction.

Depuis son introduction, la méthode de Morris a été le sujet d'améliorations, concernant principalement son efficacité, telles que l'introduction de nouvelles mesures de sensibilité et l'emploi d'autres stratégies de génération des points d'expérimentation.

Lorsque le modèle est non monotone, sa réponse peut osciller dans des régions de valeurs positives ou négatives. Par conséquent, les effets élémentaires peuvent avoir des valeurs positives ou négatives. Dans ce cas, la moyenne des effets élémentaires peut être nulle ou proche de zéro du fait de l'annulation des effets élémentaires entre eux. Dans ce contexte, Campolongo et al. [66] ont proposé de calculer la moyenne sur les valeurs absolues des effets élémentaires afin d'annuler l'effet des signes. Cette mesure notée  $\mu_i^*$  est définie par :

$$\mu_i^* = \frac{1}{N_d} \sum_{j=1}^{N_d} |EE_i^j| \quad (I.108)$$

Les mêmes auteurs ont proposé une amélioration de la technique de génération des points d'expérimentation. En effet, l'estimation des mesures de sensibilité (i.e. statistiques des effets élémentaires) est basée sur la construction d'un nombre  $N_d$  de matrices d'orientation ou trajectoires (généralement 10 à 50 trajectoires sont choisies). La technique employée pour la construction d'une trajectoire est de choisir dans un premier temps un point aléatoire de référence ; ensuite les autres points d'expérimentation sont obtenus en perturbant une seule composante à la fois dans un ordre aléatoire. Cette stratégie peut mener à un recouvrement non optimal du domaine de définition des paramètres d'entrée, particulièrement lorsque le nombre de paramètres du modèle est élevé. L'idée proposée par Campolongo et al. [66] est de choisir les  $N_d$  trajectoires de telle manière que la dispersion entre elles soit maximale. Dans un premier temps, on construit un nombre  $N_T \gg N_d$  (ex.  $N_T \sim 500 - 1000$ ) de trajectoires, ensuite on choisit parmi celles-ci les  $N_d$  trajectoires qui assurent la meilleure dispersion. Ce concept de dispersion est basé sur le calcul de la distance  $d_{ml}$  entre un couple de trajectoires distinctes  $m$  et  $l$ , qui est définie par :

$$d_{ml} = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} \sqrt{\sum_{k=1}^N [x_i^m(k) - x_j^l(k)]^2} \quad (I.109)$$

où  $x_i^m(k)$  est la  $k^{\text{ème}}$  composante du  $i^{\text{ème}}$  point de la trajectoire  $m$ . Autrement dit,  $d_{ml}$  est la somme des distances géométriques entre tous les couples de points de deux trajectoires différentes. Ainsi, les  $N_d$  meilleures trajectoires sélectionnées sont celles qui ont les distances  $d_{ml}$  maximales entre les couples de trajectoires parmi  $N_T$ . Dans un premier temps, on choisit  $N_d$  trajectoires parmi  $N_T$ , et on calcule la distance  $D$  qui est la somme des distances  $d_{ml}$  entre les différents couples de trajectoires de cette combinaison. A titre illustratif, si le nombre de trajectoires à sélectionner est  $N_d = 4$  et le nombre total de trajectoires est  $N_T = 10$ , et si on s'intéresse à la combinaison formée par les trajectoires  $\{4, 6, 7, 9\}$ , on définit la distance

$D_{4,6,7,9}$  telle que  $D_{4,6,7,9} = \sqrt{d_{4,6}^2 + d_{4,7}^2 + d_{4,9}^2 + d_{6,7}^2 + d_{6,9}^2 + d_{7,9}^2}$ . Ensuite, parmi toutes les combinaisons possibles, on choisit celle dont la distance  $D$  est maximale. L'inconvénient de cette stratégie de sélection des trajectoires, est purement technique. En effet, si le nombre total de trajectoires  $N_T$  est très élevé, ce qui est souhaitable pour obtenir une grande gamme de choix des  $N_d$  trajectoires, les calculs effectués exigent une importante capacité de mémoire. Cette stratégie de construction des points d'expérimentation peut être remplacée par une

procédure plus simple, qui fournit les mêmes résultats (i.e. le recouvrement optimal du domaine de définition des paramètres d'entrée). Elle consiste à générer les points de référence des différentes trajectoires en utilisant la technique de tirage en hyper-cube latin.

Bien que le coût de calcul exigé par la méthode de Morris soit relativement faible par rapport à d'autres méthodes d'analyse de sensibilité, le nombre d'appels au modèle reste important dans le cas où le nombre de paramètres d'entrée est élevé. Dans ce contexte, l'efficacité de la méthode de Morris peut être améliorée en travaillant avec des groupes de paramètres, tout en gardant la même procédure d'expérimentation. Le choix des groupes peut être effectué, sur la base d'informations préalablement disponibles (par exemple des observations expérimentales), sur le comportement des paramètres du modèle. Cela est possible aussi, en effectuant un calcul préliminaire avec un nombre limité de trajectoires pour déterminer les paramètres qui occupent, a priori, la même zone d'importance. Puisque le nombre de groupes  $N_G$  sera inférieur au nombre de paramètres  $N$ , le nombre d'appels au modèle diminue et il est alors  $N_s = N_d(N_G + 1)$ . Ainsi, les effets élémentaires ne sont pas calculés pour chaque paramètre mais pour chaque groupe de paramètres. L'emploi de cette procédure est intéressant si on connaît préalablement les groupes de paramètres ou si on dispose de critères de groupement. Le seul risque dans le groupement de paramètres est de noyer les effets élémentaires individuels associés à chaque paramètre, ce qui peut mener à de fausses interprétations concernant la sensibilité de la réponse du modèle vis-à-vis des différents paramètres d'entrée. En plus des améliorations concernant l'efficacité et la précision, la méthode de Morris a été développée une fois de plus par Campolongo et Braddock [67] pour pouvoir estimer les effets d'interaction entre les couples de paramètres. Dans ce contexte (i.e. l'évaluation de l'effet des interactions), la méthode a été analysée par Cropp et Braddock [68] sur un cas test sur lequel ils ont prouvé son efficacité.

On se propose dans ce qui suit d'illustrer au moyen d'une application numérique les différents aspects de la procédure de calcul des effets élémentaires proposée par Morris. On considère un modèle analytique contenant 20 paramètres d'entrée, défini par [50] :

$$y = \beta_0 + \sum_{i=1}^{20} \beta_i \omega_i + \sum_{i<j}^{20} \beta_{ij} \omega_i \omega_j + \sum_{i<j<k}^{20} \beta_{ijk} \omega_i \omega_j \omega_k + \sum_{i<j<k<l}^{20} \beta_{ijkl} \omega_i \omega_j \omega_k \omega_l$$

avec  $\omega_i = 2(x_i - 0,5)$  sauf pour  $i = 3, 5$  et  $7$  où  $\omega_i = 2(1,1x_i/(x_i + 1) - 0,5)$ . Les paramètres  $\{x_i\}_{i=1}^{20}$  sont uniformément distribués sur l'intervalle  $[0,1]$ . Les coefficients  $\beta_i$ ,  $\beta_{ij}$ ,  $\beta_{ijk}$  et  $\beta_{ijkl}$  sont définis comme suit :

$$\beta_i = 20, \quad i = 1, \dots, 10$$

$$\beta_{ij} = -15, \quad i, j = 1, \dots, 6$$

$$\beta_{ijk} = -10, \quad i, j, k = 1, \dots, 5$$

$$\beta_{ijkl} = 5, \quad i, j, k, l = 1, \dots, 4$$

Les coefficients d'ordre 1 et d'ordre 2 restant sont obtenus à partir de tirages indépendants suivant une loi Gaussienne standard, alors que les coefficients d'ordre 3 et d'ordre 4 restant sont nuls.

Les caractéristiques statistiques  $\mu_i$ ,  $\sigma_i$  et  $\mu_i^*$  sont estimées à partir de 10 trajectoires (i.e. matrices d'expérimentation) obtenues par la procédure décrite auparavant en choisissant un nombre de niveau  $p = 4$  pour la discrétisation des domaines de définition des 20 paramètres d'entrée (ce choix correspond à une perturbation  $\Delta = 2/3$ ). Ainsi, le nombre total d'appels au modèle est 210.

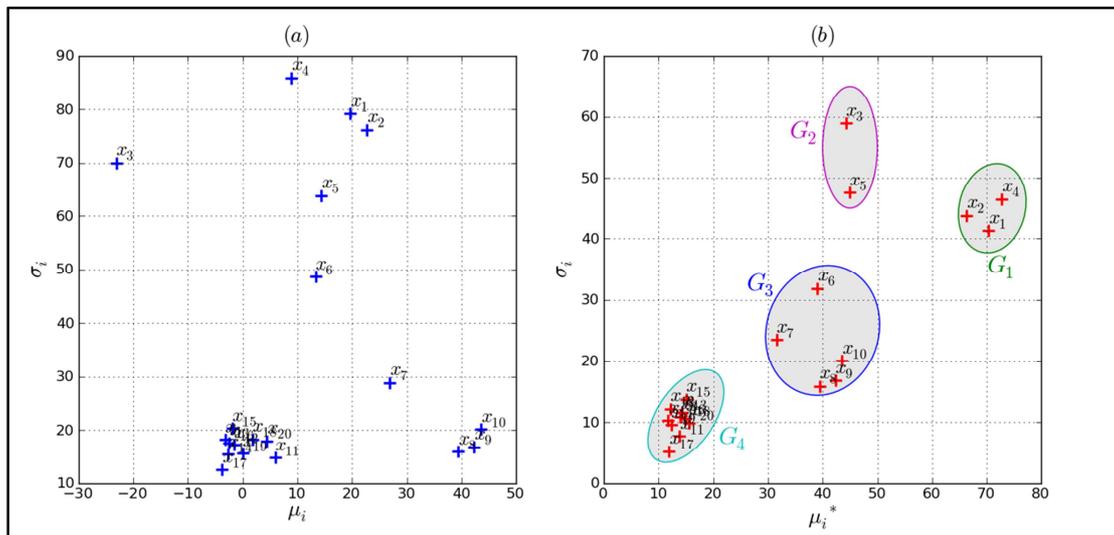
**Tableau I. 6 :** Statistiques  $\mu_i^*$  et  $\sigma_i$  des paramètres d'entrée pour le modèle de Morris

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
$\mu_i^*$	70,35	66,30	44,23	72,76	44,88	39,02	31,58	39,39	42,31	43,55
$\sigma_i$	41,31	43,77	58,93	46,48	47,57	31,83	23,48	15,87	16,81	20,10
ordre	2	3	5	1	4	9	10	8	7	6

---

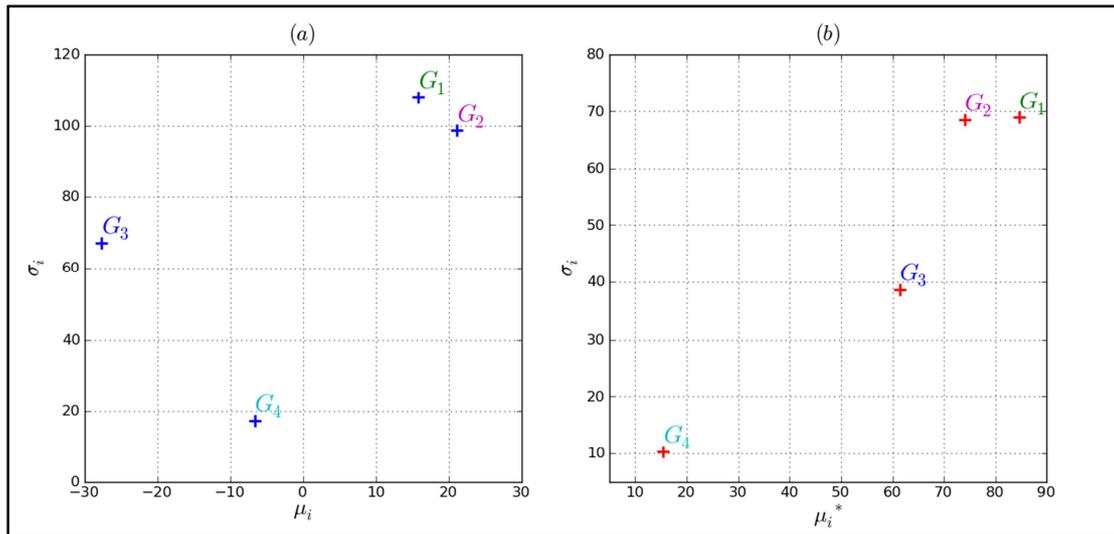
	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$	$x_{16}$	$x_{17}$	$x_{18}$	$x_{19}$	$x_{20}$
$\mu_i^*$	13,87	12,14	14,28	11,75	15,00	14,01	11,92	14,93	12,28	15,57
$\sigma_i$	7,68	12,23	11,65	10,30	13,76	10,94	5,23	10,64	9,56	9,90
ordre	16	18	14	20	12	15	19	13	17	11

Le tableau I.6 donne les valeurs de la moyenne  $\mu_i^*$  et de l'écart-type  $\sigma_i$  des effets élémentaires, ainsi que l'ordre d'importance des paramètres d'entrée (l'ordre d'importance est basé sur la mesure de sensibilité  $\mu_i^*$ ). En se référant à la valeur de la moyenne  $\mu_i^*$  des effets élémentaires, on constate que les paramètres  $\{x_1, \dots, x_{10}\}$  sont les plus influents sur la réponse du modèle, alors que, si on considère l'écart-type  $\sigma_i$  des effets élémentaires, on constate que les paramètres  $\{x_1, x_2, x_3, x_4, x_5\}$  contribuent à des effets de non linéarité et d'interaction.



**Figure I. 22 :** Hiérarchisation des différents paramètres du modèle basée sur les mesures (a)  $\mu_i$  et  $\sigma_i$ , (b)  $\mu_i^*$  et  $\sigma_i$  pour le modèle de Morris

Afin de mieux interpréter les résultats obtenus, les figures I.22a et I.22b illustrent respectivement la moyenne  $\mu_i$  et l'écart-type  $\sigma_i$ , et la moyenne  $\mu_i^*$  et l'écart-type  $\sigma_i$  des effets élémentaires associés à chaque paramètre d'entrée. Le but de fournir les deux présentations des résultats est de mettre en évidence l'importance d'utiliser la mesure  $\mu_i^*$  au lieu de la mesure  $\mu_i$  pour hiérarchiser les paramètres du modèle. En effet, on constate, d'après la figure I.22b, qu'on peut mieux distinguer les paramètres appartenant à chaque zone de hiérarchisation (i.e. zone d'effets négligeables, zone d'effets linéaires et additifs et zone d'effets non linéaires et d'interactions). De plus, on peut observer que les paramètres de faibles influences sont nettement séparés des autres paramètres. Ils sont situés au voisinage de l'origine du référentiel formé par les mesures  $\mu_i^*$  et  $\sigma_i$ . En se basant sur ces résultats, les 20 paramètres d'entrée de départ peuvent être réduits à 10.



**Figure I. 23 :** Hierarchisation des groupes de paramètres  $G_1 = \{x_1, x_2, x_4\}$ ,  $G_2 = \{x_3, x_5\}$ ,  $G_3 = \{x_6, x_7, x_8, x_9, x_{10}\}$  et  $G_4 = \{x_{11}, \dots, x_{20}\}$  du modèle basée sur les mesures (a)  $\mu_i$  et  $\sigma_i$ , (b)  $\mu_i^*$  et  $\sigma_i$

Mieux encore, en se référant à la figure I.22b, les paramètres les plus influents peuvent être eux aussi regroupés en trois groupes distincts  $G_1 = \{x_1, x_2, x_4\}$ ,  $G_2 = \{x_3, x_5\}$  et  $G_3 = \{x_6, x_7, x_8, x_9, x_{10}\}$ . Ainsi, on peut envisager par la suite l'utilisation des méthodes d'analyse de sensibilité permettant une mesure quantitative de la sensibilité telle que les indices de sensibilité de Sobol, en ne considérant que trois groupes de paramètres et non pas tous les paramètres d'entrée. Cette idée est motivée par le fait que le nombre d'appels au modèle peut être significativement réduit en choisissant la procédure de calcul par groupes de paramètres. En effet, le nombre d'évaluations du modèle passe de 210 à 50. On constate la grande correspondance entre les zones de hiérarchisations auxquelles appartiennent les différents paramètres du modèle, obtenues respectivement par un calcul sans regroupement (figures I.22a et I.22b) et un calcul avec regroupement (figures I.23a et I.23b) des paramètres du modèle. Ainsi, on conclut que la méthode de Morris fournit un moyen de diagnostic qui permet de détecter les paramètres les moins influents sur la réponse du modèle.

### I.5.2.5 Mesures de sensibilité basées sur les dérivées partielles

La façon la plus simple pour mesurer la sensibilité de la réponse d'un modèle vis-à-vis de ses paramètres d'entrée est de calculer les dérivées partielles (i.e. les dérivées par rapport aux variables représentant les paramètres d'entrée) de la fonction générique  $f$  représentant le modèle. Dans la pratique, la fonction  $f$  est rarement disponible sous forme explicite. En effet, elle est souvent issue d'une modélisation numérique (par exemple un modèle éléments finis). Dans ce cas, différentes méthodes sont disponibles pour déterminer les dérivées partielles. Par exemple, par différences finies, la dérivée partielle  $E_i(x^*)$  associée à un paramètre d'entrée, évaluée au point de référence  $x^*$ , est donnée par :

$$E_i(x^*) = \left. \frac{\partial f}{\partial x_i} \right|_{x_i=x^*} \approx \frac{f(x^* + \Delta) - f(x^*)}{\Delta}, \quad i = 1, \dots, N \quad (I.110)$$

où  $\Delta$  est une perturbation.

L'avantage de cette méthode est qu'elle ne nécessite pas de modification du modèle original. En effet, le calcul d'une dérivée partielle fait appel aux évaluations du modèle au point  $x^*$  et au point  $x^* + \Delta$ . Ainsi, on a besoin à  $N + 1$  appels au modèle pour obtenir les  $N$  mesures de sensibilité  $\{E_i\}_{i=1}^N$  associées aux différents paramètres d'entrée. La précision de ces mesures

dépend principalement de la perturbation  $\Delta$ . Le choix de l'amplitude de la perturbation  $\Delta$  doit garantir que le modèle est quasi-linéaire entre le point  $x^*$  et le point  $x^* + \Delta$ . Dans le cas où le modèle est non monotone, si on considère une perturbation  $\Delta$  relativement élevé ( $> 5\%$ ), on risque de sortir du cadre de l'hypothèse de linéarité au voisinage du point  $x^*$ . Par contre, si l'amplitude de la perturbation est très faible, l'écart entre les réponses du modèle calculées aux points  $x^*$  et  $x^* + \Delta$  sera lui aussi faible, et par conséquent l'erreur d'arrondi sera importante. Dans la plupart des applications, la perturbation  $\Delta$  est prise égale à 1%.

Les mesures de sensibilité  $\{E_i\}_{i=1}^N$  dépendent aussi du choix du point de référence  $x^*$ . En effet, en changeant le point de référence, ces mesures de sensibilité changent. Pour cette raison, les quantités  $\{E_i\}_{i=1}^N$  sont désignées par mesures de sensibilité locales, et les informations qu'elles fournissent sont très limitées. Ainsi, dans la pratique, la mesure de sensibilité  $E_i$  définie par l'équation (I.110) est normalisée le plus souvent par l'écart-type  $\sigma_{x_i}$  du paramètre  $x_i$  auquel elle est associée et l'écart-type  $\sigma_Y$  de la réponse du modèle. On peut écrire alors :

$$E_i^{norm}(x^*) = \frac{\sigma_{x_i}}{\sigma_Y} E_i(x^*), \quad i = 1, \dots, N \quad (I.111)$$

Cette mesure est souvent désignée par mesure hybride du fait de son caractère local donné par la mesure  $E_i(x^*)$  et de son caractère global introduit par les paramètres  $\sigma_{x_i}$  et  $\sigma_Y$ . Son avantage est de quantifier la part de l'aléa relatif à chaque paramètre d'entrée dans la variance totale  $V_Y = \sigma_Y^2$  de la réponse du modèle.

Une autre façon de procéder, afin que  $E_i(x^*)$  soit une mesure de sensibilité globale sans passer par la procédure de normalisation décrite ci-dessus, est de calculer les statistiques (i.e. moyenne et écart-type) de cette mesure à partir de plusieurs réalisations (i.e. calcul sur plusieurs points de référence) comme pour le cas de la méthode de Morris. Si on suppose que l'ensemble des paramètres d'entrée est défini sur l'hyper-cube unitaire  $[0, 1]^N$ , la moyenne  $\mu_i$  et l'écart-type  $\sigma_i$  de la mesure  $E_i(x^*)$  sont définis par :

$$\mu_i = \int_{[0,1]^N} E_i dx \quad (I.112)$$

$$\sigma_i = \sqrt{\int_{[0,1]^N} (E_i - \mu_i)^2 dx} \quad (I.113)$$

Ainsi, la sensibilité de la réponse du modèle vis-à-vis du paramètre  $x_i$  n'est pas mesurée directement au moyen de la dérivée partielle  $E_i$  mais par sa moyenne  $\mu_i$  et son écart-type  $\sigma_i$ .

En combinant ces deux mesures (i.e.  $\mu_i$  et  $\sigma_i$ ), Sobol et Gresham [69] ont proposé une autre mesure de sensibilité qu'ils ont appelé *mesure alternative de sensibilité globale* ; elle est définie par :

$$G_i^* = \frac{G_i}{\sum_{i=1}^N G_i} \quad (I.114)$$

où  $G_i = \mu_i^2 + \sigma_i^2$ .

De manière similaire à la méthode de Morris, le problème d'annulation de l'effet des mesures  $E_i$  entre elles lorsqu'on calcule la moyenne ou l'écart-type existe. Ainsi, Kucherenko et al. [70] ont proposé de calculer ces quantités en considérant les valeurs absolues des mesures  $E_i$ .

$$\mu_i^* = \int_{[0,1]^N} |E_i| dx \quad (I.115)$$

$$\sigma_i^* = \sqrt{\int_{[0,1]^N} (|E_i| - \mu_i)^2 dx} \quad (I.116)$$

On note que ce problème ne se pose pas pour la mesure de sensibilité  $G_i^*$  définie par l'équation (I.114). Les mesures de sensibilité  $\mu_i^*$ ,  $\sigma_i^*$  et  $G_i^*$  peuvent être obtenues au moyen de simulations de Monte-Carlo basées sur des séquences de points aléatoires ou quasi-aléatoires.

Afin de justifier la validité de ces mesures de sensibilité, Sobol et Kucherenko [71] ont montré qu'elles ont un lien avec les indices de sensibilité totaux introduit par Sobol. Dans un premier lieu, ils ont défini un critère de mesure d'importance  $\nu_i$ ,  $i = 1, \dots, N$ , basé sur le carré de la dérivée partielle  $\partial f / \partial x_i$  :

$$\nu_i = \int_{[0,1]^N} \left( \frac{\partial f}{\partial x_i} \right)^2 dx \quad (I.117)$$

Ils ont montré que si la valeur de  $\nu_i$  est faible, il s'ensuit que le paramètre  $x_i$  auquel il est associé n'a pas d'influence sur la réponse du modèle. De plus, ils ont montré que  $\nu_i$  correspond à une borne supérieure de l'indice de sensibilité total  $S_{T_i}$  :

$$S_{T_i} \leq \frac{\nu_i}{\pi V_Y} \quad (I.118)$$

où  $V_Y$  est la variance totale de la réponse du modèle.

Par conséquent, il est clair, à partir des équations (I.117) et (I.118), que la dérivée partielle  $\partial f / \partial x_i$  associée à un paramètre d'entrée  $x_i$  peut fournir une mesure de sensibilité fiable (i.e. elle peut donner une image complète sur la sensibilité de la réponse du modèle vis-à-vis des différents paramètres d'entrée). Cependant, pour les modèles fortement non linéaires, l'emploi des dérivées partielles comme mesures de sensibilité peut mener à de fausses interprétations.

### I.5.3 Etude comparative des méthodes d'analyse de sensibilité

#### I.5.3.1 Introduction

Dans le paragraphe précédent, nous avons présenté les principales méthodes d'analyse de sensibilité : la méthode FAST et sa version améliorée EFAST, la méthode de Sobol, la méthode de Morris et la méthode basée sur le calcul des dérivées partielles (MSDP). Nous nous sommes restreints à l'exposition de l'essentiel de leurs formulations mathématiques ainsi qu'à leurs implémentations en nous appuyant parfois sur des modèles analytiques simples. Dans ce paragraphe, nous nous proposons d'effectuer une analyse comparative de ces différentes méthodes sur des modèles analytiques particuliers couramment utilisés dans les études de *benchmark* [70]. En se basant sur la nature des mesures de sensibilité fournies, les méthodes présentées auparavant peuvent être réparties en deux groupes : celles basées sur le calcul des effets élémentaires telles que les méthodes de Morris et MSDP, et celles basées sur la décomposition de la variance telles que les méthodes FAST, EFAST et Sobol. Le but de cette étude comparative est de déterminer la méthode qui offre le meilleur compromis entre la précision, évaluée en terme d'hierarchisation des paramètres d'un modèle par ordre d'importance, et l'efficacité, mesurée par le nombre d'appels au modèle.

Les mesures de sensibilité relatives à la méthode MSDP sont obtenues par différentes techniques de calcul telles que les simulations de Monte-Carlo (MSDP-MC), les tirages hyper-cube latin (MSDP-LHS) et les séquences quasi-aléatoires de Halton (MSDP-Halton) [56] et de Hammersley (MSDP-Hammersley) [57]. Il est important de signaler que les tirages hyper-cube latin et les séquences quasi-aléatoires sont utilisés ici dans un but illustratif, et seront présentés plus en détails dans le deuxième chapitre de ce manuscrit. Les estimations sont obtenues en moyennant les résultats de 128 simulations et pour une perturbation  $\Delta = 10^{-5}$ . Elles sont comparées, dans un premier temps, aux mesures de sensibilité relatives à la méthode de Morris obtenues pour un nombre de niveaux  $p = 10$  et pour 128 trajectoires. On note ainsi que les estimations des mesures de sensibilité  $\mu_i^*$  et  $\sigma_i^*$  sont obtenues par la méthode de Morris et la méthode MSDP à partir du même nombre d'appels au modèle (i.e. 1408 appels) afin de pouvoir comparer la précision des deux méthodes.

Dans un deuxième temps, elles sont comparées aux indices de Sobol afin de voir l'analogie entre les deux types de mesures de sensibilité. Les indices de Sobol obtenus par simulations de Monte-Carlo, sont comparés aux mesures de sensibilité relatives à la méthode FAST et à sa version améliorée EFAST. Le calcul par simulations de Monte-Carlo est basé sur un échantillon de 16384 points générés par hyper-cube latin (Sobol-LHS), ce qui correspond à 196608 appels au modèle (en plus de garantir la convergence des indices de sensibilité de Sobol, le choix d'un échantillon de 16384 points est justifié par le fait qu'il est multiple des 128 points utilisés par la méthode de Morris et la méthode MSDP, et ainsi de servir comme élément de comparaison de l'efficacité), alors que pour la méthode EFAST, les résultats sont obtenus pour différentes combinaisons de la fréquence maximale  $\omega_i$  et du nombre de répétitions  $N_r$ , le choix de ces paramètres étant basé sur le graphique de la figure I.17.

Les modèles analytiques utilisés dans notre étude peuvent être répartis en trois catégories pour lesquelles les paramètres d'entrée sont définis sur l'intervalle  $[0, 1]$ .

- Modèle A : nombre limité de paramètres influents,

$$A(\mathbf{x}) = \sum_{i=1}^N (-1)^i \prod_{j=1}^i x_j$$

- Modèle B : interactions d'ordre faible,

$$B(\mathbf{x}) = \prod_{i=1}^N \frac{N - x_i}{N - 0,5}$$

- Modèle C : interactions d'ordre élevé.

$$C(\mathbf{x}) = 2^N \prod_{i=1}^N x_i$$

### I.5.3.2 Modèle à nombre limité de paramètres influents

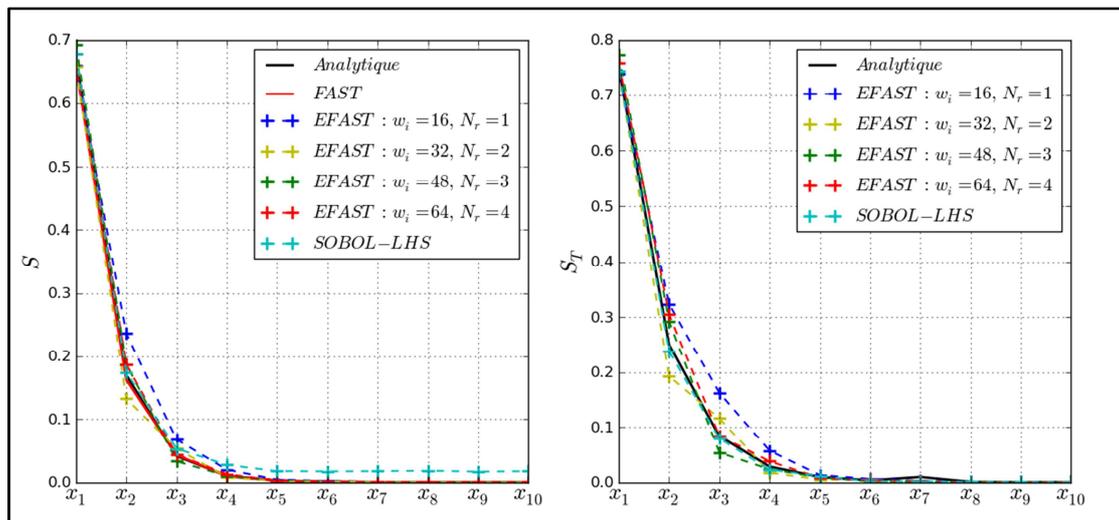
Nous nous intéressons au modèle A avec un nombre de paramètres d'entrée  $N = 10$ . Le tableau I.7 fournit les indices de sensibilité de premier ordre  $\{S_i\}_{i=1}^N$  et totaux  $\{S_{T_i}\}_{i=1}^N$  obtenus par la méthode de Sobol-LHS et la méthode EFAST. On note que, pour la méthode EFAST, les résultats présentés correspondent au calcul pour  $\omega_i = 64$  et  $N_r = 4$  (l'utilisation de cette combinaison de  $\omega_i$  et de  $N_r$  est justifiée par le fait qu'elle fournit les meilleures estimations en terme de précision). En se référant aux valeurs des indices de sensibilité totaux, on constate que les paramètres  $\{x_i\}_{i=1}^3$  sont les plus influents sur la réponse du modèle. De plus, en se référant aux valeurs du rapport entre les indices de sensibilité du premier ordre et les indices

de sensibilité totaux  $\{S_i/S_{T_i}\}_{i=1}^2 \geq 0,68$  et  $S_3/S_{T_3} \geq 0,49$ , on peut conclure que les interactions entre ces paramètres n'ont pas un effet significatif. Les estimations fournies par les tirages hyper-cube latin et la méthode EFAST sont en bon accord avec les solutions analytiques. Cependant la méthode EFAST est la plus efficace puisqu'elle nécessite un nombre limité d'appels au modèle.

**Tableau I. 7 :** Indices de sensibilité de premier ordre  $S_i$  et totaux  $S_{T_i}$  obtenus par la méthode de Sobol-LHS et la méthode EFAST pour la fonction A

$x_i$	<i>Analytique</i>		<i>Sobol-LHS</i>		<i>EFAST</i> ( $\omega_i = 64, N_r = 4$ )	
	$S_i$	$S_{T_i}$	$S_i$	$S_{T_i}$	$S_i$	$S_{T_i}$
1	0,67	0,75	0,678	0,743	0,661	0,752
2	0,17	0,25	0,174	0,238	0,166	0,244
3	$4,1 \cdot 10^{-2}$	$8,3 \cdot 10^{-2}$	0,052	0,0786	$5,9 \cdot 10^{-2}$	0,138
4	$1,0 \cdot 10^{-2}$	$2,8 \cdot 10^{-2}$	0,027	0,0222	$0,9 \cdot 10^{-3}$	0,018
5	$2,5 \cdot 10^{-3}$	$9,2 \cdot 10^{-3}$	0,017	0,0109	$7,7 \cdot 10^{-3}$	0,044
6	$6,9 \cdot 10^{-4}$	$3,2 \cdot 10^{-3}$	0,016	0,0009	$4 \cdot 10^{-4}$	0,001
7	$1,4 \cdot 10^{-4}$	$9,3 \cdot 10^{-4}$	0,017	0,0013	$2 \cdot 10^{-4}$	$6,7 \cdot 10^{-4}$
8	$5,1 \cdot 10^{-5}$	$4,2 \cdot 10^{-4}$	0,017	0,0002	$2,7 \cdot 10^{-5}$	$2,9 \cdot 10^{-4}$
9	$5,7 \cdot 10^{-6}$	$7,6 \cdot 10^{-5}$	0,016	0,0002	$1,1 \cdot 10^{-5}$	$1,2 \cdot 10^{-4}$
10	$5,7 \cdot 10^{-6}$	$7,6 \cdot 10^{-5}$	0,017	0	$3,7 \cdot 10^{-7}$	$5,9 \cdot 10^{-4}$
$N_s$	-		196608		8208	

La figure I.24 illustre une comparaison entre les estimations fournies par les différentes méthodes de calcul. On constate, d'une manière générale, que les résultats obtenus sont en bon accord avec les résultats du calcul analytique. Cependant, on note que la méthode EFAST est moins performante pour le cas où  $w_i = 16$  et  $N_r = 1$ . On constate aussi que les estimations obtenues par simulations de Monte-Carlo basées sur des points générés dans un hyper-cube latin sont moins précises dans le cas des indices de sensibilité du premier ordre de faibles amplitudes.



**Figure I. 24 :** Comparaison entre les différentes estimations des indices de sensibilité de premier ordre et les indices de sensibilité totaux pour la fonction A

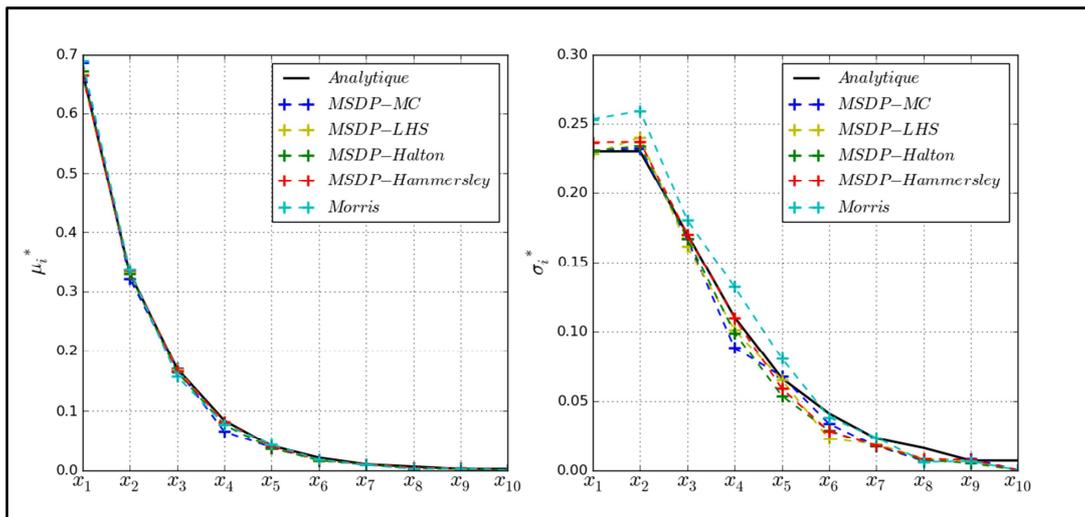
Si on s'intéresse maintenant à la comparaison des mesures de sensibilité obtenues par la méthode MSDP et la méthode de Morris, le tableau I.8 donne les estimations concernant la moyenne  $\mu_i^*$  et l'écart-type  $\sigma_i^*$  des effets élémentaires associés à la méthode de Morris et la

méthode MSDP. Globalement, on constate que les estimations fournies par ces deux méthodes sont en bon accord avec les résultats analytiques.

**Tableau I. 8 :** Mesures de sensibilité  $\mu_i^*$  et  $\sigma_i^*$  et  $G_i^*$  obtenus par la méthode MSDP-MC et la méthode de Morris pour la fonction A

$x_i$	Analytique			MSDP-MC			Morris	
	$\mu_i^*$	$\sigma_i^*$	$G_i^*$	$\mu_i^*$	$\sigma_i^*$	$G_i^*$	$\mu_i^*$	$\sigma_i^*$
1	0,67	0,23	0,67	0,675	0,241	0,647	0,674	0,293
2	0,33	0,23	0,22	0,372	0,242	0,248	0,345	0,273
3	0,17	0,17	$7,3 \cdot 10^{-2}$	0,172	0,180	0,0784	0,175	0,209
4	$8,4 \cdot 10^{-2}$	0,11	$2,5 \cdot 10^{-2}$	0,0726	0,0989	0,0189	0,0755	0,123
5	$4,1 \cdot 10^{-2}$	$6,6 \cdot 10^{-2}$	$8,0 \cdot 10^{-3}$	0,0341	0,0464	0,0041	0,0414	0,0752
6	$2,1 \cdot 10^{-2}$	$4,1 \cdot 10^{-2}$	$2,9 \cdot 10^{-3}$	0,0169	0,0255	0,0011	0,0169	0,0338
7	$9,8 \cdot 10^{-3}$	$2,3 \cdot 10^{-2}$	$8,3 \cdot 10^{-4}$	0,0098	0,0178	0,0005	0,0121	0,0422
8	$5,9 \cdot 10^{-3}$	$1,6 \cdot 10^{-2}$	$3,7 \cdot 10^{-4}$	0,0031	0,0075	$8,43 \cdot 10^{-5}$	0,0050	0,0175
9	$2,0 \cdot 10^{-3}$	$6,9 \cdot 10^{-3}$	$6,7 \cdot 10^{-5}$	0,0028	0,0080	$9,20 \cdot 10^{-5}$	0,0052	0,0228
10	$2,0 \cdot 10^{-3}$	$6,9 \cdot 10^{-3}$	$6,7 \cdot 10^{-5}$	0	0	0	0	0
$N_s$	-	-	-	1408			1408	

On note que le nombre d'appels au modèle est le même dans le cas de la méthode de Morris et la méthode MSDP basée sur les simulations de Monte-Carlo. On constate aussi que la mesure de sensibilité définie par la moyenne  $\mu_i^*$  des effets élémentaires correspond aux indices de sensibilité du premier ordre fournis par les méthodes basées sur la décomposition de la variance de la réponse du modèle (cf. tableau I.7). Par conséquent, la moyenne des effets élémentaires  $\mu_i^*$ , qu'on a qualifié auparavant de mesure de sensibilité qualitative, peut être considérée comme une mesure de sensibilité quantitative, si on cible seulement la mesure de la contribution individuelle de l'aléa associé à chaque paramètre sur la variabilité de la réponse du modèle. L'effet de l'interaction entre les paramètres est fourni implicitement par l'écart type des effets élémentaires puisqu'il ne permet pas de dissocier la contribution des interactions de différents ordres. Cependant, pour chaque paramètre, il peut mesurer l'intensité de cette interaction : en effet, plus  $\sigma_i^*$  est élevé, plus l'effet de l'interaction du paramètre  $x_i$  avec les autres paramètres est significative.



**Figure I. 25 :** Comparaison entre les différentes estimations des mesures de sensibilité  $\mu_i^*$  et  $\sigma_i^*$  pour la fonction A

La figure I.25 compare les estimations obtenues par la méthode de Morris et la méthode MSDP concernant la moyenne et l'écart-type des effets élémentaires associés à chaque paramètre d'entrée. On note que les estimations fournies par la méthode MSDP sont basées sur différentes techniques de simulations. Concernant la moyenne  $\mu_i^*$  des effets élémentaires, on constate que toutes les méthodes donnent les mêmes résultats. Cependant, concernant l'écart-type  $\sigma_i^*$ , on observe que les estimations fournies par la méthode de Morris, dans le cas des paramètres d'entrée  $x_1$  et  $x_2$ , ne s'accordent pas avec les résultats donnés par la méthode MSDP. Ce manque de précision de la méthode de Morris ne doit pas être considéré comme une limite, puisque, d'une part le moment statistique  $\sigma_i^*$  est une mesure de sensibilité qualitative qui ne fournit pas la contribution exacte d'un paramètre d'un modèle sur la variabilité de sa réponse mais il donne une idée sur une possible interaction avec d'autres paramètres ; d'autre part, les estimations restent du même ordre que la solution de référence (calcul analytique) ce qui ne perturbe pas l'ordre d'importance des différents paramètres.

Selon Saltelli et al. [52] et Campolongo et al. [72], la mesure de sensibilité  $\mu_i^*$  peut être considérée comme un bon indicateur de l'indice de sensibilité total  $S_{T_i}$ . Cependant, si on examine le rapport  $\mu_i^*/S_{T_i}$  et le rapport  $\mu_i^*/G_i^*$  pour les différents paramètres d'entrée, on constate qu'il existe une relation directe entre la mesure de sensibilité  $G_i^*$  et l'indice de sensibilité total  $S_{T_i}$  ( $G_i^* = 0,89S_{T_i}$ ,  $i = 1, 2, \dots, 10$ ) alors que la relation entre  $\mu_i^*$  et  $S_{T_i}$  est beaucoup plus complexe à déterminer. Cette constatation confirme les résultats trouvés par Sobol et Gresham [69] qui montrent que  $S_i \leq G_i^* \leq S_{T_i}$ . De plus, on peut montrer l'équivalence entre le rapport  $\mu_i^*/\sigma_i^*$  et le rapport  $S_i/S_{T_i}$ . Ainsi, cet indicateur de sensibilité  $\mu_i^*/\sigma_i^*$  peut être utilisé de la même manière que  $S_i/S_{T_i}$  pour l'évaluation de l'effet des interactions entre les différents paramètres (i.e. plus ce rapport est proche de l'unité, plus les effets des interactions entre les paramètres sont négligeables).

### I.5.3.3 Modèle aux interactions d'ordre faible

Nous nous intéressons maintenant au modèle défini par les fonctions  $B$ . La figure I.26 compare les estimations des indices de sensibilité du premier ordre et les indices de sensibilité totaux obtenues par simulations de Monte-Carlo, et les méthodes FAST et EFAST.

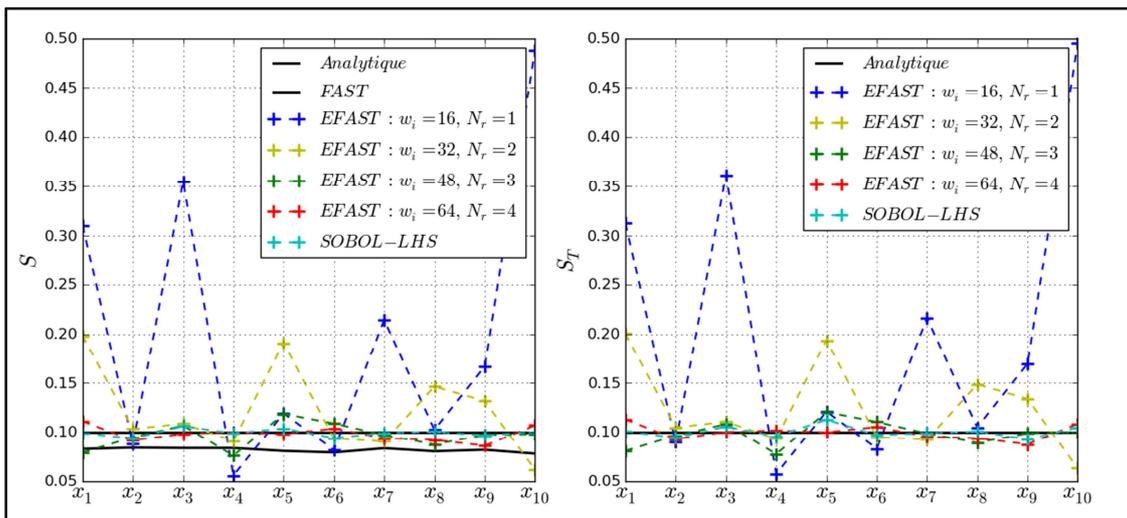


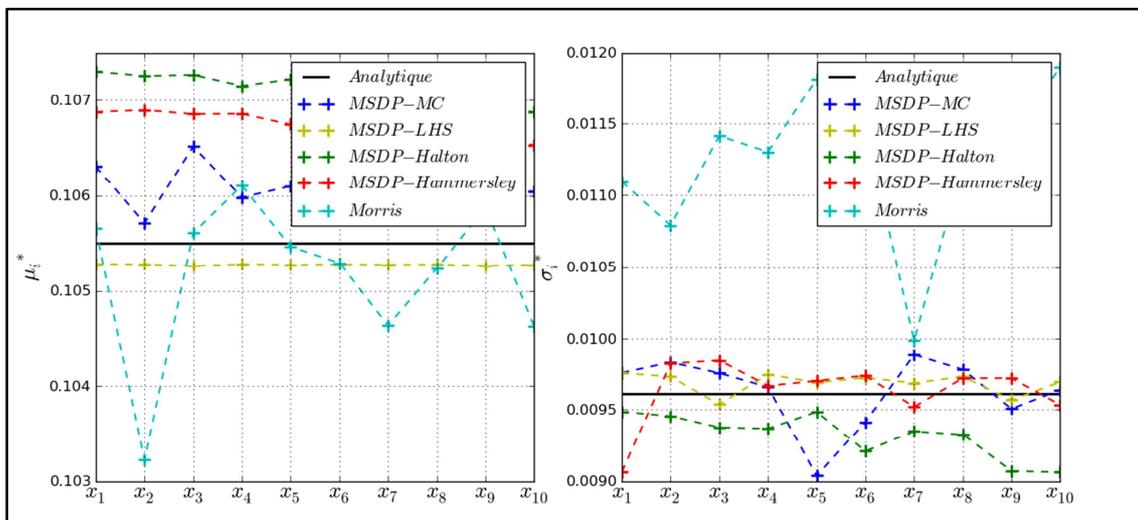
Figure I. 26 : Comparaison entre les différentes estimations des indices de sensibilité de premier ordre et les indices de sensibilité totaux pour la fonction  $B$

La particularité de ce modèle est que sa réponse est équi-sensible vis-à-vis des différents paramètres d'entrée. De plus, les indices de sensibilité du premier ordre ont les mêmes valeurs

que les indices de sensibilité totaux. Cette constatation est justifiée par les résultats illustrés par la figure I.26. On remarque que les estimations fournies par la méthode EFAST sont en bon accord avec les résultats issus du calcul analytique, en particulier, lorsque  $w_i = 64$  et  $N_r = 4$ . Cependant, on a besoin de plus de points d'expérimentation pour améliorer les estimations fournies par les simulations de Monte-Carlo.

En comparant les mesures de sensibilité obtenues par la méthode MSDP et la méthode de Morris, la figure I.27 montre que les estimations fournies par les simulations de Monte-Carlo basées sur des points générés par hyper-cube latin sont les plus proches des résultats donnés par le calcul analytique. Cependant, les fluctuations observées sur les estimations données par les différentes méthodes de calcul ne reflètent pas l'écart réel vis-à-vis de la solution analytique, mais sont dues au choix de l'échelle des ordonnées.

En se référant au rapport  $\mu_i^*/\sigma_i^* > 1$ , on peut conclure que l'effet des interactions entre les différents paramètres d'entrée est négligeable. Ce résultat est en accord avec la constatation de départ, stipulant que le rapport  $S_i/S_{T_i}$  est égal à l'unité.



**Figure I. 27 :** Comparaison entre les différentes estimations des mesures de sensibilité  $\mu_i^*$  et  $\sigma_i^*$  pour la fonction  $B$

Afin de mieux analyser la précision de la méthode de Morris et de la méthode MSDP, nous nous intéressons à l'évaluation de l'erreur  $\varepsilon_i$  au sens des moindres carrés sur les estimations de la moyenne  $\mu_i^*$  des effets élémentaires. Dans cette étude, l'erreur  $\varepsilon_i$  est calculée pour différents nombres de simulations et en moyennant, pour chacun d'eux, les résultats de  $L = 50$  calculs indépendants (i.e. les 50 échantillons de points sont générés indépendamment les uns des autres), et ce pour atténuer la dispersion. Ainsi, l'erreur  $\varepsilon_i$  est donnée par :

$$\varepsilon_i = \sqrt{\frac{1}{L} \sum_{l=1}^L \left( \frac{\mu_i^* - \mu_{i,0}^*}{\mu_{i,0}^*} \right)^2}$$

où  $\mu_{i,0}^*$  est la moyenne des effets élémentaires calculée analytiquement.

La figure I.28 illustre l'évolution de l'erreur  $\varepsilon_1$  en fonction du nombre de simulations sur l'estimation donnée par la méthode de Morris et la méthode MSDP de la moyenne  $\mu_1^*$  de l'effet élémentaire associé au paramètre  $X_1$ . On note que la présentation est effectuée dans une échelle logarithmique de base 2, d'une part pour mieux dissocier les valeurs et d'autre part puisque les nombres de simulations adoptés sont des puissances de 2.

On constate que les simulations de Monte-Carlo basées sur les séquences de nombres quasi-aléatoires de Hammersley fournissent l'erreur la plus faible sur l'estimation de la mesure de sensibilité  $\mu_1^*$ . Globalement, l'erreur enregistrée sur les simulations de Monte-Carlo basées sur les séquences de nombres quasi-aléatoires décroît rapidement en fonction du nombre de simulations par comparaison aux autres méthodes. La vitesse de convergence de la méthode de Morris est la plus faible parmi toutes les méthodes. De plus, il faut signaler que les simulations de Monte-Carlo basées sur les nombres pseudo-aléatoires et celles basées sur les tirages en hyper-cube latin exhibent les mêmes propriétés de convergence.

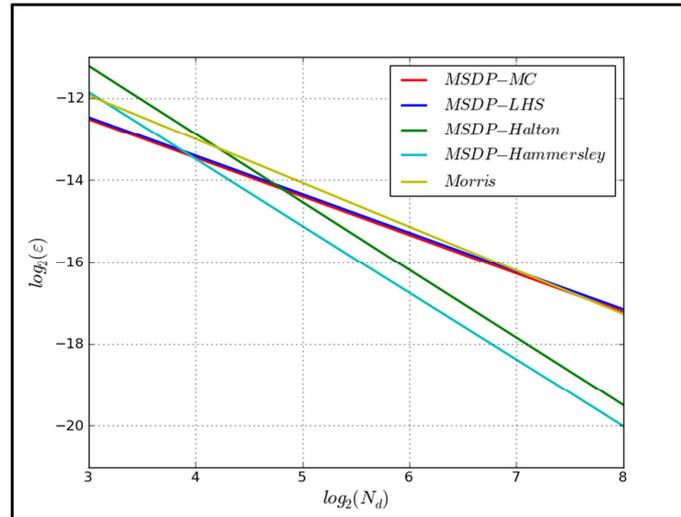


Figure I. 28 : Evolutions de l'erreur sur les estimations de  $\mu_1^*$  données par la méthode de Morris et la méthode MSDP dans le cas du modèle B

### I.5.3.4 Modèle aux interactions d'ordre élevé

Nous nous proposons de déterminer les paramètres les plus influents sur la réponse du modèle C. Pour ce faire, les indices de sensibilité du premier ordre  $\{S_i\}_{i=1}^N$  et les indices de sensibilité totaux  $\{S_{T_i}\}_{i=1}^N$  sont calculés par la méthode de Sobol basée sur les simulations de Monte-Carlo, la méthode FAST et sa version améliorée EFAST.

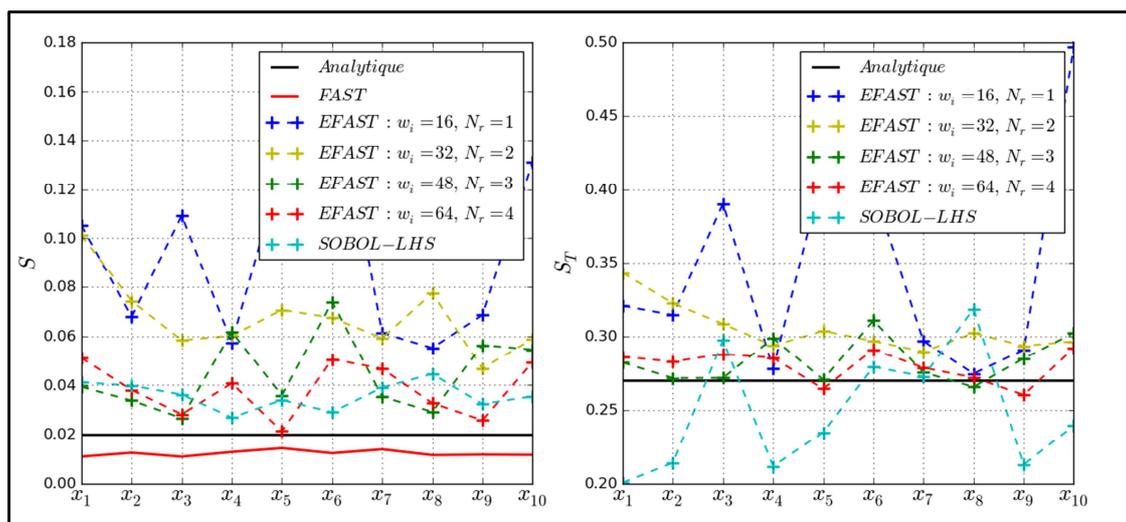
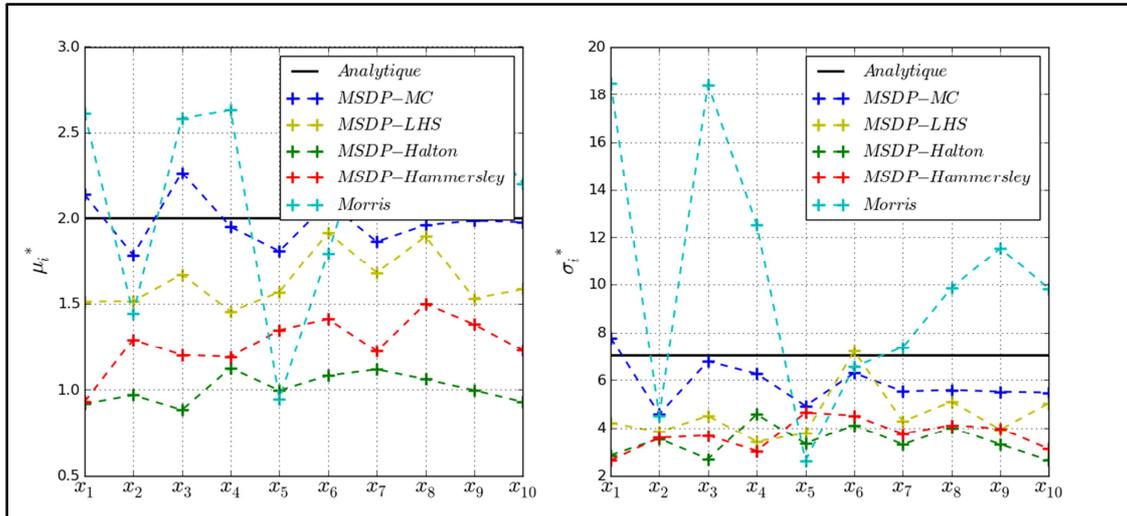


Figure I. 29 : Comparaison entre les différentes estimations des indices de sensibilité de premier ordre et les indices de sensibilité totaux pour la fonction C

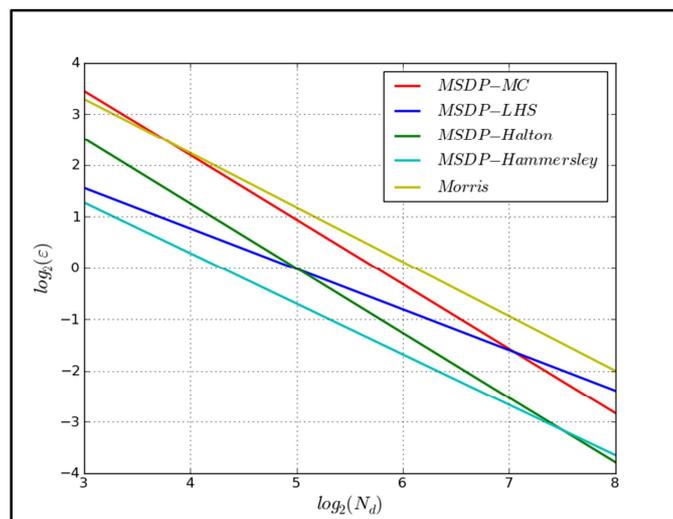
La figure I.29 illustre les résultats obtenus. On constate que les meilleures estimations concernant les indices de sensibilité du premier ordre sont données par la méthode FAST et les simulations de Monte-Carlo basées sur les points générés par hyper-cube latin. Cependant, les meilleures estimations des indices de sensibilité totaux sont données par la méthode EFAST telle que  $w_i = 64$  et  $N_r = 4$ . On note aussi, en comparant les résultats fournis pour le modèle C et le modèle B, que l'effet des interactions entre les paramètres d'entrée sur la réponse du modèle est très significatif ( $S_i \ll S_{T_i}$ ).



**Figure I. 30 :** Comparaison entre les différentes estimations des mesures de sensibilité  $\mu_i^*$  et  $\sigma_i^*$  pour la fonction C

Pour comparer les mesures de sensibilité obtenues par la méthode MSDP et la méthode de Morris, la figure I.30 montre que, pour ce type de modèle, les simulations de Monte-Carlo basées sur les séquences quasi-aléatoires perdent leur efficacité. En effet, on constate que les simulations de Monte-Carlo standard fournissent les meilleures estimations.

En se référant au rapport  $\mu_i^*/\sigma_i^* < 1$  on confirme l'effet significatif des interactions entre les paramètres d'entrée sur la réponse du modèle. On montre de plus l'analogie entre le rapport  $\mu_i^*/\sigma_i^*$  et le rapport  $S_i/S_{T_i}$ .



**Figure I. 31 :** Evolutions de l'erreur sur les estimations de  $\mu_1^*$  données par la méthode de Morris et la méthode MSDP dans le cas du modèle C

En complément, la figure I.31 illustre l'évolution de l'erreur  $\varepsilon_1$  sur les estimations de la mesure de sensibilité  $\mu_1^*$  associée au paramètre  $X_1$ , pour différentes méthodes de calculs en fonction du nombre de simulations. On constate que les simulations de Monte-Carlo basées sur les séquences quasi-aléatoires de Hammersley exhibent la meilleure convergence (i.e. la vitesse de convergence la plus élevée) par rapport aux autres méthodes. Généralement, toutes les méthodes ont presque le même taux de convergence. Il faut signaler aussi que l'ordre de l'erreur est plus élevé que dans le cas du modèle  $B$ .

### I.5.3.5 Critiques et conclusions sur les méthodes d'analyse de sensibilité

Dans le paragraphe précédent, différents modèles analytiques ont été utilisés pour analyser et comparer plusieurs méthodes d'analyses de sensibilité. Ces méthodes peuvent être classées en deux catégories : les méthodes d'analyses de sensibilité locale telles que la méthode de Morris et la méthode MSDP, et les méthodes d'analyses de sensibilité globale telles que la méthode FAST, sa version améliorée EFAST, et la méthode de Sobol.

Dans un premier temps nous nous proposons de raisonner sur les méthodes d'analyses de sensibilité globale. La méthode FAST, facile à implémenter, donne des estimations précises des indices de sensibilité de premier ordre tout en nécessitant un nombre minimal d'appels au modèle. Cependant, son efficacité se dégrade dans le cas de modèles pour lesquels l'effet des interactions entre les paramètres d'entrée n'est pas négligeable. On a vu que cette limitation est levée par la version améliorée EFAST. En effet, elle permet de déterminer, en plus des indices de sensibilité de premier ordre, les indices de sensibilité totaux qui mesurent non seulement l'effet principal de chaque paramètre d'entrée mais aussi l'effet de son interaction avec les autres paramètres. On note tout de même que les effets principaux et les effets des interactions ne peuvent pas être découplés les uns des autres. Cette limitation, bien qu'elle ne soit pas majeure, est surmontée par la méthode de Sobol au détriment de l'efficacité, surtout dans les problèmes impliquant un nombre élevé de paramètres d'entrée. Suite aux différentes constatations présentées ci-dessus, on peut conclure que la méthode EFAST est la mieux adaptée pour effectuer des analyses de sensibilité globale.

En raisonnant maintenant sur les méthodes d'analyses de sensibilité locale, on a montré que, dans certaines situations, la méthode MSDP peut être plus efficace et plus précise que la méthode de Morris. La convergence de la méthode MSDP est plus rapide lorsque les simulations de Monte-Carlo sont basées sur les tirages en hyper-cube latin ou les séquences quasi-aléatoires, que lorsqu'elles utilisent les nombres pseudo-aléatoires. Cependant, il faut signaler que, globalement, la méthode de Morris garantit le meilleur compromis entre précision des estimations et coût de calcul, bien qu'elle ne soit pas stable dans le cas de modèles non monotones. En les comparant aux méthodes d'analyses de sensibilité globale, la méthode MSDP et la méthode de Morris, peuvent fournir des mesures de sensibilité quantitatives fiables seulement lorsque l'effet des interactions entre les paramètres d'entrée est faible. En effet, dans cette situation, les valeurs des indices de sensibilité totaux correspondent en grande partie aux indices de sensibilité du premier ordre. Généralement, la méthode de Morris fournit des mesures de sensibilité qualitatives qui ne permettent pas de quantifier la contribution de chaque paramètre d'entrée dans la variance de la réponse du modèle. Mais, ces mesures permettent cependant de classer les paramètres d'entrée par ordre d'importance. De plus, la méthode de Morris permet d'effectuer un calcul des statistiques des effets élémentaires en considérant des groupes de paramètres ce qui contribue à améliorer significativement son efficacité. Par conséquent, la méthode de Morris est utilisée dans la suite de notre travail, comme un outil de hiérarchisation des paramètres afin de ne considérer que les plus importants dans les analyses de sensibilité globale, les analyses de fiabilité ou les analyses de tendance centrale. Cette aspect est traité avec plus de détails dans le chapitre

suyant où les méthodes d'analyse de sensibilité, en particulier la méthode de Morris, sont considérées comme un moyen de résoudre le problème de dimension stochastique élevée.

### I.6 Conclusion

Dans ce chapitre, nous nous sommes intéressés à la notion d'incertitude qui peut se manifester dans le contexte d'analyse de modèles représentant un processus physique. Le traitement de ces différentes sources d'incertitude implique l'utilisation de méthodes de calcul stochastique. Nous avons vu que trois finalités sont possibles : l'analyse de tendance centrale qui se focalise sur la quantification de l'effet de l'aléa associé aux paramètres d'entrée sur la réponse d'un modèle en déterminant ses caractéristiques statistiques, l'analyse de la fiabilité qui consiste, dans le contexte de génie mécanique ou civil, à évaluer l'intégrité d'une structure vis-à-vis d'un scénario de défaillance, et l'analyse de sensibilité qui permet de quantifier l'effet de l'aléa associé à chaque paramètre incertain sur l'aléa de la réponse du modèle. Nous avons présenté les méthodes classiques qui permettent de satisfaire les trois finalités du calcul stochastique. La présentation de la formulation et l'implémentation des différentes catégories des méthodes de calcul stochastique a été soutenue par des exemples d'illustration, qui ont permis, de développer une analyse critique pour déterminer leurs avantages et leurs limites, ainsi que l'aptitude à l'extension de chaque méthode.

Concernant, les méthodes d'analyse de tendance centrale, on a montré que la méthode de collocation stochastique garantit le meilleur compromis entre la précision et le coût de calcul, tant que la dimension de l'espace aléatoire est modérée ( $N < 15$ ) : en effet, le nombre d'appels au modèle augmente d'une manière exponentielle. L'efficacité de ces méthodes peut être améliorée en utilisant des règles de quadrature adaptées telles que celle proposée par Stroud ou celle basée sur les grilles de Smolyak.

Concernant l'analyse de fiabilité, nous avons montré que les méthodes d'approximation FORM/SORM sont efficaces et précises tant que la fonction d'état limite représentant le scénario de défaillance n'est pas complexe (i.e. absence de bruit de forme ou de fortes courbures), la dimension de l'espace aléatoire n'est pas élevée et que l'algorithme de détermination du point de défaillance le plus probable est robuste. Les simulations de Monte-Carlo ne sont pas sensibles à la dimension de l'espace aléatoire et permettent de plus d'obtenir un intervalle de confiance de l'estimation de la probabilité de défaillance. Cependant, leur efficacité est mise en défaut dans l'estimation des probabilités de défaillance d'ordre faible. Ce problème peut être surmonté en utilisant les techniques de réduction de la variance telles que les tirages d'importance, les tirages conditionnés, les tirages adaptatifs ou encore la méthode de Subset simulations.

Les méthodes d'analyse de sensibilité ont été traitées d'une façon plus détaillée. Selon la mesure de sensibilité fournie, ces méthodes peuvent être réparties en deux catégories : méthodes d'analyse de sensibilité globale et méthodes d'analyse de sensibilité locale. Grâce aux différentes applications traitées, nous avons déterminé le lien entre les mesures de sensibilité locales et les mesures de sensibilité globales. La méthode de Morris, basée sur l'évaluation des statistiques des effets élémentaires pour hiérarchiser les paramètres d'un modèle, permet de déterminer les paramètres importants, en étant la plus économique comparée aux autres méthodes d'analyse de sensibilité. Cependant, cette méthode ne permet qu'une analyse qualitative de la sensibilité d'un modèle vis-à-vis de ses paramètres d'entrée. Ainsi, elle peut être un moyen pour réduire la dimension de l'espace aléatoire avant d'entamer une analyse de sensibilité proprement dite, une analyse de fiabilité ou une analyse de tendance centrale.



---

## **Chapitre II : problèmes à dimension stochastique élevée**

---

## II.1 Introduction

Dans le chapitre précédent, nous avons passé en revue les principales méthodes de calcul stochastique qui permettent d'effectuer des analyses de tendance centrale, des analyses de fiabilité ou des analyses de sensibilité. Nous avons constaté que seule l'efficacité des simulations de Monte-Carlo et ses dérivées est indépendante de la dimension de l'espace aléatoire. Cependant, elles sont quasi inapplicables dans le cas où le modèle implicite représentant le phénomène physique à étudier est exigeant en temps de calcul, et en particulier pour l'évaluation des moments statistiques d'ordres élevés (i.e. coefficient d'aplatissement et coefficient d'asymétrie) et l'estimation des probabilités de défaillance d'ordre faible. Malgré leur efficacité et leur aptitude à être appliquées à des problèmes mécaniques complexes, elles ne permettent pas de traiter les différentes finalités du calcul stochastique. Malheureusement, dans les problèmes d'ingénierie, nous sommes le plus souvent amenés à développer des modèles numériques exigeants en temps de calcul pour simuler le comportement des structures réelles. De plus, les sources d'incertitude sont multiples, et puisque nous ne disposons pas d'informations préalables sur le poids de chacune d'elles, nous sommes contraints à toutes les prendre en compte dans le calcul stochastique ce qui engendre un nombre élevé de paramètres incertains. Dans le lexique du calcul stochastique, ces situations (grand nombre de paramètres incertains) sont regroupées sous l'appellation de problèmes à dimension stochastique élevée. Ces deux problèmes sont prépondérants dans notre travail qui consiste à analyser l'effet de l'aléa sur le comportement des structures à ossatures bois sous sollicitation sismique. En effet, le calcul mécanique est basé sur une analyse dynamique non linéaire exigeante en temps de calcul et le nombre de paramètres incertains peut avoisiner une centaine. Ainsi, l'application des méthodes de calcul stochastique présentées jusqu'ici n'est donc pas envisageable.

Afin de surmonter cette difficulté, nous proposons de poser le problème de la dimension stochastique élevée d'une façon plus intuitive. En effet, au lieu de chercher à développer une approche permettant d'aborder directement les différentes finalités du calcul stochastique tout en prenant en compte l'aléa associé aux différents paramètres incertains, il est plus pertinent de chercher, dans un premier temps, un moyen de réduire le nombre de paramètres incertains, et de ne considérer dans un deuxième temps que les paramètres les plus importants dans le calcul stochastique proprement dit. Ce raisonnement est soutenu par le fait que l'expérience a montré, en particulier pour les problèmes d'ingénierie, que le plus souvent, parmi l'ensemble des paramètres incertains considérés dans le calcul stochastique, seulement un nombre limité entre eux contribue efficacement à l'aléa observé sur la réponse du modèle ou sur la probabilité de défaillance. Cette observation conduit à distinguer la dimension stochastique *nominale* (nombre de paramètres incertains d'un modèle) et la dimension stochastique *efficace* (nombre de paramètres incertains qui contribuent effectivement à la variabilité de la réponse d'un modèle). Une fois cette dernière déterminée, nous ne considérons comme paramètres incertains que ceux jugés importants, les autres paramètres étant fixés à leurs valeurs moyennes respectives. Dans ce contexte, les méthodes de criblage présentées dans le premier chapitre, en particulier la méthode de Morris, sont des solutions intéressantes pour déterminer la dimension stochastique efficace. En effet, la méthode de Morris fournit des mesures de sensibilité qualitatives appelées *effets élémentaires* permettant de hiérarchiser les paramètres incertains considérés dans un problème. On a montré que ces mesures de sensibilité sont étroitement liées aux mesures de sensibilité globale telles que les indices de Sobol, ce qui justifie leur pertinence pour détecter les paramètres les plus importants. Ainsi, la procédure de calcul stochastique que nous proposons dans notre travail est composée de deux étapes : la première consiste à déterminer la dimension stochastique efficace au moyen de la

méthode de Morris, et la deuxième consiste à effectuer le calcul stochastique proprement dit en utilisant une technique de construction de *surface de réponse* ou de *méta-modèle*.

Dans ce chapitre, nous présentons dans un premier temps la stratégie de résolution du problème de dimension stochastique élevée, en rappelant au passage la notion de dimension stochastique efficace, et tout en proposant une approche efficace permettant son identification. Ensuite, nous nous intéressons à deux techniques de construction de méta-modèle : le *Développement en Chaos Polynômial* (DCP) [75] et la *Méthode de Décomposition de la Dimension* (MDD) [76]. La formulation mathématique ainsi que les différents aspects d'implémentation sont présentés en s'appuyant sur des exemples d'illustration, afin d'analyser et de comparer l'efficacité des deux approches. Enfin la stratégie de calcul stochastique est validée à travers un problème traitant le comportement d'une structure à cinq étages sollicitée par des charges latérales.

## II.2 Vaincre le problème de dimension stochastique élevée

### II.2.1 Dimension stochastique efficace

La notion de dimension stochastique efficace a été proposée par Caflisch et al. [77], et ils en ont proposé deux définitions distinctes, une au sens de la *superposition* et une au sens de la *troncature*, en utilisant la décomposition de la variance ou la décomposition de la dimension. Soit  $\mathfrak{X} = \{1, 2, \dots, N\}$  un ensemble d'entiers naturels et  $|u|$  la cardinalité du sous-ensemble  $u \subseteq \mathfrak{X}$ . La dimension efficace d'une fonction  $f$  au sens de la superposition est le plus petit entier  $d_s$  tel que :

$$\sum_{|u| \leq d_s} \sigma_u^2 \geq p, 0 < p < 1 \quad (II.1)$$

où  $\sigma_u^2$  est la variance partielle qui mesure l'effet de l'aléa associé aux paramètres d'entrée (dont les indices sont dans l'ensemble  $u$ ), sur la variance totale  $\sigma^2$  de la fonction  $f$ , et  $p$  un quantile seuil pris égal à  $0,99\sigma^2$ . L'inégalité définie par l'équation (II.1) veut dire que la fonction  $f$  peut être écrite comme une somme de fonctions  $d_s$ -dimensionnelles. Ainsi, la dimension efficace  $d_s$  est l'ordre le plus élevé de la fonction qu'il faut inclure dans la somme  $\sum_{|u| \leq d_s} \sigma_u^2 \geq p$  pour atteindre le seuil  $p$ .

La dimension efficace de la fonction  $f$  au sens de la troncature est le plus petit entier  $d_T$  tel que :

$$\sum_{u \subseteq \{1, 2, \dots, d_T\}} \sigma_u^2 \geq p, 0 < p < 1 \quad (II.2)$$

Ainsi,  $d_T$  n'est autre que le nombre de variables qu'il faut considérer dans la somme  $\sum_{u \subseteq \{1, 2, \dots, d_T\}} \sigma_u^2 \geq p$  pour atteindre le seuil  $p$ . La dimension efficace au sens de la troncature  $d_T$  ne dépend pas de l'ordre selon lequel les variables ont été échantillonnées, au contraire de la dimension efficace au sens de la superposition  $d_s$ . On note que, pour un même quantile seuil  $p$ ,  $d_s$  est toujours inférieure ou égale à  $d_T$ .

Caflisch et al. [77] ont utilisé la notion de dimension efficace pour expliquer l'efficacité des simulations quasi-Monte Carlo dans le cas où la dimension stochastique nominale est élevée. Ils ont attribué cette performance au fait que, dans certains problèmes, la dimension efficace de l'intégrande, selon l'un ou les deux sens, est de beaucoup inférieure à la dimension nominale. Cette conclusion, qui n'est pas justifiée d'une façon formelle, a été soutenue par

des constatations sur un exemple d'illustration traitant un processus brownien. Ils ont montré en plus, qu'en changeant l'ordre selon lequel les variables étaient échantillonnées, la dimension efficace peut être réduite. Par conséquent, la précision des estimations fournies par les simulations quasi-Monte Carlo, peut être significativement améliorée tout en étant économe en terme d'évaluation du modèle ou de la fonction d'état limite. Bien que Caflisch et al. [77] aient été les premiers à l'introduire, la notion de la dimension efficace est apparue dans les travaux de Paskov et Traub [78] dédiés au domaine de la finance. Ils ont remarqué que les fonctions qu'ils étudiaient ne pouvaient pas être définies en considérant seulement un nombre restreint des principaux paramètres d'entrée. Sloan et Wazniakwski [79] ont introduit et étudié des fonctions pour lesquelles le poids ou l'importance d'un paramètre  $x_i$  diminue lorsque l'indice  $i$  représentant son rang sur l'ensemble des paramètres d'entrée augmente. Ils ont montré que, pour une telle variété de fonctions, la dimension efficace au sens de la troncature est très inférieure à la dimension nominale. Généralement, l'évaluation de la dimension efficace telle qu'elle est définie par les équations (II.1) et (II.2) n'est pas abordable, puisqu'elle nécessite la détermination de  $2^N$  variances partielles  $\sigma_u^2$ .

Afin de surmonter ce problème, Owen [80] a introduit la notion de distribution de la dimension pour les fonctions de carré intégrable sur  $[0,1]^\infty$ . Cette idée est soutenue par le fait que cette classe de fonctions a nécessairement une dimension efficace finie au sens de la troncature pour tout seuil  $p$  inférieur à 100% de la variance totale  $\sigma^2$ . La dimension efficace peut être ainsi définie comme un quantile de la distribution de la dimension. Soit  $\mathcal{U} \subseteq \{1, \dots, s\}$  un ensemble choisi d'une façon aléatoire et  $P(\mathcal{U} = u) = \mu(u) = \sigma_u^2 / \sigma^2$ , la taille de l'ensemble *aléatoire*  $\mathcal{U}$  peut être mesurée soit par sa cardinalité soit par le maximum des indices de ses éléments. Ces deux grandeurs sont des variables aléatoires à partir desquelles la distribution de la dimension de la fonction  $f$  peut être définie. En se référant à Owen [80], la distribution de la dimension de la fonction  $f$ , dans le sens de la superposition, est la distribution de probabilité de la variable aléatoire discrète  $|\mathcal{U}|$  (cardinalité de l'ensemble  $\mathcal{U}$ ) dont la fonction de masse de probabilité  $\nu_s(d)$  est définie par :

$$\nu_s(d) = \sum_{|u|=d} \frac{\sigma_u^2}{\sigma^2}, d = 1, \dots, s \quad (II.3)$$

Par ailleurs, la distribution de la dimension au sens de la troncature est la distribution de probabilité de la variable aléatoire discrète  $\max\{i \mid i \in \mathcal{U}\}$ , dont la fonction de masse de probabilité est définie par :

$$\nu_T(d) = \sum_{\max\{i \mid i \in u\}=d} \frac{\sigma_u^2}{\sigma^2}, d = 1, \dots, s \quad (II.4)$$

Ainsi,  $d$  est la dimension efficace de la fonction  $f$  au sens de la superposition ou au sens de la troncature, respectivement si le  $p^{\text{ème}}$  quantile de la masse de probabilité  $\nu_s$  ou  $\nu_T$ , est au plus égal à  $d$ . Bien que Owen [80] ait pu déterminer ce quantile d'une manière explicite pour certains cas tests portant sur des fonctions additives et des fonctions multiplicatives, son estimation est généralement non triviale. Afin de simplifier le problème, le même auteur a introduit la dimension moyenne  $\mu_{d_s}$  au sens de la superposition et la dimension moyenne  $\mu_{d_T}$  au sens de la troncature, définies respectivement par :

$$\mu_{d_s} = \frac{\sum_{|u|>0} \sigma_u^2 |u|}{\sum_{|u|>0} \sigma_u^2} \quad (II.5)$$

$$\mu_{d_T} = \frac{\sum_{|u|>0} \sigma_u^2 \max\{i \mid i \in u\}}{\sum_{|u|>0} \sigma_u^2} \quad (II.6)$$

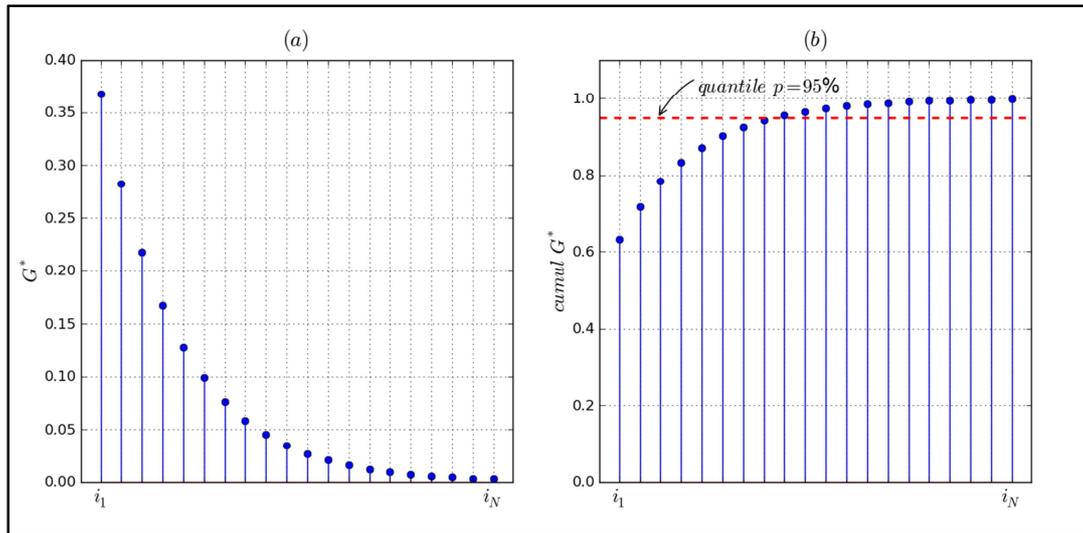
Ce même concept a été aussi proposé et étudié par Asotsky et al. [81]. Ces dimensions moyennes (au sens de la superposition et au sens de la troncature) sont respectivement les espérances mathématiques des variables aléatoires discrètes  $|u|$  et  $\max\{i \mid i \in u\}$ . Le principal avantage de l'utilisation des dimensions moyennes est qu'elles ne dépendent pas du quantile seuil  $p$ . De plus, elles sont plus faciles à calculer que le quantile de la distribution de la dimension suggéré au début.

Owen [80] a relié la distribution de la dimension aux indices de sensibilité de Sobol, pour étudier des fonctions isotropes, pour lesquelles il a montré qu'elles peuvent être définies comme la superposition de fonctions au plus tridimensionnelles. Dans ce contexte, Kucherenko et al. [82] ont récemment proposé une méthodologie basée sur l'évaluation du rapport entre les indices de sensibilité totaux et les indices de sensibilité du premier ordre pour évaluer la dimension efficace. En effet, si le rapport  $S_i/S_{T_i}$ ,  $i = 1, \dots, N$  est proche de l'unité, on peut conclure que les effets principaux associés aux différents paramètres sont les plus importants et par conséquent que la dimension efficace au sens de la superposition est  $d_s \approx 1$ . Si  $S_i/S_{T_i} < 1$  alors la dimension efficace  $d_s > 1$ , mais malheureusement on peut pas déterminer sa valeur exacte sans calculer les indices de sensibilité d'ordre supérieur à 1, ce qui induit un coût de calcul supplémentaire. La dimension efficace au sens de la troncature est par contre plus facile à évaluer. En effet, soit  $S_u$  l'indice de Sobol qui mesure l'importance des paramètres dont les indices qui leurs sont associés sont dans l'ensemble  $u \subseteq \{1, \dots, s\}$ , si  $S_u$  est inférieur ou égal au seuil  $p$ , alors la dimension efficace au sens de la troncature est  $d_T = s$ .

## II.2.2 Stratégie de résolution

Dans le chapitre précédent, on a démontré le lien étroit entre les mesures de sensibilité locale associées à la méthode de Morris et les indices de sensibilité de Sobol, ainsi que leur pertinence pour hiérarchiser les paramètres incertains d'un modèle par ordre d'importance. Par analogie, le concept basé sur les indices de Sobol, proposé par Kucherenko et al. [82] pour déterminer la dimension efficace d'un modèle, peut être appliqué en utilisant les mesures de sensibilité locale. En effet, en se référant à la moyenne  $\mu_i^*$  des effets élémentaires, on peut arranger les paramètres par ordre décroissant d'importance. Par conséquent, en se basant sur cet arrangement, on peut construire une distribution de la dimension, et définir la dimension efficace  $d_u$  par les indices  $u \subseteq \{1, \dots, s\}$  des paramètres dont la somme des mesures de sensibilité qui leur sont associées est inférieure ou égale à  $p \sum_{i=1}^N \mu_i^*$ . En réalité, les méthodes de criblage, telle que la méthode de Morris, permettent de déterminer les paramètres les moins influents sur la réponse d'un modèle. Leur identification s'effectue en traçant le graphique illustré sur la figure I.24, basé sur la moyenne et l'écart-type des effets élémentaires. Les paramètres non importants sont ceux situés au voisinage de l'origine du graphique. Ainsi, la dimension efficace est déterminée par l'élimination des paramètres non importants de l'ensemble des paramètres incertains du modèle. Malheureusement, en se basant seulement sur le graphe construit à partir des statistiques des effets élémentaires, il est parfois difficile de dissocier les paramètres importants et non importants. Afin de pallier ce problème, nous proposons, dans les situations qui l'exigent, de se baser, en plus des statistiques des effets élémentaires, sur la mesure de sensibilité normalisée  $G_i^*$  définie par l'équation (I.113) pour hiérarchiser les paramètres d'un modèle. Ce raisonnement est soutenu par le fait que la mesure de sensibilité  $G_i^*$ , dont le cumul est égal à l'unité, peut être considérée comme une

variable aléatoire discrète dont la distribution n'est autre que la distribution de la dimension stochastique nominale. En fixant un quantile  $p$ , la dimension stochastique efficace peut être définie par le nombre de paramètres dont la somme des mesures de sensibilité  $G_i^*$  qui leurs sont associés est inférieure ou égale à  $p$ . Ce concept est illustré sur les figures II.1a et II.1b.



**Figure II. 1 :** (a) hiérarchisation des paramètres incertains en se référant à la mesure de sensibilité  $G_i^*$ , (b) identification des paramètres les plus importants pour un quantile seuil  $p = 95\%$

En se référant à l'étude comparative menée dans le paragraphe I.5.3 du premier chapitre, on a constaté que, sur les cas tests étudiés, la dimension efficace est de beaucoup inférieure à la dimension nominale. Cette situation est récurrente dans les problèmes d'ingénierie. Ainsi, dans la perspective de résoudre les problèmes à dimension stochastique élevée, on se propose d'utiliser la méthode de Morris, en conjonction au critère de sélection des paramètres incertains présenté ci-dessus, pour déterminer la dimension efficace du modèle à étudier, avant d'entamer le calcul stochastique proprement dit. Par conséquent, en ne modélisant que les paramètres importants par des variables aléatoires et en fixant les paramètres restant à leurs valeurs moyennes ou leurs valeurs nominales respectives, le problème de manque d'efficacité des méthodes de calcul stochastique classiques est contourné, et on peut mener des analyses de sensibilité, des analyses de fiabilité ou des analyses de tendance centrale, tout en restant économe en terme d'appels au modèle. Afin de gagner plus en efficacité, étant donné que le modèle éléments finis est très exigeant en temps de calcul (i.e. analyse dynamique non linéaire du comportement d'une toiture à ossature bois sous sollicitation sismique), nous nous proposons d'utiliser dans le calcul stochastique proprement dit une méthode de couplage mécano-probabiliste basée sur la construction d'une représentation explicite du modèle mécanique. Dans cette perspective, nous présentons dans la suite deux techniques de construction d'un méta-modèle : le développement en chaos polynômial (DCP) et la méthode de décomposition de la dimension (MDD). Le choix de l'une d'entre elles sera motivé par des analyses comparatives ciblant l'efficacité, la flexibilité et la précision.

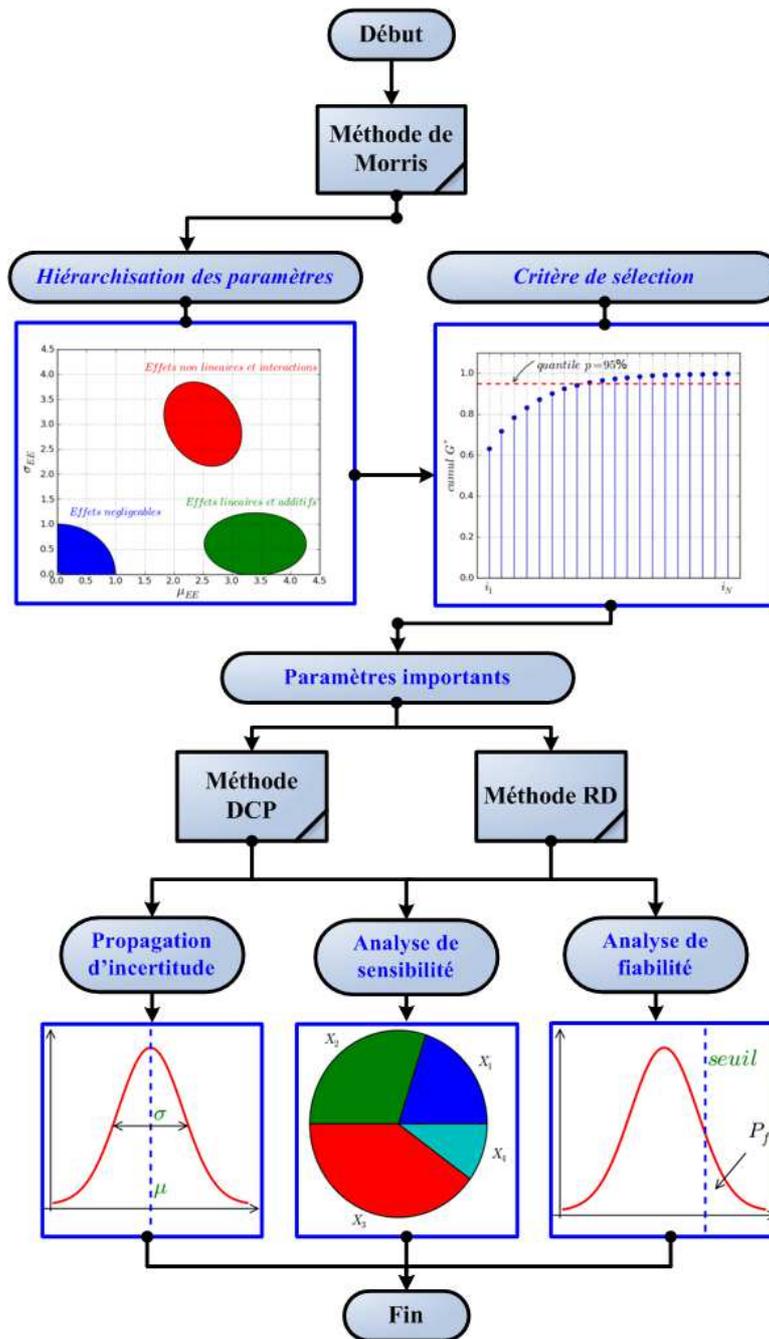


Figure II. 2 : Stratégie de résolution du problème de dimension stochastique élevée

En résumé, la stratégie que nous proposons pour la résolution d'un problème à dimension stochastique élevée est composée de deux étapes : la première consiste à déterminer la dimension stochastique efficace du modèle en se basant sur une hiérarchisation des paramètres incertains par la méthode de Morris, et la deuxième consiste à effectuer le calcul stochastique proprement dit en utilisant des techniques de construction de méta-modèles. L'organigramme de cette stratégie est illustré sur la figure II.2.

## II.3 Développement en chaos polynômial

### II.3.1 Chaos polynômial de Weiner

Le développement en chaos polynômial a été introduit par Weiner [83]. Il permet d'approximer une variable aléatoire du second ordre (i.e. une variable aléatoire dont la

variance est finie) par un ensemble de polynômes, dans l'espace aléatoire standard formé par des variables Gaussiennes indépendantes (i.e. les variables Gaussiennes de moyenne nulle et d'écart-type unitaire). Avant d'entamer la présentation du développement en chaos polynômial de Wiener, on se propose de définir quelques notions mathématiques de base.

Soit  $\{\xi_i\}_{i=1}^{\infty}$  un ensemble infini mais dénombrable de variables Gaussiennes standard indépendantes, on définit par  $\hat{\Gamma}_p$  l'espace de tous les polynômes dont le degré est inférieur ou égal à  $p$ ,  $\Gamma_p$  l'ensemble des polynômes de  $\hat{\Gamma}_p$  qui sont orthogonaux aux polynômes de l'espace  $\hat{\Gamma}_{p-1}$ , et  $\tilde{\Gamma}_p$  l'espace formé par  $\Gamma_p$  tel que :

$$\hat{\Gamma}_p = \hat{\Gamma}_{p-1} \oplus \tilde{\Gamma}_p, \quad \Theta = \bigoplus_{i=0}^{\infty} \tilde{\Gamma}_i \quad (II.7)$$

Le sous-espace  $\tilde{\Gamma}_p$  de  $\Theta$  est désigné par  $p^{\text{ème}}$  chaos homogène et  $\Gamma_p$  est désigné par chaos polynômial d'ordre  $p$ . Ce dernier est l'ensemble des polynômes d'ordre  $p$  résultant de toutes les combinaisons possibles des variables aléatoires  $\{\xi_i\}_{i=1}^{\infty}$ . En se basant sur les différents postulats introduits ci-dessus, le développement en chaos polynômial de la variable aléatoire  $\theta$  est donné par :

$$\begin{aligned} \theta &= a_0 \Gamma_0 + \sum_{i_1=1}^{\infty} a_{i_1} \Gamma_1(\xi_{i_1}) \\ &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} a_{i_1 i_2} \Gamma_2(\xi_{i_1}, \xi_{i_2}) \\ &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3} \Gamma_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) + \dots \end{aligned} \quad (II.8)$$

Cameron et Martin [84] ont montré que ce développement converge au sens de la norme de  $L_2$  ( $L_2$  est l'ensemble des fonctions de carré intégrable).

Afin de simplifier la notation et de faciliter ainsi la manipulation du développement en chaos polynômial, on introduit généralement dans la littérature [75] une nouvelle fonctionnelle  $\Psi(\cdot)$  exprimée en fonction des fonctionnelles  $\Gamma(\cdot)$ . Dans ce contexte, le développement en chaos polynômial défini par l'équation (II.8) peut être réécrit sous une formulation plus condensée :

$$\theta = \sum_{k=0}^{\infty} \theta_k \Psi_k(\xi), \quad \xi = \{\xi_1, \xi_2, \dots\} \quad (II.9)$$

Par convention  $\Psi_0$  désigne le polynôme d'ordre zéro (i.e.  $\Psi_0 = 1$ ),  $\{\theta_i\}_{i=0}^{\infty}$  sont des coefficients déterministes, désignés souvent par coefficients du développement en chaos polynômial de la variable aléatoire  $\theta$  et  $\{\Psi_i\}_{i=1}^{\infty}$  sont des polynômes multidimensionnels orthogonaux dans l'espace  $L_2$ , par rapport au produit scalaire noté  $\langle, \rangle$ , basé sur la mesure Gaussienne :

$$\langle \Psi_i, \Psi_j \rangle = \int \Psi_i(\xi) \Psi_j(\xi) f_{\xi}(\xi) d\xi = \delta_{ij} \Psi_i \Psi_j \quad (II.10)$$

où  $\delta_{ij}$  est l'opérateur de Kronecker et  $f_{\xi}$  est la fonction de densité de probabilité jointe. Puisque les variables aléatoires  $\xi = \{\xi_1, \xi_2, \dots\}$  sont des variables Gaussiennes standard indépendantes, la fonction de densité de probabilité jointe s'écrit :

$$f_{\xi}(\xi) = \prod_i \frac{\exp[-\xi_i^2/2]}{\sqrt{2\pi}} \quad (II.11)$$

On note que les polynômes  $\{\Psi_i\}_{i=1}^{\infty}$  correspondent aux polynômes multidimensionnels d'Hermite qui peuvent être obtenus par le produit de polynômes unidimensionnels d'Hermite.

Dans la pratique, le développement en chaos polynômial est exprimé en fonction d'un nombre fini  $N$  de variables Gaussiennes standard. Ainsi, le développement en chaos polynômial défini par l'équation (II.9) peut être réécrit sous la forme suivante :

$$\theta(\xi_1, \xi_2, \dots, \xi_N) = \sum_{k=0}^{\infty} \theta_k \Psi_k(\xi_1, \xi_2, \dots, \xi_N) \quad (II.12)$$

La construction du développement en chaos polynômial sur la base d'un nombre fini de variables aléatoires ne doit pas être considéré comme une limitation, puisque dans la pratique, les systèmes d'ingénierie ou les systèmes physiques font généralement intervenir un nombre fini de paramètres incertains. De plus, ce développement converge, tant qu'on travaille dans l'espace de Hilbert de dimension finie et que les polynômes d'Hermite forment une base de Hilbert.

Pour des raisons pratiques, le chaos polynômial doit être tronqué au sens de l'ordre du polynôme global de l'équation (II.12). Autrement dit, le développement en chaos polynômial de la variable aléatoire  $\theta$  est construit en sommant un nombre fini de termes  $\theta_k \Psi_k$ . Ainsi, le développement en chaos polynômial de dimension finie  $N$  et d'ordre  $p$  fini de la variable aléatoire  $\theta$  est donné par :

$$\theta(\xi_1, \xi_2, \dots, \xi_N) = \sum_{k=0}^{P-1} \theta_k \Psi_k(\xi_1, \xi_2, \dots, \xi_N) \quad (II.13)$$

où  $P$ , le nombre de termes présents dans la sommation, est obtenu en fonction du nombre  $N$  de variables aléatoires et de l'ordre  $p$  du développement en chaos polynômial par :

$$P + 1 = \frac{(p + N)!}{p! N!} \quad (II.14)$$

### II.3.2 Chaos polynômial généralisé

Dans sa première version, présentée ci-dessus, le développement en chaos polynômial est construit au moyen de variables aléatoires Gaussiennes standard et de polynômes d'Hermite. Or, dans la pratique, les distributions des paramètres incertains sont loin d'être toujours Gaussiennes. Aussi, est-il convenable d'introduire d'autres familles de distributions, telle que la distribution Gamma ou la distribution exponentielle qui sont souvent plus adaptées que la distribution Gaussienne pour modéliser certains phénomènes rencontrés dans les problèmes d'ingénierie. Dans ce contexte, Xiu et Karniadakis [85] ont proposé l'utilisation du schéma d'Askey qui permet de classer les polynômes hypergéométriques orthogonaux. De plus, ils ont montré que le développement en chaos polynômial tel qu'il est défini par l'équation (II.13) peut avoir une convergence lente dans le cas de variables non Gaussiennes. Dans la version qu'ils proposent, les polynômes multidimensionnels  $\Psi_k$  ne sont pas limités aux polynômes d'Hermite, mais n'importe quel polynôme orthogonal du schéma d'Askey peut être employé. Dans la littérature, cette version est désignée par développement en *chaos polynômial généralisé*. Le tableau II.1 donne la correspondance entre certains types de distribution de variables aléatoires et familles de polynômes orthogonaux.

**Tableau II. 1 :** Correspondance (schéma d'Askey) entre les distributions de variables aléatoires et les polynômes orthogonaux

Type de VA	Distribution de VA	Support	Polynôme
Continue	Gaussienne	$\mathbb{R}$	Hermite
	Uniforme	$[-1,1]$	Legendre
	Gamma	$\mathbb{R}^+$	Laguerre
	Beta	$[-1,1]$	Jacobi
Discrète	Poisson	$\{0,1,2, \dots\}$	Charlier
	Binomiale	$\{0,1,2, \dots, M\}$	Krawtchouk

Ces polynômes orthogonaux dans l'espace de Hilbert correspondent au support de la densité de probabilité de la variable aléatoire. On note que, pour les problèmes faisant intervenir d'autres types de distributions de variables aléatoires (i.e. autres que celles données dans le tableau II.1), il est possible d'employer les transformations iso-probabilistes [73, 74] afin que le développement en chaos polynômial généralisé puisse être construit. A titre illustratif, une variable aléatoire lognormale peut être écrite en fonction d'une variable aléatoire Gaussienne, qui est utilisée par la suite en conjonction avec les polynômes d'Hermite.

### II.3.3 Représentation de la réponse d'un modèle

On considère un modèle ayant un nombre  $N$  de paramètres incertains groupés dans le vecteur  $\mathbf{x} = (x_1, \dots, x_N)$ , permettant de représenter d'une façon implicite un phénomène physique quelconque. Ce modèle peut être représenté par une fonction générique  $y = f(\mathbf{x})$ , où  $y$  est sa réponse. En supposant qu'on travaille dans l'espace aléatoire standard, la variable aléatoire  $Y$ , représentant l'aléa associé à la réponse du modèle, peut être approximée par un développement en chaos polynômial  $N$ -dimensionnel et d'ordre  $p$ , basé sur les polynômes d'Hermite, et défini par :

$$Y = f(\mathbf{X}) = f \circ T(\mathbf{U}) \approx \sum_{k=0}^{P-1} a_k \Psi_k(\mathbf{U}) \quad (II. 15)$$

où  $\{a_k\}_{k=0}^{P-1}$  sont les coefficients du développement en chaos polynômial et  $\{\Psi_k(\mathbf{U})\}_{k=0}^{P-1}$  sont les polynômes multidimensionnels d'Hermite. Ce développement converge au sens de la norme  $L_2$ , c'est-à-dire :

$$\lim_{P \rightarrow +\infty} \left\| Y - \sum_{k=0}^{P-1} a_k \Psi_k(\mathbf{U}) \right\|_{L_2}^2 = \lim_{P \rightarrow +\infty} \mathbb{E} \left[ \left( Y - \sum_{k=0}^{P-1} a_k \Psi_k(\mathbf{U}) \right)^2 \right] = 0 \quad (II. 16)$$

où  $\mathbb{E}[\cdot]$  désigne l'espérance mathématique.

En se référant à l'équation (II.15), il est clair que le développement en chaos polynômial nécessite la détermination des coefficients  $\{a_k\}_{k=0}^{P-1}$  et la construction des polynômes multidimensionnels d'Hermite.

Etant donné que l'utilisation de la transformation probabiliste  $T$  a permis d'écrire la variable aléatoire  $N$ -dimensionnelle  $\mathbf{X}$  en fonction de variables Gaussiennes standard indépendantes, les polynômes multidimensionnels d'Hermite  $\Psi_k$  peuvent être construits en effectuant le produit scalaire de polynômes d'Hermite unidimensionnels  $H_k$ . Soit  $\{H_n(U), n \in \mathbb{N}\}$  un ensemble de polynômes d'Hermite unidimensionnels orthogonaux par rapport à la mesure de probabilité Gaussienne, c'est-à-dire tels que :

$$\langle H_i(U), H_j(U) \rangle_{L_2} = \mathbb{E}[H_i(U) H_j(U)] = \delta_{ij} \quad (II. 17)$$

En effectuant le produit scalaire des polynômes d'Hermitte unidimensionnels  $\{H_n(U), n \in \mathbb{N}\}$ , on peut construire un ensemble de polynômes  $N$ -dimensionnels d'Hermitte  $\{\Psi_\alpha(U), \alpha \in \mathbb{N}^N\}$  de telle façon que leur ordre maximal soit inférieur ou égal à  $p$ , défini par :

$$\Psi_\alpha(U) = \prod_{i=1}^N H_{\alpha_i}(U) \quad (II.18)$$

où  $\alpha$  est un indice  $N$ -dimensionnel (i.e. une séquence de  $N$  entiers naturels) dont les composantes  $\{\alpha_i, i = 1, \dots, N\}$  représentent les ordres respectifs des polynômes d'Hermitte  $\{H_{\alpha_i}\}_{i=1}^N$  utilisés dans la construction du polynôme  $N$ -dimensionnel d'Hermitte  $\Psi_\alpha$ , tel que  $\alpha_i \geq 0$  et  $\sum_{i=1}^N \alpha_i \leq p$ .

Soit  $\alpha$  et  $\beta$  deux indices  $N$ -dimensionnels, en se basant sur l'équation (II.17) on peut écrire :

$$\mathbb{E}[\Psi_\alpha(U) \Psi_\beta(U)] = \prod_{i=1}^N \mathbb{E}[H_{\alpha_i}(U_i) H_{\beta_i}(U_i)] = \delta_{ij} \prod_{i=1}^N \mathbb{E}[H_{\alpha_i}^2(U_i)] \quad (II.19)$$

L'équation (II.19) montre clairement que la base du chaos polynômial est orthogonale.

Ainsi, il ne reste qu'à déterminer les coefficients inconnus  $\{a_k\}_{k=0}^{P-1}$  pour construire le développement en chaos polynômial de la variable aléatoire  $Y$  représentant la réponse du modèle. Dans ce contexte, la méthode de projection de Galerkin est l'une des premières méthodes utilisées. Elle est basée sur la minimisation du résidu sous la contrainte qu'il soit orthogonal à la base  $\{\Psi_i\}_{i=0}^P$  du développement en chaos polynômial, c'est-à-dire :

$$\left\langle \sum_{k=0}^{P-1} a_k \Psi_k(\mathbf{U}), \Psi_i \right\rangle = 0, \quad \forall i = 0, \dots, P-1 \quad (II.20)$$

La méthode de Galerkin implique la résolution d'un système de  $P$  équations couplées. Cette approche est souvent dite *intrusive* du fait qu'elle nécessite la modification de la formulation mathématique du modèle. Lorsque le modèle mathématique est non linéaire, ce qui est souvent le cas dans la pratique, sa modification n'est pas une tâche triviale et peut induire des difficultés dans l'implémentation. Afin de pallier ce problème, des méthodes dites *non intrusives* ont été proposées. L'avantage majeur de ces méthodes est qu'elles ne nécessitent pas la modification de la formulation du modèle. Les coefficients du développement en chaos polynômial sont déterminés à partir de l'évaluation du modèle en un certain nombre de réalisations des paramètres incertains. Dans le paragraphe suivant, nous présentons les méthodes *non intrusives* les plus populaires.

### II.3.4 Détermination des coefficients du DCP

Les méthodes non intrusives sont elles-mêmes classées en deux familles : approches de projection et approches de régression. Pour les approches de projection, chaque coefficient du développement en chaos polynômial est défini comme une intégrale multidimensionnelle qui peut être calculée par quadrature ou par simulations, alors que pour les approches de régression, les coefficients du développement sont déterminés en résolvant un système d'équations issu d'une minimisation au sens des moindres carrés.

#### II.3.4.1 Méthodes de projection

Les méthodes de projection utilisent l'orthogonalité de la base  $\{\Psi_\alpha, |\alpha| \leq p\}$  du développement en chaos polynômial pour calculer les coefficients  $\{a_\alpha, |\alpha| \leq p\}$  inconnus. En

effet, la projection du développement en chaos polynômial défini par l'équation (II.17) sur la base  $\{\Psi_\alpha, |\alpha| \leq p\}$ , en se référant au produit scalaire  $\langle g, h \rangle \mapsto \mathbb{E}[g, h]$ , donne :

$$a_\alpha = \frac{\mathbb{E}[f\Psi_\alpha]}{\mathbb{E}[\Psi_\alpha^2]} \quad (II.21)$$

Ainsi, chaque coefficient du développement en chaos polynômial peut être défini par une intégrale  $N$ -dimensionnelle :

$$a_\alpha = \frac{1}{\mathbb{E}[\Psi_\alpha^2]} \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u}) \varphi_N(\mathbf{u}) d\mathbf{u} \quad (II.22)$$

où l'espérance  $\mathbb{E}[\Psi_\alpha^2]$  peut être obtenue analytiquement et  $\varphi_N$  est la densité de probabilité d'une variable aléatoire Gaussienne standard  $N$ -dimensionnelle. Pour simplifier la notation, dans ce qui suit, désignons par  $I_\alpha^N$  l'intégrale de l'équation (II.22) :

$$I_\alpha^N = \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u}) \varphi_N(\mathbf{u}) d\mathbf{u} \quad (II.23)$$

Il est clair qu'il suffit de calculer l'intégrale  $I_\alpha^N$  pour déterminer le coefficient  $a_\alpha$  du développement en chaos polynômial. Dans la pratique, cette intégrale est calculée au moyen de méthodes d'intégration numérique, qui consistent en général à l'approximer par une somme pondérée de la forme suivante :

$$I_\alpha^N \approx \sum_{i=1}^M \omega^i f \circ T(\mathbf{u}^i) \Psi_\alpha(\mathbf{u}^i) \quad (II.24)$$

Les méthodes d'intégration numérique disponibles dans la littérature se distinguent par le choix des points  $\{\mathbf{u}^i\}_{i=1}^M$  et les pondérations  $\{\omega^i\}_{i=1}^M$ . Elles peuvent être réparties en deux catégories distinctes : les méthodes de simulations et les méthodes de quadrature. Le choix d'une méthode d'intégration est dicté par sa capacité à s'adapter aux contraintes que peut impliquer la représentation mathématique du système physique. En particulier, on peut citer la régularité de l'intégrande  $f \circ T(\mathbf{u}^i) \Psi_\alpha(\mathbf{u}^i)$  et la dimension  $N$  de l'intégrale (i.e. le nombre de paramètres incertains). Dans ce qui suit, nous présentons les méthodes d'intégration les plus populaires ainsi que certaines de leurs variantes.

### II.3.4.1.1 Méthodes de simulations

#### II.3.4.1.1.1 Simulations de Monte-Carlo

L'idée de base de l'intégration par simulations, est de générer un ensemble  $\{\mathbf{u}^i\}_{i=1}^M$  de points selon une procédure aléatoire. Les poids d'intégration sont définis par  $\omega^i = 1/M$ ,  $i = 1, \dots, M$ . L'approximation de l'intégrale  $I_\alpha^N$  est donnée par :

$$\tilde{I}_\alpha^N = \frac{1}{M} \sum_{i=1}^M f \circ T(\mathbf{u}^i) \Psi_\alpha(\mathbf{u}^i) \quad (II.25)$$

La précision de cette approximation dépend essentiellement de la procédure de génération des points d'intégration  $\{\mathbf{u}^i\}_{i=1}^M$ . Dans le cas des simulations de Monte-Carlo, les points d'intégration sont obtenus au moyen de générateurs de nombres pseudo-aléatoires en se référant à la densité de probabilité de chacun des paramètres incertains. Dans le cas où le développement en chaos polynômial est exprimé dans l'espace standard, les points

d'intégration sont générés selon la distribution d'une variable aléatoire Gaussienne standard. La prise en compte des distributions réelles des différents paramètres incertains est effectuée implicitement par l'utilisation d'une transformation iso-probabiliste.

Du fait du caractère aléatoire des points d'intégration, l'intégrande  $f \circ T(\mathbf{u}^i) \Psi_\alpha(\mathbf{u}^i)$  peut être assimilée à une variable aléatoire dont l'espérance n'est autre que l'approximation  $\tilde{I}_\alpha^N$  de l'intégrale  $I_\alpha^N$ . L'évaluation de la précision de cette estimation est effectuée au moyen de la mesure de l'erreur  $\epsilon$  définie par :

$$\epsilon = \sqrt{\frac{\mathbb{V}[f \circ T(\mathbf{U}) \Psi_\alpha(\mathbf{U})]}{M}} \quad (II.26)$$

où  $\mathbb{V}[\cdot]$  désigne la variance d'une variable aléatoire.

On constate que l'erreur  $\epsilon$  décroît avec la quantité  $1/\sqrt{M}$ , ce qui révèle la limite majeure des simulations de Monte-Carlo, à savoir leur convergence lente. De plus, la convergence est d'autant plus lente que l'ordre  $p$  du développement en chaos polynômial est élevé. En effet, la variance  $\mathbb{V}[f \circ T(\mathbf{U}) \Psi_\alpha(\mathbf{U})]$  de l'intégrande augmente avec l'ordre  $|\alpha|$  du polynôme  $N$ -dimensionnel  $\Psi_\alpha$ .

Afin de remédier à ce problème, des variantes améliorées des simulations de Monte-Carlo ont été développées. Leur particularité est qu'elles utilisent des techniques de génération de variables aléatoires plus performantes dans le sens où elles permettent une meilleure représentation de l'espace aléatoire (i.e. l'espace formé par le domaine de variabilité des paramètres incertains). Autrement dit, les points d'intégration sont tirés sans favoriser une région de l'espace aléatoire par rapport à une autre. Parmi ces techniques, on peut citer les tirages en hyper-cube latin et les séquences quasi-aléatoires, qui seront l'objet des deux paragraphes suivants, qui peuvent contribuer à résoudre les problèmes à dimension stochastique élevée en diminuant le nombre d'appels au modèle.

#### II.3.4.1.1.2 Tirages en hyper-cube latin

La technique des tirages en hyper-cube latin a été proposée dans un premier temps par McKay et al. [11], puis elle a été le sujet de plusieurs tentatives de développement dans différents contextes scientifiques [86-90]. L'idée de base des tirages en hyper-cube latin est de générer des nombres pseudo-aléatoires plus représentatifs de la distribution de probabilité jointe des paramètres incertains que ceux obtenus dans le cas des simulations de Monte-Carlo. Afin de faciliter l'exposition de ses fondements et simplifier ainsi sa compréhension, nous nous proposons de commencer par présenter sa version originale, la plus simple.

Soit  $M$  le nombre de simulations nécessaires pour garantir une bonne estimation de l'intégrale  $I_\alpha^N$  et  $N$  le nombre de paramètres incertains ; ainsi, l'espace aléatoire représentant la variabilité des différents paramètres incertains est  $N$ -dimensionnel. Supposons que le support de chaque paramètre incertain puisse être représenté (au moyen d'une transformation iso-probabiliste) par l'intervalle  $[0,1]$ . Conceptuellement parlant, la procédure des tirages en hyper-cube latin consiste à diviser l'hyper-cube  $[0,1]^N$  en un nombre  $M$  d'intervalles ou strates équiprobables. Soit  $V = \{V_1, \dots, V_N\}$  une variable aléatoire  $N$ -dimensionnelle définie sur l'intervalle  $[0,1]^N$ , représentant les paramètres incertains. Les  $M$  réalisations des composantes (i.e. variables aléatoires uniformes scalaires définies sur l'intervalle  $[0,1]$ )  $V_i, i = 1, \dots, N$  sont générées d'une façon aléatoire en sélectionnant une valeur possible dans chaque strate. Cette procédure consiste dans un premier lieu à construire une matrice  $\mathbf{P}$  de dimension  $M \times N$ , dans laquelle chaque colonne est obtenue par permutation aléatoire des entiers  $\{1, \dots, M\}$ , et une matrice  $\mathbf{R}$  de dimension  $M \times N$  dont les composantes sont des nombres aléatoires indépendants générés

selon la distribution d'une variable aléatoire uniforme sur  $[0,1]$ . Ensuite, ces deux matrices sont utilisées pour définir un plan de tirages dans l'hyper-cube  $[0,1]^N$ , représenté par la matrice d'expérimentation  $\mathbf{V}$  définie par :

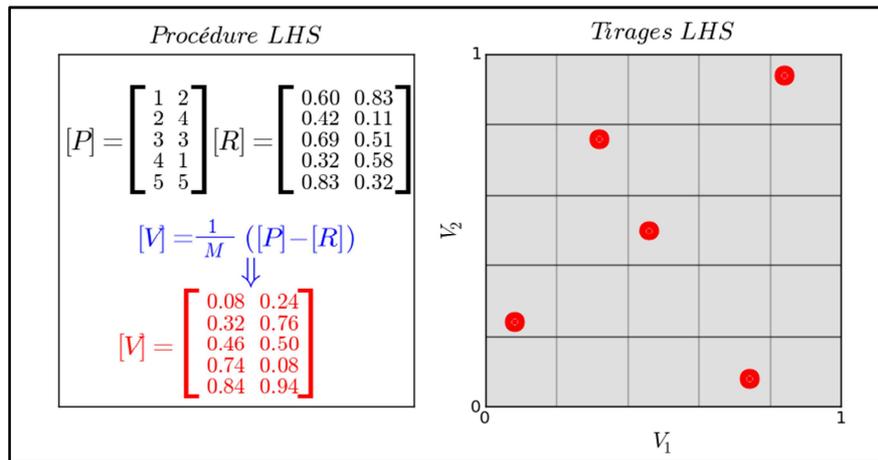
$$\mathbf{V} = \frac{1}{M}(\mathbf{P} - \mathbf{R}) \quad (II.27)$$

où chaque ligne  $\mathbf{v}^i$ ,  $i = 1, \dots, M$  de la matrice  $\mathbf{V}$  représente une réalisation des paramètres incertains dans l'espace aléatoire défini par l'hyper-cube  $[0,1]^N$ .

Enfin, l'ensemble des réalisations  $\{\mathbf{v}^1, \dots, \mathbf{v}^M\}$  contenues dans la matrice d'expérimentation  $\mathbf{V}$  est transformé en un ensemble de nombres pseudo-aléatoires  $\{\mathbf{u}^1, \dots, \mathbf{u}^M\}$  distribués dans l'espace aléatoire standard (espace des variables Gaussiennes standard). Cette transformation est définie par :

$$\mathbf{u}^{ij} = \Phi^{-1}(v^{ij}) \quad (II.28)$$

où  $\Phi^{-1}$  est l'inverse de la fonction de répartition d'une variable Gaussienne standard. Les réalisations  $\{\mathbf{u}^1, \dots, \mathbf{u}^M\}$  sont alors les points d'intégration utilisés dans l'équation (II.25) pour estimer l'intégrale  $I_{\alpha}^N$ . La figure II.3 illustre la procédure pour la génération de cinq réalisations possibles dans l'hyper-cube  $[0,1]^2$  (cas de deux paramètres incertains).



**Figure II. 3 :** Illustration de la procédure des tirages en hyper-cube latin pour la génération de cinq réalisations dans le cas de deux paramètres incertains

On constate que les réalisations sont réparties sur tout le domaine  $[0,1]^2$  puisque la construction du plan des tirages en hyper-cube latin nécessite une réalisation de chaque ligne (variable 1) et chaque colonne (variable 2). Cependant, si les  $M$  réalisations ont été choisies d'une manière aléatoire comme pour le cas des simulations de Monte-Carlo, ces réalisations peuvent être condensées dans une région particulière de l'hyper-cube  $[0,1]^N$ , menant ainsi au risque que d'autres régions représentatives de la variabilité des paramètres incertains ne soient pas exploitées. De plus, même si la distribution de chaque paramètre incertain est bien représentée, il y a un risque d'apparition d'une *fausse corrélation* entre les réalisations obtenues par la procédure de tirages en hyper-cube latin décrite ci-dessus. Afin de remédier à ce problème, Iman et Conover [86] ont montré que cette *fausse corrélation* peut être atténuée en modifiant la matrice de permutation  $\mathbf{P}$ . La procédure consiste à diviser les éléments de la matrice  $\mathbf{P}$  par le nombre de réalisations plus un (i.e.  $M + 1$ ), puis ils sont transformés au moyen de la distribution  $\Phi$  d'une variable aléatoire Gaussienne standard. Ainsi, les éléments de la nouvelle matrice  $\mathbf{P}^0$  sont obtenus par :

$$P_{ij}^0 = \Phi^{-1} \left( \frac{P_{ij}}{M+1} \right) \quad (II.29)$$

La matrice de covariance  $\mathbf{C}_{P^0}$  de la matrice  $\mathbf{P}^0$  est calculée, et factorisée selon la décomposition de Cholesky comme suit :

$$\mathbf{C}_{P^0} = \mathbf{L}\mathbf{L}^T \quad (II.30)$$

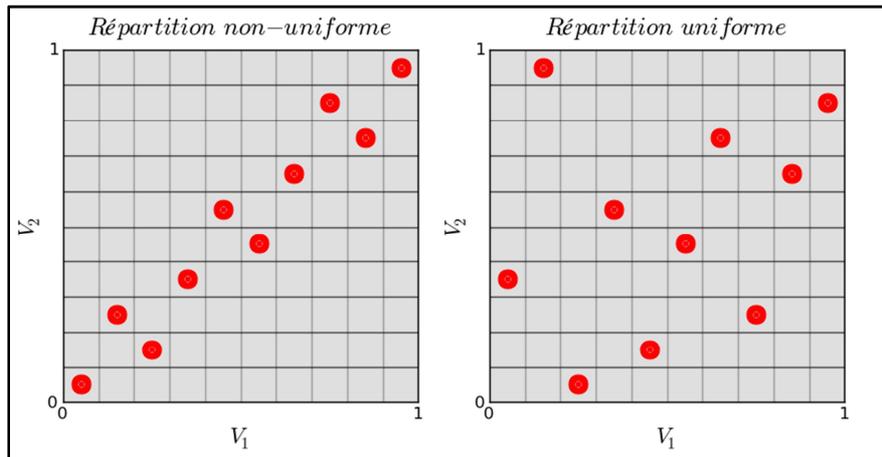
où  $\mathbf{L}$  est une matrice triangulaire inférieure.

Ainsi une nouvelle matrice  $\mathbf{P}^{0,*}$ , dont la covariance entre ses réalisations est égale à l'unité, est obtenue par :

$$\mathbf{P}^{0,*} = \mathbf{P}^0(\mathbf{L}^{-1})^T \quad (II.31)$$

Les rangs des éléments des colonnes de la matrice  $\mathbf{P}^{0,*}$  (permutations des entiers  $\{1, \dots, M\}$ ) deviennent les éléments des colonnes de la nouvelle matrice de permutation  $\mathbf{P}^*$ .

Maintenant, si les éléments de la nouvelle matrice de permutation  $\mathbf{P}^*$  sont utilisés dans l'équation (II.28), il en résulte que la dépendance (i.e. la corrélation) entre les réalisations données par la nouvelle matrice d'expérimentation  $\mathbf{V}$  est faible. La figure II.4 illustre d'une manière schématique l'effet de la procédure de réduction de corrélation décrite ci-dessus sur la répartition des réalisations obtenues par les tirages en hyper-cube latin.



**Figure II. 4 :** Illustration de l'effet de la réduction de corrélation sur la répartition des réalisations obtenues par tirages en hyper-cube latin

Dans le cas où la matrice de corrélation cible (i.e. la matrice de corrélation entre les nouvelles réalisations correspondant à la nouvelle matrice d'expérimentation  $\mathbf{V}$ ) est différente de l'unité, l'équation (II.31) est remplacée par :

$$\mathbf{P}^{0,*} = \mathbf{P}^0(\mathbf{L}^{-1})^T \hat{\mathbf{L}}^T \quad (II.32)$$

où  $\hat{\mathbf{L}}$  est une matrice triangulaire inférieure résultant de la décomposition de la matrice de corrélation cible.

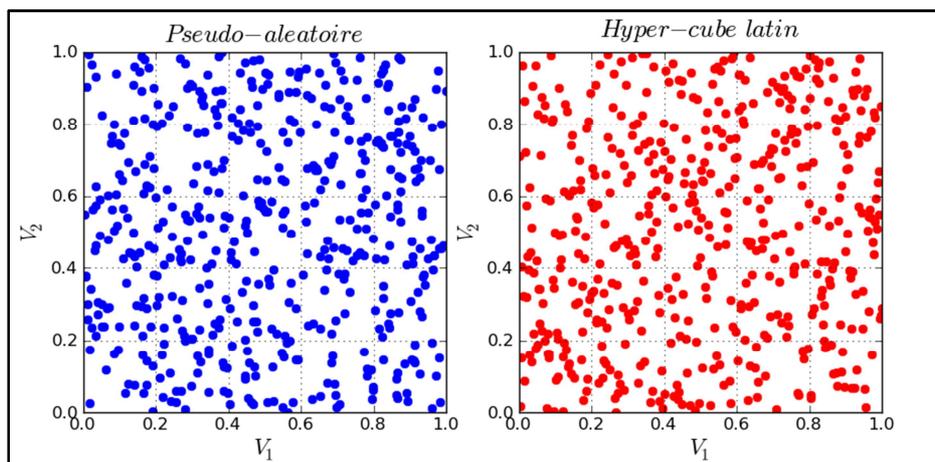
On note que la procédure de réduction de la corrélation décrite ci-dessus nécessite que la matrice de covariance soit définie positive, ce qui implique que le nombre de réalisations  $M$  soit supérieur au nombre de paramètres incertains  $N$ . Il est important de signaler aussi que la procédure de réduction de corrélation peut introduire un certain biais sur les estimations [93].

On s'attend à ce que les simulations de Monte-Carlo basées sur les tirages en hyper-cube latin soient plus performantes que celles utilisant les générateurs de nombres pseudo-aléatoires classiques. Ainsi, Owen [92] a montré que l'utilisation de points générés par tirages en hyper-

cube latin pour le calcul de la somme définie par l'équation (II.25), fournit une estimation plus précise de l'intégrale  $I_\alpha^N$  que celle donnée par les simulations de Monte-Carlo telle qu'elles sont décrites dans le paragraphe précédent. Cela est fondé sur la comparaison des variances des estimations fournies par chaque approche :

$$\mathbb{V}[f \circ T(\mathbf{U}) \Psi_\alpha(\mathbf{U})]_{HL} \leq \frac{M}{M+1} \mathbb{V}[f \circ T(\mathbf{U}) \Psi_\alpha(\mathbf{U})]_{MC} \quad (II.33)$$

L'inégalité donnée par l'équation (II.33) veut dire qu'une estimation fournie par les simulations de Monte-Carlo basée sur  $M$  points générés par hyper-cube latin est toujours meilleure que celle donnée par les simulations de Monte-Carlo basées sur  $M - 1$  points obtenus par des générateurs de nombres pseudo-aléatoires classiques. La figure II.5 compare les répartitions de 1000 points obtenus par un générateur de nombres pseudo-aléatoires et ceux donnés par les tirages en hyper-cube latin. On constate que la répartition est relativement plus uniforme dans le cas des tirages en hyper-cube latin.



**Figure II. 5 :** Répartition des points obtenus par un générateur de nombres pseudo-aléatoires et par tirages en hyper-cube latin dans le cas bidimensionnel

La technique de tirages en hyper-cube latin a été le sujet de plusieurs applications portant sur différents types de problèmes. Elle a été appliquée par Olsson et Sanderberg [87] dans le contexte des éléments finis stochastiques où l'aléa associé aux paramètres incertains est modélisé par un champs aléatoire. Giocel et Ghanem [93] l'ont utilisée pour le calcul des coefficients du développement en chaos polynômial. Olsson et al. [91] l'ont employé en conjonction avec les tirages d'importance dans le contexte d'analyse de fiabilité. Il faut signaler tout de même que les tirages en hyper-cube latin souffrent d'une limite majeure. En effet, la précision de l'estimation fournie par les tirages en hyper-cube latin ne peut pas être améliorée de façon itérative (i.e. en ajoutant des points supplémentaires aux points déjà existants), puisque les nouveaux points ne seront pas dans l'hyper-cube latin du départ. Autrement dit, il faut construire un nouveau hyper-cube latin permettant d'obtenir plus de réalisations que le premier et répéter ainsi tous les calculs. Cette difficulté a été levée par Sallaberry et al. [94] qui ont proposée une procédure qui permet de construire, à partir d'un plan d'expérimentation contenant  $M$  points obtenus par tirages en hyper-cube latin, un nouveau plan contenant  $2M$  points d'expérimentation.

#### II.3.4.1.1.3 Méthodes Quasi Monte-Carlo

Les méthodes Quasi Monte-Carlo [12] sont des versions déterministes des simulations de Monte-Carlo classiques basées sur les nombres quasi-aléatoires. Elles ont été employées en particulier dans le calcul des intégrales multidimensionnelles. Leur particularité est qu'elles

fournissent des estimations précises tout en nécessitant un nombre limité d'analyses du modèle représentant le système physique. De plus, elles sont faciles à implémenter du fait qu'elles demandent seulement de remplacer les nombres pseudo-aléatoires utilisés dans les simulations de Monte-Carlo par des séquences de points déterministes qui ont une discrépance faible, c'est-à-dire une répartition plus uniforme dans le domaine de variation des paramètres incertains. Ces séquences sont souvent désignées par séquences à faible discrépance ou séquences quasi-aléatoires. Les méthodes Quasi Monte-Carlo permettent d'obtenir une convergence plus rapide que les simulations de Monte-Carlo classiques. En effet, Schürer [95] a montré, en utilisant l'inégalité de Koksma-Hlawka [96], que le taux de convergence des méthodes Quasi Monte-Carlo est en  $\log^N(M)/M$ , ce qui est bien meilleur que le taux de convergence en  $1/M$  des simulations de Monte-Carlo classiques, et ce pour différents types d'intégrales dont la variance est bornée au sens de Hardy et Krause [96]. De plus, il a montré, à partir de résultats empiriques, que même si, en apparence, le terme  $\log^N(M)$  domine la convergence pour des valeurs raisonnables de  $M$  et  $N$ , le taux de convergence en  $1/M$  est toujours garanti. Ainsi, les méthodes Quasi Monte-Carlo sont souvent préférées aux simulations de Monte-Carlo classiques pour le calcul d'intégrales multidimensionnelles. La principale différence entre les méthodes Quasi Monte-Carlo réside dans le type de séquence à faible discrépance utilisée. Dans ce contexte, une grande variété de séquences quasi-aléatoires a été proposée dans la littérature. Parmi celles-ci, on peut citer les séquences de Halton [56], de Hammersley [57], de Sobol [97] ou de Faure [98].

Une séquence unidimensionnelle de Halton associée à un nombre premier  $k$  est définie comme le développement des entiers  $1, \dots, l$  sur la base  $k$ . Ainsi, le  $M^{\text{ème}}$  terme de la séquence est donné par :

$$v^M = \sum_{i=0}^r \frac{a_i}{k^{i+1}} \quad (II.34)$$

où les entiers  $\{a_i\}_{i=0}^r$  sont issus du développement de  $M - 1 = \sum_{i=0}^r a_i k^i$  sur la base  $k$ .

Les nombres quasi-aléatoires sont générés dans l'intervalle  $[0,1]$  suivant un cycle de  $k$  termes croissants. Dans chaque cycle, les termes sont distants d'un incrément égal à  $1/k$ .

Ainsi, une séquence de Halton  $N$ -dimensionnelle est construite à partir de  $N$  séquences de Halton unidimensionnelles basées sur un ensemble de  $N$  plus petits nombres premiers  $\{k_1, \dots, k_N\}$  deux à deux distincts. Il faut noter que l'inconvénient majeur des séquences quasi-aléatoires de Halton est que, pour  $N$  très élevé (i.e. un grand nombre de paramètres incertains), le nombre premier  $k_N$  risque d'être très élevé, ce qui fait augmenter considérablement la longueur du cycle des termes et par conséquent le nombre de tirages nécessaire pour la construction de la séquence  $N$ -dimensionnelle.

Plusieurs séquences quasi-aléatoires ont été dérivées de la séquence de Halton. Ces dernières ne sont autres que des versions améliorées en terme de discrépance (i.e. les points sont répartis plus uniformément dans le domaine de variation des paramètres incertains) de la séquence de Halton. Dans ce contexte, la séquence de Hammersley [57] peut être citée, dans laquelle le premier terme est égal à  $\frac{2k-1}{2N}$ . Ainsi, une séquence  $N$ -dimensionnelle de Hammersley est donnée par :

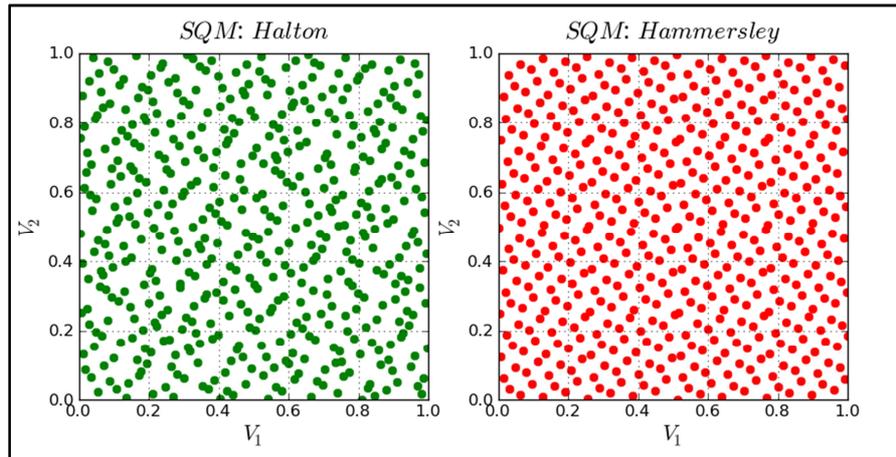
$$V = \left\{ \frac{2k-1}{2M}, v^1, \dots, v^{N-1} \right\}, \quad k = 1, \dots, M \quad (II.35)$$

où  $\{v^1, \dots, v^{N-1}\}$  sont les  $N - 1$  termes d'une séquence unidimensionnelle de Halton basée sur le nombre premier  $k$ .

Dai et Wang [57] ont montré que la discrédance  $D(M)$  de la séquence de Hammersley admet une borne supérieure définie par :

$$D(M) \leq \frac{1}{M} \prod_{i=1}^{N-1} \frac{k_i \log(Mk_i)}{\log(k_i)} \quad (II. 36)$$

La figure II.6 illustre la répartition, dans le cas bidimensionnel, de 500 points correspondant respectivement à une séquence de Halton et à une séquence de Hammersley. On constate que la répartition des points de la séquence de Hammersley est plus uniforme que celle des points générés par la séquence de Halton.



**Figure II. 6 :** Répartition des points obtenus par les séquences de Halton et Hammersley dans le cas bidimensionnel

La séquence de Sobol [97] permet de surmonter la limite de la séquence de Halton dans le cas de problèmes à grande dimension en choisissant  $k_i = 2$ ,  $i = 1, \dots, N$ . Chacune des  $N$  séquences unidimensionnelles de Sobol associée à chaque dimension est une permutation d'une séquence de Halton basée sur le nombre premier  $k = 2$ . Cependant, lorsque le nombre de paramètres incertains augmente, la probabilité d'avoir des permutations identiques devient élevée, et ainsi l'uniformité des points générés est médiocre. La séquence de Faure [98] est obtenue de la même manière que celle de Sobol par  $N$  permutations de séquences unidimensionnelles de Halton, en choisissant le nombre premier  $k$  comme le plus petit nombre premier supérieur ou égal au nombre de paramètres incertains  $N$ .

Le paramètre permettant d'évaluer l'efficacité d'une séquence de nombres quasi aléatoires est la discrédance  $D(M)$  ou la mesure d'uniformité de la répartition des points dans l'espace aléatoire. Elle est généralement définie par :

$$D(M) = \sup_{J \subset I^N} \left| \frac{\text{card}(\{v^1, \dots, v^M\} \cap J)}{M} - \mathcal{V}(J) \right| \quad (II. 37)$$

où  $J$  est un hyper-cube inclus dans l'hyper-cube unité  $I^N = [0,1]^N$ ,  $\mathcal{V}(J)$  désigne son volume et  $\{v^1, \dots, v^M\}$  sont les points correspondants à la séquence  $N$ -dimensionnelle. Cette mesure de la discrédance admet une borne supérieure donnée par :

$$D(M) \leq C_N \frac{\log^N(M)}{M} + o\left(\frac{\log^{N-1}(M)}{M}\right) \quad (II. 38)$$

L'efficacité des séquences de nombres quasi aléatoires a été analysée dans le contexte de l'intégration de fonctions multidimensionnelles par Sobol [99] et Schürer [95]. Ils ont montré

qu'elles sont plus performantes que les simulations de Monte-Carlo classiques. Dai et Wang [57] ont effectué une étude comparative dans le contexte d'analyse de fiabilité, dans laquelle ils ont montré, à travers différents types de fonctions d'état limite, que les estimations fournies par les séquences de Halton et Hammersley sont plus précises que celles données par les simulations de Monte-Carlo basées sur les nombres pseudo-aléatoires ou sur les tirages en hyper-cube latin. Sudret et al. [100] ont employé les séquences de nombres quasi aléatoires de Sobol pour l'estimation des coefficients du développement en chaos polynômial. Ils ont montré que ce type de séquence peut réduire le coût de calcul par un facteur de 10, et ce par rapport à celui des simulations de Monte-Carlo classiques.

Généralement, la convergence des simulations de Monte-Carlo basées sur les séquences quasi-aléatoires est plus rapide que celle des simulations de Monte-Carlo basées sur les nombres pseudo-aléatoires dans le cas où le nombre de paramètres incertains est limité ( $N \leq 10$ ). Toutefois, elles peuvent perdre leur efficacité lorsque la dimension de l'espace aléatoire  $N$  augmente. Caflish et al. [77] ont expliqué ce problème par la notion de dimension stochastique efficace. Ce paramètre dépend soit du nombre limité de paramètres les plus influents sur la réponse, soit du nombre limité des interactions entre les paramètres. En effet, ils ont montré que, si la dimension efficace est inférieure à la dimension nominale, les nombres pseudo-aléatoires peuvent améliorer l'efficacité des simulations de Monte-Carlo. Ce résultat a été confirmé par Kucherenko et al. [70] dans le contexte de calcul des indices de sensibilité associés à la décomposition de la variance.

On note que les points d'intégration  $\{\mathbf{u}^1, \dots, \mathbf{u}^M\}$  à utiliser dans l'équation (II.25) pour l'estimation de l'intégrale  $N$ -dimensionnelle  $I_\alpha^N$  sont obtenus par une transformation semblable à celle définie par l'équation (II.28), des points  $\{\mathbf{v}^1, \dots, \mathbf{v}^M\}$  donnés par les séquences quasi-aléatoires.

Si les dérivées mixtes de l'intégrande  $f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u})$  existent, ce qui est une hypothèse plus forte que celle de la variance finie des simulations de Monte-Carlo classiques, en se basant sur l'inégalité Koksma-Hlawka [96], l'erreur sur l'estimation de l'intégrale  $I_\alpha^N$  est donnée par :

$$|I_\alpha^N - \tilde{I}_\alpha^N| \leq \|f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u})\|_{VB} D(M) \quad (II.39)$$

où  $\|\cdot\|_{VB}$  est la variance totale de la fonction  $f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u})$  bornée sur l'hyper-cube  $[0,1]^N$  dans le sens de Hardy et Krause [96].

### II.3.4.1.2 Méthodes de quadrature

Les méthodes de quadrature sont des alternatives aux méthodes de simulations pour le calcul des intégrales. Elles diffèrent dans la façon de sélectionner les points et les pondérations d'intégration. Dans ce qui suit, nous présentons les principales méthodes de quadrature utilisées en particulier dans le calcul des intégrales multidimensionnelles telles que la méthode de quadrature de Smolyak [17] et la méthode de quadrature de Stroud [16]. Mais, avant cela, nous commençons par la présentation de la forme la plus standard des méthodes de quadrature.

#### II.3.4.1.2.1 Quadrature par produit tensoriel complet

L'intégrale  $N$ -dimensionnelle  $I_\alpha^N$  définie par l'équation (II.25) peut être approximée par la formule de quadrature  $N$ -dimensionnelle suivante, obtenue par produit tensoriel de formules de quadrature unidimensionnelles :

$$\tilde{I}_\alpha^N = I_{M_1} \otimes \dots \otimes I_{M_N} = \sum_{i_1=1}^{M_1} \dots \sum_{i_N=1}^{M_N} \omega_1^{i_1} \dots \omega_N^{i_N} f \circ T(\mathbf{u}_1^{i_1} \dots \mathbf{u}_N^{i_N}) \Psi_\alpha(\mathbf{u}_1^{i_1} \dots \mathbf{u}_N^{i_N}) \quad (II.40)$$

où  $\otimes$  est le symbole du produit tensoriel et  $I_{M_j}$ ,  $j = 1, \dots, N$ , désignent les formules de quadrature unidimensionnelles basées respectivement sur  $M_j$  points,  $j = 1, \dots, N$ , de coordonnées  $u_j^{i_j}$  et pondérations  $\omega_j^{i_j}$ .

Généralement, on emploie une formule de quadrature *isotrope*, qui n'est autre que celle définie par l'équation (II.40) dans laquelle le nombre de points d'intégration est le même dans chaque direction de l'espace aléatoire,  $M_1 = \dots = M_N = M$ . Ce paramètre  $M$  peut être interprété comme l'ordre de la formule de quadrature. Un ordre de quadrature  $M$  élevé implique plus de termes dans les sommes définies par l'équation (II.40), et ainsi plus de précision pour les estimations. Il est dépendant principalement de la régularité de l'intégrande (i.e. dans notre cas la fonction  $f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u})$ ). Les points et les pondérations sont pris égaux aux points et pondérations de Gauss-Hermite, pour cette étude dans laquelle le développement en chaos polynômial est construit dans l'espace standard. Par conséquent, cette règle de quadrature permet de fournir la valeur exacte de l'intégrale d'un polynôme  $N$ -dimensionnel de degré inférieur à  $2M - 1$ . En remplaçant la quantité  $f \circ T$  dans l'équation (II.25) par son développement en chaos polynômial de degré  $p$  donné par l'équation (II.15), le polynôme résultant représentant l'intégrande  $f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u})$  est de degré  $p + |\alpha|$ . Puisque l'indice  $N$ -dimensionnel  $\alpha = \{\alpha_i, i = 1, \dots, N\}$  est de degré  $\sum_{i=1}^N \alpha_i$  inférieur ou égal à  $p$ , le degré maximal du polynôme représentant l'intégrande est égal à  $2p$ . Par conséquent, l'intégrale  $I_\alpha^N$  peut être calculée exactement en utilisant une formule de quadrature  $N$ -dimensionnelle isotrope basée sur  $M = p + 1$  points et pondérations d'intégration de Gauss-Hermite, ce qui implique  $(p + 1)^N$  appels au modèle représentant le système physique. Ainsi, on peut constater immédiatement la limite de l'intégration basée sur la quadrature par produit tensoriel complet. En effet, le coût de calcul est fortement dépendant de la dimension  $N$  de l'intégrale  $I_\alpha^N$ , puisqu'elle détermine le nombre nécessaire d'appels au modèle pour un ordre  $M$  de la formule de quadrature associée à une précision donnée. Lorsque la dimension  $N$  est élevée, le nombre d'appels au modèle augmente d'une façon exponentielle. Pour y remédier, la formule de quadrature de Smolyak [17] a été proposée.

#### II.3.4.1.2.2 Quadrature de Smolyak

L'idée de base de la méthode de quadrature de Smolyak [17] est aussi d'effectuer un produit tensoriel comme c'est le cas de la méthode de quadrature présentée dans le paragraphe précédent. Cependant, ce produit tensoriel est défini comme une combinaison linéaire de formules de quadratures unidimensionnelles ayant des degrés élevés dans quelques directions et des degrés bien moindres dans les dimensions restantes. Considérons des formules de quadrature unidimensionnelles de degré  $l \geq 1$ , et basées sur  $M_l$  points et pondérations d'intégration, permettant de calculer la valeur exacte de l'intégrale d'un polynôme unidimensionnel de degré inférieur ou égal à  $2M_l - 1$ . En effectuant la combinaison linéaire des produits de ces formules de quadratures unidimensionnelles, la formule de quadrature  $N$ -dimensionnelle de degré  $l \leq N$  de Smolyak  $\mathcal{S}(l, N)$  donnant une estimation de l'intégrale  $I_\alpha^N$  est définie par :

$$\tilde{I}_\alpha^N = \mathcal{S}(l, N) = \sum_{|\mathbf{i}|=l}^{l+N-1} (-1)^{l+N-1-|\mathbf{i}|} C_{N-1}^{|\mathbf{i}|-l} Q_{|\mathbf{i}|}^N \quad (II.41)$$

où  $\mathbf{i} = \{i_1, \dots, i_N\} \in \mathbb{N}_+^N$  est un multi-indice de somme  $|\mathbf{i}| = \sum_{k=1}^N i_k$ ,  $C$  désigne l'opérateur combinaison et  $Q_{|\mathbf{i}|}^N$  est la formule de quadrature  $N$ -dimensionnelle obtenue par produit tensoriel de formules de quadrature unidimensionnelles de degrés respectifs  $\{i_1, \dots, i_N\}$ , donnée par :

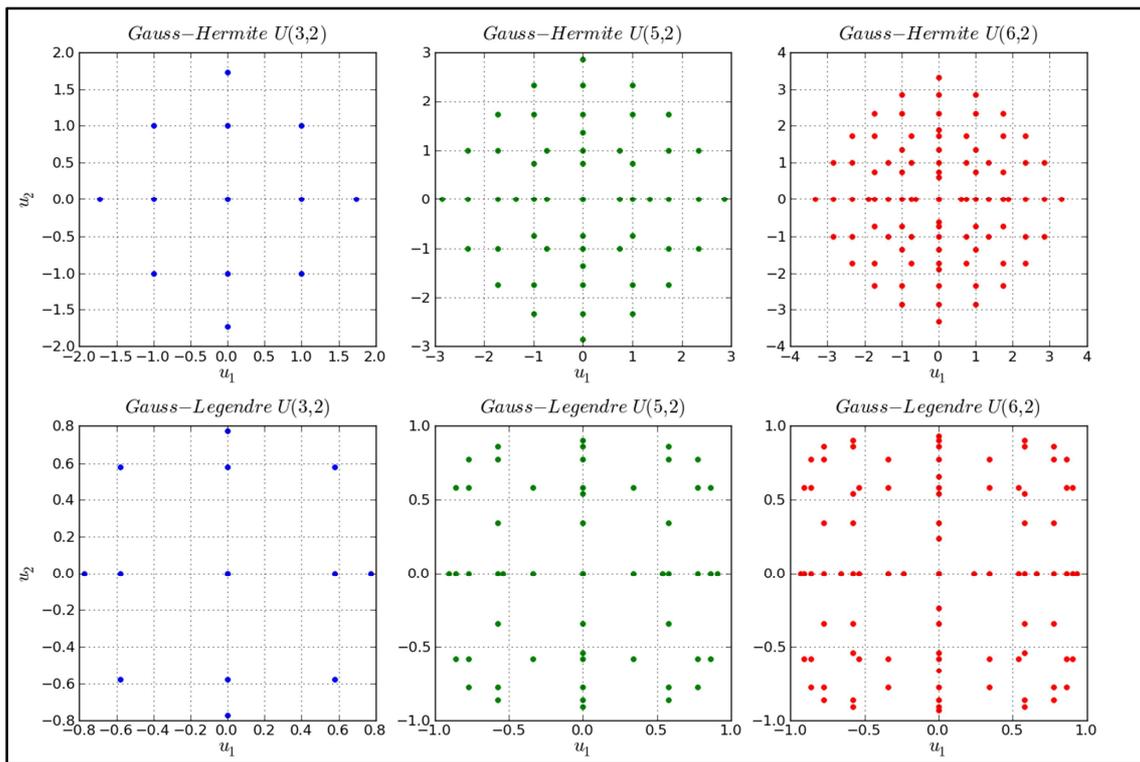
$$Q_{|i|}^N = \sum_{|i|} \sum_{k_1=1}^{i_1} \dots \sum_{k_N=1}^{i_N} \omega_{k_1}^{i_1} \dots \omega_{k_N}^{i_N} f \circ T(u_{k_1}^{i_1}, \dots, u_{k_N}^{i_N}) \Psi_\alpha(u_{k_1}^{i_1}, \dots, u_{k_N}^{i_N}) \quad (II.42)$$

La formule de quadrature de Smolyak  $\mathcal{S}(l, N)$  définie par l'équation (II.41) nécessite seulement l'évaluation du modèle dans la grille de points d'intégration  $\mathcal{U}(l, N)$  donnée par :

$$\mathcal{U}(l, N) = \bigcup_{|i|=l} (\mathbf{u}_1^{i_1} \times \dots \times \mathbf{u}_1^{i_N}) \quad (II.43)$$

où  $\mathbf{u}_1^{i_1} = (u_1^{i_1}, \dots, u_{M_i}^{i_1})$  sont les  $M_i$  points d'intégration associés à une formule d'intégration unidimensionnelle.

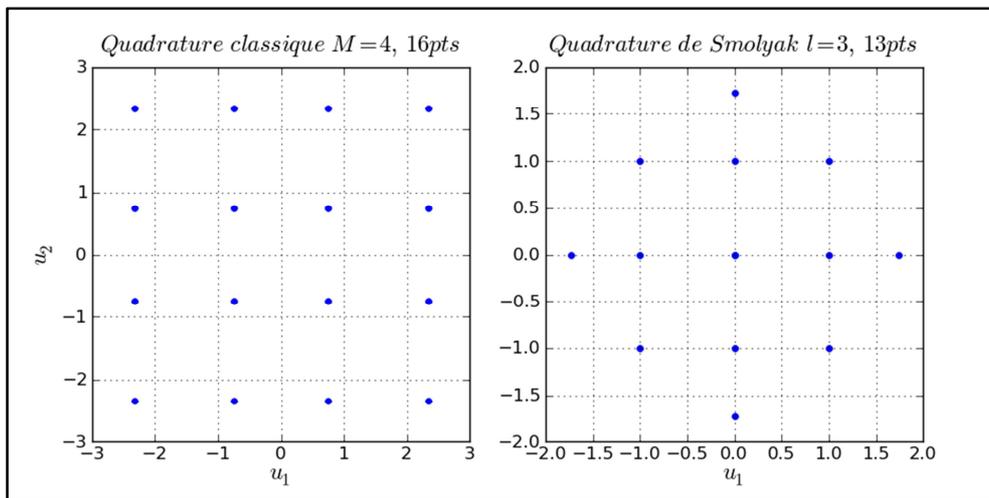
Plusieurs types de points d'intégration peuvent être utilisés, tels que ceux de Gauss-Legendre, de Clenshaw-Curtis et de Gauss-Patterson [101]. Le choix du type de points d'intégration dépend du support des variables aléatoires représentant les paramètres incertains, et de l'espace aléatoire standard (Gaussien ou uniforme) dans lequel nous souhaitons construire le développement en chaos polynômial. La figure II.7 illustre les grilles de calcul associées aux points d'intégration de Gauss-Hermite et Gauss-Legendre, utilisées dans les formules de quadrature de Smolyak d'ordre 3, 5 et 6, dans le cas de deux variables aléatoires.



**Figure II. 7 :** Grilles de calcul associées respectivement aux points d'intégration de Gauss-Hermite et Gauss-Legendre, utilisées dans les formules de quadrature de Smolyak d'ordre 3, 5 et 6, dans le cas de deux variables aléatoires

On constate que la répartition des points dans les grilles de calcul basées sur les points d'intégration de Gauss-Hermite n'est pas la même que celle des points dans la grille basée sur les points d'intégration de Gauss-Legendre. Ceci est dû au fait que les supports (i.e. les domaines de définition) des variables aléatoires associées aux polynômes d'Hermite et aux polynômes de Legendre sont différents comme indiqué dans le tableau II.1.

Dans notre étude, comme le développement en chaos polynômial est construit dans l'espace aléatoire standard, la grille de points d'intégration  $\mathcal{U}(l, N)$  est obtenue en se basant sur les points d'intégration de Gauss-Hermite. La figure II.8 illustre dans le cas bidimensionnel et en se basant sur les points d'intégration de Gauss-Hermite, une comparaison entre la grille de points utilisée par une formule de quadrature de Smolyak d'ordre  $l = 3$  et celle utilisée par une formule de quadrature construite par produit tensoriel complet d'ordre  $M = l + N - 1 = 4$ . On constate que le nombre de points (13) dans la grille utilisée par la formule de quadrature de Smolyak est inférieur à celui (16) de la grille construite à partir d'un produit tensoriel complet. Cependant, l'écart n'est pas significatif (i.e. on gagne seulement trois points dans le cas de la quadrature de Smolyak). En effet, le rapport entre le nombre de points de la grille utilisée par la quadrature de Smolyak et celui de la grille construite par un produit tensoriel complet diminue significativement lorsque la dimension  $N$  de l'intégrale augmente. Il faut noter toutefois que les estimations fournies par une quadrature de Smolyak d'ordre  $l$  sont moins précises que celles fournies par la quadrature construite par un produit tensoriel complet du même ordre  $M = l$ . Cependant, Fichtl et Prinja [102] ont montré que l'utilisation d'une formule de quadrature de Smolyak d'ordre  $l + 1$  ou même d'ordre  $l + 2$  est plus économique en terme d'appels au modèle, que l'emploi d'une formule de quadrature construite par un produit tensoriel complet d'ordre  $M = l$ , particulièrement lorsque la dimension  $N$  de l'intégrale est élevée.



**Figure II. 8 :** Comparaison dans le cas bidimensionnel entre les grilles basées sur les points d'intégration de Gauss-Hermite utilisées respectivement par une formule de quadrature d'ordre  $M = 4$  construite à partir d'un produit tensoriel complet et par une formule de quadrature de Smolyak d'ordre  $l = 3$

Une condition nécessaire dans la construction de la grille  $\mathcal{U}(l, N)$ , est que les points d'intégration  $\mathbf{u}_1^j$ ,  $j = i_1, \dots, i_N$ , associés aux formules de quadrature unidimensionnelles employées dans la quadrature  $N$ -dimensionnelle de Smolyak soient imbriqués (i.e.  $\mathbf{u}_1^j \subset \mathbf{u}_1^{j+1}$  et par conséquent  $\mathcal{U}(l, N) \subset \mathcal{U}(l + 1, N)$ ). Cette condition permet de réduire considérablement le nombre de points d'intégration qui constituent la grille  $\mathcal{U}(l, N)$  et par conséquent le nombre d'appels au modèle. Dans ce contexte, les points d'intégration de Clenshaw-Curtis sont souvent recommandés, comme le montre le tableau II.2 qui donne le nombre de points  $\mathcal{U}(l, N)$  utilisé dans une quadrature de Smolyak d'ordre  $l$ , pour différents types de points d'intégration.

**Tableau II. 2 :** Nombre de points de la grille utilisée dans la quadrature de Smolyak pour les points d'intégration de Clenshaw-Curtis, Gauss-Patterson et Gauss-Legendre

$l$	Clenshaw-Curtis	Gauss-Patterson	Gauss-Legendre
1	5	5	5
2	13	17	21
3	29	49	73
4	65	129	221
5	145	321	609
6	325	769	1573

Il faut noter que les points d'intégration de Gauss-Hermite que nous sommes censés utiliser dans cette étude ne sont pas suffisamment imbriqués, du fait qu'uniquement le point origine se répète dans les grilles construites à partir d'un nombre impair de points. Cependant, on signale tout de même que le fait que les points d'intégration ne soient pas imbriqués n'implique pas automatiquement la perte de précision sur les estimations. En effet, Liu et al. [103] ont montré que les points d'intégration de Gauss-Patterson offrent le meilleur compromis entre précision et coût de calcul, bien qu'ils sont moins imbriqués que les points d'intégration de Clenshaw-Curtis.

Novak et Ritter [104] ont montré qu'une formule de quadrature de Smolyak d'ordre  $l + p$  permet d'obtenir la valeur exacte de l'intégrale d'un polynôme de degré  $2p + 1$ . Ainsi, elle peut être utilisée pour le calcul de l'intégrale  $I_\alpha^N$ , puisque le polynôme représentant l'intégrande  $f \circ T(\mathbf{u}) \Psi_\alpha(\mathbf{u})$  est d'ordre  $2p$ . De plus, pour un ordre  $p$  donné du développement en chaos polynômial, le nombre total d'appels au modèle est donné par :

$$M \sim \frac{2^p}{p!} N^p, \quad N \rightarrow \infty \tag{II.44}$$

On constate, que le coût de calcul croît d'une manière polynômiale avec  $N$ , ce qui est bien inférieur à l'augmentation exponentielle en  $(p + 1)^N$  de la formule de quadrature construite par produit tensoriel complet. Le rapport entre le nombre d'appels au modèle donné par l'équation (II.44) et le nombre de coefficients dans le développement en chaos polynômial tend vers un facteur d'ordre  $2^p$ . Ainsi, puisque ce facteur est indépendant de  $N$ , la formule de quadrature de Smolyak est dite *optimale*. Cependant, on note que le coût de calcul peut être élevé si une grande précision (i.e. un ordre  $p$  élevé) est demandée dans le développement en chaos polynômial.

Comme pour la formule de quadrature construite par produit tensoriel complet, la précision de la formule de quadrature de Smolyak dépend de la régularité de l'intégrande. Plaskota et Wasilkowski [105] ont montré que l'erreur associée à l'estimation de l'intégrale  $\tilde{I}_\alpha^N$  par une formule de quadrature de Smolayk  $\mathcal{S}(l, N)$  est bornée par :

$$\frac{16\pi N}{(3(N - 1)(N - 2)!)^3} \frac{(\log_2 l)^{3(N-1)}}{l^2} (1 + \mathcal{O}(1)), \quad l \rightarrow \infty \tag{II.45}$$

### II.3.4.1.2.3 Quadrature de Stroud

La façon la plus simple pour calculer une intégrale  $N$ -dimensionnelle (i.e.  $N > 1$ ) est d'utiliser la méthode de quadrature de Stroud [16]. Sa première version était destinée au calcul des intégrales définies sur un domaine symétrique, en particulier dans l'hyper-cube  $[-1,1]^N$ . Dans ce contexte, il a proposé deux formules de quadrature, d'ordre 2 et d'ordre 3, désignées souvent par *Stroud-2* et *Stroud-3*. Elles sont basées respectivement sur  $N + 1$  et  $2N$  points d'intégration, et permettent respectivement d'intégrer avec une grande précision des polynômes d'ordre 2 et d'ordre 3. De plus, Stroud [16] a développé une théorie portant sur la

généralisation de cette méthode d'intégration pour différentes fonctions de pondérations. Il a montré aussi que  $N + 1$  est le nombre minimal de points d'intégration exigé par la formule de quadrature d'ordre 2. Sur ce point aussi, Mysovskih [104] a montré que, pour une précision donnée, les formules de quadrature d'ordre 2 et d'ordre 3 proposées par Stroud sont les moins coûteuses en terme de nombre de points d'intégration et sont par conséquent plus économiques en terme de temps de calcul. Xiu [107] a proposé une version améliorée des formules de quadrature de Stroud permettant de calculer des intégrales définies sur un domaine non-symétrique, ainsi que pour différentes fonctions de pondération (i.e. densité de probabilité) autres que celle associée au domaine  $[-1,1]$ . Son travail était motivé par le fait que, dans la pratique, l'aléa associé aux paramètres incertains peut avoir différents types de distributions, telles que les distributions Gaussienne, Beta ou Gamma. Pour ce faire, l'idée de base était d'utiliser une transformation affine construite à partir des polynômes d'ordre 1 associés à la distribution en question. On note que la seule hypothèse est de considérer que les paramètres incertains sont indépendants et que les distributions qui leurs sont associées sont définies positives. Les formules de quadrature de Stroud d'ordre 2 et d'ordre 3 adaptées pour les distributions Gaussiennes sont intéressantes dans notre étude puisque le développement en chaos polynômial est construit en fonction de variables Gaussiennes standard et sur la base des polynômes d'Hermite et ainsi l'intégrale  $N$ -dimensionnelle  $I_{\alpha}^N$  correspond à la mesure Gaussienne. Dans ce contexte, la formule de quadrature de Stroud d'ordre 2, désignée par *Stroud-2*, est basée sur  $N + 1$  points d'intégration  $\{\mathbf{u}^k\}_{k=0}^N$  définis par :

$$\begin{cases} u_{2r-1}^k = \sqrt{2} \cos\left(\frac{2rk\pi}{N+1}\right) \\ u_{2r}^k = \sqrt{2} \sin\left(\frac{2rk\pi}{N+1}\right) \end{cases}, \quad r = 1, 2, \dots, [N/2] \quad (II.46)$$

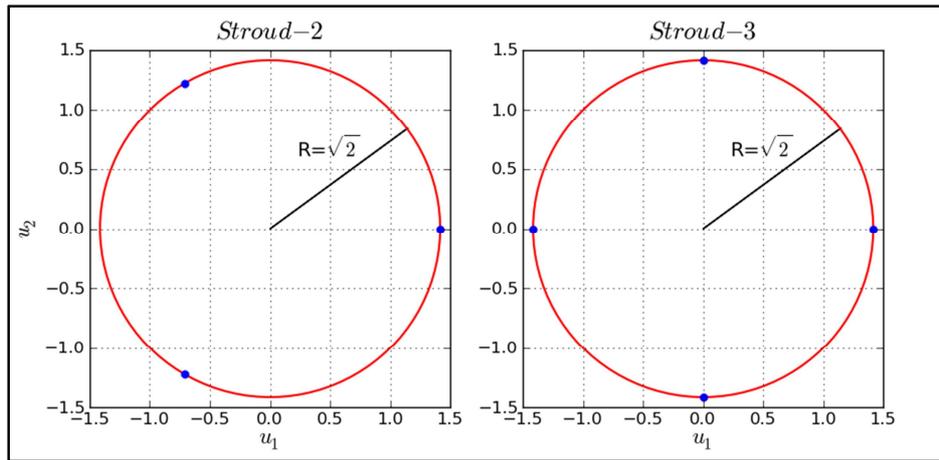
où  $[\cdot]$  désigne la partie entière, et si  $N$  est impair  $u_N^k = (-1)^k$ . Les poids d'intégration sont égaux à  $1/N + 1$ .

La formule de quadrature de Stroud d'ordre 3, désignée par *Stroud-3*, est basée sur  $2N$  points d'intégration  $\{\mathbf{u}^k\}_{k=1}^{2N}$  définis par :

$$\begin{cases} u_{2r-1}^k = \sqrt{2} \cos\left(\frac{(2r-1)k\pi}{N}\right) \\ u_{2r}^k = \sqrt{2} \sin\left(\frac{(2r-1)k\pi}{N}\right) \end{cases}, \quad r = 1, 2, \dots, [N/2] \quad (II.47)$$

et si  $N$  est impair  $u_N^k = (-1)^k$ . Les poids d'intégration sont égaux à  $1/2N$ .

Les points d'intégration utilisés dans les formules de quadrature de Stroud d'ordre 2 et d'ordre 3 sont situés sur la surface d'une hyper-sphère de rayon  $R = \sqrt{N}$ . La figure II.8 illustre, dans le cas bidimensionnel, la répartition des points d'intégration correspondant aux formules de quadrature de Stroud d'ordre 2 et d'ordre 3. On constate que les points d'intégration sont situés sur le cercle de rayon  $R = \sqrt{2}$ . De plus, le nombre de points est très limité par rapport au nombre de points d'intégration utilisés par la méthode de quadrature basée sur un produit tensoriel complet et par la méthode de quadrature de Smolyak (voir figure II.7).



**Figure II. 9 :** Points d'intégration utilisés par les formules de quadrature de Stroud d'ordre 2 et d'ordre 3 dans le cas bidimensionnel

L'expérience a montré que si les estimations n'exigent pas une grande précision, la méthode de quadrature de Stroud est très efficace, particulièrement lorsque la dimension  $N$  de l'intégrale est élevée, puisqu'elle nécessite un nombre minimal de points d'intégration et par conséquent un nombre minimal d'appels au modèle. Cependant, sa précision ne peut pas être améliorée, puisqu'elle est implicitement fixée par le nombre de points d'intégration, qui est à son tour imposé par la dimension  $N$ . De plus, l'erreur sur les estimations ne peut pas être obtenue. Ces aspects ont été démontrés dans les travaux de Xiu et Hesthaven [18] dans le contexte de résolution par la méthode de collocation stochastique d'équations différentielles ayant des paramètres d'entrée aléatoires, et par Liu et al. [103] dans le cadre de problèmes de propagation d'incertitudes.

Dans le contexte de notre travail, les points d'intégration correspondant aux formules de quadrature de Stroud d'ordre 2 et d'ordre 3 seront utilisés dans l'équation (II.24) pour fournir une estimation de l'intégrale  $N$ -dimensionnelle  $I_{\alpha}^N$ . Comme évoqué auparavant, les formules de quadrature de Stroud d'ordre 2 et d'ordre 3 permettent d'intégrer avec précision respectivement des polynômes d'ordre 2 et d'ordre 3. Or, l'intégrande  $f \circ T(\mathbf{u}) \Psi_{\alpha}(\mathbf{u})$  est représentée par un polynôme d'ordre  $2p$ , où  $p$  est l'ordre du développement en chaos polynômial. Ainsi, si l'ordre  $p$  du développement en chaos polynômial est élevé (ce qui garantit une meilleure approximation de la réponse du modèle) les formules de quadrature de Stroud ne garantissent pas une bonne précision dans l'estimation de l'intégrale  $I_{\alpha}^N$ . En effet, il est clair que, si on veut obtenir une bonne estimation de l'intégrale  $I_{\alpha}^N$ , il faut se limiter à un développement en chaos polynômial d'ordre 1, ce qui est insuffisant pour garantir une bonne approximation de la réponse du modèle sauf si celui-ci est linéaire.

### II.3.4.2 Méthode de régression

La méthode de régression a été utilisée par Isukapalli [108] pour la détermination des coefficients d'un développement en chaos polynômial. Contrairement aux méthodes décrites précédemment, elle ne nécessite pas l'évaluation d'une intégrale multidimensionnelle. L'idée de base de la méthode de régression est de déterminer les coefficients du développement en chaos polynômial en résolvant un problème de minimisation au sens des moindres carrés. Soient  $\mathbf{a} = \{a_{\alpha_0}, a_{\alpha_1}, \dots, a_{\alpha_{p-1}}\}^T$  et  $\Psi(\mathbf{U}) = \{\Psi_{\alpha_0}(\mathbf{U}), \Psi_{\alpha_1}(\mathbf{U}), \dots, \Psi_{\alpha_{p-1}}(\mathbf{U})\}^T$  respectivement les coefficients inconnus et la base de développement du chaos polynômial, le problème de régression au sens des moindres carrés est défini par :

$$\hat{\mathbf{a}} = \text{Argmin} \left\{ \mathcal{A}(\mathbf{a}) \equiv \mathbb{E} \left[ (\mathbf{a}^T \boldsymbol{\Psi}(\mathbf{U}) - f(\mathbf{X}))^2 \right] \right\} \quad (II.48)$$

La condition d'optimalité  $\frac{d\mathcal{A}}{d\mathbf{a}}(\hat{\mathbf{a}})$  (au sens de la minimisation) permet d'écrire l'égalité suivante :

$$\mathbb{E}[\boldsymbol{\Psi}(\mathbf{U})\boldsymbol{\Psi}(\mathbf{U})^T] \hat{\mathbf{a}} = \mathbb{E}[\boldsymbol{\Psi}(\mathbf{U})f(\mathbf{X})] \quad (II.49)$$

Dans le sens où le développement en chaos polynômial est construit dans l'espace aléatoire standard (dans lequel les paramètres incertains sont représentés par des variables Gaussiennes standard indépendantes), les composantes du vecteur aléatoire  $\boldsymbol{\Psi}(\mathbf{U})$  sont elles aussi indépendantes. Par conséquent, l'espérance mathématique  $\mathbb{E}[\boldsymbol{\Psi}(\mathbf{U})\boldsymbol{\Psi}(\mathbf{U})^T]$  se réduit à la matrice unité  $\mathbb{I}_{P \times P}$ , puisqu'elle représente par définition la matrice de corrélation associée au vecteur aléatoire  $\boldsymbol{\Psi}(\mathbf{U})$ . L'équation (II.49) peut être alors réécrite sous la forme suivante :

$$\hat{\mathbf{a}} = \mathbb{E}[\boldsymbol{\Psi}(\mathbf{U})f(\mathbf{X})] \quad (II.50)$$

En comparant les équations (II.50) et (II.22), on peut constater une équivalence entre les coefficients du développement en chaos polynômial obtenus par la méthode de régression et ceux fournis par la méthode de projection. Autrement dit, les coefficients du développement en chaos polynômial associés à la méthode de projection correspondent à la minimisation de l'erreur d'approximation au sens des moindres carrés.

Nous nous proposons dans ce qui suit de présenter la procédure permettant de résoudre le problème de minimisation défini par l'équation (II.48). Dans ce contexte, considérons  $M$  réalisations  $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M\}^T$  du vecteur aléatoire  $\mathbf{U}$  représentant les paramètres incertains dans l'espace aléatoire standard, et  $\mathbf{y} = \{y^1, y^2, \dots, y^M\}^T$  un vecteur de  $M$  scalaires représentant les réponses du modèle enregistrées aux différentes réalisations du vecteur  $\mathbf{U}$ , on peut écrire la forme empirique suivante de l'équation (II.49) :

$$[\boldsymbol{\Psi}]^T [\boldsymbol{\Psi}] \hat{\mathbf{a}} = [\boldsymbol{\Psi}]^T \mathbf{y} \quad (II.51)$$

où  $[\boldsymbol{\Psi}]^T [\boldsymbol{\Psi}]$  est dite *matrice d'information*, et  $[\boldsymbol{\Psi}]$  est une matrice de dimension  $(M \times P)$  dont les lignes représentent les évaluations respectives de la base  $\boldsymbol{\Psi}(\mathbf{U}) = \{\Psi_{\alpha_0}(\mathbf{U}), \Psi_{\alpha_1}(\mathbf{U}), \dots, \Psi_{\alpha_{P-1}}(\mathbf{U})\}^T$  du développement en chaos polynômial aux différentes réalisations du vecteur  $\mathbf{U}$  :

$$[\boldsymbol{\Psi}] = \begin{bmatrix} \Psi_{\alpha_0}(\mathbf{u}^1) & \dots & \Psi_{\alpha_{P-1}}(\mathbf{u}^1) \\ \vdots & \ddots & \vdots \\ \Psi_{\alpha_0}(\mathbf{u}^M) & \dots & \Psi_{\alpha_{P-1}}(\mathbf{u}^M) \end{bmatrix} \quad (II.52)$$

Ainsi, les coefficients  $\hat{\mathbf{a}}$  du développement en chaos polynômial associés à la méthode de régression sont obtenus en calculant l'inverse  $([\boldsymbol{\Psi}]^T [\boldsymbol{\Psi}])^{-1}$  de la *matrice d'information* :

$$\hat{\mathbf{a}} = ([\boldsymbol{\Psi}]^T [\boldsymbol{\Psi}])^{-1} [\boldsymbol{\Psi}]^T \mathbf{y} \quad (II.53)$$

An et Owen [109] ont désigné les coefficients du développement en chaos polynômial obtenus par la méthode de projection par des coefficients quasi-régression, du fait que les estimations ne tiennent pas compte de la matrice d'information.

La méthode de régression présentée ci-dessus est similaire à la méthode de surface de réponse [47-51] utilisée dans le contexte d'analyse de fiabilité. Dans ce contexte, les réalisations  $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M\}^T$  du vecteur  $\mathbf{U}$  représentant les paramètres incertains forment le plan d'expérimentation. Une condition nécessaire dans la construction du plan d'expérimentation est que sa taille  $M$  ne doit pas être inférieure au nombre  $P$  des coefficients à déterminer du développement en chaos polynômial. Dans la pratique, il est souvent recommandé de ne pas

calculer directement l'inverse de la matrice d'information  $[\Psi]^T[\Psi]$  comme c'est le cas dans l'équation (II.53). En effet, Berveiller [110] a constaté, dans le cas où le plan d'expérience est construit à partir d'un nombre  $M$  réduit de réalisations (i.e. la taille  $M$  du plan d'expérimentation est proche du nombre  $P$  de coefficients du développement en chaos polynômial), que la matrice d'information  $[\Psi]^T[\Psi]$  peut être *mal conditionnée*, ce qui se traduit par l'apparition d'un pivot de valeur faible lors de la résolution du système défini par l'équation (II.53), et l'obtention par la suite de solutions (i.e. coefficients du développement en chaos polynômial) de valeurs très grandes. Par conséquent, il faut utiliser une technique appropriée telle que la méthode de décomposition par valeurs singulières (i.e. Singular Value Decomposition SVD) [111] pour la résolution d'un tel type de problème. La méthode SVD consiste à fournir une décomposition de la matrice d'information  $[\Psi]^T[\Psi]$  de la forme suivante :

$$[\Psi]^T[\Psi] = [U][W][V]^T \quad (II.54)$$

où  $[U]$  et  $[V]$  sont deux matrices orthonormales (i.e.  $[U][U]^T = [V][V]^T = \mathbb{I}_{P \times P}$ ), et  $[W]$  est une matrice diagonale dont les composantes sont positives ou nulles.

Ainsi, l'inverse  $([\Psi]^T[\Psi])^{-1}$  de la matrice d'information est obtenu par :

$$([\Psi]^T[\Psi])^{-1} = [V][diag(1/W_i)][U]^T \quad (II.55)$$

En se référant à l'équation (II.55), on constate que le seul problème qu'on peut rencontrer par l'application de la méthode SVD est que une ou plusieurs composantes de la diagonale de la matrice  $[W]$  soient nulles. Dans ce contexte, Berveiller a proposé une procédure pour surmonter ce problème que nous ne détaillons pas dans ce travail ; les lecteurs intéressés sont ainsi invités à consulter la référence [110].

Le choix du plan d'expérimentation est une étape importante de la méthode de régression, et en particulier le choix de sa taille  $M$ . En effet, si  $M$  est proche du nombre  $P$  de coefficients du développement en chaos polynômial, nous avons vu qu'il y a une forte probabilité que la matrice d'information soit mal conditionnée. Cependant, si  $M$  est élevé, le nombre d'appels au modèle sera en conséquence élevé. La façon la plus simple pour construire le plan d'expérimentation est d'utiliser les simulations de Monte-Carlo basées sur les techniques d'échantillonnage présentées dans le paragraphe 3.1.1, telles que les générateurs de nombres pseudo-aléatoires, les tirages en hyper-cube latin et les séquences de nombres quasi-aléatoires. Isukapalli [108] a proposé de construire le plan d'expérimentation à partir des  $m$  racines du polynôme d'Hermite de degré  $m = p + 1$  ( $p$  est l'ordre du développement en chaos polynômial). Etant donné qu'on a besoin de calculer  $P$  coefficients du développement en chaos polynômial, le plan d'expérimentation doit être construit au minimum à partir de  $P$  points choisis parmi toutes les combinaisons possibles des  $m$  racines dans chaque direction (i.e. la direction associée à chaque variable aléatoire). Il a suggéré de choisir parmi toutes les combinaisons possibles, les  $2^P$  racines qui ont plus de probabilité d'apparaître, telles que les racines les plus proches de l'origine (si l'origine ne fait pas partie de ces points, il faut l'ajouter au plan d'expérimentation).

Owen [112] a effectué une étude statistique détaillée sur les estimations des coefficients du développement en chaos polynômial donnés par la méthode de régression. A titre illustratif, il a montré que les estimations associées aux simulations de Monte-Carlo classiques (i.e. utilisant les générateurs de nombres pseudo-aléatoires) sont asymptotiquement non biaisées et leur variance asymptotique est donnée par :

$$\mathbb{V}[\hat{\alpha}_\alpha] \sim \frac{1}{M} \mathbb{E}[\Psi(\mathbf{U})^2 \epsilon(\mathbf{U})^2], \quad M \rightarrow \infty \quad (II.56)$$

où  $\epsilon(\mathbf{U}) = f(X) - \mathbf{a}^T \boldsymbol{\Psi}(\mathbf{U})$  est le résidu de l'approximation.

La méthode de régression décrite ci-dessus consiste à construire un développement en chaos polynômial en minimisant l'erreur d'approximation au sens des moindres carrés en un ensemble de points (i.e. plan d'expérimentation) donnés dans l'espace aléatoire. Ainsi, le problème le plus récurrent qui peut se manifester dans la construction du développement en chaos polynômial est le sur-apprentissage ou sur-ajustement (overfitting). En effet, le défi principal ne consiste pas à représenter le modèle seulement en un certain nombre de points d'expérimentation mais dans tout l'espace aléatoire, d'où la nécessité d'introduire un paramètre qui contrôle le problème de sur-ajustement. Dans notre cas d'approximation de la réponse d'un modèle par un développement en chaos polynômial, le seul paramètre permettant de contrôler le problème de sur-ajustement est l'ordre total  $p$  de l'approximation polynômiale. Dans ce contexte, Lewandowski et al. [113] ont proposée trois méthodes pour la détermination de l'ordre  $p$  optimal du développement en chaos polynômial permettant d'éviter le problème de sur-ajustement : la méthode *adjusted-R<sup>2</sup>*, la méthode *early stopping* et la méthode *Wilcoxon rank sum test*. A titre illustratif, la méthode *adjusted-R<sup>2</sup>* consiste à choisir l'ordre  $p$  qui maximise le paramètre de corrélation  $R^2$  défini par :

$$R^2 = 1 - (1 - \rho^2) \frac{M - 1}{M - p - 1} \quad (II.57)$$

où  $\rho^2$  désigne le rapport de corrélation.

Le rapport de corrélation  $\rho^2$  ne doit pas être confondu avec le coefficient de corrélation qui mesure la dépendance entre des variables aléatoires. Il est défini dans le cadre de notre travail (i.e. la construction d'un développement en chaos polynômial) par :

$$\rho^2(Y, \mathbf{a}^T \boldsymbol{\Psi}(\mathbf{U})) = \frac{\mathbb{V}[\mathbb{E}[(Y | \mathbf{a}^T \boldsymbol{\Psi}(\mathbf{U}))]]}{\mathbb{V}[Y]} \quad (II.58)$$

Pour plus d'informations concernant les autres méthodes, le lecteur est invité à consulter la référence [113] dans laquelle les auteurs ont proposé une étude comparative basée sur le modèle analytique de Oakley et O'Hagan [114].

La méthode de régression basée sur la minimisation de l'erreur d'approximation au sens des moindres carrés est facile à implémenter. En effet, du fait de son caractère non intrusif, elle ne nécessite que l'évaluation du modèle aux points d'un plan d'expérimentation et l'obtention des coefficients du développement en chaos polynômial par des techniques usuelles de résolution de systèmes d'équations telles que la méthode SVD. De plus, son avantage majeur est que la précision des estimations peut être améliorée d'une façon séquentielle. Autrement dit, s'il s'est avéré, pour un plan d'expérimentation contenant  $M$  points, que la précision est inférieure à une tolérance préalablement définie, on peut ajouter d'autres points d'expérimentation pour améliorer la précision des estimations, et ce, sans perdre le bénéfice des calculs précédents. Cependant, il faut noter que des précautions sont à prendre dans le choix des points d'expérimentation ainsi que dans leur nombre.

### II.3.4.3 Analyse comparative

Les méthodes de projection consistent à estimer une intégrale par une somme pondérée et se distinguent entre elles par le choix des points et des pondérations d'intégration. Le choix d'une méthode par rapport à une autre dans une application donnée dépend de son aptitude à fournir le meilleur compromis entre la précision et le temps de calcul. Dans notre travail, nous nous sommes intéressés aux méthodes de projection les plus populaires telles que les méthodes de simulations et les méthodes de quadrature. Nous avons analysé leur

comportement (leurs avantages et leurs limites) pour pouvoir choisir par la suite la meilleure méthode, (i.e. celle qui garantit le meilleur compromis entre précision et coût de calcul) à utiliser par la suite dans l'estimation des coefficients d'un développement en chaos polynômial. Nous en avons tiré les conclusions ci-dessous.

La méthode classique des simulations de Monte-Carlo est robuste. En effet, elle converge indépendamment de la dimension  $N$  de l'intégrale et de la complexité de la fonction à intégrer. Cependant, son taux de convergence asymptotique est en  $\mathcal{O}(1/\sqrt{M})$ , ce qui se traduit par une convergence lente. Afin d'y remédier, on peut utiliser une procédure de tirages en hyper-cube latin, ce qui permet de générer des points d'intégration répartis d'une façon plus uniforme dans l'espace aléatoire que les nombres pseudo-aléatoires. Dans le même contexte, on peut employer également les séquences de nombres quasi-aléatoires qui garantissent plus d'uniformité dans la répartition des points d'intégration dans l'espace aléatoire. Cependant, on peut perdre en robustesse si la fonction à intégrer présente des irrégularités. Dans le cas contraire, la vitesse de convergence est meilleure que celle des simulations de Monte-Carlo classiques, avec un taux de convergence en  $\mathcal{O}\left(\frac{\log^N(M)}{M}\right)$ .

Les méthodes de quadrature sont des alternatives aux méthodes de simulations pour le calcul d'intégrales multidimensionnelles. Leur application nécessite la régularité de l'intégrande, ce qui limite leur robustesse. La version la plus simple des méthodes de quadrature (i.e. celle construite à partir d'un produit tensoriel complet de formules de quadratures unidimensionnelles) n'est envisageable que lorsque la dimension  $N$  de l'intégrale à calculer est faible. En effet, pour cette version, le nombre d'appels au modèle augmente d'une façon exponentielle. Afin de surmonter ce problème, on peut utiliser la méthode de quadrature de Smolyak qui est basée sur un produit tensoriel partiel de formules de quadrature unidimensionnelles. Autrement dit, les points d'intégration ne sont pas répartis d'une façon isotrope (comme dans le cas de la méthode de quadrature standard). Ainsi, le nombre d'appels au modèle augmente d'une façon polynômiale et non pas exponentielle. La méthode de quadrature de Stroud est très efficace surtout dans le cas où la dimension  $N$  de l'intégrale est élevée. Cependant, étant donné qu'une bonne précision n'est garantie que pour l'intégration d'un polynôme de degré inférieur ou égal à 3, son application n'est pas envisageable dans notre étude (i.e. calcul d'intégrales  $N$ -dimensionnelles associées aux coefficients du développement en chaos polynômial), puisque le polynôme de degré  $2p$  représentant l'intégrande  $f \circ T(\mathbf{u}) \Psi_{\alpha}(\mathbf{u})$  peut dépasser le degré 3. A l'exception du cas des modèles linéaires, on a souvent besoin d'un développement en chaos polynômial d'ordre supérieur à 1 pour garantir une bonne représentation de la réponse du modèle.

Dans notre travail, le choix de la méthode d'intégration dépend essentiellement de deux paramètres : le nombre de paramètres incertains  $N$  qui reflète la dimension de l'intégrale à évaluer et la complexité (ex. non-linéarité) du modèle qui reflète l'irrégularité de l'intégrande. Lorsque le nombre de paramètres incertains est élevé (i.e.  $> 15$ ) les simulations de Monte-Carlo sont plus adaptées pour calculer l'intégrale  $N$ -dimensionnelle  $I_{\alpha}^N$  associée aux coefficients du développement en chaos polynômial. Dans les situations pour lesquelles le nombre de paramètres incertains est modéré (i.e.  $\sim 15$ ), la méthode de quadrature de Smolyak peut fournir le meilleur compromis entre précision et coût de calcul. Cependant, il faut signaler que lorsqu'on ne dispose pas d'informations préalables sur la régularité de l'intégrande, on doit prendre beaucoup de précautions dans l'utilisation de la quadrature de Smolyak.

Contrairement aux méthodes de projection, la méthode de régression ne fait pas intervenir l'évaluation d'une intégrale multidimensionnelle. Dans ce contexte, les coefficients du développement en chaos polynômial sont déterminés en résolvant un problème de

minimisation au sens des moindres carrés. L'avantage majeur de cette méthode est qu'elle est facile à implémenter. De plus, le nombre d'appels au modèle est indépendant du nombre de paramètres incertains mais est lié à l'ordre du développement en chaos polynômial. Ainsi, si l'ordre du développement en chaos polynômial est choisi élevé pour garantir la précision, la méthode de régression peut être coûteuse en terme d'appels au modèle.

### II.3.5 Post-traitements basés sur le DCP

Une fois le développement en chaos polynômial de la réponse du modèle construit, plusieurs types de post-traitements sont possibles afin d'effectuer des analyses de tendance centrale (i.e. détermination des moments statistiques de la variable aléatoire  $Y$  et construction de sa densité de probabilité), des analyses de fiabilité (i.e. évaluation de la probabilité de défaillance vis-à-vis d'un scénario de défaillance), ou des analyses de sensibilité (i.e. calcul des indices de sensibilité globale associés à une décomposition de la variance). Dans ce qui suit, nous présentons les procédures permettant d'effectuer ces trois types d'analyses à partir d'une représentation en chaos polynômial.

#### II.3.5.1 Analyse de tendance centrale

##### II.3.5.1.1 Calcul des moments statistiques

Les moments statistiques peuvent être déterminés soit en effectuant des simulations de Monte-Carlo sur la formulation explicite obtenue par le développement en chaos polynômial, soit analytiquement en utilisant les coefficients du développement en chaos polynômial. Dans ce contexte la moyenne  $\hat{\mu}_{Y,p}$  et la variance  $\hat{\sigma}_{Y,p}^2$  sont données respectivement par :

$$\hat{\mu}_{Y,p} = \mathbb{E} \left[ \sum_{|\alpha|=0}^p a_\alpha \Psi_\alpha(\mathbf{U}) \right] \equiv \hat{a}_0 \quad (II.59)$$

$$\hat{\sigma}_{Y,p}^2 = \mathbb{V} \left[ \sum_{|\alpha|=0}^p a_\alpha \Psi_\alpha(\mathbf{U}) \right] = \sum_{|\alpha|=1}^p a_\alpha^2 \mathbb{E}[\Psi_\alpha^2] \equiv \sum_{|\alpha|=1}^p \hat{a}_\alpha^2 \quad (II.60)$$

Les estimations  $\hat{\delta}_{Y,p}$  et  $\hat{\kappa}_{Y,p}$  respectivement du coefficient d'asymétrie  $\delta_Y = \frac{1}{\sigma_Y^3} \mathbb{E}[(Y-\mu_Y)^3]$  et du coefficient d'aplatissement  $\kappa_Y = \frac{1}{\sigma_Y^4} \mathbb{E}[(Y-\mu_Y)^4]$  sont données par :

$$\hat{\delta}_{Y,p} \equiv \frac{1}{\hat{\sigma}_{Y,p}^3} \sum_{|\alpha|=1}^p \sum_{|\beta|=1}^p \sum_{|\gamma|=1}^p \hat{a}_\alpha \hat{a}_\beta \hat{a}_\gamma \mathbb{E}[\Psi_\alpha \Psi_\beta \Psi_\gamma] \quad (II.61)$$

$$\hat{\kappa}_{Y,p} \equiv \frac{1}{\hat{\sigma}_{Y,p}^4} \sum_{|\alpha|=1}^p \sum_{|\beta|=1}^p \sum_{|\gamma|=1}^p \sum_{|\delta|=1}^p \hat{a}_\alpha \hat{a}_\beta \hat{a}_\gamma \hat{a}_\delta \mathbb{E}[\Psi_\alpha \Psi_\beta \Psi_\gamma \Psi_\delta] \quad (II.62)$$

Puisque les polynômes d'Hermite unidimensionnels constituent une base orthogonale (équation (II.19)) et ont de plus la propriété :

$$H_i(u)H_j(u) = \sum_{k=|i-j|}^{i+j} C_{ijk}H_k(u) \quad (II.63)$$

avec :

$$\begin{cases} C_{ijk} = 0 & \text{si } \frac{i+j+k}{2} \notin \mathbb{N} \\ C_{ijk} = \frac{i!j!k!}{\left(\frac{i+j-k}{2}\right)! \left(\frac{i+k-j}{2}\right)! \left(\frac{j+k-i}{2}\right)!} & \text{sinon} \end{cases} \quad (II.64)$$

les espérances mathématiques  $\mathbb{E}[\Psi_\alpha\Psi_\beta\Psi_\gamma]$  et  $\mathbb{E}[\Psi_\alpha\Psi_\beta\Psi_\gamma\Psi_\delta]$  peuvent être obtenues analytiquement sans trop de difficultés. On note que la majorité des termes des sommations présentes dans les équations (II.63) et (II.64) sont nulles, ce qui simplifie considérablement le calcul. Cependant, dans le cas où l'ordre  $p$  du développement en chaos polynômial est élevé, ce qui implique un nombre  $P$  élevé de coefficients, l'estimation du coefficient d'asymétrie  $\delta_\gamma$  et du coefficient d'aplatissement  $\kappa_\gamma$  respectivement par les équations (II.63) et (II.64) peut induire un temps de calcul prohibitif. Dans ce contexte, on peut employer les techniques d'intégration par quadrature en substituant le modèle par son développement en chaos polynômial.

### II.3.5.1.2 Construction de la densité de probabilité

La densité de probabilité est un paramètre essentiel dans la caractérisation d'une variable aléatoire. Elle peut être construite à partir des moments statistiques de la variable aléatoire en utilisant la méthode de Pearson [14] ou la méthode de Johnson [15]. Cependant, l'expérience a montré que ces méthodes sont peu précises. Puisqu'on dispose d'une représentation explicite de la réponse du modèle fournie par le développement en chaos polynômial, les simulations de Monte-Carlo ne sont pas coûteuses en terme de temps de calcul et peuvent ainsi servir de base pour la construction de la densité de probabilité de la variable aléatoire  $Y$  représentant la réponse du modèle.

On considère  $M$  réalisations  $\{y^1, y^2, \dots, y^M\}$  de la variable aléatoire  $Y$  obtenues par simulations de Monte-Carlo du développement en chaos polynômial aux points  $\{\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^M\}$ . Ces réalisations sont utilisées par la suite pour la construction d'un histogramme contenant  $m$  classes. Pour cela, on se fixe tout d'abord une borne supérieure  $y_{sup} = \max(\{y^1, y^2, \dots, y^M\})$  et une borne inférieure  $y_{inf} = \min(\{y^1, y^2, \dots, y^M\})$  représentant l'étendue de l'histogramme. Ensuite, l'intervalle  $[y_{inf}, y_{sup}]$  est divisé en  $m$  classes de largeurs constantes égales à  $\Delta = y_{sup} - y_{inf} / m$ . On note que le nombre de classes  $m$  est souvent déterminé par la règle de Sturges [115] définie par :

$$m = 1 + \log_2(M) \quad (II.65)$$

où  $\log_2$  est le logarithme de base 2.

L'effectif  $M_i$  de chaque classe (i.e. le nombre de réalisations  $\{y^1, y^2, \dots, y^M\}$  appartenant à une classe) est calculé par :

$$M_i = \sum_k^M \mathbb{I}_{]y_{inf}+(i-1)\Delta, y_{inf}+i\Delta]}(y^k) \quad (II.66)$$

où  $\mathbb{I}_{]y_{inf}+(i-1)\Delta, y_{inf}+i\Delta]}$  désigne la fonction indicatrice définie sur l'intervalle  $]y_{inf} + (i-1)\Delta, y_{inf} + i\Delta]$ .

En divisant les effectifs  $\{M_i\}_{i=1}^m$  correspondants aux différentes classes par la largeur  $\Delta$  de la classe, nous obtenons  $m$  points de la courbe de la densité de probabilité de la variable aléatoire  $Y$ . Afin d'obtenir une courbe plus lisse, Yu [116] a proposé de construire la densité

de probabilité en moyennant plusieurs histogrammes basés sur le même nombre de classes, mais dont les centres sont différents.

### II.3.5.1.3 Exemple d'illustration

Dans ce paragraphe, nous nous proposons d'étudier la précision et l'efficacité de l'approche probabiliste basée sur le développement en chaos polynômial pour la résolution d'un problème de propagation d'incertitude. Pour ce faire, nous considérons le modèle défini par la fonction polynômiale suivante [117] :

$$y = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + 2x_2x_3 + 4x_1x_3 \quad (III.62)$$

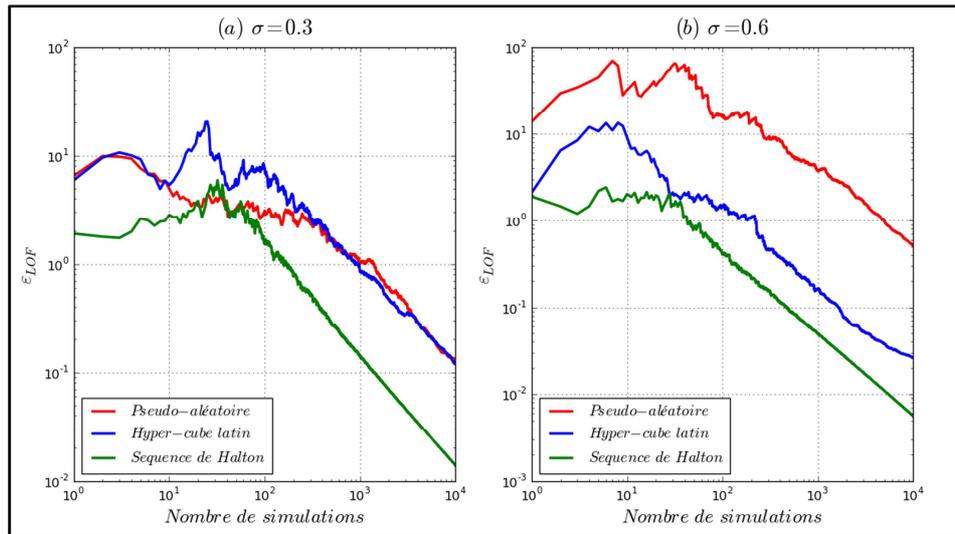
où les paramètres incertains  $\{x_i\}_{i=1}^3$  sont identiquement distribués selon une loi Gaussienne de moyenne unité et d'écart-type  $\sigma$ . Dans un premier temps, les intégrales  $N$ -dimensionnelles représentant les coefficients du développement en chaos polynômial sont calculées par simulations de Monte-Carlo basées sur des échantillons obtenus par un générateur de nombres pseudo-aléatoires, des tirages dans un hyper-cube latin et des séquences quasi-aléatoires. La précision globale du développement en chaos polynômial est évaluée par une mesure normalisée de l'erreur  $\varepsilon_{LOF}$  définie au sens des moindres carrés. Autrement dit, la mesure de l'erreur  $\varepsilon_{LOF}$  permet d'évaluer la capacité du développement en chaos polynômial à reproduire le modèle cible. Elle est définie par :

$$\varepsilon_{LOF} = \frac{\mathbb{E} \left[ \left( y(\mathbf{X}) - y_{M,p}(\mathbf{X}) \right)^2 \right]}{\mathbb{V}[y(\mathbf{X})]} \quad (II.68)$$

où  $y_{M,p}$  est le développement en chaos polynômial d'ordre  $p$  de la réponse  $y$  du modèle dont les coefficients sont obtenus par simulations de Monte-Carlo basées sur un échantillon de  $M$  points. La détermination des différentes grandeurs dans l'équation (II.68) n'est pas triviale, surtout quand le modèle n'est pas disponible sous forme explicite. Afin de surmonter ce problème, la mesure de l'erreur  $\varepsilon_{LOF}$  peut être estimée par simulations de Monte-Carlo par :

$$\hat{\varepsilon}_{LOF} = \frac{1}{M \sigma_{Y_{M,p}}^2} \sum_{i=1}^M \left( y(\mathbf{x}^i) - y_{M,p}(\mathbf{x}^i) \right)^2 \quad (II.69)$$

où  $\sigma_{Y_{M,p}}$  est l'écart-type de la réponse du modèle, calculé en utilisant les coefficients du développement en chaos polynômial. Les figures II.10a et II.10b illustrent, en échelle logarithmique, l'évolution de l'erreur  $\varepsilon_{LOF}$ , pour différents échantillons obtenus par un générateur pseudo-aléatoire, des tirages dans un hyper-cube latin et la séquence quasi-aléatoire de Halton, en fonction du nombre de simulations, respectivement pour  $\sigma = 0,3$  et  $\sigma = 0,6$ . On note que les calculs sont effectués en considérant un développement en chaos polynômial d'ordre 2, afin de reproduire exactement le modèle défini par l'équation (II.67). Ainsi, l'erreur d'approximation est due seulement à l'erreur d'intégration lors de l'estimation des coefficients du développement en chaos polynômial.



**Figure II. 10 :** Comparaison pour différents échantillons de l'évolution de l'erreur  $\varepsilon_{LOF}$  en fonction du nombre de simulations

On constate que les simulations de Monte-Carlo basées sur la séquence de nombres quasi-aléatoires de Halton convergent plus rapidement que celles basées sur les nombres pseudo-aléatoires et les tirages dans un hyper-cube latin. Le taux de convergence des simulations de Monte-Carlo n'est pas affecté par le niveau de dispersion (mesuré par l'écart-type  $\sigma$ ) des paramètres incertains du modèle. On note aussi que les tirages dans un hyper-cube latin améliorent l'efficacité des simulations de Monte-Carlo par comparaison aux nombres pseudo-aléatoires, particulièrement dans le cas où la dispersion est relativement élevée.

**Tableau II. 3 :** Comparaison des estimations des moments statistiques de la réponse du modèle défini par l'équation (II.67)

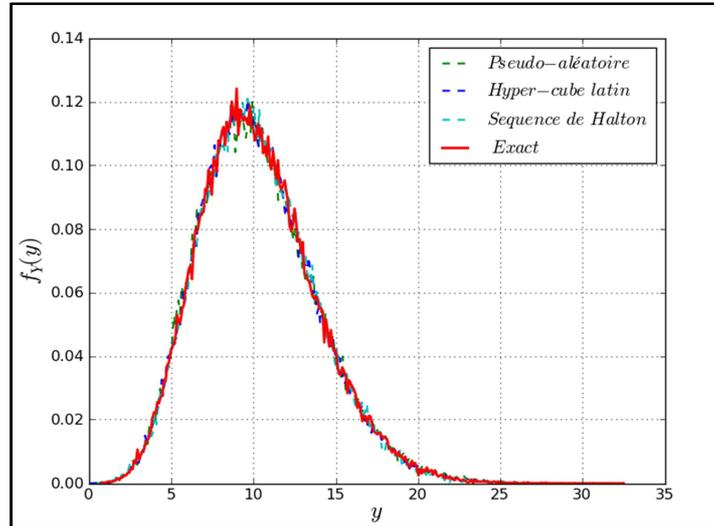
	$\sigma = 0,3$					$\sigma = 0,6$				
	$\hat{\mu}_{Y,p}$	$\hat{\sigma}_{Y,p}$	$\hat{\delta}_{Y,p}$	$\hat{\kappa}_{Y,p}$	$\hat{\varepsilon}_{LOF}(\%)$	$\hat{\mu}_{Y,p}$	$\hat{\sigma}_{Y,p}$	$\hat{\delta}_{Y,p}$	$\hat{\kappa}_{Y,p}$	$\hat{\varepsilon}_{LOF}(\%)$
<b>MC</b>	10,285	3,586	0,470	3,297	12,6	11,129	7,299	0,918	4,144	51,4
<b>HL</b>	10,262	3,561	0,543	3,407	11,9	11,063	7,325	0,947	4,412	2,62
<b>Halton</b>	10,270	3,520	0,491	3,328	1,38	11,08	7,226	1,018	4,222	0,5
<b>Exact</b>	10,27	3,545	0,507	3,363	-	11,08	7,212	1,015	4,346	-

Nous nous intéressons maintenant à la convergence des quatre premiers moments statistiques de la réponse  $y$  en fonction de la dispersion  $\sigma$ . Le tableau II.3 compare les estimations obtenues à partir des coefficients du développement en chaos polynôme dont les coefficients sont calculés par simulations de Monte-Carlo basées respectivement sur les nombres pseudo-aléatoires, les tirages dans un hyper-cube latin et les nombres quasi-aléatoires associés à la séquence de Halton.

On note que la solution de référence est obtenue par un calcul analytique. On constate que toutes les méthodes convergent vers la solution exacte. Cependant, les nombres quasi-aléatoires associés à la séquence de Halton fournissent les meilleures estimations, en terme de précision, des moments statistiques. Cela est justifié de plus par la valeur de l'erreur  $\hat{\varepsilon}_{LOF}$ , qui est de l'ordre de 1,3%. On note aussi que la précision de l'approche probabiliste proposée (i.e. basée sur un couplage entre le développement en chaos polynôme et les simulations de Monte-Carlo) n'est pas affectée par la dispersion des paramètres d'entrée.

La densité de probabilité  $f_Y(y)$  de la réponse du modèle est obtenue par différentes méthodes : d'une part, la solution de référence est obtenue par simulations de Monte-Carlo

basées sur  $10^6$  évaluations du modèle (II.67), et d'autre part, cette densité est construite à partir de  $10^6$  évaluations du développement en chaos polynômial dont les coefficients sont obtenus par différentes méthodes. On constate que les solutions basées sur les tirages dans un hyper-cube latin et les nombres quasi-aléatoires associés à la séquence de Halton sont en bon accord avec la solution de référence.



**Figure II. 11 :** Comparaison de la densité de probabilité  $f_Y(y)$  obtenue par les simulations de Monte-Carlo sur le développement en chaos polynômial de la réponse du modèle

### II.3.5.2 Analyses de fiabilité

En se référant au cas élémentaire présenté dans le paragraphe I.4.1, la fonction d'état limite peut être définie dans le contexte de problèmes d'ingénierie comme la confrontation entre deux quantités fondamentales : la « Résistance » et la « Sollicitation » du système. Elle s'écrit dans l'espace aléatoire standard par :

$$H(\mathbf{U}) = R(\mathbf{U}) - S \quad (II.70)$$

Afin de construire une représentation explicite de la fonction d'état limite, la résistance  $R$  disponible au départ sous forme implicite est substituée par son développement en chaos polynômial. Ainsi, l'équation (II.70) est réécrite sous la forme suivante :

$$H(\mathbf{U}) = \sum_{k=0}^{P-1} a_k \Psi_k(\mathbf{U}) - S \quad (II.71)$$

Cette forme explicite de la fonction d'état limite peut être utilisée par les simulations de Monte-Carlo ou par les méthodes d'approximation FORM et SORM (cf. I.4.2.2) pour fournir une estimation de la probabilité de défaillance. Ainsi, le problème d'analyse de fiabilité dont la difficulté consiste initialement à l'évaluation d'une intégrale  $N$ -dimensionnelle, est réduit à la construction d'un développement en chaos polynômial de la fonction d'état limite dont la difficulté principale est de déterminer les coefficients inconnus. On note aussi que, puisque la densité de probabilité de la variable aléatoire associée à la fonction d'état limite  $G(\mathbf{U})$  peut être construite en suivant la procédure décrite dans le paragraphe précédent, elle peut être utilisée dans l'estimation de la probabilité de défaillance (i.e. par intégration directe de la densité de probabilité). Cependant, il faut noter que cette approche n'est conseillée que pour l'estimation des probabilités de défaillance d'ordre élevé (i.e.  $> 10^{-3}$ ).

### II.3.5.3 Analyses de sensibilité globale

Le calcul des indices de sensibilité de Sobol (cf. I.5.2.2), souvent effectué par les simulations de Monte-Carlo, est coûteux en terme d'appels au modèle mécanique. Pour y remédier, les indices de Sobol peuvent être déterminés analytiquement à partir des coefficients du développement en chaos polynômial de la réponse du modèle. La procédure à suivre est présentée dans ce qui suit [118].

Soit  $\mathcal{J}_{i_1, \dots, i_s}$  l'ensemble des multi-indices  $\alpha$  tel que les indices  $i_1, \dots, i_s$  sont non nuls :

$$\mathcal{J}_{i_1, \dots, i_s} = \left\{ \alpha : \begin{array}{l} \alpha_k > 0 \quad \forall k = 1, \dots, N, k \in (i_1, \dots, i_s) \\ \alpha_k = 0 \quad \forall k = 1, \dots, N, k \notin (i_1, \dots, i_s) \end{array} \right. \quad (II.72)$$

On note que  $\mathcal{J}_i$  correspond aux polynômes du développement en chaos polynômial qui dépendent seulement du paramètre d'entrée  $x_i$ . En se basant sur cette notation, les  $P - 1$  termes correspondant à la base polynômiales  $\{\Psi_j\}_{j=0}^P$  peuvent être rassemblés en se référant aux paramètres dont ils sont dépendants. Ainsi, le développement en chaos polynômial de la variable aléatoire  $Y$  représentant la réponse du modèle peut être réécrit de la façon suivante :

$$\begin{aligned} Y &= a_0 + \sum_{i=1}^N \sum_{\alpha \in \mathcal{J}_i} a_\alpha \Psi_\alpha(u_i) \\ &+ \sum_{i_1=1}^N \sum_{i_2=i_1+1}^N \sum_{\alpha \in \mathcal{J}_{i_1, i_2}} a_\alpha \Psi_\alpha(u_{i_1}, u_{i_2}) \\ &+ \sum_{i_1=1}^N \sum_{i_2=i_1+1}^N \sum_{i_3=i_2+1}^N \sum_{\alpha \in \mathcal{J}_{i_1, i_2, i_3}} a_\alpha \Psi_\alpha(u_{i_1}, u_{i_2}, u_{i_3}) + \dots \\ &+ \sum_{\alpha \in \mathcal{J}_{i_1, \dots, i_N}} a_\alpha \Psi_\alpha(u_{i_1}, u_{i_2}, \dots, u_{i_N}) \end{aligned} \quad (II.73)$$

où le terme  $\sum_{\alpha \in \mathcal{J}_{i_1, \dots, i_N}} a_\alpha \Psi_\alpha(u_{i_1}, u_{i_2}, \dots, u_{i_N})$  est une fonction polynômiale qui dépend seulement des paramètres d'entrée correspondant aux indices  $(i_1, \dots, i_s)$ . Par identification, les termes du développement en chaos polynômial donné par l'équation (II.73) correspondent à ceux de la décomposition de Sobol de la fonction générique  $f(\mathbf{X})$  représentant le modèle. Ainsi, puisque la décomposition de Sobol est unique, on peut conclure que l'équation (II.73) est la décomposition de Sobol de la fonction  $f(\mathbf{X})$ .

Par conséquent, il est facile d'obtenir les indices de sensibilité de Sobol à partir des coefficients du développement en chaos polynômial. En effet, les indices du sensibilité de premier ordre, les indices de sensibilité d'ordre supérieur à 1 et les indices de sensibilité totaux sont respectivement donnés par :

$$\hat{S}_{i;p} = \frac{1}{\hat{\sigma}_{Y,p}^2} \sum_{\alpha \in \mathcal{J}_i} a_\alpha^2 \mathbb{E}[\Psi_\alpha^2] \quad (II.74)$$

$$\hat{S}_{i_1, \dots, i_s;p} = \frac{1}{\hat{\sigma}_{Y,p}^2} \sum_{\alpha \in \mathcal{J}_{i_1, \dots, i_s}} a_\alpha^2 \mathbb{E}[\Psi_\alpha^2] \quad (II.75)$$

$$\hat{S}_{i;p}^T = \frac{1}{\hat{\sigma}_{Y,p}^2} \sum_{\alpha \in \mathcal{J}_i^+} a_\alpha^2 \mathbb{E}[\Psi_\alpha^2] \quad (II.76)$$

où  $\mathcal{J}_i^+$  désigne l'ensemble des indices dont le  $i^{\text{ème}}$  composant est non nul, c'est-à-dire  $\mathcal{J}_i^+ \equiv \{\alpha \in \mathbb{N}^N : 0 \leq \sum_{k=1}^N \alpha_k \leq 1, \alpha_k \neq 0\}$ .

Suite au développement présenté ci-dessus, on peut conclure que, une fois le développement en chaos polynômial de la variable aléatoire  $Y$  construit, les  $2^N - 1$  indices de sensibilité de Sobol sont directement déterminés à partir des coefficients du développement en chaos polynômial sans avoir recours à des appels supplémentaires au modèle.

## II.4 Méthode de décomposition de la dimension

La méthode de décomposition de la dimension (MDD) consiste à représenter une fonction  $N$ -dimensionnelle par une somme hiérarchisée de fonctions de dimensions croissantes. La première version de cette méthode a été proposée par Hoeffding [119]. Dans le contexte des statistiques, elle est connue sous le nom de décomposition de la variance [120]. Elle a été reprise et développée par plusieurs auteurs dans différents contextes scientifiques. En effet, Sobol [51] l'a utilisée dans le contexte d'analyse de sensibilité pour quantifier la contribution de l'aléa relatif à chaque paramètre incertain d'un modèle donné sur la variabilité de sa réponse. Rabitz et Alis [76] ont employé la méthode de décomposition de la dimension dans la construction d'approximations de modèles dépendant d'un grand nombre de paramètres d'entrée, représentant des réactions chimiques. Dans le même contexte, et dans le but d'en améliorer l'efficacité, Li et al. [121] ont utilisé des polynômes orthogonaux et des B-splines cubiques comme alternatives aux simulations de Monte-Carlo pour construire les fonctionnelles présentes dans cette approximation. Plus récemment, la méthode de décomposition de la dimension a été le sujet de plusieurs applications, notamment, pour l'estimation des moments statistiques [122], et l'analyse de fiabilité [123].

### II.4.1 Représentation d'une fonction $N$ -dimensionnelle

En se référant à la méthode de décomposition de la dimension, la réponse  $y$  d'un modèle quelconque dépendant de  $N$  paramètres incertains  $\mathbf{x} = (x_1, \dots, x_N)$ , s'écrit sous la forme suivante :

$$\begin{aligned} y = f(\mathbf{x}) = & f_0 + \sum_{k=1}^N f_k(x_k) + \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^N f_{k_1,k_2}(x_{k_1}, x_{k_2}) \\ & + \dots + \sum_{k_1=1}^{N-s+1} \dots \sum_{k_s=k_{s-1}+1}^N f_{k_1,\dots,k_s}(x_{k_1}, \dots, x_{k_s}) \\ & + \dots + f_{1,\dots,N}(x_1, \dots, x_N) \end{aligned} \quad (II.77)$$

où  $f_0$  est une constante,  $f_k(x_k)$  est une fonction unidimensionnelle représentant la contribution individuelle du paramètre  $x_k$ ,  $f_{k_1,k_2}(x_{k_1}, x_{k_2})$  est une fonction bidimensionnelle représentant la contribution de l'interaction des paramètres  $x_{k_1}$  et  $x_{k_2}$ , et  $f_{k_1,\dots,k_s}(x_{k_1}, \dots, x_{k_s})$  une fonction représentant la contribution de l'interaction des paramètres  $x_{k_1}, x_{k_2}, \dots$ , et  $x_{k_s}$ , et ainsi de suite. Le dernier terme  $f_{1,\dots,N}(x_1, \dots, x_N)$  peut être considéré comme le résidu correspondant à la contribution de l'interaction entre tous les paramètres d'entrée sur la réponse du modèle. Dans la pratique, le développement donné par l'équation (II.77) est

tronqué à un ordre  $s$ . Ainsi, une approximation  $\tilde{y}_s = \tilde{f}_s(\mathbf{x})$   $s$ -dimensionnelle de la réponse  $y = f(\mathbf{x})$  s'écrit :

$$\begin{aligned} \tilde{y}_s = f_0 &+ \sum_{k=1}^N f_k(x_k) + \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^N f_{k_1, k_2}(x_{k_1}, x_{k_2}) \\ &+ \dots + \sum_{k_1=1}^{N-s+1} \dots \sum_{k_s=k_{s-1}+1}^N f_{k_1, \dots, k_s}(x_{k_1}, \dots, x_{k_s}) \end{aligned} \quad (II.78)$$

Il est clair, à partir de l'équation (II.78), qu'une fois l'ordre  $s$  de la décomposition est choisi, il reste à déterminer la constante  $f_0$ , et les différentes composantes fonctionnelles  $f_k(x_k)$ ,  $f_{k_1, k_2}(x_{k_1}, x_{k_2})$  et  $f_{k_1, \dots, k_s}(x_{k_1}, \dots, x_{k_s})$  pour construire une approximation  $s$ -dimensionnelle  $\tilde{y}_s$  de la réponse du modèle.

Dans un premier temps, il faut définir un point de référence  $\mathbf{x}^* = (x_1^*, \dots, x_N^*)$  au voisinage duquel l'approximation est construite. Dans la pratique, le choix du point de référence  $\mathbf{x}^*$  est crucial dans la construction de l'approximation fournie par la méthode de décomposition de la dimension, surtout lorsque cette dernière est limitée aux fonctionnelles d'ordre 1 ou aux fonctionnelles d'ordre 2. Sur ce point, Sobol [124] a montré que définir les coordonnées du point de référence comme les valeurs moyennes respectives des paramètres incertains peut être considéré comme un choix optimal. Ensuite, les différentes fonctionnelles impliquées par l'ordre de la décomposition  $s$  choisi, sont déterminées à partir de l'évaluation du modèle le long des droites (i.e. fonctionnelles dépendant d'un seul paramètre), des surfaces (i.e. fonctionnelles dépendant de deux paramètres) et des volumes (i.e. fonctionnelles dépendant de trois paramètres) qui sont associées au point de référence  $\mathbf{x}^*$ . Ce processus de calcul peut être représenté par les relations suivantes, entre les différentes fonctionnelles du développement (II.78) :

$$f_0 = f(\mathbf{x}^*) \quad (II.79)$$

$$f_k(x_k) = f(x_k, \mathbf{x}_k^*) - f_0 \quad (II.80)$$

$$f_{k_1, k_2}(x_{k_1}, x_{k_2}) = f(x_{k_1}, x_{k_2}, \mathbf{x}_{k_1, k_2}^*) - f(x_{k_1}, \mathbf{x}_{k_1}^*) - f(x_{k_2}, \mathbf{x}_{k_2}^*) - f_0 \quad (II.81)$$

où la notation  $(x_k, \mathbf{x}_k^*) = (x_1^*, \dots, x_{k-1}^*, x_k, x_{k+1}^*, \dots, x_N^*)$  indique que seul le paramètre  $x_k$  est considéré comme variable et que tous les paramètres  $\{x_i\}_{i=1, i \neq k}^N$  sont fixés aux coordonnées du

point de référence ; de la même façon,  $(x_{k_1}, x_{k_2}, \mathbf{x}_{k_1, k_2}^*)$  indique que tous les paramètres sont fixés aux coordonnées du point de référence à l'exception des paramètres  $x_{k_1}$  et  $x_{k_2}$  qui sont considérés comme variables. La même procédure s'applique pour les composantes d'ordre plus élevé. Ainsi, chaque fonctionnelle d'ordre 1  $f_k(x_k)$  est définie le long de la droite engendrée par la variation du paramètre  $x_k$  et chaque fonctionnelle d'ordre 2  $f_{k_1, k_2}(x_{k_1}, x_{k_2})$  est définie dans le plan formé par les deux droites engendrées respectivement par la variation des paramètres  $x_{k_1}$  et  $x_{k_2}$ .

Supposons la fonction  $f(\mathbf{x})$  continue, différentiable, et qu'elle possède un développement en série de Taylor convergent au point de référence  $\mathbf{x}^*$ , défini par :

$$f(\mathbf{x})|_{\mathbf{x}=\mathbf{x}^*} = f(\mathbf{x}^*) + \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{j=1}^N \frac{\partial^i f}{\partial x_j^i}(\mathbf{x}^*) (x_j - x_j^*)^i + \mathcal{R}_2 \quad (II.82)$$

ou

$$f(\mathbf{x})|_{\mathbf{x}=\mathbf{x}^*} = f(\mathbf{x}^*) + \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{j=1}^N \frac{\partial^i f}{\partial x_j}(\mathbf{x}^*) (x_j - x_j^*)^i \\ + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \frac{1}{i_1! i_2!} \sum_{j_1=1}^{N-1} \sum_{j_2=j_1+1}^N \frac{\partial^{i_1+i_2} f}{\partial x_{j_1} \partial x_{j_2}}(\mathbf{x}^*) (x_{j_1} - x_{j_1}^*)^{i_1} (x_{j_2} - x_{j_2}^*)^{i_2} + \mathcal{R}_3 \quad (II.83)$$

où  $\mathcal{R}_2$  désigne le résidu contenant les termes d'ordre supérieur à 1 et  $\mathcal{R}_3$  désigne le résidu contenant les termes d'ordre supérieur à 2. On note que les équations (II.82) et (II.83) représentent respectivement les développements en série de Taylor d'ordre 2 et d'ordre 3 de la fonction  $f(\mathbf{x})$ . En se référant à l'équation (II.78), l'approximation  $\tilde{y}_1$  d'ordre 1 de la réponse du modèle s'écrit :

$$\tilde{y}_1 = f_0 + \sum_{k=1}^N f_k(x_k) \quad (II.84)$$

En substituant les équations (II.79) et (II.80) dans l'équation (II.84) nous obtenons :

$$\tilde{y}_1 = \sum_{k=1}^N f_k(x_1^*, \dots, x_{k-1}^*, x_k, x_{k+1}^*, \dots, x_N^*) - (N-1)f(\mathbf{x}^*) \quad (II.85)$$

où chaque terme dans la sommation est une fonction d'un seul paramètre d'entrée. Ainsi, le développement en série de Taylor qui lui correspond est défini au point de référence  $\mathbf{x}^*$  par :

$$\tilde{y}_1|_{\mathbf{x}=\mathbf{x}^*} = f(\mathbf{x}^*) + \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{j=1}^N \frac{\partial^i f}{\partial x_j}(\mathbf{x}^*) (x_j - x_j^*)^i \quad (II.86)$$

La comparaison des équations (II.82) et (II.86) montre que l'erreur  $\epsilon = y - \tilde{y}_1$ , dans le cas où on utilise une approximation d'ordre 1, n'est autre que le résidu  $\mathcal{R}_2$  défini par les termes d'ordre supérieur ou égal à 2 du développement en série de Taylor. De plus, on note que si la fonction  $f(\mathbf{x})$  est suffisamment régulière et possède un développement en série de Taylor convergent, les termes d'ordre élevé représentant les interactions entre deux paramètres d'entrée ou plus sont la plupart du temps négligeables par rapport au terme d'ordre 1 correspondant à l'effet individuel de chaque paramètre d'entrée. Ainsi, l'approximation d'ordre 1 résultant de la méthode de décomposition de la dimension suffit pour représenter adéquatement la réponse du modèle.

De la même manière, en se basant sur l'équation (II.78), l'approximation  $\tilde{y}_2$  d'ordre 2 fournie par la méthode de décomposition de la dimension, s'écrit :

$$\tilde{y}_2 = f_0 + \sum_{k=1}^N f_k(x_k) + \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^N f_{k_1, k_2}(x_{k_1}, x_{k_2}) \quad (II.87)$$

où le terme  $f_{k_1, k_2}(x_{k_1}, x_{k_2})$  représente l'effet de l'interaction entre les deux paramètres  $x_{k_1}$  et  $x_{k_2}$ . A partir de l'équation (II.87) on constate que l'approximation  $\tilde{y}_2$  d'ordre 2 de la réponse du modèle n'est autre que la somme d'une approximation d'ordre 1 et des termes correspondant aux interactions d'ordre 2 (i.e. faisant intervenir seulement deux paramètres d'entrée). Cet aspect peut contribuer à faciliter l'implémentation de la méthode de décomposition de la dimension. En effet, si par exemple on choisit de construire une approximation d'ordre 1 et on s'est rendu compte que cette dernière n'est pas assez précise pour représenter adéquatement la réponse du modèle en question, on peut construire une

approximation d'ordre 2 simplement en déterminant les termes  $f_{k_1, k_2}(x_{k_1}, x_{k_2})$  et sans reprendre la totalité du calcul.

Le développement en série de Taylor associé à l'approximation  $\tilde{y}_2$  d'ordre 2 de la réponse du modèle, est défini au point de référence  $\mathbf{x}^*$  par :

$$\begin{aligned} \tilde{y}_2|_{\mathbf{x}=\mathbf{x}^*} &= f(\mathbf{x}^*) + \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{j=1}^N \frac{\partial^i f}{\partial x_j}(\mathbf{x}^*) (x_j - x_j^*)^i \\ &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \frac{1}{i_1! i_2!} \sum_{j_1=1}^{N-1} \sum_{j_2=j_1+1}^N \frac{\partial^{i_1+i_2} f}{\partial x_{j_1} \partial x_{j_2}}(\mathbf{x}^*) (x_{j_1} - x_{j_1}^*)^{i_1} (x_{j_2} - x_{j_2}^*)^{i_2} \end{aligned} \quad (II.88)$$

Comme dans le cas d'une approximation d'ordre 1, la comparaison des équations (II.83) et (II.88) montre que l'erreur d'approximation  $\epsilon = y - \tilde{y}_2$  n'est autre que le résidu  $\mathcal{R}_3$  qui englobe les termes représentant les interactions entre trois paramètres d'entrée ou plus. Ainsi, une approximation d'ordre 2 de la réponse du modèle est évidemment plus précise qu'une approximation d'ordre 1 puisqu'elle tient compte des interactions correspondant aux différents couples formés par les paramètres d'entrée, ce qui lui confère une grande flexibilité et par conséquent une forte probabilité de capter le comportement réel du modèle considéré.

La procédure décrite ci-dessus pour construire les approximations d'ordre 1 et d'ordre 2 peut être généralisée pour l'ordre  $s$ . Ainsi, l'approximation  $s$ -dimensionnelle,  $1 \leq s \leq N$ , associée à la méthode de décomposition de la dimension s'écrit :

$$\tilde{y}_s = \sum_{i=0}^s (-1)^i C_{N-s+i-1}^i \sum_{k_1=1}^{N-s+i+1} \dots \sum_{k_{s-i}=k_{s-i-1}+1}^N f_{k_1, \dots, k_{s-i}}(x_{k_1}, x_{k_2}, \dots, x_{k_{s-i}}) \quad (II.89)$$

où la notation  $(x_{k_1}, x_{k_2}, \dots, x_{k_s}) = (x_1^*, \dots, x_{k_1-1}^*, x_{k_1}, x_{k_1+1}^*, \dots, x_{k_{s-i}-1}^*, x_{k_{s-i}}, x_{k_{s-i}+1}^*, \dots, x_N^*)$  signifie que tous les paramètres d'entrée sont fixés aux coordonnées respectives du point de référence à l'exception des paramètres  $x_{k_1}, x_{k_2}, \dots, x_{k_s}$  qui sont considérés comme variables.

Supposons qu'on puisse construire un développement en série de Taylor convergent, pour la fonction  $N$ -dimensionnelle  $f(\mathbf{x})$ , sous la forme  $T(f, R) = \sum_{i=1}^R t_i$ , les termes  $\{t_i\}_{i=1}^N$  du développement en série de Taylor construit au point de référence  $\mathbf{x}^*$  sont donnés par :

$$\begin{aligned} t_0 &= f(\mathbf{x}^*) \\ t_1 &= \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{j=1}^N \frac{\partial^i f}{\partial x_j}(\mathbf{x}^*) (x_j - x_j^*)^i \\ t_2 &= \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \frac{1}{i_1! i_2!} \sum_{j_1=1}^{N-1} \sum_{j_2=j_1+1}^N \frac{\partial^{i_1+i_2} f}{\partial x_{j_1} \partial x_{j_2}}(\mathbf{x}^*) (x_{j_1} - x_{j_1}^*)^{i_1} (x_{j_2} - x_{j_2}^*)^{i_2} \\ &\vdots \\ t_N &= \sum_{i_1, i_2, \dots, i_N} \frac{1}{i_1! i_2! \dots i_N!} \sum_{j_1 < j_2 < \dots < j_N} \frac{\partial^{i_1+\dots+i_N} f}{\partial x_{j_1} \dots \partial x_{j_N}}(\mathbf{x}^*) (x_{j_1} - x_{j_1}^*)^{i_1} \dots (x_{j_N} - x_{j_N}^*)^{i_N} \end{aligned} \quad (II.90)$$

Soit  $f_R \equiv f_{k_1, \dots, k_R}(x_1^*, \dots, x_{k_1-1}^*, x_{k_1}, x_{k_1+1}^*, \dots, x_{k_R-1}^*, x_{k_R}, x_{k_R+1}^*, \dots, x_N^*)$ ,  $0 \leq R \leq s$ , un terme de l'approximation  $s$ -dimensionnelle définie par l'équation (II.89), qui dépend seulement de  $R$  paramètres d'entrée. Le développement en série de Taylor qui lui est associé, est exprimé au point de référence  $\mathbf{x}^*$  par :

$$f_R = \sum_{r=0}^R C_{N-r}^{R-r} t_r \quad (II.91)$$

En examinant les équations (II.90) et (II.91), et en faisant référence au théorème de décomposition des fonctions multidimensionnelles, Xu et Rahman [125] ont montré que l'approximation  $s$ -dimensionnelle  $\tilde{y}_s$ , définie par l'équation (II.89), contient tous les termes d'ordre inférieur ou égal à  $s$  du développement en série de Taylor de la fonction  $N$ -dimensionnelle  $f(\mathbf{x})$ . En effet, par analogie avec le développement effectué auparavant pour les approximations d'ordre 1 et d'ordre 2, la comparaison de la forme explicite de l'équation (II.89) et du développement en série de Taylor de la fonction  $N$ -dimensionnelle  $f(\mathbf{x})$ , montre que l'erreur résiduelle  $\epsilon = y - \tilde{y}_s$  associée à une approximation d'ordre  $s$  n'est autre que le résidu  $\mathcal{R}_{s+1}$  constitué des termes d'ordre supérieur ou égal à  $s + 1$ . Les approximations d'ordre 1 ou d'ordre 2 définies respectivement par les équations (II.85) et (II.87) peuvent être obtenues directement en remplaçant  $s$  dans l'équation (II.89) par 1 ou 2. Ainsi, des approximations de différents ordres peuvent être construites facilement en utilisant l'équation (II.89). On note que cette dernière converge vers la fonction exacte  $f(\mathbf{x})$  lorsque  $s$  tend vers  $N$ .

En fixant l'ordre  $s$  de la décomposition, la réponse d'un modèle dépendant de  $N$  paramètres peut être représentée par une approximation faisant appel, au pire des cas, à des fonctions  $s$ -dimensionnelles. Dans le cas où il existe un ordre de décomposition  $s \ll N$  garantissant une bonne approximation de la réponse du modèle, cela peut contribuer à réduire considérablement les efforts de calcul. En effet, cette représentation exploite le fait que l'effet des interactions d'ordre  $s$  (i.e. faisant intervenir  $s$  paramètres d'entrée) pourrait être négligeable quand  $s$  tend vers  $N$ , et par conséquent les termes dépendant de plus de  $s$  paramètres d'entrée peuvent être omis tout en garantissant une approximation précise de la réponse du modèle. Cette idée a été exploitée dans un premier temps par Xu et Rahman [125] pour calculer les moments statistiques de la réponse d'un système mécanique en substituant l'évaluation des intégrales  $N$ -dimensionnelles souvent complexes, par une série d'intégrales  $s$ -dimensionnelles plus simples à évaluer. Dans un deuxième temps, les mêmes auteurs ont utilisé cette idée en conjonction avec les polynômes d'interpolation de Lagrange ou la projection sur la base de polynômes orthogonaux pour la construction d'une représentation explicite de modèles dont le comportement était initialement défini par une formulation implicite (par exemple une modélisation par éléments finis). Cette surface de réponse stochastique a été utilisée par la suite dans l'analyse de fiabilité [123] ou dans le calcul des indices de sensibilité de Sobol [126]. Les stratégies d'utilisation de la méthode de décomposition dans les contextes d'analyse de tendance centrale (i.e. calcul des moments stochastiques), d'analyse de fiabilité (i.e. estimation de fiabilité vis-à-vis d'un scénario de défaillance donné) et de la décomposition de la variance (i.e. calcul des indices de sensibilité de Sobol) sont présentées en détail dans ce qui suit.

## II.4.2 Intégration par réduction de la dimension

### II.4.2.1 Présentation

Dans les problèmes d'ingénierie, on est souvent amené à manipuler des intégrales multidimensionnelles. Cette situation est fréquente dans le cadre d'une analyse stochastique du comportement d'un système physique pour lequel un certain nombre de paramètres d'entrée sont incertains, visant soit le calcul des moments statistiques de la réponse du système, soit l'évaluation de l'intégrité du système vis-à-vis d'un scénario de défaillance. La difficulté majeure dans l'estimation de ces intégrales multidimensionnelles par les méthodes

d'intégrations numériques classiques telles que les techniques de quadrature est que le nombre d'appels au modèle croît d'une façon exponentielle lorsque la dimension de l'intégrale est élevée, ce qui rend le calcul inabordable dans le cas où la simulation du modèle est exigeante en terme de temps de calcul. Afin de pallier ce problème, Xu et Rahman [125] ont proposé d'utiliser la méthode de décomposition de la dimension présentée auparavant pour réduire la dimension de l'intégrale à évaluer. En effet, l'idée de base consiste, avant de procéder à l'intégration, à remplacer l'intégrande par une approximation  $s$ -dimensionnelle telle que celle définie par l'équation (II.89). Par conséquent, une intégrale  $N$ -dimensionnelle est substituée par une série d'intégrales  $s$ -dimensionnelles ( $s \ll N$ ) dont l'évaluation est abordable au moyen d'une technique de quadrature. Cette nouvelle technique d'intégration baptisée *méthode d'intégration par réduction de la dimension* a été introduite dans un premier temps par Rahman et Xu [122] en utilisant uniquement une décomposition d'ordre 1 de l'intégrande. Ensuite elle a été généralisée par les mêmes auteurs [125] en introduisant une décomposition d'ordre  $s$ .

Dans le but de faciliter la présentation de cette méthode et d'illustrer son implémentation, nous considérons un cas bidimensionnel pour lequel la fonction générique  $f(\mathbf{x})$  représentant le modèle dépend uniquement de deux paramètres réels indépendants  $x_1$  et  $x_2$ . On se propose de calculer l'intégrale  $I$  de la fonction  $f(x_1, x_2)$  dans le domaine symétrique  $\Omega = [-a, a]^2$ , défini par :

$$I[f(x_1, x_2)] = \int_{-a}^a \int_{-a}^a f(x_1, x_2) dx_1 dx_2 \quad (II.92)$$

Supposons que la fonction  $f(x_1, x_2)$  admette un développement en série de Taylor convergent. Il est exprimé au point de référence  $\mathbf{x}^* = (0, 0)$  par :

$$\begin{aligned} f(x_1, x_2)|_{\mathbf{x}=0} &= f(\mathbf{0}) + \frac{\partial f}{\partial x_1}(\mathbf{0})x_1 + \frac{\partial f}{\partial x_2}(\mathbf{0})x_2 \\ &+ \frac{1}{2!} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{0})x_1^2 + \frac{1}{2!} \frac{\partial^2 f}{\partial x_2^2}(\mathbf{0})x_2^2 + \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{0})x_1 x_2 \\ &+ \frac{1}{3!} \frac{\partial^3 f}{\partial x_1^3}(\mathbf{0})x_1^3 + \frac{1}{3!} \frac{\partial^3 f}{\partial x_2^3}(\mathbf{0})x_2^3 + \frac{1}{2!} \frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\mathbf{0})x_1^2 x_2 + \frac{1}{2!} \frac{\partial^3 f}{\partial x_1 \partial x_2^2}(\mathbf{0})x_1 x_2^2 \\ &+ \frac{1}{4!} \frac{\partial^4 f}{\partial x_1^4}(\mathbf{0})x_1^4 + \frac{1}{4!} \frac{\partial^4 f}{\partial x_2^4}(\mathbf{0})x_2^4 + \frac{1}{3!} \frac{\partial^4 f}{\partial x_1^3 \partial x_2}(\mathbf{0})x_1^3 x_2 + \frac{1}{3!} \frac{\partial^4 f}{\partial x_1 \partial x_2^3}(\mathbf{0})x_1 x_2^3 \\ &+ \frac{1}{2! 2!} \frac{\partial^4 f}{\partial x_1^2 \partial x_2^2}(\mathbf{0})x_1^2 x_2^2 + \dots \end{aligned} \quad (II.93)$$

Si on substitue la fonction  $f(x_1, x_2)$  dans l'équation (II.92) par son développement en série de Taylor défini par l'équation (II.93), l'intégrale  $I$  s'écrit :

$$\begin{aligned} I[f(x_1, x_2)] &= I[f(\mathbf{0})] + \frac{1}{2!} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{0})I[x_1^2] + \frac{1}{2!} \frac{\partial^2 f}{\partial x_2^2}(\mathbf{0})I[x_2^2] \\ &+ \frac{1}{4!} \frac{\partial^4 f}{\partial x_1^4}(\mathbf{0})I[x_1^4] + \frac{1}{4!} \frac{\partial^4 f}{\partial x_2^4}(\mathbf{0})I[x_2^4] + \frac{1}{2! 2!} \frac{\partial^4 f}{\partial x_1^2 \partial x_2^2}(\mathbf{0})I[x_1^2 x_2^2] + \dots \end{aligned} \quad (II.94)$$

où les termes  $I[x_1^{l_1} x_2^{l_2}] = \int_{-a}^a x_1^{l_1} dx_1 \times \int_{-a}^a x_2^{l_2} dx_2$  sont nuls lorsque un des exposants  $l_1$  et  $l_2$  est impair, puisque, dans chaque direction, l'intervalle d'intégration est symétrique.

En se référant à la méthode de décomposition de la dimension, l'approximation d'ordre 1 de la fonction  $f(x_1, x_2)$  est donnée par :

$$\tilde{y}_1 = \tilde{f}_1(x_1, x_2) = f(x_1, 0) + f(0, x_2) - f(\mathbf{0}) \quad (II.95)$$

Ainsi, l'intégration de l'équation (II.93) sur le domaine  $\Omega = [-a, a]^2$  fournit une approximation d'ordre 1 de l'intégrale  $I$ , qui s'écrit :

$$I[\tilde{f}_1(x_1, x_2)] = I[f(x_1, 0)] + I[f(0, x_2)] - I[f(\mathbf{0})] \quad (II.96)$$

Nous constatons que l'évaluation de l'intégrale bidimensionnelle définie par l'équation (II.93) est réduite à l'estimation de deux intégrales unidimensionnelles  $I[f(x_1, 0)]$  et  $I[f(0, x_2)]$ , et l'intégrale d'une constante  $I[f(\mathbf{0})]$ . Afin d'évaluer la précision de l'approximation  $I[\tilde{f}_1(x_1, x_2)]$ , nous nous proposons de substituer, dans l'équation (II.96), les fonctions unidimensionnelles  $f(x_1, 0)$  et  $f(0, x_2)$  par leurs développements en série de Taylor respectifs exprimés au point  $x_1 = 0$  et au point  $x_2 = 0$ . Ainsi, nous obtenons :

$$\begin{aligned} I[\tilde{f}_1(x_1, x_2)] = I[f(\mathbf{0})] &+ \frac{1}{2!} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{0}) I[x_1^2] + \frac{1}{2!} \frac{\partial^2 f}{\partial x_2^2}(\mathbf{0}) I[x_2^2] \\ &+ \frac{1}{4!} \frac{\partial^4 f}{\partial x_1^4}(\mathbf{0}) I[x_1^4] + \frac{1}{4!} \frac{\partial^4 f}{\partial x_2^4}(\mathbf{0}) I[x_2^4] + \dots \end{aligned} \quad (II.97)$$

Nous constatons que l'équation (II.97) n'est autre qu'une approximation d'ordre 1 de l'intégrale  $I$ , puisqu'elle est constituée seulement des intégrales unidimensionnelles présentes dans l'équation (II.94). L'erreur d'approximation qui lui est associée est donnée par :

$$I[f(x_1, x_2)] - I[\tilde{f}_1(x_1, x_2)] = \frac{1}{2!2!} \frac{\partial^4 f}{\partial x_1^2 \partial x_2^2}(\mathbf{0}) I[x_1^2 x_2^2] + \dots \quad (II.98)$$

Lorsque l'intégrande est suffisamment régulière, le terme  $\frac{1}{2!2!} \frac{\partial^4 f}{\partial x_1^2 \partial x_2^2}(\mathbf{0}) I[x_1^2 x_2^2]$  ainsi que les termes d'ordre supérieur sont très faibles par rapport au terme d'ordre 1, et ils peuvent être négligés. Par conséquent, l'approximation d'ordre 1 de l'intégrale  $I[f(x_1, x_2)]$  est précise.

Nous nous proposons maintenant de généraliser l'intégration par la technique de réduction de la dimension d'ordre 1 pour le cas où l'intégrande dépend de  $N$  paramètres (i.e.  $N > 2$ ). Dans ce cas, l'intégrale  $I$  définie par l'équation (II.92) dans le cas bidimensionnel, s'écrit :

$$I[f(\mathbf{x})] = \int_{-a}^a \dots \int_{-a}^a f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N \quad (II.99)$$

En remplaçant la fonction  $f(x_1, x_2, \dots, x_N)$  dans l'équation (II.99) par son approximation d'ordre 1 (Eq. (II.85)) en prenant  $\mathbf{x}^* = (0, 0, \dots, 0)$  comme point de référence, le calcul de l'intégrale  $N$ -dimensionnelle  $I[f(\mathbf{x})]$  est remplacé par le calcul d'une série d'intégrales unidimensionnelles. Ainsi, la réduction d'ordre 1 de l'intégrale  $I[f(\mathbf{x})]$  est donnée par :

$$\begin{aligned} I[\tilde{f}_1(\mathbf{x})] = \sum_{k=1}^N \int_{-a}^a f_k(x_1^*, \dots, x_{k-1}^*, x_k, x_{k+1}^*, \dots, x_N^*) dx_k \\ - (N-1) f(\mathbf{x}^*) \int_{-a}^a \dots \int_{-a}^a d\mathbf{x} \end{aligned} \quad (II.100)$$

Comme on l'a montré dans le cas bidimensionnel traité ci-dessus, l'erreur résiduelle d'intégration (Eq. (II.98)) inclut les termes d'intégration par rapport à deux paramètres ou plus. Ainsi, dans le cas où ces termes sont négligeables, l'équation (II.100) peut fournir une

estimation acceptable en terme de précision de l'intégrale  $I[f(\mathbf{x})]$ . De plus, si l'intégrande peut être exprimée comme une somme de fonctionnelles qui dépendent d'un seul paramètre à la fois, l'équation (II.100) fournit la valeur exacte de l'intégrale  $I[f(\mathbf{x})]$ . Afin d'améliorer la précision de l'approximation de l'intégrale  $I[f(\mathbf{x})]$ , on peut augmenter l'ordre de la réduction de l'intégration, autrement dit, remplacer la fonction  $f(x_1, x_2, \dots, x_N)$  par une approximation d'ordre supérieur à 1. A titre illustratif, en prenant  $s = 2$  dans l'équation (II.89), l'approximation d'ordre 2 de la fonction  $f(x_1, x_2, \dots, x_N)$  s'écrit :

$$\begin{aligned} \tilde{y}_2 = \tilde{f}_2(\mathbf{x}) = & \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^N f_{k_1, k_2}(x_1^*, \dots, x_{k_1-1}^*, x_{k_1}, x_{k_1+1}^*, \dots, x_{k_2-1}^*, x_{k_2}, x_{k_2+1}^*, \dots, x_N^*) \\ & - (N-2) \sum_{k=1}^N f_k(x_1^*, \dots, x_{k-1}^*, x_k, x_{k+1}^*, \dots, x_N^*) \\ & + \frac{(N-1)(N-2)}{2} f(\mathbf{x}^*) \end{aligned} \quad (II.101)$$

En intégrant l'équation (II.101) sur le domaine  $\Omega = [-a, a]^N$ , nous obtenons une approximation d'ordre 2 de l'intégrale  $I[f(\mathbf{x})]$ , donnée par :

$$\begin{aligned} I[\tilde{f}_2(\mathbf{x})] = & \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^N \int_{-a}^a \int_{-a}^a f_{k_1, k_2}(x_{k_1}, x_{k_2}) dx_{k_1} dx_{k_2} \\ & - (N-2) \sum_{k=1}^N \int_{-a}^a f_k(x_k) dx_k + \frac{(N-1)(N-2)}{2} f(\mathbf{x}^*) \int_{-a}^a \dots \int_{-a}^a dx \end{aligned} \quad (II.102)$$

où les quantités  $(x_k)$  et  $(x_{k_1}, x_{k_2})$  désignent respectivement  $(x_1^*, \dots, x_{k-1}^*, x_k, x_{k+1}^*, \dots, x_N^*)$  et  $(x_1^*, \dots, x_{k_1-1}^*, x_{k_1}, x_{k_1+1}^*, \dots, x_{k_2-1}^*, x_{k_2}, x_{k_2+1}^*, \dots, x_N^*)$ . Cette approximation est dite réduction d'ordre 2 de la dimension d'intégration puisqu'elle implique au maximum l'évaluation d'intégrales bidimensionnelles. On note que l'équation (II.102) fournit la valeur exacte de l'intégrale  $I[f(\mathbf{x})]$  si l'intégrande  $f(\mathbf{x})$  peut être décomposé en une somme de fonctionnelles qui dépendent au maximum de deux paramètres.

La précision de l'intégration par la technique de réduction de la dimension peut être améliorée en augmentant l'ordre  $s$  de l'approximation de l'intégrande. En suivant la procédure adoptée pour la construction des réductions d'ordre 1 et d'ordre 2, l'intégrale  $I[f(\mathbf{x})]$  peut être approximée par une réduction d'ordre  $s$  en remplaçant l'intégrande  $f(\mathbf{x})$  par son approximation  $\tilde{f}_s(\mathbf{x})$  d'ordre  $s$  définie par l'équation (II.89), ce qui donne :

$$\begin{aligned} I[\tilde{f}_s(\mathbf{x})] = & \sum_{i=0}^s (-1)^i C_{N-s+i-1}^i \sum_{k_1=1}^{N-s+i+1} \dots \sum_{k_{s-i}=k_{s-i-1}+1}^N \\ & \int_{[-a, a]^{s-i}} f_{k_1, \dots, k_{s-i}}(x_{k_1, \dots, k_{s-i}}) \prod_{l=1}^{s-i} dx_{k_l} \end{aligned} \quad (II.103)$$

Cette approximation de l'intégrale  $N$ -dimensionnelle  $I[f(\mathbf{x})]$  consiste à évaluer un ensemble d'intégrales dont la dimension ne dépasse pas l'ordre  $s - i$  qui sont faciles à calculer par une technique de quadrature. Ainsi, elle est plus efficace que l'évaluation d'une seule intégrale  $N$ -dimensionnelle particulièrement lorsque  $s \ll N$ . De plus,  $I[\tilde{f}_s(\mathbf{x})]$  s'approche de la valeur exacte de l'intégrale  $I[f(\mathbf{x})]$  lorsque l'ordre  $s$  de la réduction de la dimension d'intégration

tend vers  $N$ . On note que le fait de présenter la formulation de la méthode de réduction de la dimension d'intégration en considérant un domaine d'intégration symétrique ne doit pas être perçu comme une limitation de la méthode puisqu'on peut se mettre facilement dans cette configuration dans le cas où le domaine d'intégration est non symétrique au moyen d'une transformation linéaire. A titre illustratif, si l'intégrale  $N$ -dimensionnelle à évaluer  $I[f(\mathbf{x})]$  est définie sur le domaine non symétrique  $\mathcal{X} = \prod_{i=1}^N [a_i, b_i]$ , (i.e.  $a_i \in \mathbb{R}$  et  $b_i \in \mathbb{R}$ ) en utilisant le changement de variable  $u_i = \frac{2x_i - (a_i + b_i)}{b_i - a_i}$ , l'intégration peut se faire sur le domaine symétrique  $\mathcal{U} = [-1, 1]^N$  sans affecter la précision des estimations.

En se basant sur les développements déjà présentés, on constate que la méthode d'intégration par réduction de la dimension est très efficace pour le calcul des intégrales  $N$ -dimensionnelles et particulièrement dans le cas où  $N$  est élevé (i.e.  $N > 15$ ). Par conséquent, cette méthode semble permettre de relever certains défis, en particulier, dans l'évaluation des moments statistiques dans le cadre d'analyse de propagation d'incertitude : cet aspect sera l'objet du paragraphe suivant. On note aussi que cette méthode peut être employée pour la détermination des coefficients du développement en chaos polynômial traité au début de ce travail. Nous nous proposons d'illustrer l'efficacité et la précision de la méthode d'intégration par réduction de la dimension sur un cas test analytique.

#### II.4.2.2 Exemple d'illustration

Nous nous proposons d'étudier l'effet de la dimension  $N$  de l'intégrale à évaluer sur la précision de la méthode de réduction de la dimension d'intégration. Pour ce faire, nous considérons l'intégrale  $N$ -dimensionnelle suivante :

$$I_N = \int_{[0,1]^N} \sqrt{\sum_{i=1}^N x_i} dx_1 \dots dx_N \quad (II.104)$$

Cette intégrale a été étudiée par Entacher [127] au moyen d'une approche d'intégration quasi-Monte-Carlo pour laquelle l'erreur d'estimation a été évaluée par l'intermédiaire d'un développement en séries de Haar généralisé. Le tableau II.4 donne la solution exacte de l'intégrale (II.104), et compare l'erreur d'estimation absolue associée à la méthode de réduction de la dimension d'intégration avec celle associée à la méthode quasi-Monte Carlo, pour différentes valeurs de la dimension d'intégration. On note que la règle de quadrature de Gauss-Legendre adoptée pour l'évaluation des intégrales élémentaires impliquées par la méthode de réduction de la dimension d'intégration, est d'ordre  $n = 3$ .

**Tableau II. 4 :** Erreur d'intégration associée à la méthode quasi-Monte Carlo et à la méthode de réduction de la dimension pour différentes valeurs de la dimension  $N$  de l'intégrale

$N$	Quasi-MC	Ordre de la méthode de réduction		
		$s = 1$	$s = 2$	$s = 3$
4	$8,9 \cdot 10^{-6}$	$3,42 \cdot 10^{-3}$	$2,71 \cdot 10^{-3}$	$2,61 \cdot 10^{-3}$
5	$1,9 \cdot 10^{-3}$	$2,84 \cdot 10^{-3}$	$2,33 \cdot 10^{-3}$	$2,27 \cdot 10^{-3}$
6	$4,0 \cdot 10^{-8}$	$2,56 \cdot 10^{-3}$	$2,17 \cdot 10^{-3}$	$2,13 \cdot 10^{-3}$
7	$1,5 \cdot 10^{-5}$	$2,21 \cdot 10^{-3}$	$1,90 \cdot 10^{-3}$	$1,87 \cdot 10^{-3}$
8	$1,7 \cdot 10^{-4}$	$2,00 \cdot 10^{-3}$	$1,74 \cdot 10^{-3}$	$1,72 \cdot 10^{-3}$
9	$3,4 \cdot 10^{-1}$	$1,93 \cdot 10^{-3}$	$1,71 \cdot 10^{-3}$	$1,70 \cdot 10^{-3}$
<i>Nb appels au modèle</i>		32768	13-28	67-352
				175-2620

On constate que l'ordre de l'erreur d'estimation associée à la méthode quasi-Monte Carlo est fortement dépendant de la dimension  $N$  de l'intégrale. En effet, lorsque  $N$  passe de 6 à 9 l'ordre de l'erreur d'estimation passe de  $10^{-8}$  à  $10^{-1}$ . Cependant, quand cette intégrale est calculée par la méthode de réduction de la dimension d'intégration, l'erreur d'estimation est relativement indépendante de la dimension  $N$  de l'intégrale. En effet, pour différentes valeurs de  $N$ , l'ordre de l'erreur reste approximativement constant à  $10^{-3}$ . De plus, en comparant le nombre d'appels requis par les différentes méthodes pour le calcul de l'intégrale (II.104), on constate que la méthode de réduction de la dimension est la plus efficace. A titre d'exemple, lorsque on emploie une réduction d'ordre  $s = 1$ , le nombre d'évaluations de l'intégrande lorsque la dimension de l'intégrale varie entre 4 et 9 passe de 13 à 28, qui est significativement inférieur aux 32768 appels exigés par la méthode quasi-Monte Carlo. Ainsi, la méthode de réduction de la dimension n'est pas seulement efficace, mais mieux encore, sa précision est indépendante de la dimension  $N$  de l'intégrale. On note tout de même, pour cette application, que la précision de la méthode de réduction de la dimension est légèrement améliorée en augmentant l'ordre  $s$  de la réduction. En effet, l'intégrande dans l'équation (II.104) n'est autre qu'une somme un par un des paramètres, ce qui explique que les réductions d'ordres  $s = 2$  et  $s = 3$ , qui permettent de représenter les interactions respectivement entre deux et trois paramètres, n'améliore presque pas la précision puisque ces termes n'existent pas dans le modèle.

## II.4.3 Résolution de problèmes d'analyse de tendance centrale

### II.4.3.1 Evaluation des moments statistiques

En se référant à la méthode d'intégration par réduction de la dimension, ou plus précisément en se référant à l'équation (II.103), le moment statistique  $m_Y^l$  peut être approximé par :

$$m_Y^l \cong \sum_{i=0}^s (-1)^i C_{N-s+i-1}^i \sum_{k_1=1}^{N-s+i+1} \dots \sum_{k_{s-i}=k_{s-i-1}+1}^N \mathbb{E}[\tilde{g}^l(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}})] \quad (II.105)$$

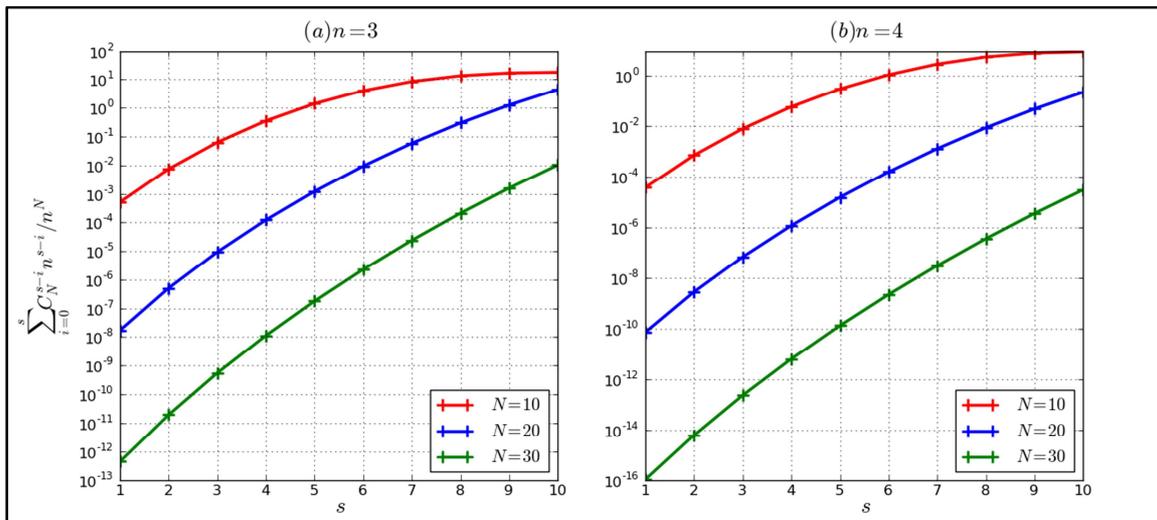
où  $\tilde{g}$  est l'approximation par décomposition de la dimension de la fonction générique  $f$  dans l'espace standard et la notation  $\mathbf{u}_{k_1, k_2, \dots, k_{s-i}}$  désigne le vecteur  $(u_1^*, \dots, u_{k_1-1}^*, u_{k_1}, u_{k_1+1}^*, \dots, u_{k_{s-i}-1}^*, u_{k_{s-i}}, u_{k_{s-i}+1}^*, \dots, u_N^*)$  où seuls les paramètres d'indices  $k_1, k_2, \dots, k_{s-i}$  sont variables et les autres sont fixés aux coordonnées du point de référence  $\mathbf{u}^*$ .

Par définition la quantité  $\mathbb{E}[\tilde{g}^l(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}})]$  s'écrit :

$$\mathbb{E}[\tilde{g}^l(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}})] = \int_{\mathbb{R}^{s-i}} \tilde{g}^l(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}}) \phi(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}}) du_{k_1} \dots du_{k_{s-i}} \quad (II.106)$$

On peut constater que le calcul du moment statistique d'ordre  $l$  est réduit au calcul d'une série d'intégrales dont la dimension ne dépasse pas  $s - i$ , qui peuvent être facilement calculées par une technique de quadrature. Ainsi cette méthode sera plus efficace que le calcul d'une seule intégrale  $N$ -dimensionnelle, en particulier lorsque le nombre  $N$  de paramètres incertains est élevé. On note que, plus l'ordre  $s$  de la réduction de la dimension d'intégration augmente, plus l'erreur d'approximation diminue ; cependant les efforts de calcul deviennent considérables. Ainsi, il faut trouver l'ordre de réduction qui garantit le meilleur compromis entre la précision des estimations et le nombre d'appels au modèle. A titre illustratif, si l'intégrale définie par l'équation (II.106) est calculée par quadrature de Gauss-Hermite basée sur  $n$  points et pondérations d'intégration, le nombre total d'appels au modèle nécessité par la

méthode de réduction de la dimension est  $\sum_{i=0}^s C_N^{s-i} n^{s-i}$ . Par contre,  $n^N$  appels au modèle sont nécessaires si le moment statistique est évaluée directement par quadrature de Gauss-Hermite. Les figures II.12a et II.12b illustrent l'évolution du rapport  $\sum_{i=0}^s C_N^{s-i} n^{s-i} / n^N$  en fonction de l'ordre  $s$  de la réduction de la dimension et pour différents nombres  $N$  de paramètres incertains, respectivement quand l'ordre de la quadrature de Gauss-Hermite employé pour le calcul des grandeurs  $\mathbb{E}[\tilde{g}^l(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}})]$  vaut 3 et 4. La méthode d'intégration par réduction de la dimension est jugée plus efficace quand le rapport  $\sum_{i=0}^s C_N^{s-i} n^{s-i} / n^N$  tend vers 0. On constate à partir des figures II.12a et II.12b que l'efficacité de la méthode d'intégration par réduction de la dimension est fortement dépendante des deux paramètres  $n$  et  $N$ , représentant respectivement le nombre de points d'intégration employé dans le schéma d'intégration par quadrature de Gauss-Hermite et la dimension de l'intégrale.



**Figure II. 12 :** Evolutions du rapport du nombre d'appels au modèle demandé par la méthode d'intégration par réduction de la dimension et par la méthode d'intégration directe en fonction des paramètres  $s$  et  $N$  : (a)  $n = 3$  et (b)  $n = 4$

A titre d'exemple, quand la dimension de l'intégrale à évaluer est  $N = 10$ , les méthodes de réduction d'ordres  $s = 1$ ,  $s = 2$  et  $s = 3$  permettent de réduire le nombre d'appels au modèle par rapport à la méthode d'intégration directe respectivement par un facteur de 1969, 136, 17 lorsque le nombre de points d'intégrations  $n = 3$ , et respectivement par un facteur de 26215, 1380, 125 lorsque le nombre de points d'intégration  $n = 4$ . Ainsi, on constate que les efforts de calcul en terme d'appels au modèle sont considérablement réduits lorsque le nombre de points d'intégration  $n$  augmente, et l'efficacité de la méthode de réduction est plus remarquable lorsque la dimension  $N$  de l'intégrale augmente. Il faut signaler tout de même que cette efficacité reste relative puisqu'on compare à la méthode d'intégration directe. On note que grâce à la formulation de la méthode de réduction de la dimension dans l'espace standard, les intégrales élémentaires  $\mathbb{E}[\tilde{g}^l(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}})]$  sont toujours évaluées par quadrature de Gauss-Hermite quelle que soit la distribution  $F_X$  des paramètres d'entrée  $\mathbf{X}$ . On signale aussi que le fait de supposer la réponse  $y$  du modèle scalaire était une simplification pour la présentation de la méthode et ne doit pas être vu comme une limite. En effet, si on s'intéresse à une réponse vectorielle du modèle  $\mathbf{y} = \{y_1, y_2, \dots, y_M\}^T$ , l'efficacité de la méthode d'intégration par réduction de la dimension ne sera pas affectée puisque le nombre d'appels au modèle est le même que lorsque la réponse du modèle est scalaire. Les deux premiers moments statistiques  $\mu_Y$  et  $C_Y$ , respectivement la moyenne et la matrice de covariance de la

variable aléatoire  $M$ -dimensionnelle  $\mathbf{Y}$  associée à l'aléa de la réponse  $\mathbf{y}$  sont donnés respectivement par :

$$\mu_Y = \mathbb{E}[\mathbf{Y}] \cong \sum_{i=0}^s (-1)^i C_{N-s+i-1}^i \sum_{k_1=1}^{N-s+i+1} \dots \sum_{k_{s-i}=k_{s-i-1}+1}^N \mathbb{E}[\tilde{\mathbf{g}}(u_{k_1}, u_{k_2}, \dots, u_{k_{s-i}})] \quad (II.107)$$

$$\begin{aligned} C_Y &= \mathbb{E}[\mathbf{Y}\mathbf{Y}^T] - \mu_Y \mu_Y^T \\ &\cong \sum_{i=0}^s (-1)^i C_{N-s+i-1}^i \sum_{k_1=1}^{N-s+i+1} \dots \sum_{k_{s-i}=k_{s-i-1}+1}^N \mathbb{E}[\tilde{\mathbf{g}}(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}}) \tilde{\mathbf{g}}^T(\mathbf{u}_{k_1, k_2, \dots, k_{s-i}})] \\ &\quad - \mu_Y \mu_Y^T \end{aligned} \quad (II.108)$$

Les espérances mathématiques figurant dans les équations (II.107) et (II.108) sont calculées par quadrature de Gauss-Hermite puisqu'elles correspondent à des intégrales dont la dimension ne dépasse pas  $s - i$ .

### II.4.3.2 Exemple d'illustration

Dans ce paragraphe, nous nous proposons de mettre en évidence les différentes constatations exposées précédemment. Dans cette perspective, nous reprenons le modèle polynômial défini par l'équation (II.67), étudié dans le paragraphe II.3.5.1.3. Nous nous intéressons à la convergence des quatre premiers moments statistiques de la réponse  $\mathbf{y}$  en fonction de la dispersion  $\sigma$  des paramètres d'entrée et de l'ordre  $s$  de la méthode d'intégration par réduction de la dimension. Le tableau II.5 donne les résultats obtenus pour différents ordres de la méthode de réduction de la dimension et pour  $\sigma = 0,4$ , sachant que les intégrales élémentaires sont calculées par quadrature de Gauss-Hermite d'ordre  $n = 3$ .

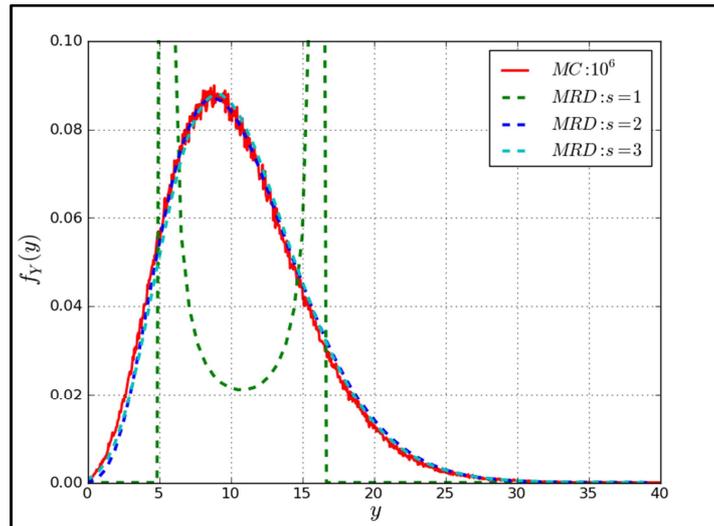
**Tableau II. 5 :** Moments statistiques de la réponse du modèle pour  $\sigma = 0,4$

<i>Moments statistiques</i>	<i>Ordre de la méthode de réduction</i>			<i>Exacte</i>
	$s = 1$	$s = 2$	$s = 3$	
$\mu_Y$	10,48	10,48	10,48	10,48
$\sigma_Y$	4,698	4,771	4,771	4,777
$\delta_Y$	0,0007	0,682	0,686	0,687
$\kappa_Y$	1,163	3,416	3,635	3,634

On constate que la méthode de réduction d'ordre 1 permet d'estimer les deux premiers moments statistiques avec une assez bonne précision. Cependant, les moments statistiques d'ordre plus élevé tels que le coefficient d'asymétrie  $\delta_Y$  et le coefficient d'aplatissement  $\kappa_Y$  exigent l'augmentation de l'ordre de réduction. En effet, pour  $s = 3$ , les estimations de tous les moments statistiques sont en bon accord avec la solution exacte obtenue par un calcul analytique.

Si nous nous intéressons à la convergence de la densité de probabilité  $f_Y(\mathbf{y})$  de la variable aléatoire scalaire  $Y$ , la figure II.13 permet de comparer les évolutions de la densité de probabilité  $f_Y(\mathbf{y})$  construites par la méthode de Pearson [14] à partir des moments statistiques obtenus par différents ordres de réduction et celle obtenue par  $10^6$  simulations de Monte-Carlo sur l'équation (II.67). On constate que la méthode des moments converge vers la solution de référence, obtenue par les simulations de Monte-Carlo, lorsque les moments statistiques correspondent aux estimations données par des réductions d'ordre 2 et d'ordre 3. Cependant, lorsque les moments statistiques sont issus d'une estimation par une réduction d'ordre 1, la densité de probabilité est peu précise, car seuls la moyenne et l'écart-type de la réponse sont

précis. En effet, la précision de la méthode des moments est fortement liée à la précision des moments statistiques.



**Figure II. 13 :** Comparaison de la densité de probabilité  $f_Y(y)$  de la réponse du modèle obtenue par la méthode des moments et les simulations de Monte-Carlo

#### II.4.4 Méta-modèle basé sur les polynômes de Lagrange

Le but principal de l'utilisation de la méthode de décomposition de la dimension est de construire une approximation de la réponse d'un modèle. Cependant, dans les développements présentés jusqu'ici, cette approximation a été utilisée implicitement (i.e. sans construire vraiment la surface de réponse) dans le calcul des intégrales  $N$ -dimensionnelles et en particulier le calcul des moments statistiques de la réponse d'un système stochastique. En effet, nous avons bénéficié seulement de l'architecture de cette approximation (i.e. somme hiérarchisée de fonctionnelles de dimensions croissantes) pour réduire la dimension des intégrales  $N$ -dimensionnelles, sans déterminer explicitement les différents termes fonctionnels  $f_k$ ,  $f_{k_1, k_2}$  et  $f_{k_1, \dots, k_s}$  qui la composent. Dans ce qui suit, nous continuons à raisonner dans l'espace standard dans lequel l'aléa associé aux paramètres incertains  $\mathbf{X}$  est représenté par une variable aléatoire Gaussienne standard  $N$ -dimensionnelle  $\mathbf{U}$ ,  $\mathbf{u}^*$  désigne le point de référence et  $g_k = f_k \circ T$ ,  $g_{k_1, k_2} = f_{k_1, k_2} \circ T$  et  $g_{k_1, \dots, k_s} = f_{k_1, \dots, k_s} \circ T$  définissent respectivement les termes fonctionnels unidimensionnel, bidimensionnel et  $s$ -dimensionnel. Ainsi, il est clair que la construction d'une approximation  $\tilde{g}$  de la réponse d'un modèle ne nécessite que la détermination des fonctionnelles  $g_k$ ,  $g_{k_1, k_2}$  et  $g_{k_1, \dots, k_s}$ . Pour ce faire, il est possible d'utiliser une technique d'interpolation basée sur les polynômes de Lagrange.

Nous nous intéressons à la construction des fonctionnelles unidimensionnelles  $g_k(u_k) = g_k(u_1^*, \dots, u_{k-1}^*, u_k, u_{k+1}^*, \dots, u_N^*)$ . Supposons qu'il existe un ensemble de points d'expérimentation  $\{u_k^j\}_{j=1}^n$  dans la direction  $u_k$  pour lesquels on connaît les réponses respectives du modèle  $\{g(u_k^j) = g(u_1^*, \dots, u_{k-1}^*, u_k^j, u_{k+1}^*, \dots, u_N^*)\}_{j=1}^n$ . En se référant à la technique d'interpolation par polynômes de Lagrange, la fonctionnelle  $g_k(u_k)$  associée au paramètre  $u_k$  s'écrit :

$$g_k(u_k) = \sum_{j=1}^n g(u_k^j) L_j(u_k) \quad (II.109)$$

où la fonction  $L_j(u_k)$  n'est autre que le polynôme de Lagrange construit autour de la coordonnée  $u_k^j$ , défini par :

$$L_j(u_k) = \prod_{l=1, l \neq j}^n \frac{u_k - u_k^l}{u_k^j - u_k^l} \quad (II.110)$$

En utilisant les équation (II.109) et (II.110), les différentes fonctionnelles unidimensionnelles  $g_k(u_k)$ ,  $k = 1, \dots, N$  peuvent être construites en effectuant à chaque fois  $n$  appels au modèle. Ainsi, le coût de calcul nécessaire pour la construction d'une approximation d'ordre 1  $\tilde{y}_1 = \tilde{g}_1(\mathbf{u})$  est évalué à  $nN + 1$  appels. Cette approximation est définie dans l'espace aléatoire standard par :

$$\tilde{g}_1(\mathbf{u}) = \sum_{k=1}^N \sum_{j=1}^n g(u_k^j) L_j(u_k) - (N-1)g_0 \quad (II.111)$$

où  $g_0 = g(\mathbf{u}^*)$  est la réponse du modèle au point de référence  $\mathbf{u}^*$ .

Si nous nous intéressons maintenant à la construction des fonctionnelles bidimensionnelles  $g_{k_1, k_2}(u_{k_1}, u_{k_2}) = g_{k_1, k_2}(u_1^*, \dots, u_{k_1-1}^*, u_{k_1}, u_{k_1+1}^*, \dots, u_{k_2-1}^*, u_{k_2}, u_{k_2+1}^*, \dots, u_N^*)$ , le même processus est utilisé dans les directions  $u_{k_1}$  et  $u_{k_2}$ . Soient  $\{u_{k_1}^{j_1}\}_{j_1=1}^n$  et  $\{u_{k_2}^{j_2}\}_{j_2=1}^n$  les points d'expérimentation respectivement dans les directions  $u_{k_1}$  et  $u_{k_2}$ , et  $g(u_{k_1}^{j_1}, u_{k_2}^{j_2}) = g(u_1^*, \dots, u_{k_1-1}^*, u_{k_1}^{j_1}, u_{k_1+1}^*, \dots, u_{k_2-1}^*, u_{k_2}^{j_2}, u_{k_2+1}^*, \dots, u_N^*)$  les réponses du modèle qui leurs sont associées, la fonctionnelle bidimensionnelle  $g_{k_1, k_2}(u_{k_1}, u_{k_2})$  est alors donnée par :

$$g_{k_1, k_2}(u_{k_1}, u_{k_2}) = \sum_{j_1=1}^n \sum_{j_2=1}^n g(u_{k_1}^{j_1}, u_{k_2}^{j_2}) L_{j_1}(u_{k_1}) L_{j_2}(u_{k_2}) \quad (II.112)$$

où les fonctions  $L_{j_1}(u_{k_1})$  et  $L_{j_2}(u_{k_2})$  ne sont autres que les polynômes de Lagrange construits respectivement autour des coordonnées  $u_{k_1}^{j_1}$  et  $u_{k_2}^{j_2}$ . En utilisant les équations (II.110) et (II.112), toutes les fonctionnelles bidimensionnelles peuvent être construites en nécessitant à chaque fois (i.e. chaque combinaison de deux directions)  $n^2$  appels au modèle. Ainsi, le coût de calcul nécessaire pour la construction d'une approximation d'ordre 2  $\tilde{y}_2 = \tilde{g}_2(\mathbf{u})$  est évalué à  $N(N-1)n^2/2 + nN + 1$  appels. Cette approximation est définie par :

$$\begin{aligned} \tilde{g}_2(\mathbf{u}) = & \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^N \sum_{j_1=1}^n \sum_{j_2=1}^n g(u_{k_1}^{j_1}, u_{k_2}^{j_2}) L_{j_1}(u_{k_1}) L_{j_2}(u_{k_2}) \\ & - (N-2) \sum_{k=1}^N \sum_{j=1}^n g(u_k^j) L_j(u_k) + \frac{(N-1)(N-2)}{2} g_0 \end{aligned} \quad (II.113)$$

En suivant le même procédé, l'approximation d'ordre  $s$   $\tilde{y}_s = \tilde{g}_s(\mathbf{u})$  s'écrit :

$$\begin{aligned} \tilde{g}_s(\mathbf{u}) = & \sum_{i=0}^s (-1)^i C_{N-s+i-1}^i \\ & \times \sum_{k_1=1}^{N-s+i+1} \dots \sum_{k_{s-i}=k_{s-i-1}+1}^N \sum_{j_1=1}^n \dots \sum_{j_{s-i}=1}^n g(u_{k_1}^{j_1}, \dots, u_{k_{s-i}}^{j_{s-i}}) L_{j_1}(u_{k_1}) \dots L_{j_{s-i}}(u_{k_{s-i}}) \end{aligned} \quad (II.114)$$

où  $u_{k_{s-i}}^{j_{s-i}}$  est le  $j_{s-i}$ ème point d'expérimentation associé à la  $k_{s-i}$ ème direction, la fonction  $L_{j_{s-i}}(u_{k_{s-i}})$  n'est autre que le polynôme de Lagrange construit autour de la coordonnée  $u_{k_{s-i}}^{j_{s-i}}$ , et  $g(u_{k_1}^{j_1}, \dots, u_{k_{s-i}}^{j_{s-i}})$  est un coefficient réel obtenu par évaluation du modèle au point  $(u_{k_1}^{j_1}, \dots, u_{k_{s-i}}^{j_{s-i}}) = (u_1^*, \dots, u_{k_1-1}^*, u_{k_1}^{j_1}, u_{k_1+1}^*, \dots, u_{k_{s-i}-1}^*, u_{k_{s-i}}^{j_{s-i}}, u_{k_{s-i}+1}^*, \dots, u_N^*)$ .

Dans les problèmes d'ingénierie dans lesquels le comportement d'une structure ou d'un système est souvent décrit d'une façon implicite (par exemple par modélisation par éléments finis), l'approximation définie par l'équation (II.114) peut constituer une représentation explicite du modèle, qui peut être utilisée en conjonction avec les simulations de Monte-Carlo dans un contexte d'analyses stochastiques tel que le calcul des moments statistiques ou l'estimation de la fiabilité vis-à-vis d'un scénario de défaillance. Ainsi, le coût de calcul en terme d'appels au modèle implicite provient principalement de la construction de l'approximation  $\tilde{g}_s(\mathbf{u})$  correspondant à la méthode de décomposition de la dimension. Cette approximation est souvent désignée par surface de réponse stochastique. Une étape importante dans la construction de cette surface de réponse stochastique, est le choix des points d'expérimentation. Ils peuvent être choisis comme les points d'intégration de Gauss-Hermite ou uniformément distribués dans l'espace aléatoire standard. Dans ce dernier cas, les points d'expérimentation dans une direction donnée  $u_k$  sont donnés par :

$$\left\{ u_k^* - \frac{(n-1)}{2\delta_k}, u_k^* - \frac{(n-3)}{2\delta_k}, \dots, u_k^*, \dots, u_k^* + \frac{(n-3)}{2\delta_k}, u_k^* + \frac{(n-1)}{2\delta_k} \right\} \quad (II.115)$$

où  $n = 3, 5, 7, 9$  est le nombre de points d'expérimentation,  $u_k^*$  la coordonnée du point de référence associée à la direction  $u_k$  et  $\delta_k$  est une perturbation qui commande l'écart entre les points d'expérimentation. Les figures II.14a, II.14b et II.14c illustrent les plans d'expérimentation utilisés respectivement pour une approximation d'ordre 1 de fonctions unidimensionnelle et bidimensionnelle, et pour une approximation d'ordre 2 d'une fonction bidimensionnelle. On note que d'autre type de plans d'expérimentation peuvent être utilisés. Cependant, plus les points sont uniformément répartis plus l'approximation est précise.

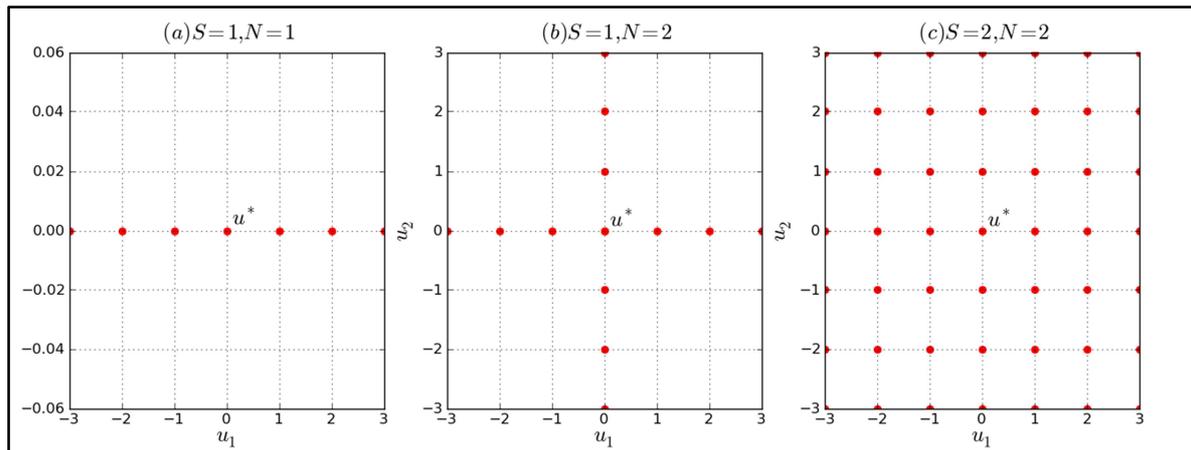
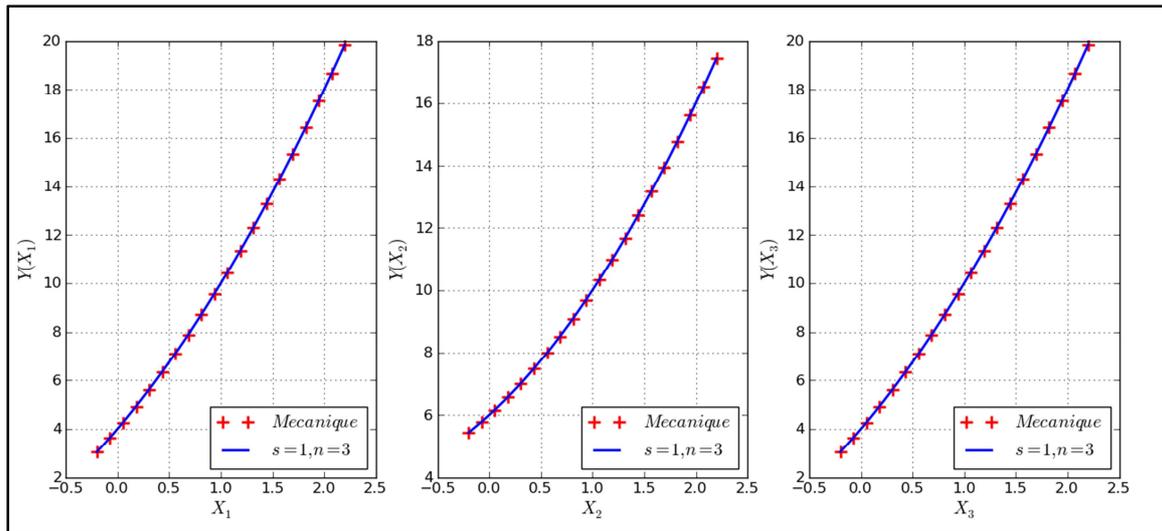


Figure II. 14 : Plan d'expérimentation : (a)  $s = 1, N = 1$ , (b)  $s = 1, N = 2$ , (c)  $s = 2, N = 2$

Nous nous proposons de valider la procédure de construction de la surface de réponse stochastique décrite ci-dessus, ainsi que son éventuelle utilisation en conjonction avec les simulations de Monte-Carlo pour l'estimation des caractéristiques statistiques de la réponse d'un modèle.



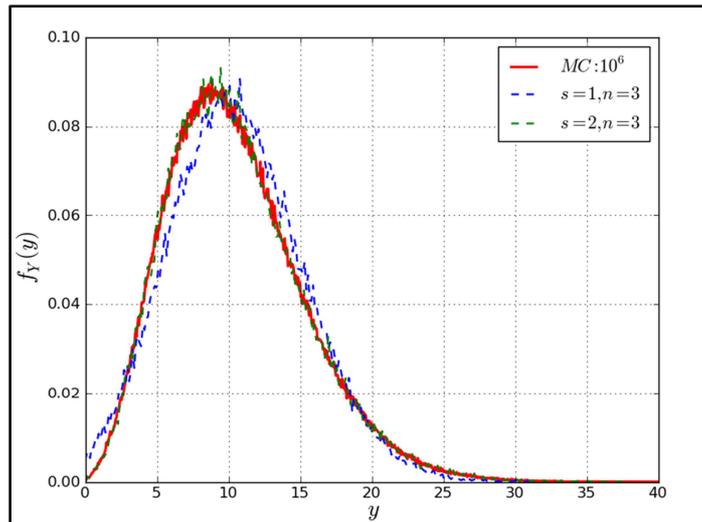
**Figure II. 15 :** Evolution de l'approximation de la réponse du modèle par décomposition de la dimension en fonction des différents paramètres incertains

Nous considérons une fois de plus le modèle polynômial représenté par l'équation (II.67). Nous nous intéressons dans un premier temps à la précision de la surface de réponse construite par la méthode de décomposition de la dimension. La figure II.15 illustre l'évolution de l'approximation de la réponse du modèle défini par l'équation (II.114) en fonction des différents paramètres incertains dans un intervalle  $[\mu_{X_i} \pm 3\sigma_{X_i}]$ . Sachant que la décomposition de la dimension est d'ordre  $s = 1$  et le nombre de points d'expérimentation est  $n = 3$ , on constate que l'approximation reproduit exactement la réponse du modèle. Par conséquent, cette dernière peut se substituer au modèle exact pour effectuer des analyses probabilistes. Dans ce contexte, nous nous intéressons dans ce qui suit au calcul des caractéristiques statistiques de la réponse du modèle. Pour ce faire,  $10^6$  simulations de Monte-Carlo ont été effectuées sur la surface de réponse obtenue par la méthode de décomposition de la dimension. Le tableau II.6 donne les estimations des quatre premiers moments statistiques de la réponse du modèle pour différents ordres  $s$  de décomposition de la dimension et pour un nombre de points d'expérimentation  $n = 3$ .

**Tableau II. 6 :** Moments statistiques de la réponse du modèle obtenus par simulations de Monte-Carlo sur la surface de réponse construite par la méthode de décomposition de la dimension

Moments statistiques	Ordre de la méthode de réduction	
	$s = 1, n = 3$	$s = 2, n = 3$
$\mu_Y$	10,48	10,48
$\sigma_Y$	4,707	4,768
$\delta_Y$	0,208	0,691
$\kappa_Y$	3,046	3,615

On constate que les estimations obtenues par la surface de réponse d'ordre  $s = 2$  en conjonction avec les simulations de Monte-Carlo s'accordent bien avec la solution de référence (cf. §I.4.3.2), en ne nécessitant que 37 appels au modèle. De plus, on constate que l'erreur d'estimation des moments statistiques obtenus par simulations de Monte-Carlo en conjonction avec une approximation d'ordre  $s = 1$  du modèle, est moins significative que celle observée sur les estimations obtenues par intégration basée sur la méthode de réduction de la dimension du même d'ordre.



**Figure II. 16 :** *Comparaison de la densité de probabilité  $f_Y(y)$  de la réponse du modèle obtenue par la méthode indirecte (simulations de Monte-Carlo sur la surface de réponse) et la méthode directe (simulations de Monte-Carlo sur le modèle)*

La densité de probabilité  $f_Y(y)$  de la réponse du modèle a été construite par deux approches : la première consiste à effectuer des simulations de Monte-Carlo sur la surface de réponse obtenue par la méthode de décomposition et la deuxième, prise comme solution de référence, consiste à effectuer des simulations de Monte-Carlo directement sur le modèle défini par l'équation (II.67). La figure II.16 compare les densités de probabilité obtenues par  $10^6$  simulations de Monte-Carlo sur la surface de réponse construite par la méthode de décomposition de la dimension d'ordre  $s = 1$  et d'ordre  $s = 2$ , avec la solution de référence. On constate que la densité de probabilité construite par la surface de réponse obtenue à partir de la méthode de décomposition de la dimension d'ordre  $s = 2$  s'accorde bien avec la solution de référence. De plus, on constate qu'en utilisant une approximation d'ordre  $s = 1$ , la densité de probabilité obtenue par simulations de Monte-Carlo est plus précise que celle construite à partir de la méthode des moments (cf. figure II.13). Ainsi, on peut conclure que la méthode de décomposition de la dimension est plus efficace pour l'approximation de fonction  $N$ -dimensionnelle que pour l'estimation d'intégrale  $N$ -dimensionnelle.

#### II.4.5 Estimation de la fiabilité

Les deux paragraphes précédents ont montré l'efficacité et la précision de la méthode de décomposition de la dimension dans la résolution des problèmes d'analyse de tendance centrale. En effet, nous avons vu que cette méthode offre deux techniques distinctes pour le calcul des caractéristiques statistiques (i.e. les moments statistiques et la densité de probabilité) de la réponse d'un modèle stochastique : la première consiste à calculer les intégrales  $N$ -dimensionnelles définissant les moments statistiques par réduction de la dimension et à construire la densité de probabilité par une méthode de moments, et la deuxième consiste à construire une surface de réponse par projection sur la base de polynômes de Lagrange, et à déterminer les moments statistiques et la densité de probabilité par des simulations de Monte-Carlo sur la surface de réponse. Nous avons montré précédemment que les deux techniques permettent une construction précise de la densité de probabilité, à condition que, pour la première technique, les estimations des quatre premiers moments statistiques soient précises, ce qui est très intéressant dans le cadre d'analyse de fiabilité du fait que la densité de probabilité peut être directement utilisée dans l'estimation de la probabilité de défaillance. Cependant, il faut signaler que cette technique n'est efficace

que pour l'estimation des probabilités de défaillance d'ordre élevé (i.e.  $\geq 10^{-3}$ ). Cette situation est rarement rencontrée dans la pratique dans la mesure où on conçoit les structures les plus robustes et les plus durables possibles, ce qui se traduit par des probabilités de défaillance d'ordre faible. Ainsi, l'utilisation d'une surface de réponse représentant le scénario de défaillance en conjonction avec les simulations de Monte-Carlo peut être une solution à ce problème (i.e. estimation des probabilités de défaillance faibles). Dans ce contexte, Xu et Rahman [123] ont utilisé cette technique en construisant la surface de réponse dans l'espace physique. Ensuite Wei et Rahman [128] ont suggéré de construire la surface de réponse dans l'espace standard en prenant comme point de référence, le point de défaillance le plus probable. Cette amélioration est fondée sur la détermination du point de défaillance le plus probable par le méthode d'approximation FORM. Si on raisonne en terme d'appels au modèle, un coût de calcul supplémentaire s'ajoute au coût nécessaire pour la construction de la surface de réponse. De plus, lorsque la fonction d'état limite est complexe (par exemple présence de bruit de fond, de fortes courbures, points de défaillance multiples), la convergence de la méthode FORM n'est pas garantie et par conséquent la surface de réponse obtenue peut être sujette à caution. Ce problème peut être surmonté par l'utilisation d'un algorithme robuste pour la détermination du point de défaillance le plus probable. Dans ce contexte, Wei et Rahman [129] ont proposé, dans le cas où la fonction d'état limite possède plusieurs optima, d'utiliser un algorithme spécifique [31] pour la détermination de tous les points de défaillance possibles. Dans ce cas, ils ont suggéré de construire autant de surfaces de réponse que de points de défaillance.

Afin d'illustrer cette approche, nous nous proposons d'estimer la probabilité de défaillance associée au scénario de défaillance représenté par la fonction d'état limite suivante [128] :

$$g(x_1, x_2) = 2,2257 - \frac{0,025\sqrt{2}}{27}(x_1 + x_2 - 20)^3 + \frac{33}{140}(x_1 - x_2) \quad (II.116)$$

où  $x_1$  et  $x_2$  sont deux paramètres incertains représentés par deux variables Gaussiennes indépendantes de moyenne  $\mu = 10$  et d'écart-type  $\sigma = 3$ . La particularité de cette fonction d'état limite cubique est l'existence d'un point d'inflexion qui servira dans la suite pour évaluer la précision de la surface de réponse obtenue par la méthode de décomposition de la dimension. Afin de déterminer les coordonnées du point de défaillance le plus probable, nous avons utilisé l'algorithme de Rackwitz-Fiessler [44]. Le tableau II.7 donne les coordonnées du point de défaillance le plus probable  $\mathbf{u}^*$ , l'indice de fiabilité  $\beta$  de Hasofer-Lind et l'estimation de la probabilité de défaillance  $P_f$  par les méthodes d'approximation FORM et SORM, et les simulations de Monte-Carlo.

**Tableau II. 7 : Résultats d'analyse de fiabilité par les méthodes d'approximation FORM et SORM**

$\mathbf{u}^*$	$\mathbf{x}^*$	$\beta = \ \mathbf{u}^*\ $	FORM	SORM			MC
				Breitung	Hohenbichler	Tvedt	
-1,5737 1,5737	5,2788 14,7211	2,2255	0,01302	0,01302	0,01302	0,01302	0,01891

On note que l'algorithme de Rackwitz-Fiessler nécessite 14 appels au modèle pour la détermination du point de défaillance le plus probable. Les résultats donnés dans le tableau II.7 montrent que la méthode d'approximation de second ordre (SORM) basée sur les formulations de Breitung [28], de Tvedt [30] et de Hohenbichler [130] donne le même résultat que la méthode d'approximation d'ordre 1 (FORM), ce qui indique un état limite linéaire au point de défaillance le plus probable. Toutefois, les estimations de la probabilité de défaillance ne s'accordent pas avec le résultat donné par les simulations de Monte-Carlo.

Nous nous proposons maintenant d'utiliser la méthode de réduction de la dimension pour construire une surface de réponse de la fonction d'état limite (II.116). Le tableau II.8 donne l'estimation de la probabilité de défaillance en utilisant  $10^6$  simulations de Monte-Carlo sur la surface de réponse. Nous notons que la surface de réponse est construite à partir de cinq points d'interpolation dans chaque direction, et en utilisant le point moyen (noté  $\{\mathbf{0}\}$  dans le tableau II.8) ou le point de défaillance le plus probable (noté  $\{\mathbf{u}^*\}$  ou  $\{\mathbf{v}^*\}$  dans le tableau II.8) comme point de référence ; dans ce cas, l'approximation de la fonction d'état limite est construite soit dans l'espace aléatoire standard  $\mathbf{u}$ , soit dans l'espace aléatoire standard  $\mathbf{v}$  obtenu par la rotation de l'espace  $\mathbf{u}$  vers la direction du point de défaillance le plus probable. Ce choix est motivé, d'une part, par le fait que, dans les problèmes d'analyse de fiabilité, la masse de probabilité est localisée au voisinage du point de défaillance le plus probable, ce qui peut améliorer la précision de l'estimation de la probabilité de défaillance, et, d'autre part, par le fait que nous pouvons mieux capter la non linéarité qui peut se manifester dans la fonction d'état limite.

**Tableau II. 8** : Estimations de la probabilité de défaillance par couplage de la méthode de décomposition de la dimension et les simulations de Monte-Carlo

	Décomposition d'ordre 1			Décomposition d'ordre 2		
	$\{\mathbf{0}\}$	$\{\mathbf{u}^*\}$	$\{\mathbf{v}^*\}$	$\{\mathbf{0}\}$	$\{\mathbf{u}^*\}$	$\{\mathbf{v}^*\}$
$P_f$	$1,408 \cdot 10^{-2}$	$1,556 \cdot 10^{-2}$	$1,920 \cdot 10^{-2}$	$1,858 \cdot 10^{-2}$	$1,858 \cdot 10^{-2}$	$1,858 \cdot 10^{-2}$
<i>Nb appels au modèle</i>	11	18	18	36	43	43

On constate que le couplage entre la surface de réponse obtenue par la méthode de décomposition de la dimension et les simulations de Monte-Carlo fournit une estimation précise de la probabilité de défaillance sans avoir recours à un nombre élevé d'évaluations de la fonction d'état limite. En particulier, dans le cas où on utilise une décomposition d'ordre 2, l'estimation de la probabilité de défaillance est insensible au point de référence autour duquel est construite l'approximation de la fonction d'état limite. En effet une décomposition de la dimension d'ordre 2 permet de fournir une approximation suffisamment précise et le point de référence ne joue pas un grand rôle pour l'amélioration de cette précision. Cependant, dans le cas où une décomposition d'ordre 1 est utilisée, le point de référence joue un rôle important sur la précision de l'estimation de la probabilité de défaillance. On peut conclure que, si la précision associée à l'ordre de la méthode de réduction n'est pas optimale, le point de référence autour duquel est construite la fonction d'état limite peut améliorer cette précision. On note aussi que la meilleure précision sur l'estimation de la probabilité de défaillance est obtenue lorsque l'approximation est construite autour du point de défaillance le plus probable  $\mathbf{v}^*$ .

## II.4.6 Méthode de décomposition polynômiale de la dimension

### II.4.6.1 Motivations

En se référant à la méthode de décomposition de la dimension, la construction d'une approximation de la réponse d'un modèle nécessite le choix d'un point de référence et la construction d'un plan d'expérimentation autour de ce point de référence. A l'aide de l'ensemble de ces points, les différentes fonctionnelles impliquées par la méthode de décomposition sont considérées comme des fonctions de forme construites par projection sur la base des polynômes de Lagrange. A l'aide d'une application illustrative traitée dans le paragraphe précédent, nous avons montré que cette surface de réponse en conjonction avec les simulations de Monte-Carlo fournit de bons résultats en particulier dans la résolution de

problèmes de propagation d'incertitude ou d'analyse de fiabilité. Cependant, nous signalons que cette procédure basée sur les polynômes d'interpolation de Lagrange pour la construction de la surface de réponse peut avoir deux limites majeures.

La première est définie par la forte dépendance de la précision de la surface de réponse au point de référence choisi. Si ce choix est inadéquat, cela peut mener à des résultats erronés. En effet, dans le paragraphe précédent, nous avons montré que l'estimation de la probabilité de défaillance par simulations de Monte-Carlo est plus précise lorsque la surface de réponse associée à la méthode de décomposition de la dimension est construite autour du point de défaillance le plus probable. Dans le cas où ce dernier correspond au point moyen, on a constaté que l'estimation de la probabilité de défaillance n'est que peu améliorée par comparaison aux résultats donnés par les méthodes d'approximation FORM et SORM.

La deuxième limite est liée au choix des points d'expérimentation (i.e. points d'interpolation autour desquels sont construits les polynômes de Lagrange), qui est généralement effectué de manière arbitraire sans consistance physique. En effet, nous avons vu que la procédure classique proposée par Xu et Rahman [123] consiste à construire ces points en imposant une perturbation dans chaque direction de l'espace aléatoire dépendant de l'écart-type de la variable aléatoire associée à cette direction. Dans le cas où les variables aléatoires représentant les paramètres incertains du modèle sont strictement positives ou négatives ou leurs densités de probabilité sont définies sur un support compact, les points d'expérimentation obtenus par la procédure ci-dessus peuvent être en dehors du domaine physique. Dans cette situation, il y a une forte probabilité que l'approximation générée par la méthode de décomposition de la dimension basée sur les polynômes d'interpolation de Lagrange ne soit pas représentative du comportement réel du modèle. Pour pallier ces deux problèmes, Rahman [131] a proposé le développement sur la base des polynômes de Fourier pour construire les fonctionnelles associées à la méthode de décomposition de la dimension. Cette approche, baptisée méthode de décomposition polynômiale de la dimension (MDPD), fera l'objet du paragraphe suivant.

#### II.4.6.2 Méta-modèle par projection sur une base polynômiale orthonormale

Si  $\mathbf{U}$  est une variable Gaussienne standard  $N$ -dimensionnelle, représentant dans l'espace aléatoire standard les paramètres incertains  $\mathbf{X}$  du modèle, l'approximation d'ordre  $s$ , construite par la méthode de décomposition polynômiale de la dimension, de la réponse scalaire  $y$  du modèle s'écrit dans l'espace aléatoire standard :

$$\begin{aligned} \tilde{y}_s = y_0 &+ \sum_{i=1}^N \sum_{j=1}^m \alpha_{ij} \Psi_j(u_i) + \sum_{i_1=1}^{N-1} \sum_{i_2=i_1+1}^N \sum_{j_1=1}^m \sum_{j_2=1}^m \beta_{i_1 i_2 j_1 j_2} \Psi_{j_1}(u_{i_1}) \Psi_{j_2}(u_{i_2}) \\ &+ \sum_{i_1=1}^{N-2} \sum_{i_2=i_1+1}^{N-1} \sum_{i_3=i_2+1}^N \sum_{j_1=1}^m \sum_{j_2=1}^m \sum_{j_3=1}^m \gamma_{i_1 i_2 i_3 j_1 j_2 j_3} \Psi_{j_1}(u_{i_1}) \Psi_{j_2}(u_{i_2}) \Psi_{j_3}(u_{i_3}) \\ &+ \dots + \sum_{i_1=1}^{N-s+1} \dots \sum_{i_s=i_{s-1}+1}^N \sum_{j_1=1}^m \dots \sum_{j_s=1}^m \zeta_{i_1 \dots i_s j_1 \dots j_s} \Psi_{j_1}(u_{i_1}) \dots \Psi_{j_s}(u_{i_s}) \end{aligned} \quad (II.117)$$

où les polynômes unidimensionnels  $\{\Psi_j\}_{j=0}^m$  forment une base orthonormale et  $\alpha_{ij}$ ,  $\beta_{i_1 i_2 j_1 j_2}$ ,  $\gamma_{i_1 i_2 i_3 j_1 j_2 j_3}$  et  $\zeta_{i_1 \dots i_s j_1 \dots j_s}$  sont des coefficients réels à déterminer.

Les polynômes unidimensionnels  $\{\Psi_j\}_{j=0}^m$  sont choisis de telle manière à correspondre aux distributions des paramètres incertains. Le tableau II.1 donne la correspondance entre les familles de polynômes orthogonaux et les différentes distributions de probabilité. Etant donné que l'approximation est définie dans l'espace aléatoire standard, ces polynômes ne sont autres que les polynômes d'Hermite normalisés, obtenus par :

$$\Psi_j = \frac{H_j}{\|H_j\|} = \frac{H_j}{\sqrt{j!}}, j = 1, \dots, m \quad (II.118)$$

En se référant à l'équation (II.117), la construction d'une approximation d'ordre  $s$  de la réponse d'un modèle nécessite la détermination des coefficients réels  $y_0$ ,  $\alpha_{ij}$ ,  $\beta_{i_1 i_2 j_1 j_2}$ ,  $\gamma_{i_1 i_2 i_3 j_1 j_2 j_3}$  et  $\zeta_{i_1 \dots i_s j_1 \dots j_s}$ , définis par :

$$y_0 = \mathbb{E}[f \circ T(\mathbf{U})] = \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \phi(\mathbf{u}) d\mathbf{u} \quad (II.119)$$

$$\alpha_{ij} = \mathbb{E}[f \circ T(\mathbf{U}) \Psi_j(U_i)] = \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \Psi_j(u_i) \phi(\mathbf{u}) d\mathbf{u} \quad (II.120)$$

$$\beta_{i_1 i_2 j_1 j_2} = \mathbb{E}\left[f \circ T(\mathbf{U}) \prod_{l=1}^2 \Psi_{j_l}(U_{i_l})\right] = \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \prod_{l=1}^2 \Psi_{j_l}(u_{i_l}) \phi(\mathbf{u}) d\mathbf{u} \quad (II.121)$$

$$\gamma_{i_1 i_2 i_3 j_1 j_2 j_3} = \mathbb{E}\left[f \circ T(\mathbf{U}) \prod_{l=1}^3 \Psi_{j_l}(U_{i_l})\right] = \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \prod_{l=1}^3 \Psi_{j_l}(u_{i_l}) \phi(\mathbf{u}) d\mathbf{u} \quad (II.122)$$

$$\zeta_{i_1 \dots i_s j_1 \dots j_s} = \mathbb{E}\left[f \circ T(\mathbf{U}) \prod_{l=1}^s \Psi_{j_l}(U_{i_l})\right] = \int_{\mathbb{R}^N} f \circ T(\mathbf{u}) \prod_{l=1}^s \Psi_{j_l}(u_{i_l}) \phi(\mathbf{u}) d\mathbf{u} \quad (II.123)$$

où  $T$  désigne la transformation iso-probabiliste,  $\phi$  est la densité de probabilité d'une variable Gaussienne standard  $N$ -dimensionnelle, et  $\mathbb{E}[\cdot]$  est l'espérance mathématique associée à la mesure de probabilité Gaussienne. Il est clair que la détermination des coefficients inconnus de l'approximation n'est autre que l'évaluation d'un ensemble d'intégrales  $N$ -dimensionnelles. Ceci peut être effectué, comme on l'a montré dans le cas de la détermination des coefficients inconnus du développement en chaos polynômial, par les simulations de Monte-Carlo ou par la quadrature de Gauss-Hermite. Cependant, dans le cas où la dimension  $N$  de l'espace aléatoire est élevée, ces techniques sont inabordables du fait du nombre important d'appels au modèle. Cette difficulté peut être levée d'une façon partielle, dans le cas où la dimension  $N$  de l'espace aléatoire n'est pas trop élevée, en utilisant la technique de quadrature de Smolyak. Dans le même contexte, Rahman [131] a suggéré d'utiliser la méthode de réduction de la dimension d'intégration. En se référant à l'équation (II.103) qui représente la généralisation de la méthode d'intégration par réduction de la dimension pour le calcul d'une intégrale  $N$ -dimensionnelle, et notant  $r$  l'ordre de la réduction, les coefficients  $y_0$  et  $\zeta_{i_1 \dots i_s j_1 \dots j_s}$  sont respectivement donnés par :

$$y_0 \cong \sum_{k=0}^r (-1)^k C_{N-r+k-1}^k \times \sum_{k_1=1}^{N-r+k+1} \dots \sum_{k_{r-i}=k_{r-i-1}+1}^N \int_{\mathbb{R}^{r-k}} f \circ T(u_{k_1}, \dots, u_{k_{r-k}}) \prod_{l=1}^{r-k} \phi_{k_l}(u_{k_l}) du_{k_l} \quad (II.124)$$

$$\zeta_{i_1 \dots i_s j_1 \dots j_s} \cong \sum_{k=0}^r (-1)^k C_{N-r+k-1}^k$$

$$\times \sum_{k_1=1}^{N-r+k+1} \dots \sum_{k_{r-i}=k_{r-i-1}+1}^N \int_{\mathbb{R}^{r-k}} f \circ T(u_{k_1}, \dots, u_{k_{r-k}}) \prod_{l=1}^s \Psi_{j_l}(u_{i_l}) \prod_{l=1}^{r-k} \phi_{k_l}(u_{k_l}) du_{k_l} \quad (II.125)$$

Les équations (II.124) et (II.125) impliquent, au pire des cas, l'évaluation d'intégrales  $r$ -dimensionnelles. Dans le cas où  $r \ll N$ , cela contribue à réduire considérablement le nombre d'évaluations de l'intégrande ou le nombre d'appels au modèle implicite. En effet, si  $r = 1$ ,  $r = 2$  ou  $r = 3$ , cela signifie que les intégrales élémentaires (i.e. résultant de la méthode de réduction de la dimension d'intégration) qui figurent dans les équations (II.124) et (II.125), sont respectivement unidimensionnelles, bidimensionnelles ou tridimensionnelles, et peuvent être calculées facilement par quadrature de Gauss-Hermite. Il faut noter que l'application de la méthode de réduction de la dimension exige que l'ordre  $r$  de la réduction soit choisi de telle manière à être supérieur ou égal à l'ordre de la décomposition  $s$  dans l'équation (II.117).

### II.4.6.3 Post-traitements basés sur la MDPD

Une fois la surface de réponse obtenue par la méthode de décomposition polynômiale de la dimension, différents types de post-traitement peuvent être effectués, tels que l'analyse de tendance centrale, analyse de fiabilité et analyse de sensibilité. Dans ce qui suit, nous nous intéressons en particulier au calcul des moments statistiques et des indices de sensibilité de Sobol. Il est à noter que dans le cas de l'analyse de fiabilité, la procédure à suivre ne diffère pas de celle présentée dans le cas du développement en chaos polynômial sauf dans l'étape de construction de la surface de réponse représentant la fonction d'état limite associée au scénario de défaillance. Une fois cette tâche effectuée, la probabilité de défaillance peut être estimée par simulations de Monte-Carlo.

#### II.4.6.3.1 Calcul des moments statistiques

Les moments statistiques de la variable aléatoire scalaire représentant l'aléa associé à la réponse d'un modèle peuvent être calculés soit par simulations de Monte-Carlo sur la surface de réponse, soit en utilisant les coefficients  $y_0$  et  $\zeta_{i_1 \dots i_s j_1 \dots j_s}$  de l'approximation. Nous nous intéressons à la deuxième solution. Dans ce contexte, les moments statistiques sont facilement obtenus en appliquant l'espérance mathématique  $\mathbb{E}[\cdot]$  à l'approximation donnée par l'équation (II.117) et en exploitant l'orthogonalité des polynômes d'Hermite. La moyenne  $\hat{\mu}_{Y,S}$  et la variance  $\hat{\sigma}_{Y,S}^2$  sont aussi données respectivement par :

$$\hat{\mu}_{Y,S} = \mathbb{E}[\tilde{y}_S] = y_0 \quad (II.126)$$

$$\hat{\sigma}_{Y,S}^2 = \mathbb{E}[(\tilde{y}_S - y_0)^2] = \sum_{k=1}^s \sum_{i_1=1}^{N-s+1} \dots \sum_{i_s=i_{s-1}+1}^N \sum_{j_1=1}^m \dots \sum_{j_s=1}^m \zeta_{i_1 \dots i_s j_1 \dots j_s}^2 \quad (II.127)$$

Les moments statistiques d'ordre supérieur à deux peuvent être obtenus de la même manière. Cependant, malgré la simplification engendrée par l'orthogonalité des polynômes d'Hermite, le calcul des espérances élémentaires résultantes après décomposition de la dimension n'est pas trivial puisqu'il implique le produit de plus de quatre polynômes d'Hermite unidimensionnels. On note que, dans le cas d'un développement en chaos polynômial, cette difficulté n'est pas observée puisque le moment d'ordre 4 exige seulement l'évaluation d'espérances mathématiques du produit de quatre polynômes d'Hermite.

### II.4.6.3.2 Calcul des indices de Sobol

Dans le cadre de la méthode de décomposition polynômiale, la détermination des indices de sensibilité de Sobol est plus simple que dans le cas du développement en chaos polynômial. En effet, si on normalise l'approximation fournie par la méthode de décomposition polynômiale par la variance totale  $\hat{\sigma}_{Y,S}^2$ , on obtient la même architecture (i.e. les mêmes termes) que la décomposition de Sobol [51], et par conséquent les indices de Sobol de différents ordres sont directement identifiés. Les indices d'ordre 1, d'ordre 2, et d'ordre 3 sont aussi donnés respectivement par :

$$\hat{S}_{i;s} = \frac{1}{\hat{\sigma}_{Y,S}^2} \sum_{j=1}^m \alpha_{ij}^2 \quad (II.128)$$

$$\hat{S}_{i_1 i_2 ;s} = \frac{1}{\hat{\sigma}_{Y,S}^2} \sum_{j_1=1}^m \sum_{j_2=1}^m \beta_{i_1 i_2 j_1 j_2}^2 \quad (II.129)$$

$$\hat{S}_{i_1 i_2 i_3 ;s} = \frac{1}{\hat{\sigma}_{Y,S}^2} \sum_{j_1=1}^m \sum_{j_2=1}^m \sum_{j_3=1}^m \gamma_{i_1 i_2 i_3 j_1 j_2 j_3}^2 \quad (II.130)$$

où les termes  $\sum_{j=1}^m \alpha_{ij}^2$ ,  $\sum_{j_1=1}^m \sum_{j_2=1}^m \beta_{i_1 i_2 j_1 j_2}^2$  et  $\sum_{j_1=1}^m \sum_{j_2=1}^m \sum_{j_3=1}^m \gamma_{i_1 i_2 i_3 j_1 j_2 j_3}^2$  peuvent être interprétés comme les variances partielles quantifiant les parts respectives de l'effet individuel du paramètre  $X_i$ , de l'effet de l'interaction entre les deux paramètres  $X_{i_1}$  et  $X_{i_2}$ , et de l'effet de l'interaction entre les trois paramètres  $X_{i_1}$ ,  $X_{i_2}$  et  $X_{i_3}$  sur la variance totale  $\hat{\sigma}_{Y,S}^2$  de la réponse du modèle. D'une manière générale, l'indice de sensibilité de Sobol d'ordre  $s$  est défini par :

$$\hat{S}_{i_1, \dots, i_s ;s} = \frac{1}{\hat{\sigma}_{Y,S}^2} \sum_{j_1=1}^m \dots \sum_{j_s=1}^m \zeta_{i_1 \dots i_s j_1 \dots j_s}^2 \quad (II.131)$$

Il est clair que la convergence des indices de Sobol  $\hat{S}_{i_1, \dots, i_s ;s}$  dépend essentiellement de l'ordre  $s$  de la décomposition polynômiale et du paramètre  $m$  représentant l'ordre de troncature adopté dans la construction par projection sur la base des polynômes d'Hermite des différentes fonctionnelles dans l'équation (II.117). En effet, il est évident que  $\hat{S}_{i_1, \dots, i_s ;s}$  converge vers la valeur exacte lorsque  $s$  tend vers la dimension stochastique  $N$ . Cependant, la convergence par rapport au paramètre  $m$  est plus compliquée à évaluer, et elle est fortement dépendante de la régularité de la réponse du modèle. L'ordre des indices de Sobol qu'on peut évaluer est imposé par l'ordre  $s$  choisi pour construire la surface de réponse. A titre illustratif, si  $s = 3$ , cela signifie que la méthode de décomposition polynômiale permet de déterminer les indices de Sobol d'ordre au plus égal à 3. Cela peut être considéré comme une limite dans le cas où la dimension stochastique efficace est supérieure à trois du fait que l'effet des interactions entre quatre paramètres ne peut pas être quantifié bien qu'il puisse être significatif. On note, de plus, que l'indice de Sobol total, associé à un paramètre donné, peut être évalué à partir des indices de Sobol de différents ordres impliquant ce paramètre. On peut écrire alors :

$$\hat{S}_{i;s}^T = \hat{S}_{i;s} + \sum_{i \neq j} \hat{S}_{ij;s} + \sum_{j \neq i, k \neq i, j < k} \hat{S}_{ijk;s} + \dots \quad (II.132)$$

### II.4.6.4 Exemples d'illustration

Dans ce paragraphe nous nous intéressons à la comparaison de la méthode de développement en chaos polynômial et de la méthode de décomposition polynômiale dans l'estimation des indices de sensibilité de Sobol. Pour ce faire, nous étudions deux exemples d'illustrations. Le premier, relativement simple, traite un problème usuel dans le contexte de la mécanique de la rupture. Ce problème a été traité par Lemaire et al. [132] dans un contexte d'analyse de fiabilité, pour lequel le scénario de défaillance est défini par l'atteinte de l'intégrale  $J$  caractérisant la rupture du matériau d'une valeur seuil  $J_{Ic}$  liée à la ténacité du matériau. Ce scénario de défaillance est représenté par la fonction d'état limite suivante :

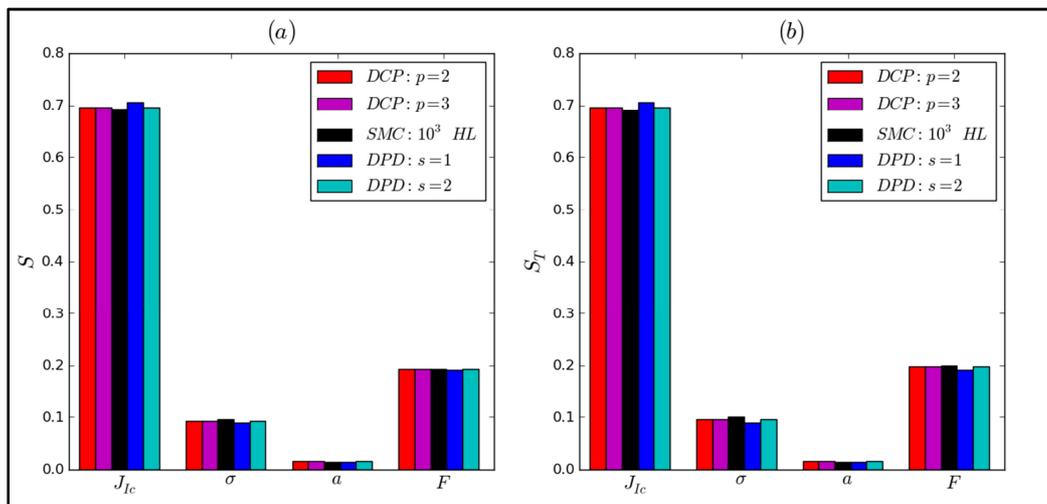
$$g(\mathbf{x}) = J_{Ic} - \frac{1 - \nu^2}{E} (F\sigma\sqrt{\pi a})^2 \quad (II.133)$$

où  $E$  et  $\nu$  sont respectivement le module de Young et le coefficient de Poisson du matériau,  $\sigma$  la contrainte appliquée,  $a$  la taille de la fissure initiale et  $F$  le facteur de correction de forme. Dans notre étude, l'aléa concerne la résistance à la rupture  $J_{Ic}$ , la charge appliquée  $\sigma$ , la taille initiale de la fissure  $a$  et le facteur de correction de forme  $F$ . Ils sont modélisés par des variables aléatoires lognormales indépendantes dont les caractéristiques statistiques sont données dans le tableau II.9.

**Tableau II. 9 :** Caractéristiques statistiques des paramètres incertains du problème de la mécanique de la rupture

Paramètre	$J_{Ic}$	$\sigma$	$a$	$F$	$E$	$\nu$
Distribution		lognormale			déterministe	
$\mu$	55	145	90	1	210000	0,3
$\sigma$	10	10	5	0,1		

Nous nous proposons de déterminer les paramètres les plus influents sur la fonction d'état limite définie par l'équation (II.133). Dans cette perspective, les indices de Sobol du premier ordre et totaux sont calculés d'une part en se basant sur un développement en chaos polynômial, les coefficients étant déterminés par quadrature de Smolyak et, d'autre part, en se basant sur un développement par décomposition polynômiale. On note que la solution de référence est obtenue par simulations de Monte-Carlo basées sur les tirages dans un hypercube Latin.



**Figure II. 17 :** Comparaison des résultats obtenus par développement en chaos polynômial et par décomposition polynômiale de la dimension pour (a) les indices de Sobol de premier ordre et (b) les indices de Sobol totaux

Les figures II.17a et II.17b comparent les résultats obtenus respectivement pour les indices de Sobol de premier ordre et totaux. On constate que l'écart entre les indices de Sobol de premier ordre et les indices de Sobol totaux est négligeable. Ainsi, on peut conclure que l'effet des interactions entre les différents paramètres incertains est négligeable par rapport aux effets principaux. De plus, on constate que la variabilité de la fonction d'état limite définie par l'équation (II.133) est due principalement à l'aléa associé à la résistance à la rupture  $J_{Ic}$ .

Globalement, les estimations données par les différentes méthodes sont en bon accord avec la solution de référence. Par ailleurs on peut signaler que la méthode de décomposition de la dimension d'ordre  $s = 2$  garantit le meilleur compromis entre la précision et le temps de calcul, puisqu'elle nécessite 67 évaluations de la fonction d'état limite, tandis que le développement en chaos polynômial nécessite 125 évaluations de la fonction d'état limite, ce qui reste tout de même plus économique que les simulations de Monte-Carlo basées sur les tirages en hyper-cube Latin. On note que, pour la méthode de décomposition polynômiale, les sommes sont tronquées à l'ordre  $m = 3$  et les intégrales élémentaires sont calculées par quadrature de Gauss-Hermite basée sur trois points d'intégration.

Nous considérons maintenant un second exemple caractérisé par la fonction d'état limite suivante, associée au comportement d'un oscillateur à deux degrés de liberté :

$$g(\mathbf{x}) = F_s - 3k_s \sqrt{\frac{\pi S_0}{4\zeta_s \omega_s^3} \frac{\zeta_a \zeta_s}{\zeta_p \zeta_s (4\zeta_a^2 + \theta^2) + \gamma \zeta_a^2} \frac{(\zeta_p \omega_p^3 + \zeta_s \omega_s^3) \omega_p}{4\zeta_a^2 \omega_a^4}} \quad (II.134)$$

où,  $\omega_p = \sqrt{\frac{k_p}{m_p}}$ ,  $\omega_s = \sqrt{\frac{k_s}{m_s}}$ ,  $\omega_a = \frac{1}{2}(\omega_p + \omega_s)$ ,  $\zeta_a = \frac{1}{2}(\zeta_p + \zeta_s)$  et  $\theta = \frac{1}{\omega_a}(\omega_p - \omega_s)$ .

La particularité de ce problème, traité à l'origine par Der Kiureghian et De Stefano [133] dans le contexte d'analyse de fiabilité, est que la fonction d'état limite est fortement non linéaire. Dans cette étude, nous cherchons à déterminer les paramètres les plus influents sur la fonction d'état limite. Les paramètres incertains sont modélisés par des variables aléatoires lognormales indépendantes, dont les caractéristiques statistiques sont données dans le tableau II.10.

**Tableau II. 10 :** Caractéristiques statistiques des paramètres incertains du problème de d'oscillateur à deux degrés de liberté

Paramètre	$m_p$	$m_s$	$k_p$	$k_s$	$\zeta_p$	$\zeta_s$	$F_s$	$S_0$
Distribution	lognormale							
$\mu$	1,5	0,01	1	0,01	0,05	0,02	15	100
CV(%)	10	10	20	20	40	50	10	10

Les indices de sensibilité de Sobol de premier ordre et totaux sont calculés d'une part en se basant sur les coefficients d'un développement en chaos polynômial d'ordres 2 et 3, et d'autre part en utilisant les coefficients d'une décomposition polynômiale de second ordre mais pour différents ordres de troncature  $m = \{2, 3\}$ .

Les résultats obtenus sont donnés dans le tableau II.11, la solution de référence étant obtenue par simulations de Monte-Carlo basées sur un échantillon de 50000 points générés dans un hyper-cube Latin. D'une façon générale, on constate que les estimations fournies par les deux méthodes sont en bon accord avec la solution de référence. On constate, de plus, que la précision des estimations est améliorée soit en augmentant l'ordre  $p$  du développement en chaos polynômial, soit en augmentant le nombre de termes retenus  $m$  dans les différentes sommes présentes dans la décomposition polynômiale. On note tout de même que la

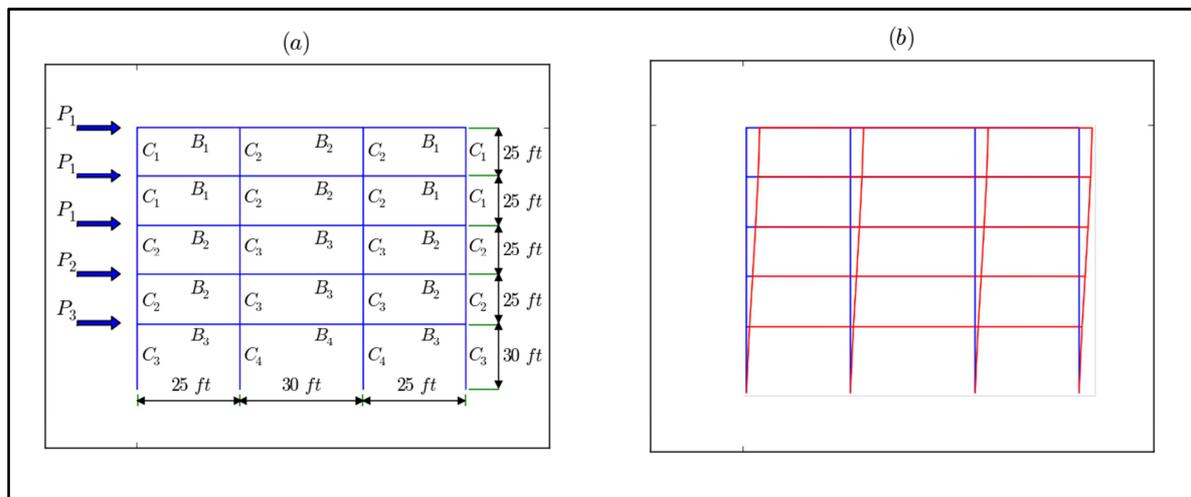
précision des indices de Sobol totaux calculés par la méthode de décomposition de la dimension est relativement meilleure.

**Tableau II. 11** : Indice de Sobol totaux obtenus par le développement en chaos polynômial et la décomposition polynômiale de la dimension pour le problème de l'oscillateur à deux degrés de liberté

Paramètre	chaos polynômial		décomposition polynômiale		SMC-HL
	$p = 2$	$p = 3$	$m = 2$	$m = 3$	
$m_p$	0,0489	0,0668	0,0856	0,1006	0,0957
$m_s$	0,1260	0,1408	0,1481	0,1729	0,1668
$k_p$	0,2630	0,2954	0,3335	0,3335	0,3255
$k_s$	0,1751	0,2204	0,2741	0,2713	0,2673
$\zeta_p$	0,0837	0,0781	0,1386	0,0869	0,0823
$\zeta_s$	0,0707	0,0658	0,0744	0,0614	0,0659
$F_s$	0,2606	0,2392	0,1776	0,2014	0,2244
$S_0$	0,0095	0,0088	0,0106	0,0085	0,0084

### II.5 Validation de la stratégie de résolution

Dans ce paragraphe, nous nous proposons d'illustrer l'implémentation de la procédure de résolution du problème de la dimension stochastique élevée sur un problème académique traitant l'évaluation de la résistance d'une structure à cinq niveaux sollicitée par des charges latérales (cf. figure II.18a). Ce problème a été traité par Liu et Der Kiureghian [134] pour évaluer l'intégrité de la structure vis-à-vis d'un scénario de défaillance défini par le dépassement du déplacement horizontal de l'étage supérieur de la structure d'un seuil critique. Ensuite, il a été repris par Blatman et Sudret [135] dans un contexte d'analyse de sensibilité basée sur le calcul des indices de Sobol et le développement en chaos polynômial.



**Figure II. 18** : Structure à cinq niveaux sollicitée par des charges latérales (a) géométrie (b) déformée

Nous nous intéressons au déplacement horizontal de l'étage supérieur. Un modèle éléments finis a été développé avec le logiciel Cast3m pour lequel la structure est modélisée en utilisant des éléments de poutre. Les paramètres incertains du problème sont les trois charges  $P_1$ ,  $P_2$  et  $P_3$ , deux modules de Young  $E_4$  et  $E_5$ , huit moments d'inertie  $\{I_i\}_{i=6}^{13}$  et huit sections  $\{A_i\}_{i=14}^{21}$ . Les propriétés géométriques et matérielles associées à chaque élément de la structure sont données dans le tableau II.12. Ils sont modélisés par des variables aléatoires normales et

lognormales indépendantes dont les caractéristiques statistiques sont données dans le tableau II.13.

**Tableau II. 12 :** Propriétés géométriques et matérielles des éléments de la structure à cinq niveaux

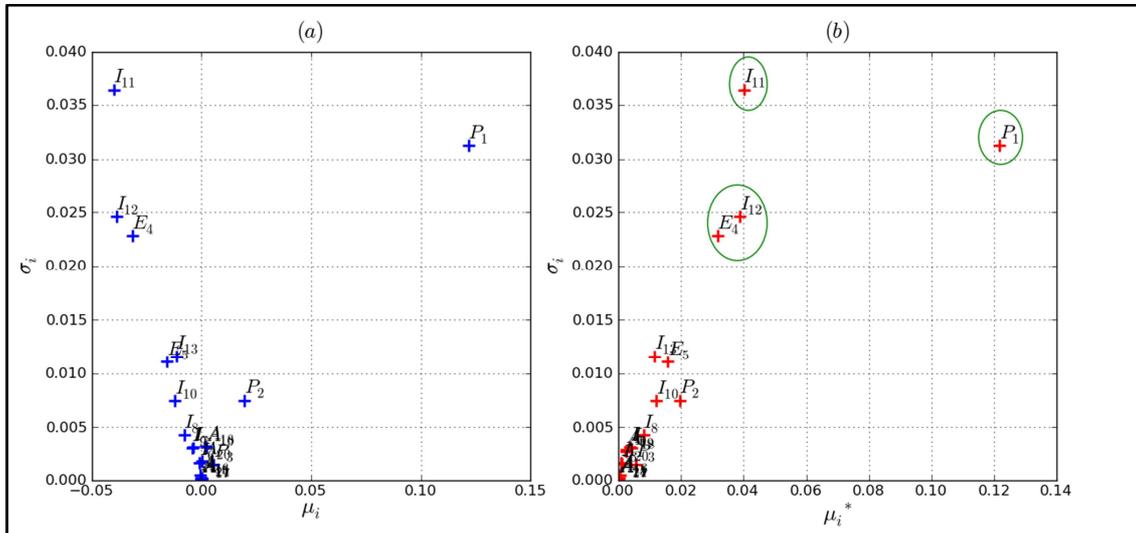
<i>Elément</i>	<i>Module de Young</i>	<i>Moment d'inertie</i>	<i>Section</i>
$B_1$	$E_4$	$I_{10}$	$A_{18}$
$B_2$	$E_4$	$I_{11}$	$A_{19}$
$B_3$	$E_4$	$I_{12}$	$A_{20}$
$B_4$	$E_4$	$I_{13}$	$A_{21}$
$C_1$	$E_5$	$I_6$	$A_{14}$
$C_2$	$E_5$	$I_7$	$A_{15}$
$C_3$	$E_5$	$I_8$	$A_{16}$
$C_4$	$E_5$	$I_9$	$A_{17}$

Dans un premier temps, nous nous proposons de déterminer la dimension stochastique efficace du problème. Autrement dit, nous cherchons, parmi les 21 paramètres incertains, les plus influents sur la réponse mécanique de la structure. Pour cela, nous appliquons la méthode de criblage de Morris, en considérant que chaque paramètre du modèle varie dans l'intervalle  $[\mu_{X_i} \pm 3\sigma_{X_i}]$ , où  $\mu_{X_i}$  et  $\sigma_{X_i}$  sont respectivement la moyenne et l'écart-type relatifs au paramètre  $X_i$ . Les figures II.19a et II.19b illustrent la hiérarchisation des paramètres incertains en se référant aux caractéristiques statistiques respectivement non normalisées et normalisées des effets élémentaires fournies par la méthode de Morris, pour un nombre de niveaux  $p = 4$  et dix trajectoires. On constate que la charge  $P_1$  est le paramètre le plus influent sur la réponse mécanique. Puis viennent, le module de Young  $E_4$  affecté aux éléments horizontaux, et les moments d'inertie  $I_{11}$  et  $I_{12}$  affectés respectivement aux éléments  $B_2$  et  $B_3$ . Les autres paramètres n'ont pas d'effet significatif sur la variabilité de la réponse du modèle puisqu'ils sont localisés au voisinage de l'origine du repère.

**Tableau II. 13 :** Caractéristiques statistiques des paramètres incertains pour le problème de la structure à cinq niveaux

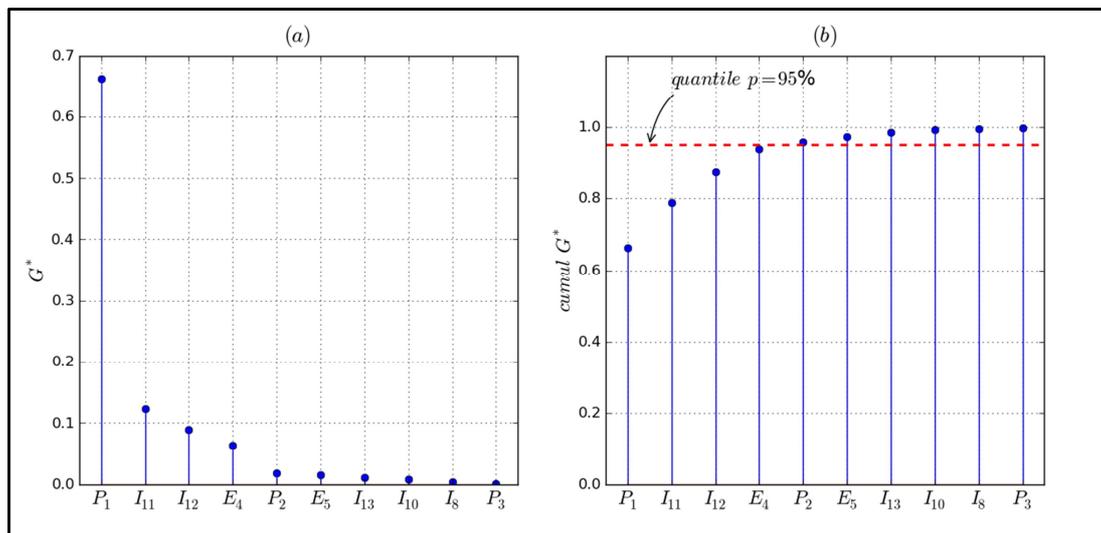
<i>Paramètre</i>	<i>Nature</i>	$\mu$	$\sigma$
$P_1$	lognormal	30	9
$P_2$	lognormal	20	8
$P_3$	lognormal	16	6,40
$E_4$	normal	454000	40000
$E_5$	normal	497000	40000
$I_6$	normal	0,94	0,12
$I_7$	normal	1,33	0,15
$I_8$	normal	2,47	0,30
$I_9$	normal	3,00	0,35
$I_{10}$	normal	1,25	0,30
$I_{11}$	normal	1,63	0,40
$I_{12}$	normal	2,69	0,65
$I_{13}$	normal	3,00	0,75
$A_{14}$	normal	3,36	0,60
$A_{15}$	normal	4,00	0,80
$A_{16}$	normal	5,44	1,00
$A_{17}$	normal	6,00	1,20
$A_{18}$	normal	2,72	1,00
$A_{19}$	normal	3,13	1,10
$A_{20}$	normal	4,01	1,30
$A_{21}$	normal	4,50	1,50

Pour certaines applications, il est parfois difficile de dissocier les paramètres en se référant seulement au graphique basé sur les mesures de sensibilité  $\mu_i^*$  et  $\sigma_i$ . En effet, dans le cas des paramètres  $P_2$ ,  $I_{10}$ ,  $I_{13}$  et  $E_5$ , on ne peut pas affirmer d'une façon définitive si l'aléa qui leur est associé est significatif ou pas sur la variabilité de la réponse mécanique de la structure.



**Figure II. 19 :** Hiérarchisation des paramètres incertains pour le problème de la structure à 5 étages, basée sur les caractéristiques statistiques des effets élémentaires, (a) non normalisées, (b) normalisées

Afin de pallier cette difficulté nous nous proposons d'ajouter un autre critère de sélection des paramètres importants. Dans ce contexte, nous nous basons, en plus des mesures de sensibilité  $\mu_i^*$  et  $\sigma_i$ , sur la mesure de sensibilité globale  $G_i^*$  (globale dans le sens où elle intègre en même temps les mesures de sensibilité  $\mu_i^*$  et  $\sigma_i$ ). Ainsi, si on a des difficultés pour dissocier certains paramètres, les plus importants seront ceux dont le cumul des mesures de sensibilité  $G_i^*$  correspond à un quantile  $p$ , préalablement établi, du cumul total.



**Figure II. 20 :** (a) hiérarchisation des paramètres incertains en se référant à la mesure de sensibilité  $G_i^*$  (b) identification des paramètres les plus importants pour un quantile seuil  $p = 95\%$

Les figures II.20a et II.20b illustrent respectivement la hiérarchisation des 10 premiers paramètres incertains en se référant à la mesure de sensibilité  $G_i^*$ , et l'identification des paramètres les plus importants pour un quantile seuil  $p = 95\%$ . Ce critère donne plus de consistance à la sélection des paramètres importants comme on peut le constater sur la figure

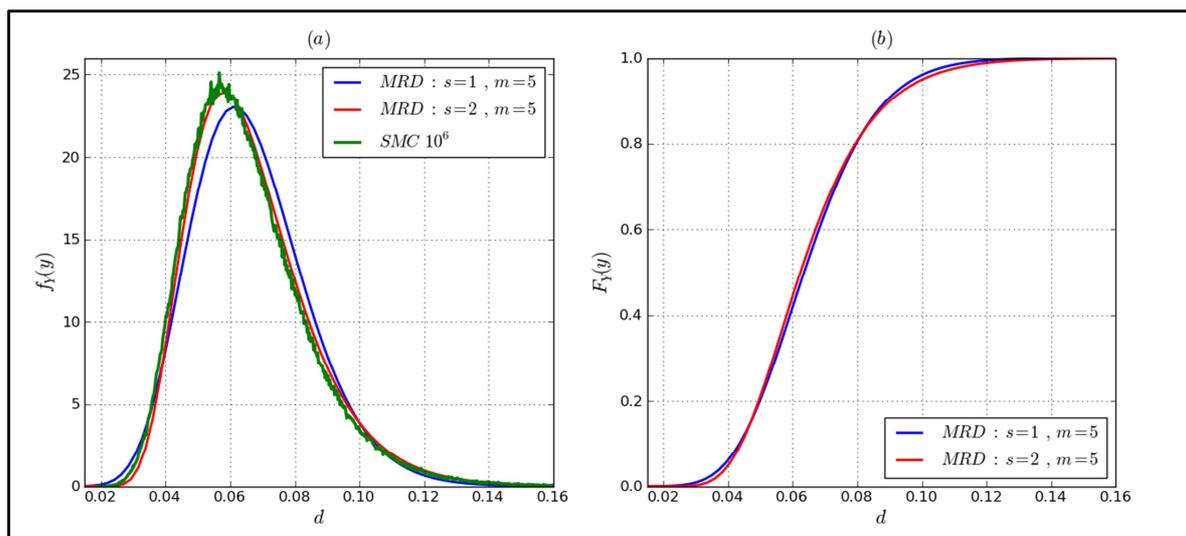
II.19b. En effet, en plus des informations qu'on peut recueillir du graphique de la figure II.19a, ce critère permet de hiérarchiser les paramètres incertains par ordre d'importance croissant. Grâce à cette procédure, la dimension stochastique efficace du problème est 4, ce qui correspond aux quatre premiers paramètres importants  $P_1, I_{11}, I_{12}$  et  $E_4$ . Elle est inférieure à la dimension stochastique nominale (21 paramètres incertains). Le calcul stochastique proprement dit en sera simplifié et rendu plus abordable. Dans ce contexte, nous nous proposons d'évaluer la variabilité de la réponse mécanique de la structure induite par l'aléa associé aux paramètres incertains les plus importants. Ainsi, nous ne modélisons par des variables aléatoires que les paramètres  $P_1, I_{11}, I_{12}$  et  $E_4$ , et nous utilisons les développements en chaos polynômial et la méthode de réduction de la dimension comme méthodes de calcul stochastique.

Nous nous intéressons à la détermination des moments statistiques de la réponse mécanique ainsi qu'à la construction de sa densité de probabilité. Le tableau II.14 donne les estimations des quatre premiers moments statistiques du déplacement horizontal de l'étage supérieur de la structure.

**Tableau II. 14 :** Estimations des moments statistiques – problème de la structure à cinq niveaux

Moments statistiques	s = 1		s = 2		SMC
	m = 4	m = 5	m = 4	m = 5	
$\mu_Y$	0,0651	0,0651	0,0652	0,0652	0,0653
$\sigma_Y$	0,0179	0,0179	0,0187	0,0187	0,0187
$\delta_Y$	0,5527	0,5588	0,9672	0,9741	1,0151
$\kappa_Y$	3,3455	3,4911	4,4226	4,5749	4,9620
Nb d'appels au MEF	17	21	113	171	$10^5$

On constate que la convergence des deux premiers moments statistiques est garantie en utilisant une intégration par réduction de la dimension d'ordre 1. Cependant, la convergence des coefficients d'aplatissement et d'asymétrie nécessite l'augmentation de l'ordre de la réduction à 2. De plus, on constate que les estimations sont en bon accord avec les résultats donnés par les simulations de Monte-Carlo basées sur un échantillon de  $10^5$  points obtenus par un générateur de nombres pseudo-aléatoires.



**Figure II. 21 :** Comparaison des densités de probabilité obtenues par la méthode de Pearson et les simulations de Monte-Carlo

Cette constatation est confirmée par les résultats donnés sur la figure II.21a qui illustre une comparaison des densités de probabilités construites d'une part par la méthode de Pearson basée sur les moments statistiques et d'autre part par les simulations de Monte-Carlo. En effet, la distribution de densité de probabilité construite en se basant sur les estimations des moments statistiques fournies par la méthode de réduction de la dimension d'ordre 2, est en bon accord avec la solution de référence donnée par les simulations de Monte-Carlo. Dans un but illustratif, nous donnons également sur la figure II.21b la fonction de répartition de la réponse mécanique, qui peut être utilisée dans les analyses de fiabilité.

La stratégie à deux étapes, proposée pour effectuer le calcul stochastique, permet de réduire significativement le nombre d'évaluations du modèle mécanique. En effet, en supposant qu'on a considéré dans le calcul des moments statistiques les 21 paramètres incertains de base, en utilisant la méthode de réduction de la dimension d'ordre 2 basée sur une quadrature de Gauss-Hermite d'ordre 5, le nombre d'appels au modèle éléments finis nécessaire pour atteindre la convergence est de 5336, tandis que la stratégie proposée nécessite seulement 391 appels au modèle éléments finis. Ainsi, l'économie en temps de calcul est de l'ordre de 92%.

### II.6 Conclusion

Une des principales sources de difficulté dans la résolution des problèmes de calcul stochastique est la dimension stochastique, définie par le nombre de variables aléatoires utilisées pour représenter l'aléa associé aux paramètres incertains. En effet, plus la dimension stochastique est élevée, plus le temps de calcul pour déterminer les grandeurs d'intérêt est conséquent. Dans ce contexte, nous avons proposé une stratégie de résolution composée de deux étapes. La première consiste à identifier la dimension efficace, représentant le nombre de paramètres importants du modèle, avant d'entamer le calcul stochastique proprement dit. La dimension efficace, qui peut être très inférieure à la dimension nominale dans le cas de problèmes d'ingénierie, est déterminée en se basant sur une hiérarchisation des paramètres incertains par la méthode de criblage de Morris. Dans le cas où les mesures de sensibilités qualitatives fournies par la méthode de Morris ne suffisent pas pour distinguer facilement les paramètres importants, nous avons proposé un critère de sélection basé sur l'évaluation d'une mesure de sensibilité normalisée, dans le sens où son cumul est égal à l'unité. Autrement dit, les paramètres importants sont ceux dont le cumul des mesures de sensibilité normalisées est égal à un quantile préalablement fixé. Une fois les paramètres importants identifiés, ils sont modélisés par des variables aléatoires et les paramètres restant sont fixés à leurs valeurs moyennes ou nominales respectives.

La deuxième étape consiste à effectuer le calcul stochastique proprement dit. Dans ce contexte, nous avons proposé deux approches permettant de construire une représentation explicite (ou méta-modèle) de la réponse d'un modèle ou d'une fonction d'état limite disponible au départ sous forme implicite. Ces deux approches sont basées sur deux techniques distinctes qui sont le développement en chaos polynômial et la décomposition de la dimension de l'espace aléatoire. Nous avons montré qu'une fois le méta-modèle construit, différents post-traitements peuvent être menés pour effectuer des analyses de tendance centrale, des analyses de sensibilité basées sur la décomposition de la variance ou des analyses de fiabilité. Ces post-traitements ne nécessitent pas d'appels supplémentaires au modèle ou à la fonction d'état limite explicite, que ceux nécessaires à la construction du méta-modèle.

En ce qui concerne la méthode de calcul stochastique basée sur le développement en chaos polynômial, nous avons proposé les simulations de Monte-Carlo basées sur les séquences de nombres quasi-aléatoires pour gagner en efficacité lors de la détermination des coefficients

inconnus du développement en chaos polynômial. En nous appuyant sur un exemple illustratif, nous avons montré que la séquence de Halton offre le meilleur compromis entre précision et coût de calcul. Dans la même perspective, nous avons retenu également la quadrature de Stroud et la quadrature basée sur les grilles de Smolyak. Cette dernière permet de réduire considérablement le nombre d'appels au modèle implicite en comparaison à la quadrature classique construite à partir d'un produit tensoriel complet. Cependant, sa précision est étroitement liée à la régularité de l'intégrande. La quadrature de Stroud est la plus économique en terme d'appels au modèle implicite, mais son application n'est préconisée que lorsque les grandeurs à appréhender n'exigent pas un niveau de précision élevé.

La méthode de réduction de la dimension permet de calculer une intégrale multidimensionnelle en la décomposant en une somme d'intégrales uni, bi et/ou tridimensionnelles faciles à calculer par quadrature classique. Elle offre une bonne alternative pour déterminer directement les moments statistiques de la réponse d'un modèle avec précision tout en étant économe en terme d'appels au modèle. En conjonction avec la projection sur la base des polynômes d'interpolation de Lagrange ou une base polynômiale associée aux distributions des paramètres incertains, cette méthode offre la possibilité de fournir une approximation analytique de la réponse du modèle ou de la fonction d'état limite. A partir des coefficients de ce méta-modèle, on peut déterminer directement les moments statistiques ou les indices de sensibilité de Sobol. La probabilité de défaillance peut être estimée, dans le contexte d'analyse de fiabilité, en effectuant des simulations de Monte-Carlo sur le méta-modèle ou par l'intégration de la densité de probabilité associée à la fonction d'état limite. Ces aspects ont été illustrés par l'étude de cas tests, pour lesquels on a montré qu'un bon compromis entre précision et coût de calcul est assuré.

A la fin du chapitre, nous avons illustré l'implémentation de la stratégie proposée pour résoudre le problème de dimension stochastique élevée à travers une application portant sur la résistance d'une structure à cinq niveaux sollicitée par des charges latérales. On a montré que cette stratégie permet une économie de l'ordre 92% du temps de calcul. Cette efficacité est due principalement à la réduction du nombre de paramètres incertains effectuée en identifiant la dimension efficace du problème. Elle sera appliquée dans le chapitre suivant pour évaluer l'effet des incertitudes sur le comportement des toitures à ossatures en bois sollicitées par une action sismique. Le choix entre le développement en chaos polynômial et la méthode de décomposition de la dimension, dans la deuxième étape de la stratégie proposée, revient à l'utilisateur. Il est dicté par la finalité du calcul stochastique effectué. Néanmoins, nous notons tout de même que la méthode de décomposition de la dimension est la plus adaptée en terme d'efficacité, surtout dans le cas des analyses de tendance centrale ou d'analyse de fiabilité.



---

## **Chapitre III : application aux structures bois sous sollicitation sismique**

---

### III.1 Introduction

Dans le chapitre précédent, nous nous sommes focalisés sur une méthodologie de résolution de problèmes à dimension stochastique élevée. Les méthodes de calcul stochastique ont été présentées dans un contexte général sans spécifier le problème mécanique à traiter. Cependant, dans ce travail, nous sommes confrontés à un problème mécanique spécifique portant sur l'analyse du comportement de structures bois sous sollicitation sismique. Le traitement de ce type de problème implique de relever un certain nombre de défis.

Les structures en bois sont généralement constituées d'éléments assemblés ou fixés à d'autres éléments structuraux par des organes métalliques qui peuvent être de simples clous, des pointes, des agrafes, des boulons, des équerrres, ou des plaques métalliques embouties. L'expérience a montré que la modélisation de leur comportement mécanique réel est non triviale. En effet, le comportement de ce type d'assemblage est fortement non linéaire, et ce principalement en raison de la conjonction du comportement orthotrope des éléments en bois et de la ductilité des connecteurs métalliques. Dans ce contexte, divers modèles non linéaires [136-138] visant la description du comportement mécanique de ces assemblages ont été développés au cours des dernières années. Bien qu'ils soient largement utilisés, ces modèles souffrent de certaines limitations puisqu'ils sont basés sur une loi unidimensionnelle avec hystérésis pour décrire la relation force-déplacement. Or, à cause du contact entre les éléments bois lors de charges de compression, une loi unidimensionnelle ne peut pas décrire de manière adéquate le comportement bidimensionnel des assemblages bois. Pour cette raison, un des objectifs majeurs du projet SISBAT [7] était de développer une loi de comportement bidimensionnelle, basée sur les modèles décrits auparavant, permettant de décrire le comportement réel des assemblages bois avec connecteurs métalliques. Cette tâche a été développée au laboratoire 3SR, de l'Université Joseph Fourier, dans le cadre de la thèse de Jérôme Humbert [139].

Par ailleurs, à cause de son origine naturelle, des défauts sous forme d'irrégularités ou de nœuds sont inhérents au matériau bois. Ces défauts induisent une variabilité des propriétés mécaniques du matériau bois telles que le module de Young ou la densité. Aussi, s'il est utilisé dans la construction, il est primordial de prendre en compte l'effet de cet aléa sur le comportement des structures. Dans ce contexte, les propriétés du matériau bois sont représentées par des variables ou champs aléatoires au lieu de les fixer à leurs valeurs moyennes ou nominales respectives. De plus, les essais expérimentaux réalisés dans le cadre du projet SISBAT ont montré que certaines propriétés mécaniques des assemblages bois, ou plus précisément certains paramètres de la loi avec hystérésis utilisée pour la modélisation de leur comportement, sont entachés d'incertitudes. Aussi, faire l'hypothèse que ces paramètres sont des grandeurs déterministes peut mener à des décisions non conservatives voire erronées. Cela peut avoir de graves conséquences quand on conçoit des structures susceptibles d'être installées dans des zones de sismicité élevée. Par conséquent, les méthodes de calculs stochastiques offrent une alternative intéressante pour le traitement rationnel de ces incertitudes. Il s'agit là d'un des principaux objectifs du projet SISBAT et la problématique sur laquelle est articulé ce travail, dont la finalité est non seulement de prendre en compte l'effet des incertitudes dans l'analyse du comportement des structures bois sous sollicitations sismiques, mais aussi de fournir des éléments d'analyse des codes de conception. Comme on l'a déjà évoqué tout au long de ce manuscrit, le choix d'une méthode de calcul stochastique est principalement dicté par le type d'analyse qu'on souhaite effectuer (i.e. analyse de tendance centrale, analyse de fiabilité ou analyse de sensibilité). Cependant, dans ce travail, la contrainte principale dans le choix de la méthode de calcul stochastique est la dimension de l'espace aléatoire (i.e. le nombre de paramètres incertains). En effet, nous souhaitons analyser

le comportement sous l'effet d'une sollicitation sismique de la toiture complète d'une maison individuelle. Cette structure est composée de plusieurs assemblages réalisés principalement par des connecteurs métalliques à dents ou des pointes métalliques, mais aussi avec des équerrres. En intégrant l'aléa associé aux paramètres de la loi avec hystérésis modélisant le comportement de ces assemblages, ainsi que l'aléa associé aux propriétés mécaniques du matériau bois, le nombre de paramètres incertains est trop élevé pour être traité par les méthodes de calcul stochastique classiques telles que celles présentées dans le premier chapitre de ce manuscrit. Afin de surmonter cette difficulté, nous nous proposons d'appliquer la stratégie développée dans le chapitre précédent. Cette dernière est constituée de deux étapes de calcul distinctes, la première consiste à réduire la dimension stochastique en utilisant la méthode de criblage de Morris, et la deuxième consiste à aborder le calcul stochastique proprement dit en ne modélisant par des variables aléatoires que les paramètres significatifs.

Dans ce chapitre, nous nous intéressons dans un premier lieu à la présentation de la loi de comportement bidimensionnelle avec hystérésis, développée par Jérôme Humbert [139] dans le cadre du projet SISBAT, en se basant sur l'analyse d'un modèle éléments finis représentant un assemblage bois élémentaire. En conjonction, nous présentons une analyse de sensibilité basée sur le calcul des indices de Sobol et sur le développement en chaos polynômial pour quantifier l'effet de l'aléa associé aux paramètres de la loi de comportement sur la variabilité de la dissipation de l'énergie et la résistance de l'assemblage bois. Ensuite, des analyses de fiabilité sont menées sur une structure en treillis représentant une ferme en bois sollicitée par une action sismique. De plus, nous présentons une procédure de sélection des signaux sismiques utilisés dans ce travail, en nous référant aux consignes fournies par le BRGM (*Bureau de Recherches Géologiques et Minières*). Enfin, nous présentons l'analyse du comportement d'une toiture complète d'une maison individuelle sollicitée par une action sismique, ainsi que les différents critères adoptés dans le cadre du projet SISBAT pour évaluer la criticité de l'endommagement accumulé dans cette structure.

### **III.2 Modèle éléments finis d'un assemblage bois**

#### **III.2.1 Comportement des assemblages bois par plaques métalliques**

Les plaques métalliques embouties sont souvent utilisées comme éléments d'assemblage dans les fermes en bois présentes dans les toitures des maisons individuelles. Leur abondance est justifiée par le fait qu'elles permettent de réaliser des assemblages robustes et bon marché, en un temps optimal. D'un point de vue conceptuel, les assemblages bois réalisés avec ce type de connecteur sont composés généralement de deux ou plusieurs éléments bois connectés de part et d'autre par des plaques métalliques embouties à dents de tailles égales et orientées de la même façon. Les fermes en bois sont considérées comme des structures en treillis dont la conception est effectuée par rapport aux charges verticales. Ainsi, le comportement mécanique des assemblages bois fabriqués en utilisant des plaques métalliques embouties est souvent représenté par des modèles linéaires. Cependant, des études soutenues par des essais expérimentaux, en particulier sous chargements cycliques [1] ont révélé que le comportement réel de ces assemblages est non linéaire et présente une forte anisotropie. La non linéarité est due à la plastification de la plaque métallique ainsi qu'aux dents assurant l'assemblage des éléments bois. L'anisotropie est induite d'une part par l'anisotropie dans le comportement des plaques métalliques, et d'autre part par le contact entre les éléments bois produisant une certaine dissymétrie. Une fois encore, ces constatations ont été confirmées par les résultats expérimentaux obtenus dans le cadre du projet de recherche SISBAT. De plus, sur l'ensemble des assemblages testés sous chargements monotones et cycliques suivant différentes directions, plusieurs modes de défaillances ont été révélés. Le premier mode de défaillance

consiste en l'arrachement des dents de la plaque métallique des éléments bois (figure III.1a). Ceci est observé lorsque la plaque métallique montre une résistance suffisante alors que la surface d'ancrage est trop limitée pour pouvoir supporter le chargement. En revanche, lorsque l'ancrage est suffisant, on constate la rupture de la plaque métallique, ou la rupture fragile de l'un des éléments bois (figure III.1c). Finalement, sous l'effet de charges de compression, le flambage de la plaque métallique induit un écrasement du bois dû au contact entre les éléments bois de l'assemblage, entre lesquels un espacement est normalement prévu (figure III.1b).



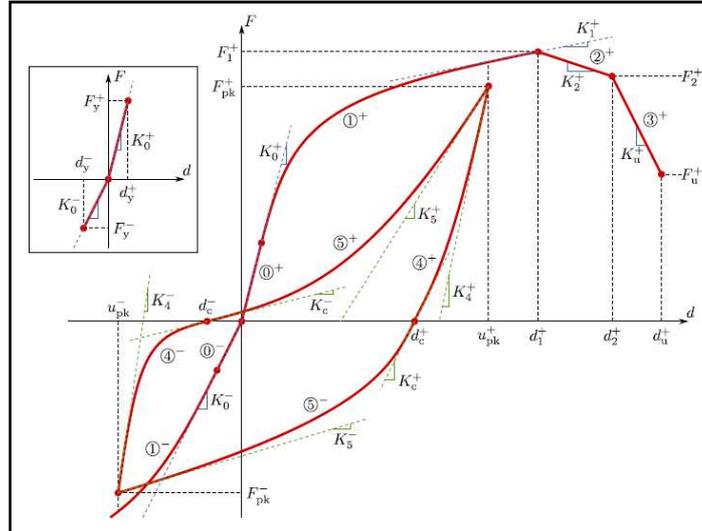
**Figure III. 1 :** Modes de défaillance des assemblages bois par plaques métalliques embouties : (a) arrachement des dents de la plaque du bois, (b) plastification de la plaque, (c) rupture d'un élément bois [140]

En se référant à l'Eurocode 5 [4], les assemblages bois réalisés par des plaques métalliques embouties sont caractérisés par plusieurs paramètres tels que les dimensions de la plaque métallique, son orientation à la fois par rapport à la direction du chargement et la direction des fibres des éléments bois, et les propriétés mécaniques des matériaux. On note de plus que la résistance de la plaque métallique elle-même dépend fortement de la disposition des dents et des trous oblongs qu'elle contient. Ainsi, on définit souvent deux directions suivant lesquelles la plaque métallique a des capacités de résistance distinctes. En effet, la direction suivant laquelle la plaque métallique exhibe une forte résistance correspond à la direction principale des trous oblongs résultants du processus de fabrication des dents. La résistance des assemblages conçus par des plaques métalliques à dents est évaluée, comme prescrit dans l'Eurocode 5, vis-à-vis des deux modes de défaillance : l'ancrage des dents dans les éléments bois et le cisaillement du corps de la plaque métallique. Ces modes de défaillance sont vérifiés en comparant la réponse de l'assemblage vis-à-vis d'un chargement donné à des valeurs caractéristiques des paramètres mécaniques. Ces dernières sont obtenues à partir des essais expérimentaux, et en se référant à la norme EN 1075 [141]. La définition de ces valeurs caractéristiques peut avoir recours à des coefficients de sécurité ou à des modèles probabilistes qui permettent de prendre en compte les incertitudes qui peuvent se manifester.

### III.2.2 Loi anisotrope avec hystérésis

Le comportement des assemblages bois réalisés par des plaques métalliques à dents, comme on l'a déjà évoqué dans le paragraphe précédent, est fortement non linéaire. Ainsi, dans la pratique, on a souvent recours à une modélisation par éléments finis pour décrire d'une façon implicite le comportement de ces assemblages. Afin qu'elle puisse être représentative du comportement réel de l'assemblage, la modélisation par éléments finis doit être très raffinée et prendre en compte les différents détails géométriques et les caractéristiques mécaniques des différents matériaux. Bien que cette alternative paraisse pertinente, elle engendre souvent un temps de calcul prohibitif. Par conséquent, son emploi en conjonction avec une procédure de calcul stochastique n'est pas envisageable. En effet, quand on mène des analyses

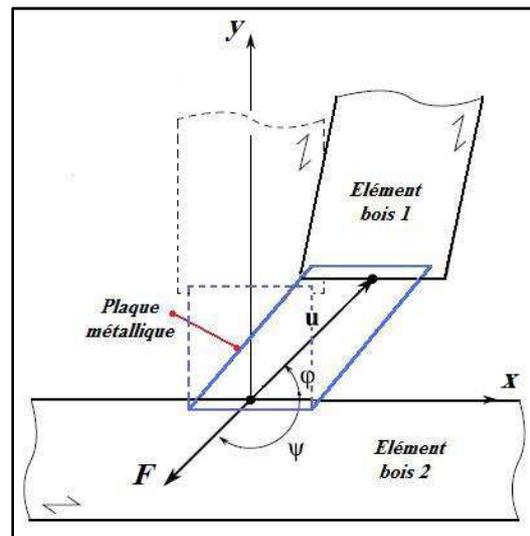
stochastiques, une des préoccupations principales est d'utiliser des modèles mécaniques à la fois robustes et non exigeants en terme de temps de calcul. Pour pallier ce problème, des macro-modèles sont alors employés pour décrire le comportement global de l'assemblage bois. Ces derniers sont basés sur des lois de comportements qui décrivent l'évolution de la force en fonction du déplacement dans l'assemblage. Dans ce contexte, une loi bidimensionnelle avec hystérésis [139], développée par le laboratoire 3S-R, est utilisée dans cette étude. Cette dernière est basée sur une loi unidimensionnelle avec hystérésis développée sur la base du modèle de Yasumura et Yasui [137] et du modèle LMT [138]. La correspondance entre le déplacement scalaire  $d$  et la force résultante  $F$  est donnée par la courbe force-déplacement présentée sur la figure III.2.



**Figure III. 2 :** Loi d'hystérésis unidimensionnelle [139] développée dans le cadre du projet SISBAT

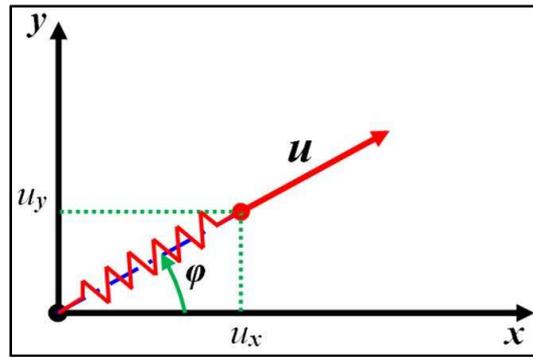
Elle est composée d'une enveloppe monotone (branche 1) jusqu'à la force maximale  $F_1$ , définie à l'aide de polynômes de Bézier, suivie de deux fonctions linéaires (branches 2 et 3) représentant les pertes de rigidité post-pic. L'emploi des polynômes de Bézier permet d'assurer une continuité analytique de la courbe force-déplacement, ce qui n'est pas le cas des modèles proposés dans la littérature, pour lesquels cette continuité est seulement établie d'une façon asymptotique. A partir de la force au pic, les boucles d'hystérésis sont composées d'une décharge élastique (branche 4) jusqu'à  $F = 0$ , suivie d'une recharge (branche 5). Afin d'assurer la continuité, le point de départ de la branche de recharge et le point de fin de la branche de décharge sont les mêmes. Les coordonnées  $u_{pk}$  et  $F_{pk}$  du point pic peuvent avoir différentes valeurs dépendant de l'histoire du chargement. Ainsi, la taille de la boucle d'hystérésis formée par les branches de recharge et décharge peut être variable. La rigidité  $K_4$  peut être liée au comportement élastique, et, pour certains types d'assemblage, elle peut être égale à la rigidité initiale  $K_0$ . Le déplacement critique  $d_c$  représente la limite de la décharge élastique non linéaire. On note que les boucles d'hystérésis sont délimitées d'une part par l'axe des déplacements, et d'autre par une courbe enveloppe obtenue par une homothétie de rapport  $\alpha \times \beta$  sur la courbe monotone correspondante en traction ou en compression. En effet, les résultats des essais expérimentaux ont montré que la courbe monotone n'est pas une enveloppe des essais cycliques, mais, pour un cycle donné, elle peut être une enveloppe de ce dernier en effectuant une homothétie sur la force. Cette loi unidimensionnelle avec hystérésis permet de prendre en compte la perte de rigidité après le pic, l'évolution de l'endommagement de l'assemblage durant des cycles de déplacement à amplitude constante et de modéliser des boucles cycliques de tailles variables. En effet, les coefficients  $\alpha$  et  $\beta$ , définissant le rapport de l'homothétie, représentent en réalité l'évolution de

l'endommagement au sein de l'assemblage. Ils sont obtenus à partir des résultats des essais expérimentaux. Le coefficient  $\alpha$  représente un endommagement modéré piloté par le déplacement maximal des cycles, tandis que le coefficient  $\beta$  correspond à un endommagement significatif activé après le pic d'effort et piloté à son tour par le déplacement. Cependant, à cause du contact qui peut se manifester entre les éléments bois ainsi que des différentes directions que peut avoir le chargement, cette loi unidimensionnelle avec hystérésis ne permet pas de décrire le comportement réel (anisotrope) de l'assemblage. Ainsi, en se basant sur l'hypothèse d'un modèle ressort stipulant que la force est colinéaire au déplacement, cette loi unidimensionnelle avec hystérésis a été généralisée [139] pour le cas bidimensionnel, en considérant un assemblage composé par deux éléments bois connectés de part et d'autre par deux plaques métalliques à dents (figure III.3).



**Figure III. 3 :** Définition du déplacement relatif  $u$ , de la force résultante  $F$  et du plan d'anisotropie  $(x, y)$

La direction  $y$ , définie comme la normale au plan de contact entre les deux éléments bois, est parallèle aux plans contenant les plaques métalliques. Cette configuration est réaliste puisque les éléments bois ont des sections rectangulaires de même épaisseur. L'orientation de la direction  $y$  est conventionnellement choisie pour correspondre à la traction. Le contact entre les éléments bois n'aura lieu que si le déplacement relatif suivant la direction  $y$  est négatif. La direction  $x$  est définie par la tangente au plan de contact et elle est toujours parallèle aux plans des plaques métalliques. Ainsi, le plan d'anisotropie est défini par les directions  $x$  et  $y$ . Le déplacement relatif  $u$  entre les deux éléments bois est orienté par l'angle  $\varphi$  qui représente la direction du chargement par rapport à la direction  $x$ . La loi bidimensionnelle avec hystérésis est définie comme un ressort non linéaire pour lequel le vecteur force  $F$  et le vecteur déplacement  $u$  sont colinéaires dans le plan d'anisotropie. Autrement dit, la loi bidimensionnelle avec hystérésis n'est autre que l'application de la loi unidimensionnelle donnée par la figure III.2 dans la direction définie par l'angle  $\varphi$ . Ainsi, la force résultante pour un déplacement donné est facilement obtenue en utilisant la loi unidimensionnelle dont les paramètres sont déterminés par une interpolation par rapport à l'angle  $\varphi$  dans les deux directions  $x$  et  $y$ . La figure III.4 illustre le modèle ressort utilisé dans la modélisation du comportement anisotrope des assemblages bois.

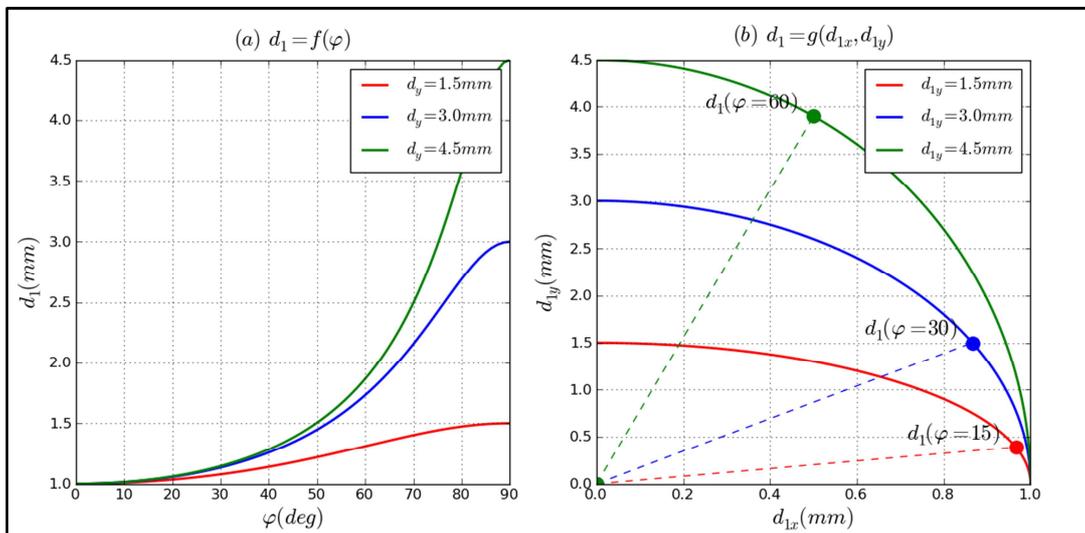


**Figure III. 4 :** *Modèle ressort utilisé dans la modélisation du comportement anisotrope des assemblages bois*

Le déplacement  $u$  du ressort avec hystérésis se déplaçant dans le plan d'anisotropie  $(x, y)$  suivant la direction définie par l'angle  $\varphi$  est obtenu comme la norme des déplacements  $u_x$  et  $u_y$ . Les déplacements limites  $d_y$ ,  $d_1$ ,  $d_2$  et  $d_u$ , représentant les paramètres de la loi de unidimensionnelle modélisant le comportement du ressort, sont obtenus à partir de leurs équivalents dans les directions  $x$  et  $y$  à l'aide d'une interpolation elliptique de la forme suivante :

$$d_y = d_{i,x} d_{i,y} \sqrt{d_{i,x}^2 \cos(\varphi)^2 + d_{i,y}^2 \sin(\varphi)^2}, \quad i = \{y, 1, 2, u\} \quad (III. 1)$$

Les figures III.5a et III.5b illustrent respectivement l'évolution du déplacement au pic  $d_1$  en fonction de l'angle de chargement  $\varphi$  et pour différentes combinaisons des déplacements au pic  $d_{1,x}$  et  $d_{1,y}$ , définis respectivement dans les directions  $x$  et  $y$ .



**Figure III. 5 :** *Evolution du déplacement au pic  $d_1$  en fonction (a) de l'angle de chargement et (b) des déplacements au pic équivalent dans le plan d'anisotropie  $(x, y)$*

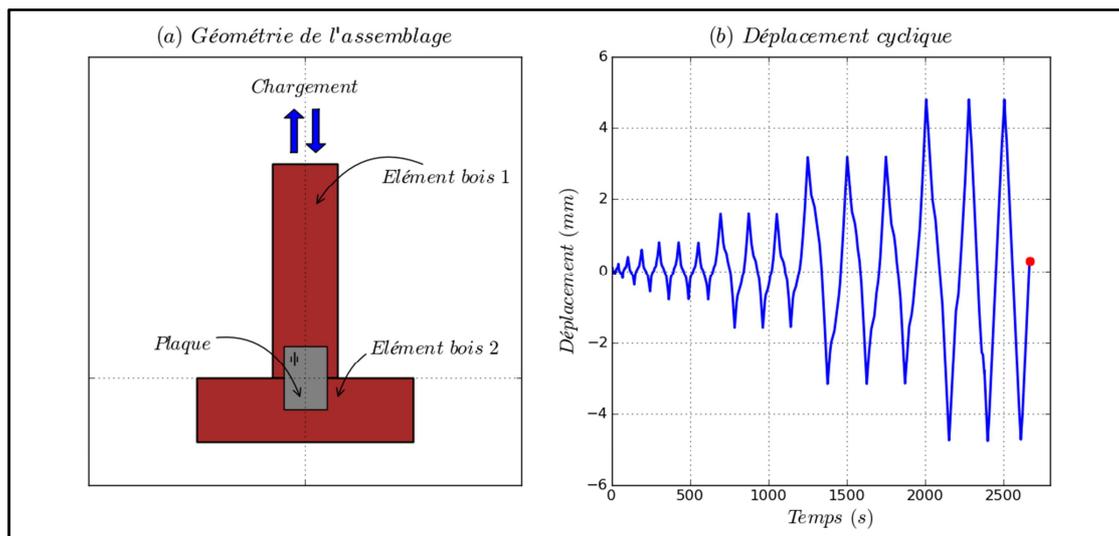
Le choix d'une interpolation elliptique par rapport à d'autres types d'interpolation (en particulier l'interpolation linéaire) est soutenu par le fait qu'elle fournit des résultats compatibles avec la physique, plus précisément avec les observations expérimentales. On note que, parmi toutes les forces définissant la courbe force-déplacement, seule la force au pic  $F_1$  est à interpoler. De plus, certains déplacements tels que  $u_{pk}$  doivent être enregistrés pour l'incrément de temps suivant. Etant donné que les paramètres sont interpolés à partir des paramètres équivalents dans les deux directions  $x$  et  $y$  du plan d'anisotropie, il faut extrapoler

les paramètres équivalents de  $u_{pk}$  dans ces directions à partir de sa valeur suivant la direction de déplacement du ressort avec hystérésis.

Bien que cette loi anisotrope avec hystérésis soit présentée ici pour la modélisation du comportement des assemblages bois avec plaques métalliques à dents, notons qu'elle peut être utilisée également dans la modélisation du comportement des assemblages bois réalisés avec d'autres types de connecteurs métalliques, tels que les clous, les pointes ou les équerres, que nous les aborderons plus tard dans ce chapitre.

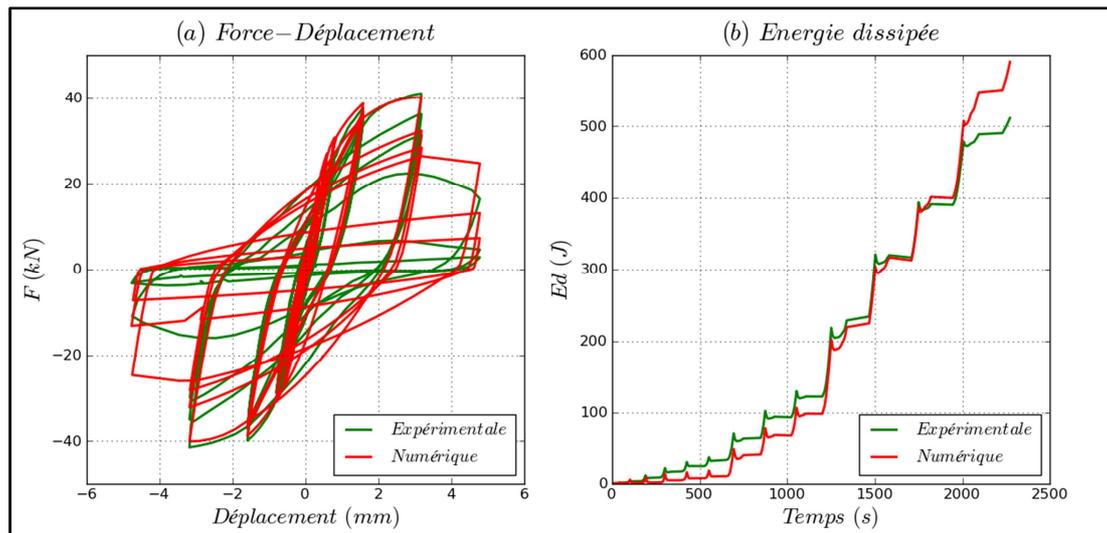
### III.2.3 Analyse du comportement d'un assemblage bois

Un modèle numérique représentant un assemblage constitué de deux éléments bois connectés par une plaque métallique à dents est implémenté dans un code de calcul par éléments finis. Le maillage représentant l'assemblage est constitué de deux nœuds, dont un est encastré et l'autre libre en déplacement seulement suivant la direction  $x$ . Le comportement du connecteur est modélisé par la loi avec hystérésis présentée dans le paragraphe précédent, dont les paramètres sont déterminés à partir des résultats expérimentaux collectés sur des assemblages bois sollicités par des chargements monotones et cycliques. Afin de valider le modèle éléments finis, on se propose de confronter les prévisions numériques aux résultats expérimentaux obtenus à partir d'un essai réalisé sur un assemblage bois sollicité en déplacement cyclique. La figure III.6 illustre la configuration géométrique de l'assemblage testé, ainsi que le chargement qui lui est appliqué.



**Figure III. 6 :** (a) Configuration géométrique de l'assemblage testé, (b) évolution du chargement

La validation de la loi unidimensionnelle est basée sur deux critères : le tracé de la courbe force-déplacement et l'évolution du cumul de l'énergie dissipée par l'assemblage durant l'histoire du chargement. On note que l'énergie dissipée est définie par le produit de l'incrément du déplacement par la force résultante. La figure III.7 illustre une comparaison entre les prévisions données par le modèle éléments finis et les résultats expérimentaux concernant la courbe force-déplacement et l'évolution du cumul de l'énergie dissipée au cours de la durée du chargement. On constate que la courbe force-déplacement est en bon accord avec les résultats collectés durant les essais expérimentaux, en particulier pour les déplacements de faibles amplitudes avant d'atteindre le pic (cf. figure III.2). De plus, on constate que le modèle éléments finis permet de prédire avec une grande précision la force maximale, ainsi que la charge cyclique pour un déplacement de faible amplitude.



**Figure III. 7 :** Comparaison entre les prévisions données par le modèle éléments finis et les résultats expérimentaux : la courbe force-déplacement et évolution du cumul de l'énergie dissipée

La comparaison des résultats concernant l'évolution du cumul d'énergie dissipée au cours du chargement, montre que, globalement, l'écart entre les résultats numériques et expérimentaux n'est pas significatif. Cependant, il faut signaler que, lorsque le niveau de dégradation de l'assemblage est élevé, le modèle éléments finis tend à surestimer la dissipation de l'énergie à la fin de la durée du chargement. Suite aux différentes constatations présentées ci-dessus, on conclut que la loi décrite dans le paragraphe précédent permet de prédire le comportement réel des assemblages bois par connecteurs métalliques à dents.

### III.3 Analyse stochastique d'un assemblage bois

#### III.3.1 Etude de sensibilité

La loi avec hystérésis présentée dans le paragraphe 2 pour modéliser le comportement semi-rigide d'un assemblage bois dépend de plusieurs paramètres qui sont déterminés à partir de résultats expérimentaux. L'analyse statistique des données collectées à partir des essais effectuées sur différents assemblages bois [7] a révélé l'existence d'un aléa sur certains paramètres de la loi avec hystérésis. Par conséquent, il est nécessaire d'effectuer une analyse de sensibilité pour déterminer les paramètres les plus influents sur la réponse mécanique de l'assemblage, telle que sa capacité résistante (i.e. la force maximale que l'assemblage peut supporter) et la dissipation d'énergie. Le tableau III.1 donne les caractéristiques statistiques (i.e. moyenne, écart-type et coefficient de variation) des différents paramètres incertains. Nous rappelons que nous étudions l'assemblage bois représenté par la figure III.2, sollicité par les conditions de chargement présentées dans le paragraphe précédent.

**Tableau III. 1 :** Caractéristiques statistiques des paramètres incertains [139]

	$K_0$ [kN/mm]	$d_y$ [mm]	$d_1$ [mm]	$F_1$ [kN]	$d_2$ [mm]	$d_u$ [mm]
$\mu$	57,11	0,32	3,15	45,67	8,10	11,02
$\sigma$	5,7	0,043	0,5	2,8	2,9	3,3
CV (%)	10,0	13,4	15,7	6,2	35,7	30,0

D'une façon générale, l'analyse de sensibilité est composée de deux étapes principales. La première consiste à choisir le modèle probabiliste permettant de représenter l'aléa associé aux

paramètres entachés d'incertitudes. Autrement dit, pour chaque paramètre incertain, il faut définir la distribution de la variable aléatoire qui le modélise. Cette tâche nécessite un nombre relativement important de données expérimentales, ce qui n'est pas le cas dans notre travail pour lequel le nombre d'essais est limité. Nous sommes donc obligés de poser quelques hypothèses. Afin de simplifier les calculs, il est commode de modéliser les paramètres incertains par des distributions Gaussiennes. Cependant, quand on manipule des grandeurs physiques telles que le module de Young ou les rigidités, ce choix peut induire des incompatibilités (i.e. valeurs négatives). Afin de surmonter ce problème, les paramètres incertains dans ce travail sont modélisés par des variables aléatoires lognormales ayant les caractéristiques statistiques données dans le tableau III.1. Ce choix est conforme à la littérature [132], puisque les grandeurs représentant la résistance et les propriétés mécaniques des matériaux sont bien représentées par des lois lognormales.

La deuxième étape consiste à déterminer les paramètres les plus influents sur la réponse mécanique. La façon la plus pertinente pour aborder ce genre de problème est d'utiliser une méthode permettant de décomposer la variance de la réponse mécanique en fonction des variances associées aux paramètres incertains. Dans ce contexte, la méthode de décomposition de Sobol peut être utilisée. Nous avons utilisé l'approche développée dans le chapitre précédent, qui consiste, dans un premier temps, à construire une représentation explicite du modèle mécanique basée sur un développement en chaos polynômial et de déduire par la suite les indices de sensibilité à partir des coefficients de ce développement en se référant aux équations (II.75) et (II.77). Le tableau III.2 donne les estimations des indices de Sobol de premier ordre et totaux associés à chaque paramètre incertain, en considérant comme réponses mécaniques la capacité résistante de l'assemblage bois  $F_{max}$  et la dissipation d'énergie  $E_d$ . Ces estimations ont été obtenues en considérant un développement en chaos polynômial d'ordre  $p = 3$ , et les coefficients  $\{a_\alpha, |\alpha| \leq p\}$  de ce développement ont été déterminés en évaluant l'intégrale  $N$ -dimensionnelle  $I_\alpha^N$  définie par l'équation (II.19) avec la formule de quadrature de Smolyak (Eq. (II.37)) d'ordre  $l = 6$ .

**Tableau III. 2 :** Indices de Sobol de premier ordre  $S$  et totaux  $S_T$  associés aux paramètres incertains de l'assemblage bois pour un déplacement appliqué inférieur à  $d_y$

	$F_{max}(N)$		$E_d(J)$	
	$S$	$S_T$	$S$	$S_T$
$K_0$	0,663	0,781	0,416	0,450
$d_y$	0,114	0,161	0,081	0,120
$d_1$	0,046	0,086	0,116	0,184
$F_1$	0,064	0,097	0,271	0,350
$d_2$	0	0	0	0
$d_u$	0	0	0	0

En se référant aux indices de Sobol totaux ( $S_T$ ), on constate que la rigidité initiale  $K_0$  et le déplacement limite  $d_y$  (i.e. déplacement délimitant le comportement élastique de l'assemblage à partir duquel l'endommagement se manifeste) sont les paramètres les plus influents sur la capacité résistante  $F_{max}$  de l'assemblage. En effet, une variation de 10% de la rigidité initiale contribue à 78% de la variance de la capacité résistante de l'assemblage, tandis qu'une variation de 13,4% du déplacement limite  $y$  contribue seulement à 16%. L'aléa associé aux autres paramètres incertains ne contribue qu'à 2% de la variance de la capacité résistante de l'assemblage. En comparant les valeurs des indices de sensibilité du premier ordre et les indices de sensibilité totaux, on constate qu'il existe des interactions entre certains paramètres de la loi avec hystérésis. Ceci peut être confirmé par l'évaluation des indices de Sobol du second ordre. En effet, l'indice de Sobol du second ordre associé à la rigidité initiale  $K_0$  et le

déplacement limite  $d_y$  est d'environ 4%, ce qui montre l'existence d'une interaction négligeable entre ces deux paramètres.

La variance de la dissipation d'énergie  $E_d$  est principalement due à l'aléa de la rigidité initiale  $K_0$ . En effet, l'indice de Sobol total qui lui est associé indique que la variation de 10% de ce paramètre contribue à 45% de la dispersion de la dissipation d'énergie. La force au pic  $F_1$  vient en second ordre, puisque l'aléa qui lui est associé contribue uniquement à 35% de la variance de la dissipation d'énergie. L'importance de ces deux paramètres peut être expliquée par le fait qu'ils contrôlent la forme des boucles d'hystérésis. Les paramètres  $d_y$  et  $d_1$  de la loi ont un effet non significatif sur la dissipation d'énergie puisque les deux ensembles contribuent à environ 20% de la variabilité de cette dernière. En comparant les indices de Sobol du premier ordre et les indices totaux, on conclut que l'effet de l'interaction entre les différents paramètres incertains n'est pas négligeable. On note que l'étude de sensibilité qui vient d'être présentée a été effectuée pour un déplacement appliqué inférieur au déplacement limite élastique  $d_y$  (i.e. la comparaison est effectuée par rapport à la valeur moyenne). Cela peut expliquer la dépendance de la capacité résistante de l'assemblage (i.e. la force maximale enregistrée dans l'assemblage) du déplacement limite élastique  $d_y$  et justifie aussi la faible dissipation d'énergie, observée durant les quelques premiers cycles du chargement (cf. figure III.7). En effet, l'endommagement s'amorce dans l'assemblage lorsque le déplacement appliqué est supérieur à  $d_y$ .

Aussi, on se propose de refaire cette analyse de sensibilité pour un déplacement appliqué supérieur à  $d_y$ , afin de voir si la contribution de l'aléa associé aux différents paramètres incertains dans la variabilité de la capacité résistante de l'assemblage et de l'énergie dissipée change de distribution. Le tableau III.3 donne les indices de Sobol du premier ordre et totaux respectifs aux paramètres incertains.

**Tableau III. 3 :** Indices de Sobol de premier ordre  $S$  et totaux  $S_T$  associés aux paramètres incertains de l'assemblage bois pour un déplacement appliqué supérieur à  $d_y$

	$F_{max}(N)$		$E_d(J)$	
	$S$	$S_T$	$S$	$S_T$
$K_0$	0,279	0,284	0,786	0,861
$d_y$	0,060	0,064	0,042	0,066
$d_1$	0,126	0,133	0,031	0,057
$F_1$	0,524	0,528	0,038	0,116
$d_2$	0	0	0	0
$d_u$	0	0	0	0

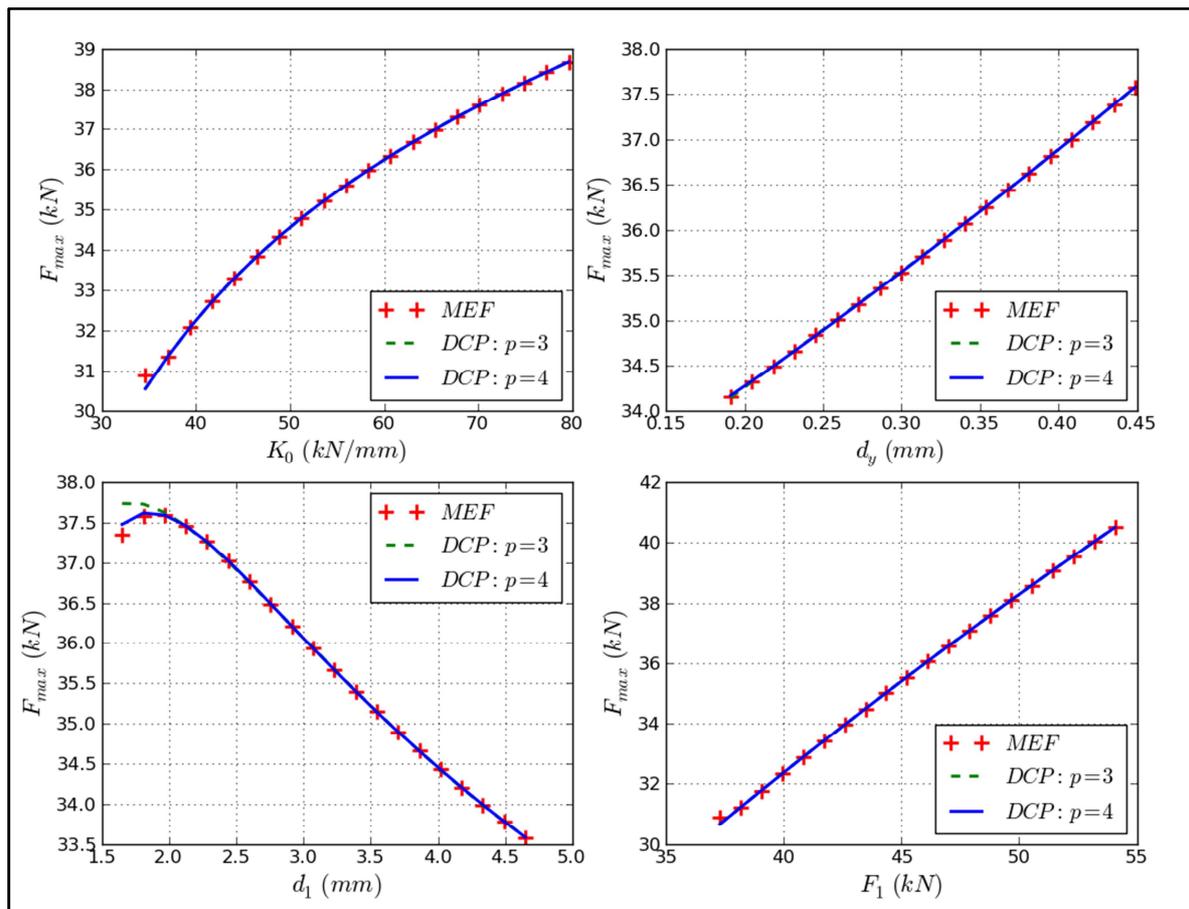
Les résultats donnés dans le tableau III.3 confirment le changement de l'ordre d'importance des paramètres incertains sur la variabilité des réponses mécaniques. Ceci est encore plus vrai pour la capacité résistante de l'assemblage. En effet, en se référant aux indices de Sobol totaux, on constate que la variabilité de la force maximale enregistrée dans l'assemblage est principalement due à l'aléa associé à la force au pic  $F_1$ . Or, on a montré précédemment que la rigidité initiale de l'assemblage  $K_0$  jouait ce rôle, pour un déplacement appliqué inférieur au déplacement à la limite élastique  $d_y$ . De plus, on constate que l'effet de l'aléa associé au déplacement au pic  $d_1$  est plus visible. En effet, une fois le déplacement limite élastique dépassé, les paramètres  $F_1$  et  $d_1$  commandent l'évolution de l'endommagement dans l'assemblage. Ces deux paramètres, comme on l'a déjà vu lors de la présentation de la loi avec hystérésis, permettent une certaine flexibilité du modèle afin qu'il puisse s'adapter à l'évolution de l'endommagement. En comparant les indices de Sobol du premier ordre et les indices de Sobol totaux, on constate que l'écart est non significatif, ce qui montre que l'effet

d'interaction entre les paramètres incertains est faible. Si nous nous intéressons maintenant à la dissipation de l'énergie, nous constatons que sa variabilité est induite exclusivement par l'aléa associé à la rigidité initiale  $K_0$  de l'assemblage, ce qui peut être expliqué par le fait que le taux d'endommagement dans l'assemblage n'a pas atteint un seuil critique.

Afin de réduire le nombre de paramètres incertains dans les analyses stochastiques que nous souhaitons mener dans la suite de notre travail, notamment dans le cas de la toiture à ossature bois sollicitée par une action sismique, nous ne modéliserons à l'échelle d'un assemblage élémentaire que l'aléa associé à la force au pic  $F_1$  puisqu'il a l'effet le plus significatif sur son intégrité. Les autres paramètres de la loi de comportement avec hystérésis seront fixés à leurs valeurs moyennes.

### III.3.2 Evaluation de l'aléa de la réponse de l'assemblage

Dans le paragraphe précédent, nous avons effectué une étude de sensibilité, basée sur la notion de décomposition de la variance, qui nous a permis d'identifier les paramètres importants sur la capacité résistante de l'assemblage bois et sur l'énergie dissipée. Cependant, l'aléa associé à ces réponses mécaniques n'est pas évalué ; autrement dit, nous n'avons pas déterminé leurs caractéristiques statistiques (i.e. moments statistiques et densité de probabilité). Pour ce faire, nous nous intéressons seulement à la capacité résistante de l'assemblage, et nous nous proposons de calculer ses premiers moments statistiques et de construire sa densité de probabilité, au moyen de l'approche basée sur le développement en chaos polynômial présenté dans le paragraphe II.3.5.1 du chapitre précédent.



**Figure III. 8 :** Evolution de l'approximation de la capacité résistante de l'assemblage bois par un développement en chaos polynômial en fonction des différents paramètres incertains

En se référant à l'étude de sensibilité effectuée précédemment, on ne modélise que la rigidité initiale  $K_0$ , la force au pic  $F_1$ , le déplacement limite  $d_y$  et le déplacement au pic  $d_1$ , par des variables aléatoires. Les déplacements limites  $d_2$  et  $d_u$  sont fixés à leurs valeurs moyennes respectives. Les quatre premiers moments statistiques sont déterminés respectivement par les équations (II.59 à II.62). La densité de probabilité est construite à partir de  $10^6$  simulations de Monte-Carlo effectuées sur le modèle explicite basé sur un développement en chaos polynômial d'ordre  $p = \{3,4\}$ , dont les coefficients sont déterminés par quadrature de Smolyak.

Nous nous intéressons dans un premier lieu à la convergence du modèle explicite fourni par le développement en chaos polynômial. Dans ce contexte, la figure III.8 compare le modèle obtenu pour différents ordres du développement en chaos polynômial avec le modèle éléments finis.

On note que les graphes tracés sur la figure III.8 sont obtenus en faisant varier les paramètres incertains un par un dans un intervalle  $[\mu_{X_i} \pm 3\sigma_{X_i}]$  (i.e.  $\mu_{X_i}$  et  $\sigma_{X_i}$  sont respectivement la moyenne et l'écart-type associés au paramètre  $X_i$ ). On constate que le développement en chaos polynômial d'ordre  $p = 4$  reproduit avec une grande précision la réponse mécanique obtenue par le modèle éléments finis. Ainsi, il peut être couplé aux simulations de Monte-Carlo pour effectuer des calculs stochastiques tels que l'analyse de fiabilité.

Si nous nous intéressons maintenant à la convergence des moments statistiques obtenus à partir des coefficients du développement en chaos polynômial, les résultats obtenus, ainsi que le nombre d'appels au modèle éléments finis, pour les ordres 3 et 4 du développement en chaos polynômial, sont donnés dans le tableau III.4.

**Tableau III. 4 :** Estimations des moments statistiques de la capacité résistante de l'assemblage

Ordre du DCP	Moments statistiques				Nombre d'appels au MEF
	$\hat{\mu}_{Y,p}$	$\hat{\sigma}_{Y,p}$	$\hat{\delta}_{Y,p}$	$\hat{\kappa}_{Y,p}$	
$p = 3$	35693	2251,5	0,1283	3,0337	85
$p = 4$	35693	2251,4	0,1285	3,0362	170

On constate que l'écart entre les estimations des moments statistiques fournies par un développement en chaos polynômial respectivement d'ordre  $p = 3$  et  $p = 4$  est de l'ordre de  $10^{-3}$ . Ainsi, on peut conclure que la convergence des moments statistiques est garantie par un développement en chaos polynômial d'ordre 3. Cette grande précision est obtenue à un faible coût de calcul puisqu'elle ne nécessite que 85 appels au modèle éléments finis. La densité de probabilité est construite par deux méthodes : la première consiste à effectuer  $10^6$  simulations de Monte-Carlo sur le développement en chaos polynômial et la deuxième consiste à utiliser une méthode des moments de Pearson. Les figures III.9a et III.9b illustrent respectivement la densité de probabilité  $f_{F_{max}}$  et la distribution de probabilité  $F_{F_{max}}$  de la capacité résistante  $F_{max}$  de l'assemblage bois.

On constate que les résultats donnés par les deux méthodes de construction de la densité de probabilité sont en bon accord. De plus, on remarque que l'aléa associé à la capacité résistante de l'assemblage peut être représenté par une variable Gaussienne de moyenne  $\mu_{F_{max}} = 35693$  et d'écart-type  $\sigma_{F_{max}} = 2251,4$ . Ce modèle probabiliste (variable Gaussienne) peut être expliqué par le fait que les paramètres d'entrée sont indépendants et que les sources d'aléa ont différentes origines. Cela est très intéressant si on suggère d'effectuer des analyses de fiabilité puisqu'elle peut être déterminée directement par l'intégration de la densité de probabilité. En effet, si on considère un scénario de défaillance défini par la fonction d'état limite :

$$G(X) = F_{critique} - F_{max}(X) \quad (III.2)$$

où  $F_{critique}$  est une grandeur déterministe représentant l'effort critique que l'assemblage peut supporter, l'évolution de la probabilité de défaillance n'est autre que la fonction de répartition de la capacité résistante de l'assemblage représentée sur la figure III.9b.

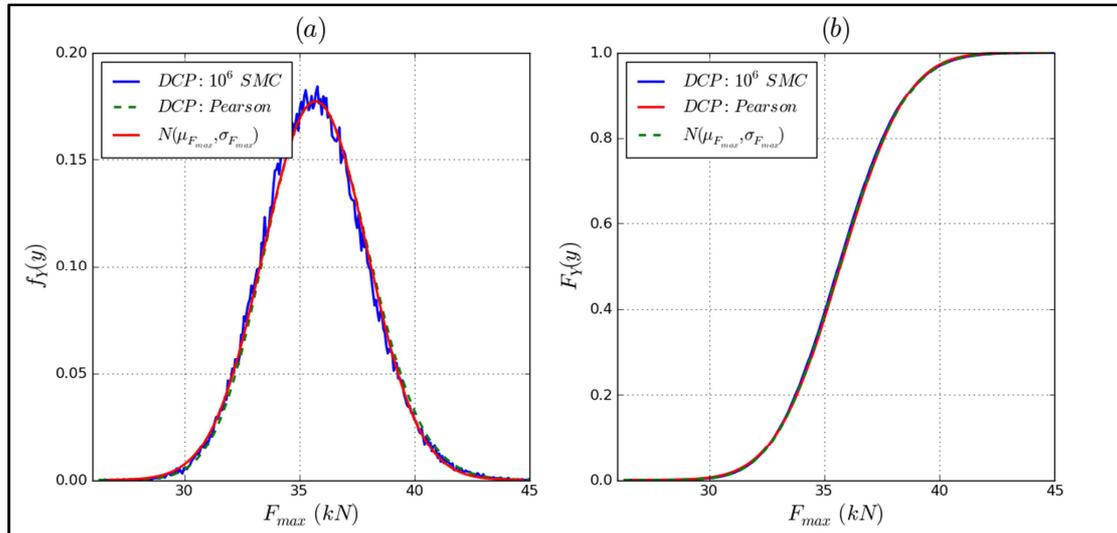


Figure III.9 : Comparaison des évolutions de (a) la densité de probabilité et (b) la répartition de probabilité

### III.4 Analyse de fiabilité d'une fermette en bois

#### III.4.1 Comportement des fermettes en bois

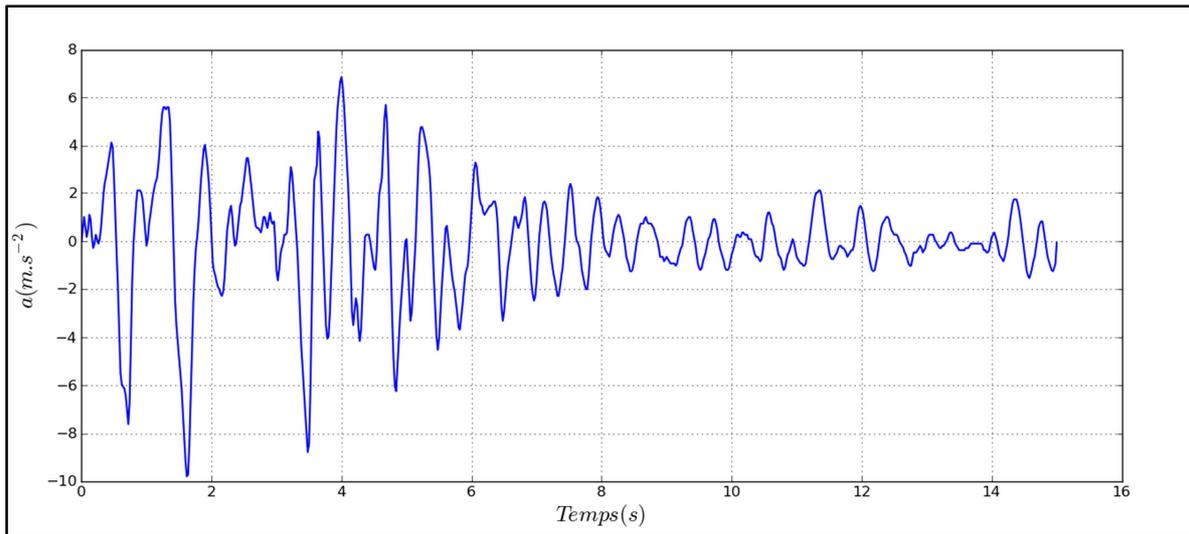
Les fermettes en bois sont utilisées dans la construction de toitures des habitats individuels. D'un point de vue conceptuel, ces structures sont constituées d'assemblages bois réalisés par des plaques métalliques embouties à dents. Les fermettes en bois sont généralement considérées comme des structures en treillis dont la résistance est vérifiée uniquement vis-à-vis des charges verticales. Dans ce contexte, on s'intéresse uniquement au comportement global de la fermette pour laquelle les éléments en bois sont supposés subir des charges de traction ou de compression le long de leurs axes respectifs. Cette hypothèse est loin d'être réaliste. D'une part, dans le cas de charges latérales telles que les séismes, des moments de flexion et des efforts de cisaillement peuvent se manifester dans les éléments bois. D'autre part, comme on l'a montré dans le paragraphe précédent, le comportement des assemblages bois réalisés par des plaques métalliques embouties à dents est fortement non linéaire, ce qui contredit l'utilisation courante de modèles linéaires. De plus, on a vu que ce type d'assemblage permet de dissiper l'énergie lors du chargement appliqué. Bien que cette énergie dissipée puisse améliorer le comportement de la structure maîtresse (i.e. fermette en bois ou toiture en bois d'une habitation individuelle), en particulier vis-à-vis des chargements cycliques tels que les séismes, elle n'est pas prise en compte d'une manière explicite par la réglementation. La plupart des études existantes dans la littérature concernant les structures bois sont focalisées sur le comportement de ces dernières vis-à-vis des charges verticales. De plus, peu d'études s'intéressent à l'analyse des effets des incertitudes sur le comportement des structures bois. Néanmoins, on peut citer quelques tentatives pour intégrer les différentes sources d'incertitudes. En effet, Rosowsky [142] a proposé une approche stochastique pour effectuer dans un premier temps des analyses de sensibilité pour évaluer la contribution de plusieurs sources d'incertitudes sur la réponse mécanique de murs de contreventement en bois caractérisée par le déplacement maximal enregistré durant le passage d'un séisme. Ensuite, la

distribution du déplacement maximal est construite à partir de la réponse des structures bois vis-à-vis de plusieurs accélérogrammes synthétiques utilisés pour caractériser l'aléa sismique en Californie. Enfin, en se basant sur le calcul de la probabilité de défaillance vis-à-vis du franchissement d'un seuil critique de déplacement, il a proposé des abaques pour concevoir les murs de contreventement en bois. Dans le même contexte, van de Lindt et Walz [143] ont étudié l'effet des incertitudes qui entachent les paramètres de la loi de comportement avec hystérésis utilisée pour la modélisation du comportement des murs de contreventement. La distribution du déplacement maximal, enregistré durant l'application de signaux sismiques représentant l'aléa sismique de plusieurs villes des Etats-Unis, a été approximée par une distribution de Weibull. Cette dernière a été utilisée pour l'estimation de la probabilité de défaillance ainsi que les indices de fiabilité qui lui sont associés, pour différents seuils critiques. En se référant aux estimations de la probabilité de défaillance, ils ont montré que l'aléa associé aux paramètres de la loi de comportement avec hystérésis est significatif sur l'intégrité des structures bois installées sur des sites à Seattle et Los Angeles. Par contre, son effet est négligeable sur la fiabilité des structures bois installées sur des sites à Boston. Gupta et Gebremedhin [144] ont étudié plusieurs types d'assemblage bois en prenant en compte la variabilité de la résistance du matériau bois. Ils ont montré que la fiabilité d'un assemblage bois est étroitement liée à sa rigidité. Dans une autre étude [145], les mêmes auteurs se sont focalisés sur la construction de la distribution de la résistance d'une structure bois en treillis vis-à-vis des charges induites par la neige. Le critère de défaillance est représenté par la rupture des éléments bois. Hansson et Ellegaard [146] ont utilisé les simulations de Monte-Carlo pour l'analyse de la fiabilité de fermettes bois contenant des assemblages bois réalisés par des plaques métalliques embouties à dents. Ils ont montré que 10 à 20% des défaillances de ce type de structures bois sont induites par la rupture des éléments bois, tandis que la rupture des connecteurs métalliques contribue à 10% des défaillances. Song et Lam [147] ont utilisé la méthode de surface de réponse quadratique pour analyser la fiabilité d'une toiture d'un habitat individuel composées de quinze fermettes en bois, vis-à-vis de charges horizontales appliquées verticalement aux plans des fermettes. L'aléa associé aux propriétés des matériaux, à la géométrie de la toiture et au chargement est modélisé par une distribution de Weibull. En se basant sur un scénario de défaillance représentant le flambage des éléments bois centraux sur lesquels sont fixées les fermettes constituant la toiture, ils ont proposé un facteur de stabilité, pour caractériser la rigidité latérale de la toiture en bois. De plus, ils ont montré que l'intégrité de cette structure peut être améliorée en ajoutant des éléments de contreventement.

Dans ce qui suit, nous proposons une approche alternative basée sur le développement en chaos polynômial, en conjonction avec les méthodes de fiabilité du premier ordre et les simulations de Monte-Carlo, pour l'analyse de fiabilité d'une fermette en bois sollicitée par une action sismique. Nous nous intéressons en particulier à la défaillance induite par la rupture des assemblages bois présents dans la fermette. Mais, avant de procéder à cette étude, nous nous focalisons dans un premier lieu sur la présentation du modèle éléments finis de la fermette en bois.

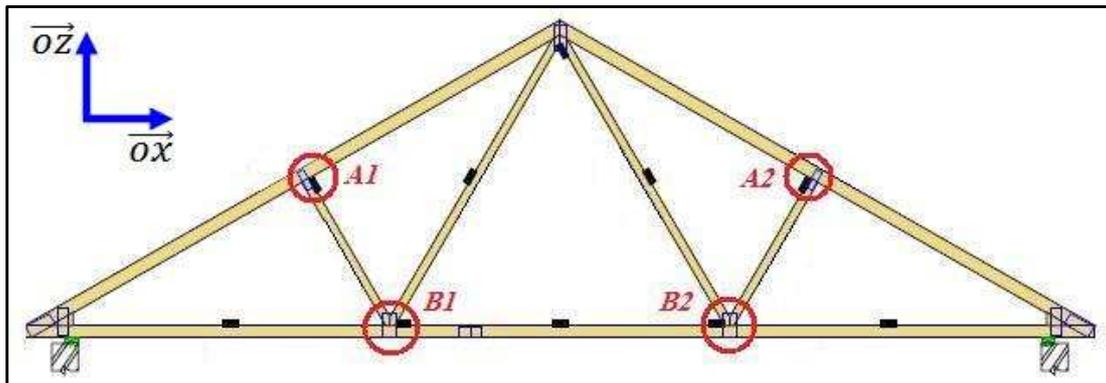
### **III.4.2 Modèle éléments finis de la fermette en bois**

Nous nous proposons d'étudier le comportement d'une fermette en bois sollicitée par une action sismique représentée par l'accélérogramme enregistré lors du séisme qui a eu lieu le 17 janvier 1995 à Kobe au Japon [148] (figure III.10).



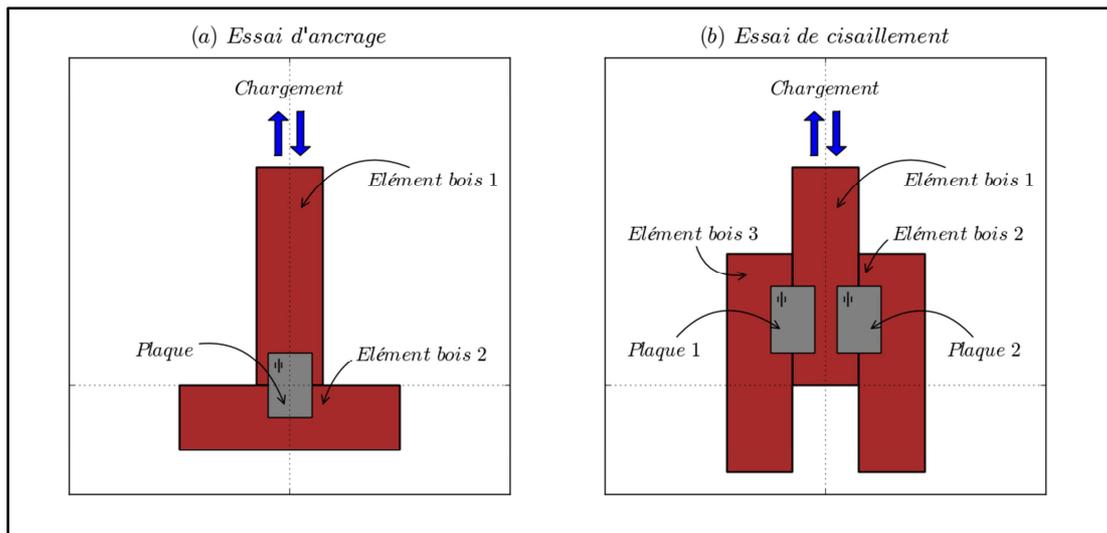
**Figure III. 10 :** Accélérogramme enregistré lors du séisme de Kobe au Japon [148]

La ferme est composée de plusieurs assemblages par plaques métalliques à dents comme le montre la figure III.11. Le modèle éléments finis représentant la ferme est composé d'éléments poutres modélisant les éléments en bois. Le comportement de chacun des assemblages constituant la ferme est modélisé par la loi anisotrope avec hystérésis présentée auparavant, dont les paramètres sont déterminés à partir des résultats des essais expérimentaux réalisés sur des assemblages bois élémentaires.



**Figure III. 11 :** Géométrie de la ferme en bois

Cependant, puisque les différentes configurations géométriques des assemblages dans la ferme ne sont pas toutes représentées dans les essais expérimentaux, un facteur d'ajustement est appliqué pour prendre en compte toutes les configurations. Dans ce contexte, on note que seules quatre configurations d'assemblages élémentaires étudiées expérimentalement ont été utilisées pour déterminer les paramètres de la loi anisotrope avec hystérésis. Elles correspondent respectivement à deux essais de cisaillement à 0° et à 45°, et deux essais d'ancrage à 0° et à 45°. Les figures III.12a et III.12b illustrent les géométries des assemblages bois élémentaires utilisés pour les essais d'ancrage et les essais de cisaillement.



**Figure III. 12 :** Géométrie des assemblages bois utilisés pour (a) les essais d'ancrage et (b) les essais de cisaillements

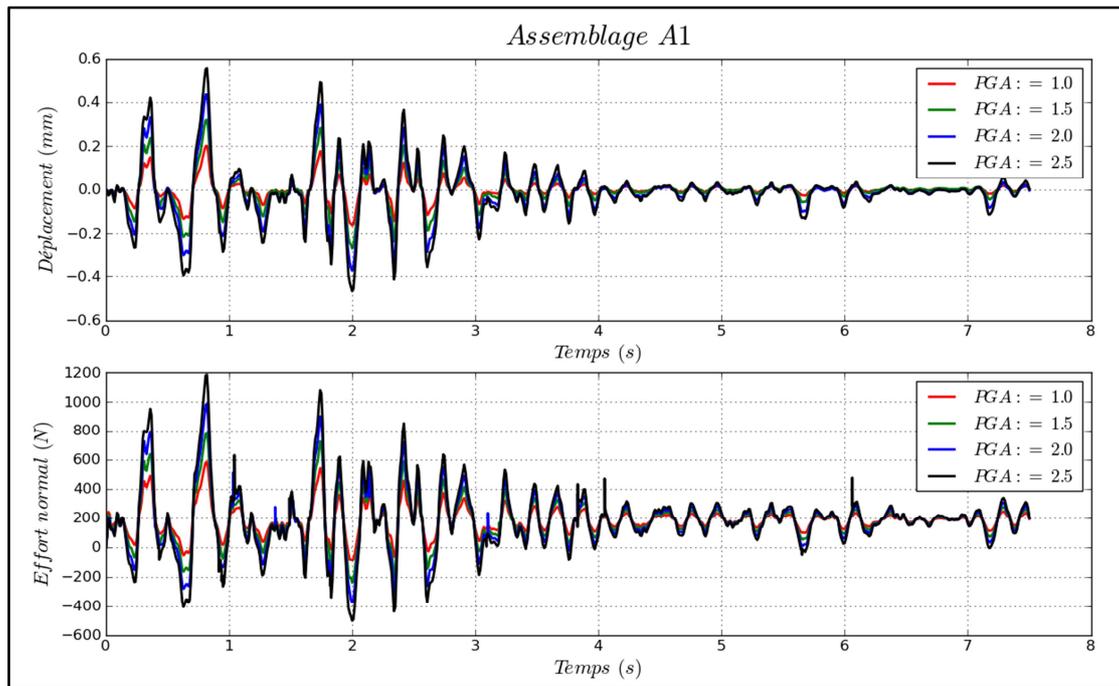
On suppose que la non linéarité est due principalement au comportement des plaques métalliques utilisées dans les assemblages. Ainsi, le comportement des éléments bois est supposé élastique linéaire. De plus, puisque les éléments bois sont élancés, les charges qui leur sont appliquées restent colinéaires à leurs directions principales et les efforts de cisaillement dans le bois sont négligeables. Par conséquent, le bois est considéré comme un matériau isotrope dans la modélisation. Les essais expérimentaux sous chargements cycliques, réalisés dans le cadre du projet SISBAT sur des fermettes en bois utilisées dans les toitures à comble perdu des habitations individuelles, ont révélés deux modes de défaillance : le premier consiste à la rupture du bois au niveau des assemblages réalisés entre les arbalétriers et les contrefiches (assemblages A1 et A2 sur la figure III.11), tandis que le deuxième mode de défaillance consiste à l'arrachement de l'ancrage au niveau des assemblages, en bas de la fermette, entre les fiches et l'entrait (assemblages B1 et B2). Ces deux modes de défaillance sont illustrés sur la figure III.13.



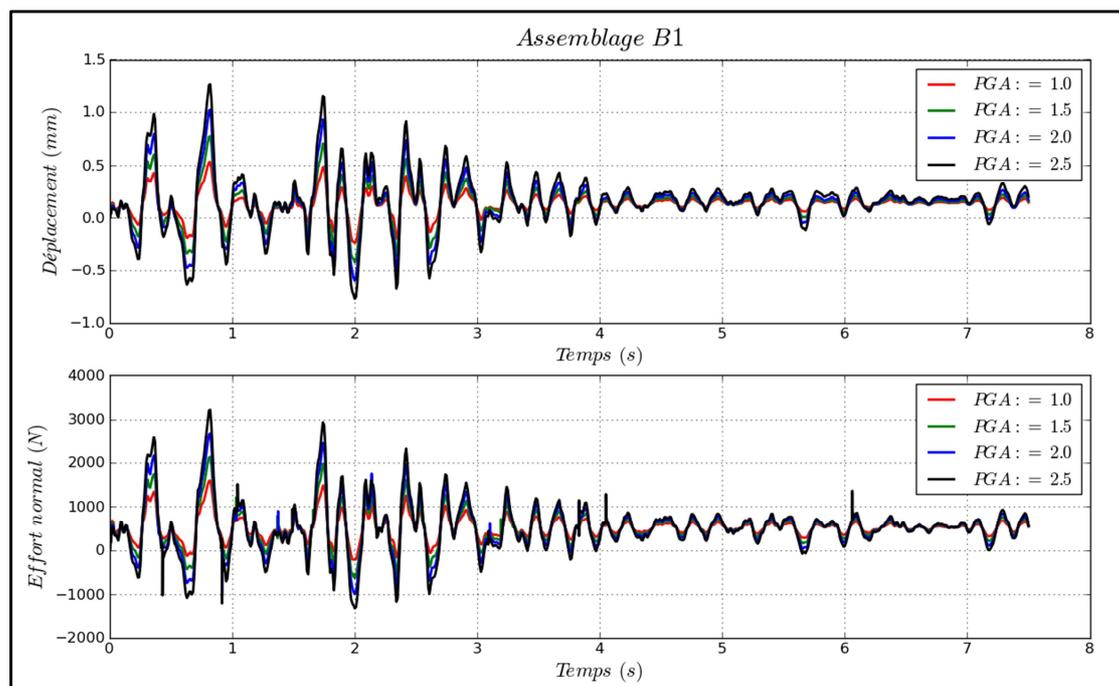
**Figure III. 13 :** Modes de défaillances observés au niveau des assemblages de la fermette en bois [139]

La rupture du bois au niveau des arbalétriers de la fermette est principalement induite par les charges de traction perpendiculaire aux fibres. Ce scénario de défaillance est en bon accord avec les observations recueillies lors des expérimentations sur des assemblages bois élémentaires. Elles ont révélé la faible résistance des assemblages vis-à-vis des charges de traction perpendiculaire aux fibres, et ce quelle que soit l'étendue de la surface d'ancrage (i.e.

les dimensions de la plaque métallique à dents utilisée dans l'assemblage). L'arrachement de l'ancrage est dû probablement à la surface limitée d'ancrage, puisque dans les assemblages *B1* et *B2*, plus de deux éléments bois sont connectés les uns aux autres tandis que des plaques métalliques de mêmes dimensions sont utilisées dans les différents assemblages qui constituaient la fermette. Aussi, nous nous focalisons dans ce qui suit sur l'étude du comportement des assemblages au sein desquels un des modes de défaillance est observé. Les figures III.14 et III.15 illustrent l'évolution du déplacement global et l'effort normal enregistrés au droit des assemblages *A1* et *B1* respectivement, pour différentes amplifications du signal séismique.



**Figure III. 14 :** Evolutions du déplacement global et de l'effort normal dans l'assemblage *A1*



**Figure III. 15 :** Evolutions du déplacement global et de l'effort normal dans l'assemblage *B1*

En se référant aux valeurs maximales du déplacement global et de l'effort normal, on constate que l'assemblage *BI* est plus sollicité que l'assemblage *AI*. A titre illustratif, si on considère une amplification de 150% de l'amplitude de l'action sismique, l'effort maximal enregistré dans l'assemblage *BI* est 2129,43 *N*, alors que l'assemblage *AI* encaisse seulement un effort maximal de 782,2 *N*. On note que le calcul dynamique est effectué à l'aide d'une procédure incrémentale, et que l'action sismique est appliquée dans la direction  $\vec{OX}$  (cf. figure III.11).

### III.4.3 Evaluation de l'intégrité de la fermette en bois

Dans ce paragraphe, nous nous proposons d'évaluer l'intégrité de la fermette en bois sollicitée par une action sismique vis-à-vis des deux modes de défaillance présentés auparavant. Pour ce faire, nous nous intéressons seulement à la réponse mécanique des assemblages *AI* et *BI* représentée par l'effort maximal enregistré durant l'action séismique. Ainsi, la défaillance de la fermette est définie par le franchissement de l'effort maximal enregistré aux assemblages *AI* ou *BI* d'une valeur critique  $F_{crit}$ . La figure III.16 illustre d'une façon schématique ce scénario de défaillance.

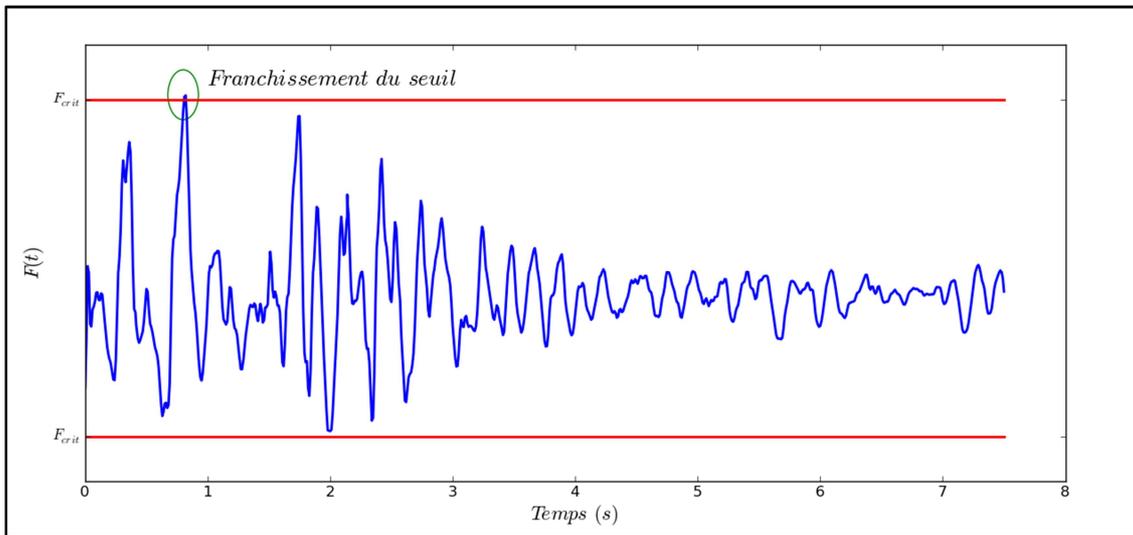


Figure III. 16 : Illustration schématique du critère de défaillance de la fermette en bois

Mathématiquement parlant, ces deux scénarios de défaillance peuvent être représentés par la fonction d'état limite suivante :

$$G(\mathbf{X}) = F_{crit} - \max_{t \in [0, T]} \{|F(\mathbf{X}; t)|\} \quad (III. 3)$$

où  $F_{crit}$  est l'effort critique représentant la résistance de l'assemblage, défini dans cette étude comme une grandeur déterministe, et  $\mathbf{X}$  est un vecteur représentant les paramètres incertains.

Grâce à l'étude de sensibilité d'un assemblage bois élémentaire présentée dans le paragraphe 3.1, nous avons montré que parmi les paramètres incertains représentant la loi anisotrope avec hystérésis utilisée pour la modélisation du comportement de l'assemblage, l'aléa associé à la force au pic  $F_1$  contribue en majeure partie à la variabilité de la capacité résistante de l'assemblage (i.e. l'effort maximal atteint durant l'action sismique). Par conséquent, seul ce paramètre est supposé incertain dans l'analyse de fiabilité de la fermette en bois. Comme on l'a déjà évoqué auparavant, les paramètres déterministes de la loi de comportement avec hystérésis des différents assemblages constituant la fermette sont dérivés à partir des quatre configurations de référence étudiées expérimentalement. A leur tour, ces dernières peuvent être réparties en deux catégories principales : les assemblages utilisés pour les essais de cisaillement et les assemblages utilisés pour les essais d'ancrage (cf. figure III.12). Pour la

première catégorie, la même valeur de la force au pic  $F_1$  de la loi de comportement avec hystérésis est utilisée pour les chargements de traction et de compression mais en changeant seulement de signe (i.e. positif pour les charges de traction et négatif pour les charges de compression). Ainsi, on affecte la même variable aléatoire pour représenter l'aléa associé à la force au pic  $F_1$  à la fois en traction et en compression pour les assemblages de la fermette dont les paramètres de la loi de comportement avec hystérésis étaient déterminés à partir d'une configuration d'assemblage testé en cisaillement. Pour la deuxième catégorie, en plus du signe, les valeurs attribuées à la force au pic  $F_1$ , respectivement en traction et en compression, sont différentes. Ainsi, afin de réduire la dimension de l'espace aléatoire, nous supposons que la force au pic en traction  $F_1^+$  et la force au pic en compression  $F_1^-$  sont totalement corrélées. Ainsi, la variable aléatoire  $X_{F_1^-}$  représentant l'aléa associé à cette dernière (i.e. la force au pic en compression  $F_1^-$ ) est définie par  $X_{F_1^-} = X_{F_1^+} + \Delta F_1$ , où  $\Delta F_1$  est une grandeur déterministe et  $X_{F_1^+}$  est une variable aléatoire représentant l'aléa associé à la force au pic en traction  $F_1^+$ . On note que cette hypothèse simplificatrice ne doit pas être considérée comme une limite de notre étude. En effet, puisque nous partons de la même configuration pour déterminer les caractéristiques mécaniques d'un assemblage en traction ou en compression, ces dernières peuvent être différentes mais elles sont totalement corrélées. Les différents paramètres incertains sont modélisés par des variables aléatoires lognormales dont les caractéristiques statistiques sont données dans le tableau III.5.

**Tableau III. 5 :** *Caractéristiques statistiques des paramètres incertains pour l'analyse de fiabilité d'une fermette en bois sous sollicitation sismique-rupture des assemblages [149]*

Type d'assemblage	Essai de cisaillement		Essais d'ancrage	
	$\alpha = 45^\circ$	$\alpha = 90^\circ$	$\alpha = 0^\circ$	$\alpha = 45^\circ$
Paramètre	$F_{1,C}^{\alpha=45^\circ}$	$F_{1,C}^{\alpha=90^\circ}$	$F_{1,A}^{\alpha=0^\circ}$	$F_{1,A}^{\alpha=45^\circ}$
$\mu$	$48,48 \cdot 10^3$	$43,79 \cdot 10^3$	$33,10 \cdot 10^3$	$40,18 \cdot 10^3$
CV	12%	12%	12%	12%

On note que  $\alpha$  désigne l'angle de la direction principale de la plaque métallique à dents par rapport à la direction des fibres. Les valeurs moyennes de la force au pic  $F_1$  données dans le tableau III.5 correspondent aux charges de traction. Pour les assemblages d'ancrage, les réalisations de la force au pic en compression sont obtenues à partir de leurs réciproques en traction en ajoutant un incrément  $\Delta F_1 = 1950 \text{ N}$  pour  $\alpha = 0^\circ$  et un incrément  $\Delta F_1 = 12620 \text{ N}$  pour  $\alpha = 45^\circ$  et en inversant le signe. Pour les assemblages de cisaillement, on utilise les mêmes réalisations de la force au pic en traction et en compression mais en inversant le signe. On note que les valeurs des incréments  $\Delta F_1$  correspondent à l'écart entre les valeurs déterministes  $F_1^+$  et  $F_1^-$  de la force au pic respectivement en traction et en compression. Ainsi, pour chaque assemblage, les deux variables aléatoires représentant la force au pic en traction et la force au pic en compression sont totalement corrélées.

L'analyse de la fiabilité de la fermette vis-à-vis des deux modes de défaillance est effectuée par une approche basée sur le développement en chaos polynômial. En effet, dans un premier lieu, une formulation explicite de la fonction d'état limite définie par l'équation (III.3) est construite à partir d'un développement en chaos polynômial d'ordre 3. Ensuite cette dernière est couplée avec la méthode de fiabilité du premier ordre ou avec les simulations de Monte-Carlo pour estimer la probabilité de défaillance. On note que les deux modes de défaillance (i.e. rupture de l'assemblage *AI* ou de l'assemblage *BI*) sont étudiés séparément. Autrement dit, la fermette en bois est considérée comme un composant et non pas comme un système. Les résultats obtenus sont donnés dans le tableau III.6.

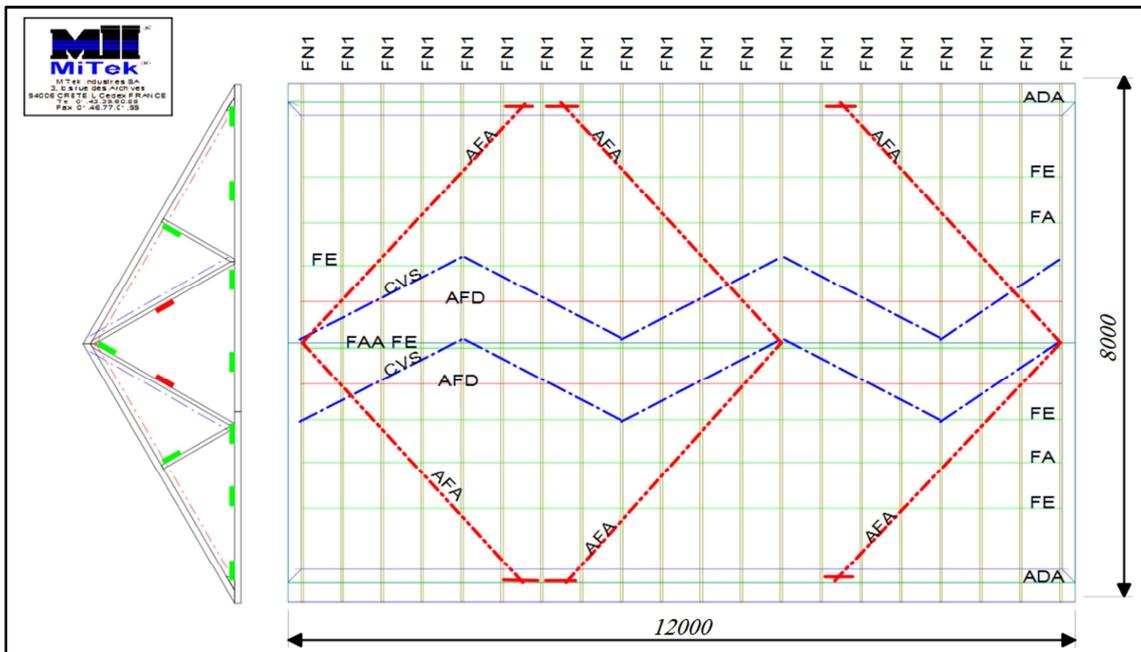
**Tableau III. 6 : Résultats de l'analyse de la fiabilité d'une fermette en bois sollicitée par une action sismique**

	Assemblage A1					Assemblage B1			
$\beta_{FORM}$	3,765					2,130			
$P_{f,FORM}$	$8,335 \cdot 10^{-5}$					$1,658 \cdot 10^{-2}$			
$P_{f,SMC}$	$1,5 \cdot 10^{-4}$					$1,99 \cdot 10^{-2}$			
$x^*$	$F_{1,C}^{\alpha=45^\circ}$	$F_{1,C}^{\alpha=90^\circ}$	$F_{1,A}^{\alpha=0^\circ}$	$F_{1,A}^{\alpha=45^\circ}$	$F_{1,C}^{\alpha=45^\circ}$	$F_{1,C}^{\alpha=90^\circ}$	$F_{1,A}^{\alpha=0^\circ}$	$F_{1,A}^{\alpha=45^\circ}$	
	41956,57	35345,2	22930,04	35696,13	45344,87	38016,27	27564,22	35696,12	

On note que les analyses de fiabilité sont effectuées pour  $F_{crit} = 783 \text{ N}$  et  $F_{crit} = 2130 \text{ N}$  respectivement pour l'assemblage A1 et l'assemblage B1. On constate que la défaillance de la fermette en bois est fort probablement induite par la défaillance de l'assemblage B1 puisque la probabilité de défaillance qui lui est associée est plus élevée que celle associée à l'assemblage A1. De plus, on constate que les estimations de la probabilité de défaillance obtenues par simulations de Monte-Carlo appliquées sur la surface de réponse construite par développement en chaos polynômial, sont plus élevées que celles fournies par la méthode de fiabilité de premier ordre (FORM). En effet, la fonction d'état limite définie par l'équation (III.3) est loin d'être linéaire par rapport aux paramètres incertains. Ainsi l'estimation de la probabilité de défaillance donnée par la méthode FORM a de fortes chances de ne pas correspondre à la solution exacte puisque l'idée de base de cette méthode est de substituer la fonction d'état limite réelle par un hyperplan qui lui est tangent au point de défaillance le plus probable  $x^*$ .

### III.5 Evaluation de l'endommagement d'une toiture à ossature bois

#### III.5.1 Modèle éléments finis de la toiture



**Figure III. 17 : Géométrie et répartition des fermettes de la toiture à combles perdus étudiée dans le cadre du projet SISBAT [7]**

Dans le cadre du projet SISBAT, il est demandé d'effectuer des calculs stochastiques (i.e. analyse de tendance centrale, analyse de sensibilité ou analyse de fiabilité) afin de cerner

l'effet des différentes sources d'incertitudes sur le comportement d'une habitation individuelle installée dans un territoire à forte sismicité. Dans cette perspective, un modèle éléments finis, développé sur le code éléments finis *Code\_Aster*, nous a été fourni par le LMT [159], partenaire du projet SISBAT. Il permet de simuler le comportement dynamique non linéaire d'une toiture à ossature bois de dimensions réelles sollicitée par une action sismique. Notre tâche consiste à y intégrer les différentes sources d'incertitudes et de quantifier leurs effets sur l'évolution de l'endommagement, en utilisant la stratégie de calcul stochastique développée dans le chapitre précédent. Mais avant d'entamer cette tâche, nous nous proposons, dans ce paragraphe, de présenter brièvement le modèle éléments finis de la charpente d'une toiture à combles perdus de dimensions  $8m \times 12m$ , dont la figure III.17 illustre la géométrie et la répartition des fermettes.

Notons que le modèle éléments finis de la toiture à l'échelle réelle (cf. figure III.17) a été validé en comparant les estimations fournies par le modèle aux résultats des essais effectués sur des maquettes de dimensions réduites  $6m \times 6m$  dont la géométrie est illustrée sur la figure III.18.

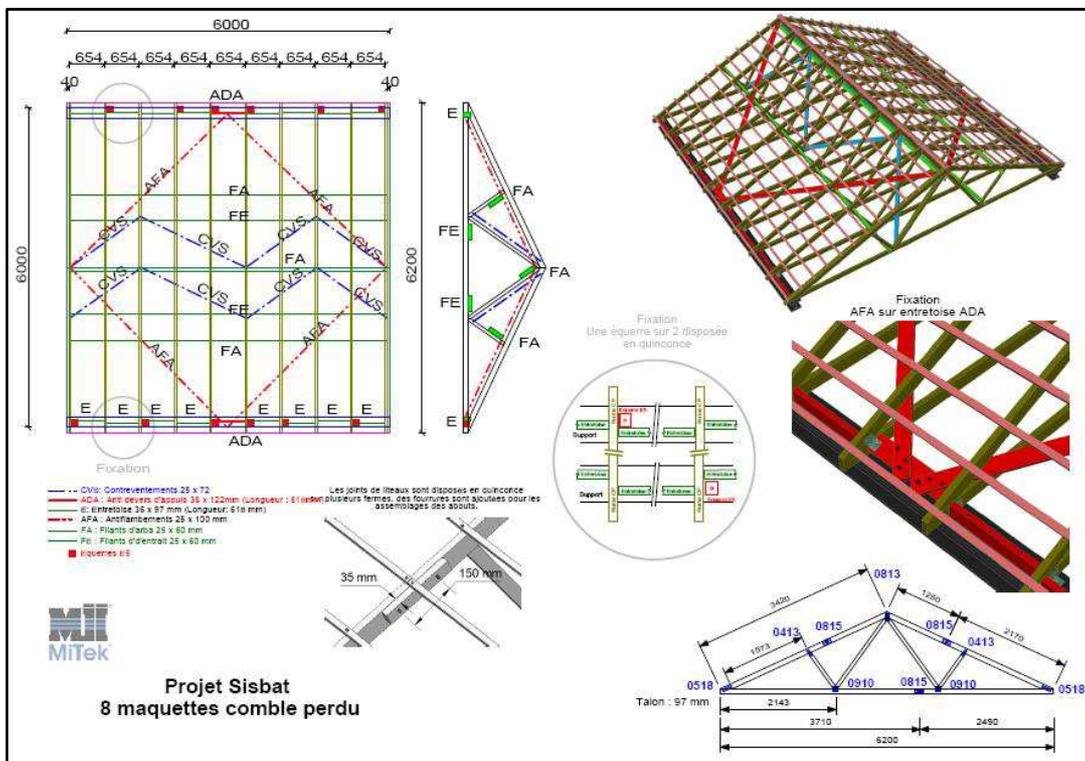


Figure III. 18 : Géométrie de la maquette de toiture à combles perdus testée dans le cadre du projet SISBAT [7]

Afin d'être le plus représentatif possible du comportement réel d'une toiture installée au dessus d'une maison individuelle, le modèle éléments finis prend en compte les chaînages représentés par des poutres en béton armé de section  $0,15m \times 0,15m$ , ainsi que les poutres au vent représentées par un treillis en bois dont la géométrie et les dimensions sont données sur la figure III.19.

Cette modélisation est motivée par le fait que la rigidité du support sur lequel est installée la toiture influence le comportement de cette dernière. On note que les poutres en béton des chaînages sont supposées articulées entre elles, et ne forment pas ainsi un cadre rectangulaire rigide. Les poutres au vent sont fixées d'une part aux poutres de chaînage par des équerres métalliques et d'autre part aux entrails des fermettes par des pointes.

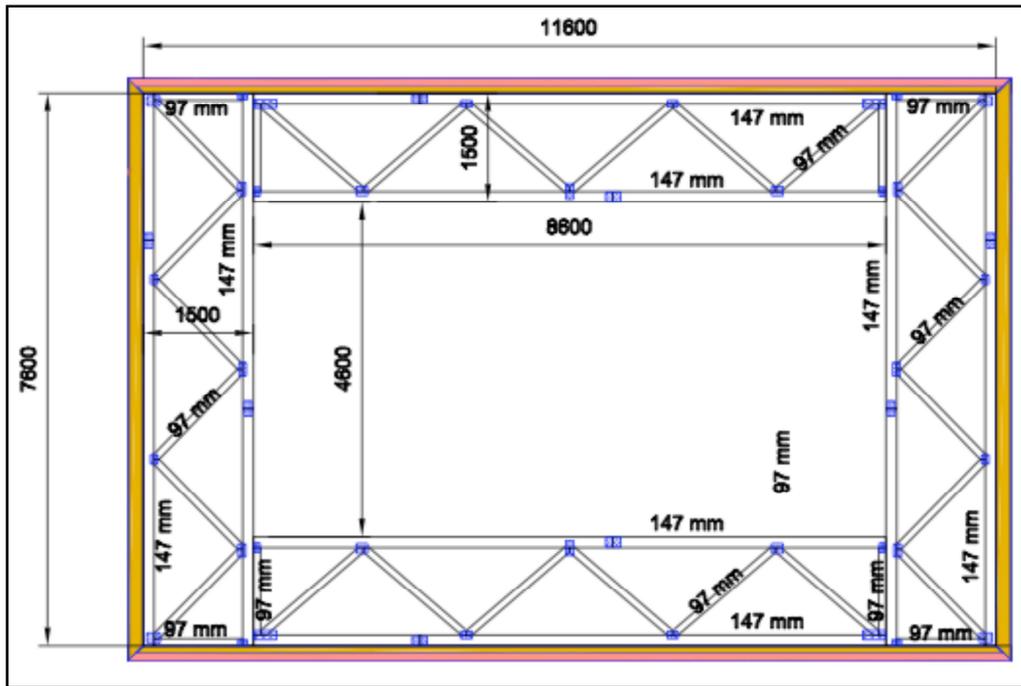


Figure III. 19 : Géométrie et dimensions des poutres au vent [7]

Les éléments en bois constituant la toiture sont représentés par des éléments poutres dont le comportement est élastique linéaire isotrope. Le matériau bois constituant les membres d'une ferme a un module d'Young  $E_f = 12 \text{ GPa}$ , tandis que celui des membres de contreventement et des liteaux est  $E_c = 9 \text{ GPa}$ . Les tuiles et les liteaux sont représentés par des éléments grilles qui ont une raideur seulement dans le sens perpendiculaire aux plans des fermes. La masse totale des tuiles est prise en compte dans le modèle éléments finis par la masse volumique  $\rho_{grille} = 50 \text{ kg/m}^2$  des éléments grilles. Les liaisons entre les éléments bois eux-mêmes ou avec d'autres éléments structuraux tels que les chaînages en béton armé sont représentées par des éléments discrets à deux nœuds de mêmes coordonnées. Le comportement des assemblages par plaques métalliques à dents constituant les fermes est élastique linéaire. La raideur de l'assemblage en translation dans le plan de la plaque est  $30 \text{ kN/mm}$  et infinie hors plan. Le comportement des assemblages réalisés par des équerres ou par pointes métalliques est non linéaire : il est modélisé par la loi anisotrope avec hystérésis présentée dans le paragraphe III.2.2 dont les différents paramètres sont déterminés pour chaque type d'assemblages à partir des résultats des essais sous chargements monotones et cycliques [151, 152].

Généralement, lorsqu'on procède à un calcul dynamique transitoire, l'amortissement dans la structure est représenté par un amortissement visqueux de type Rayleigh lié à la matrice de masse  $[M]$  et à la matrice de rigidité  $[K]$  de la structure par la relation :

$$[C] = \alpha[M] + \beta[K] \quad (III. 4)$$

où  $\alpha$  et  $\beta$  sont deux coefficients réels déterminés à partir des amortissements désirés  $\xi$  sur les deux premiers modes propres de la structure, représentés par les fréquences propres  $\omega_1$  et  $\omega_2$  :

$$\{\alpha, \beta\} = \frac{2\xi}{\omega_1 + \omega_2} \{1, \omega_1 \omega_2\} \quad (III. 5)$$

Pour les éléments finis classiques tels que les éléments poutres, les coefficients  $\alpha$  et  $\beta$  sont déterminés à partir des caractéristiques des matériaux. Malheureusement, cela n'est pas le cas

pour les éléments discrets. Afin de pallier ce problème, des éléments discrets *amortisseurs* ont été ajoutés parallèlement aux éléments discrets représentant les liaisons par connecteurs métalliques. Les premières simulations du modèle éléments finis, en dynamique non linéaire, ont montré que lorsque les liaisons clouées sont fortement dégradées ou rompues, la toiture résiste encore au chargement. Il s'est révélé que cette résistance est fictive, et qu'elle résulte du mouvement des amortisseurs qui engendre des efforts presque du même ordre que la résistance des liaisons. En effet, dans *Code\_Aster*, la matrice d'amortissement est calculée à partir de la matrice de raideur élastique et des fréquences propres de la structure saine. Ainsi, il a été décidé d'utiliser un amortissement uniquement proportionnel à la masse de la structure :

$$[C] = 2\xi\omega_1[M] \quad (III.6)$$

Ce choix est soutenu par le fait que les simulations ont révélé que la toiture possède un seul mode de vibration dominant dans chaque direction de chargement (i.e. directions perpendiculaire et parallèle aux plans des fermettes).

L'action sismique à laquelle est soumise la toiture est représentée par des accélérogrammes fournis par la *BRGM*, correspondant à des événements sismiques enregistrés dans différents sites en France métropolitaine ou en territoires d'outre mer. D'un point de vue pratique, l'effet de l'action sismique est représenté par des forces d'inertie équivalentes induites par l'accélération du sol. Elles sont appliquées aux nœuds du maillage éléments finis de la toiture. La réponse de la structure est obtenue par une procédure itérative.

### III.5.2 Critère d'endommagement dans la toiture

La résistance de la toiture au cours d'un événement sismique est basée sur l'évaluation du niveau d'endommagement atteint dans les assemblages. L'endommagement  $d_i$  dans un assemblage  $i$  quelconque de la toiture est défini en fonction du déplacement maximal  $u_{max}$  (i.e. en traction ou en compression) atteint au cours de l'évènement sismique :

$$d_i = 1 + d_c \left[ \max \left( \frac{u_{max}^+}{d_u^+}, \frac{u_{max}^-}{d_u^-} \right) - 1 \right] \quad (III.7)$$

où  $d_u$  est le déplacement à la rupture pour un chargement monotone, et  $d_c$  est l'endommagement cyclique qui vaut 1 pour un assemblage sain et un chargement monotone, et tend vers 0 lorsque le nombre de cycles de chargement tend vers l'infini [139].

Si on considère un groupe constitué de  $n$  assemblages du même type (assemblages entre les éléments d'anti-flambement et les arbalétriers, assemblages entre contreventement et les fiches, assemblages des lisses sur les fiches et les entrants, assemblages des liteaux sur les arbalétriers), l'endommagement  $D$  qui lui est associé peut être défini comme la moyenne des endommagements  $d_i$  enregistrés respectivement dans les  $n$  assemblages :

$$D = \frac{1}{n} \sum_{i=1}^n d_i \quad (III.8)$$

Ainsi, l'endommagement  $D$  varie dans l'intervalle  $[0,1]$ . Il vaut 0 si aucun des déplacements enregistrés dans les  $n$  assemblages n'a atteint le déplacement à la rupture respectif, et vaut 1 si tous les déplacements ont atteint le déplacement à la rupture. Les simulations ont montré que les endommagements des assemblages cloués localisés aux bouts des contreventements ainsi que ceux des assemblages cloués entre les éléments bois d'anti-flambement et les arbalétriers des fermettes sont les plus significatifs. Par conséquent, il a été décidé que la défaillance de la toiture est enregistrée lorsque l'indicateur d'endommagement  $D$  atteint une valeur supérieure

ou égale à 50% simultanément pour ces deux groupes d'assemblages. Le choix  $D = 50\%$  peut correspondre soit à la rupture de la moitié des assemblages, soit à l'enregistrement dans tous les assemblages d'un déplacement de 50% du déplacement à la rupture. La charge correspondant à cette valeur seuil de l'endommagement est comparable à la charge de rupture observée dans les essais. De plus, le paramètre  $D$  tel que défini par les équations (III.7) et (III.8) est continu dans l'intervalle  $[0,1]$ , ce qui est compatible avec les exigences du calcul stochastique. La figure III.20 illustre schématiquement le critère de défaillance de la toiture.

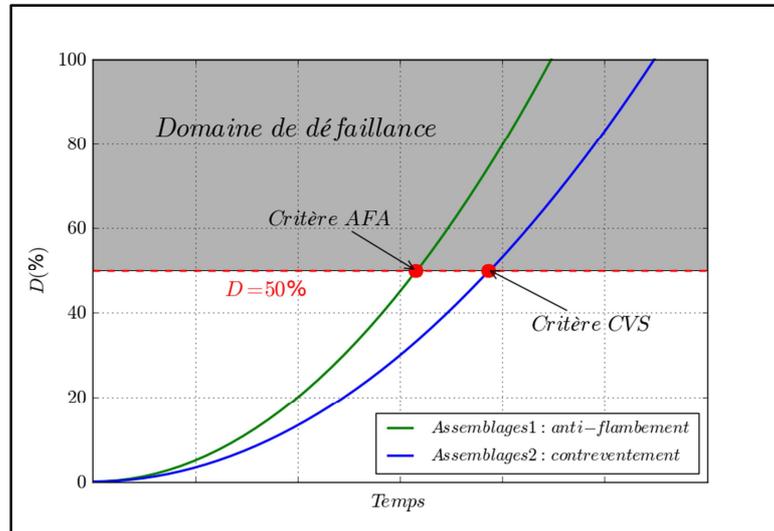
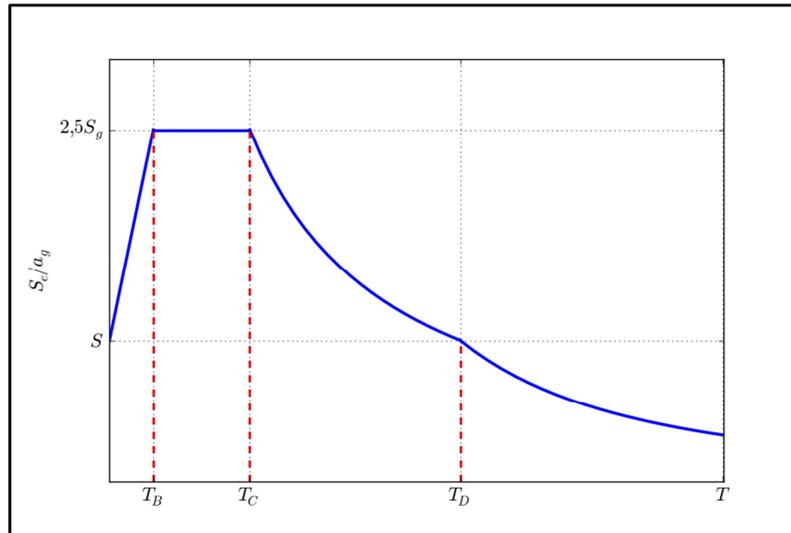


Figure III. 20 : Illustration schématique du critère de défaillance adopté pour l'évaluation de la résistance de la toiture

### III.5.3 Modélisation de l'action sismique

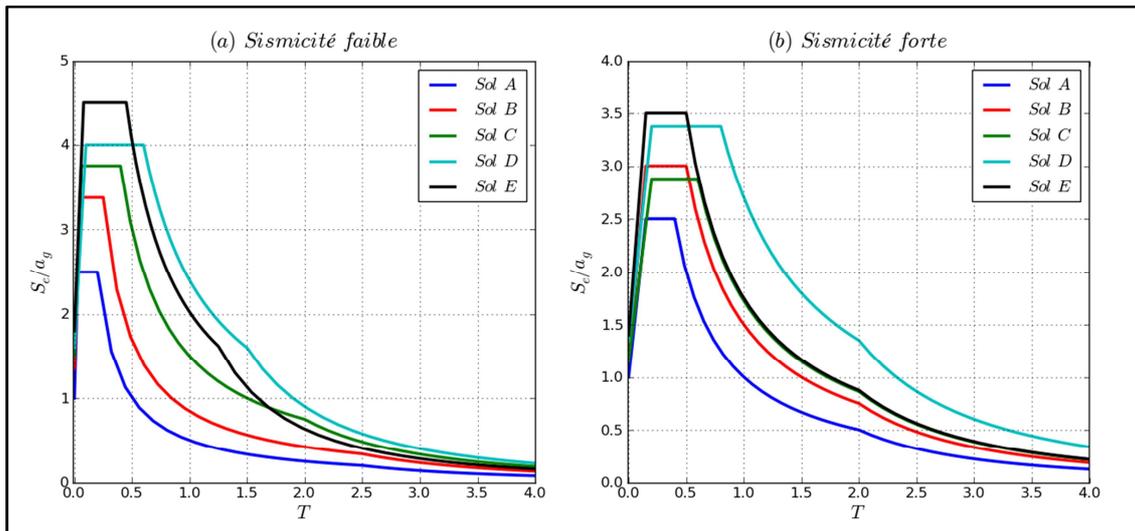
En se référant à la réglementation parasismique européenne Eurocode 8 [5], qui préconise d'utiliser des approches de calcul simplifiées pour la vérification des structures vis-à-vis des événements sismiques, l'évaluation de l'action sismique est basée sur la définition d'un spectre de réponse élastique associé à un site donné. Ce spectre est utilisé par la suite soit pour déterminer des forces statiques équivalentes, soit pour générer des accélérogrammes synthétiques ou pour modifier des accélérogrammes naturels (i.e. des enregistrements in-situ de mouvements sismiques) qui seront employés dans le cas où on souhaite effectuer une analyse chronologique non linéaire (analyse dynamique temporelle). La caractérisation d'un mouvement sismique repose sur plusieurs paramètres (la magnitude qui mesure l'énergie libérée par le séisme, la durée de l'évènement sismique, l'intensité d'Arias qui permet de quantifier l'énergie dissipée et sa répartition durant le séisme, ...). Cependant, on utilise en règle générale un seul paramètre, l'accélération maximale du sol ( $PGA$  qui est l'abréviation de la dénomination anglaise *Peak Ground Acceleration*) pour caractériser un séisme. Ce paramètre dépend du site sur lequel la structure est installée, et sa valeur est donnée par le zonage sismique fourni par la réglementation parasismique utilisée dans le pays (l'Eurocode 8 en Europe). En effet, pour un site donné, le zonage séismique fournit une valeur de référence  $a_{gr}$  du  $PGA$  pour une période de retour de référence. En se référant à l'Eurocode 8, ces valeurs de référence d'accélération  $a_{gr}$  ont été définies pour une période de retour de 475 ans. Cela signifie que la probabilité de dépasser cette accélération cible dans les 50 ans à venir est de 10%. On note tout de même que d'autres périodes de retour peuvent être facilement considérées en employant un coefficient d'importance, dont le rôle est de pondérer les accélérations de référence pour une période de retour de 475 ans. La figure III.21 illustre

l'évolution du spectre de réponse élastique associé à une accélération  $a_g$  en se référant aux indications de l'Eurocode 8.



**Figure III. 21 :** Evolution du spectre de réponse élastique selon l'Eurocode 8

On constate que le spectre de réponse élastique est défini par plusieurs paramètres modélisant l'effet du type du sol ainsi que l'amortissement de la structure. En se référant toujours à l'Eurocode 8, les sols sont répartis en cinq catégories désignées par des lettres de A jusqu'à E pour lesquels différentes valeurs des paramètres  $T_A$ ,  $T_B$ ,  $T_C$  et  $S$  sont données. Par exemple, le sol de type A représente un sol rigide et le sol de type E représente un sol souple. Les figures III.22a et III.22b illustrent l'évolution du spectre de réponse élastique respectivement pour des territoires de faible sismicité correspondant aux zones 1 à 4 sur la carte de zonage sismique et des territoires de forte sismicité correspondant à la zone 5 sur la carte de zonage sismique.



**Figure III. 22 :** Spectre de réponse élastique pour différents type de sol selon l'Eurocode 8, (a) zones de sismicité faible et (b) zones de sismicité élevée

La vérification de la résistance d'une structure vis-à-vis d'un évènement sismique doit être effectuée dans les trois directions principales, deux horizontales et une verticale. De plus, pour chaque direction, plusieurs spectres de réponse élastique doivent être utilisés (généralement on utilise trois spectres de réponse élastique pour chaque direction). Bien que les réglementations parasismiques n'excluent pas la possibilité d'effectuer un calcul non

linéaire dynamique, on a recours le plus souvent à un calcul élastique linéaire. Dans ce cas, l'effet de la ductilité qui se manifeste dans le comportement réel de la structure est pris en compte au travers d'un coefficient de comportement  $q$ . Il n'est autre que le rapport entre la force à la rupture obtenue pour un comportement élastique linéaire, correspondant à un déplacement de référence de la structure, et la force à la rupture obtenue pour un comportement non linéaire et pour le même déplacement de référence. En effet, ce calcul simplifié est basé sur un spectre de réponse de calcul ou de conception obtenu par la réduction du spectre de réponse élastique par un facteur  $q$ . La détermination de la valeur du coefficient de comportement est une étape importante dans la procédure de dimensionnement. Elle dépend de plusieurs paramètres tels que les matériaux constitutifs et les sources de non linéarité dans la structure, ainsi que de la présence de composants dissipatifs. En ce qui concerne les structures à ossatures bois, le choix de la valeur du facteur de comportement est dicté par le niveau de ductilité de la structure. En effet, pour les structures dissipatives, pour lesquelles la ductilité est élevée, une valeur supérieure à 1,5 est souvent choisie pour le coefficient de comportement, à condition que les zones ductiles soient localisées dans les assemblages. On note tout de même que, malgré la définition simpliste du coefficient du comportement, la détermination de la valeur adéquate à utiliser dans un dimensionnement n'est pas clairement explicitée dans la plupart des cas.

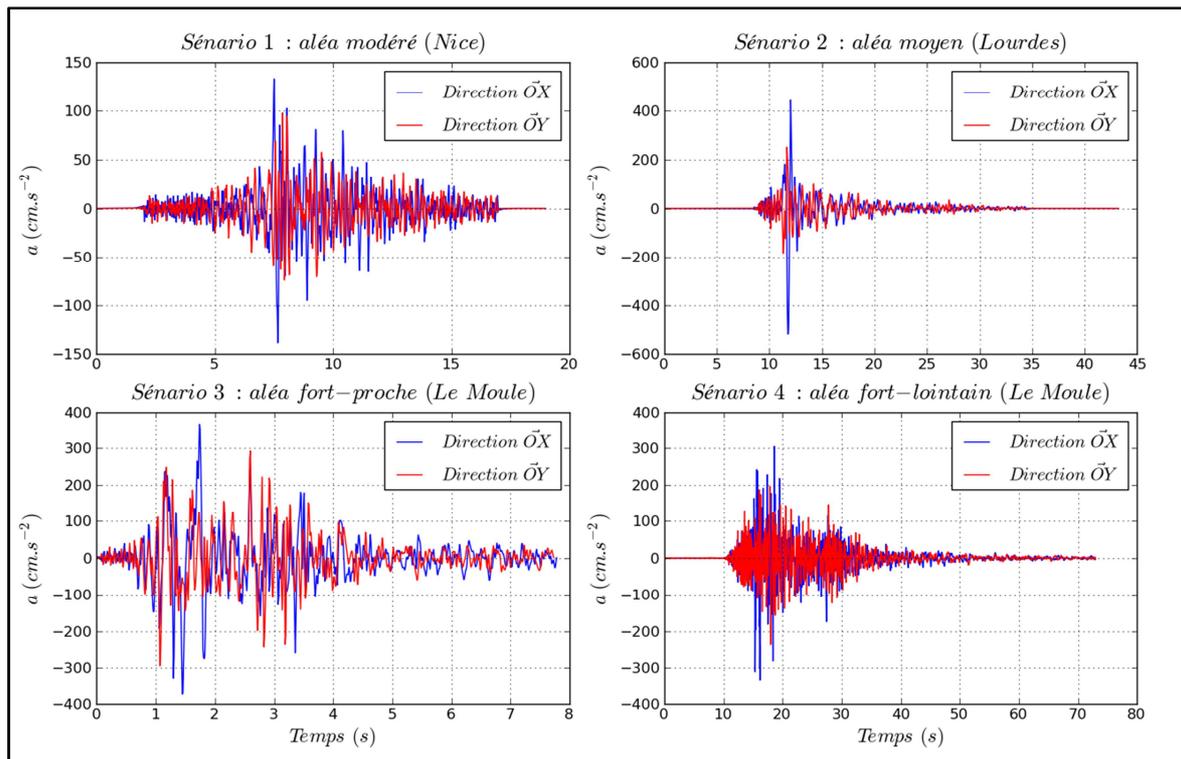
Dans le cadre du projet SISBAT, qui s'articule principalement sur la prise en compte et la modélisation réaliste des sources de non linéarité dans les structures à ossature bois, en particulier dans les assemblages bois réalisés par des connecteurs métalliques, le choix le plus pertinent pour rester fidèle à cet objectif est d'effectuer une analyse chronologique non linéaire. Dans ce contexte, la modélisation de l'action sismique peut être effectuée par des accélérogrammes synthétiques (i.e. générés par une procédure numérique), des accélérogrammes naturels (i.e. des enregistrements de mouvements sismiques réels), ou encore des accélérogrammes naturels modifiés. Ayant pour objectif d'être le plus réaliste possible dans la modélisation numérique d'une part et dans la représentation de l'aléa sismique de la France d'autre part, le choix retenu est de travailler avec des accélérogrammes naturels modifiés (i.e. des enregistrements de mouvements sismiques réels dont le contenu fréquentiel est modifié pour se caler sur un spectre cible). En effet, on a souvent tendance à travailler avec des accélérogrammes correspondant à des enregistrements de mouvements sismiques forts (par exemple Kobe) qui ont marqué l'histoire de l'humanité. Or, ces événements sismiques peuvent être des cas extrêmes voire même peu probables si on change de contexte. Afin de pallier cette difficulté, il a été décidé d'utiliser des enregistrements de mouvements sismiques représentatifs de l'aléa sismique en France métropolitaine et dans les territoires d'outre-mer. Ces derniers sont fournis par le *BRGM* [153], un des partenaires du projet SISBAT, qui dispose d'une base de données importante en termes d'enregistrements de mouvements sismiques. Les critères de choix des accélérogrammes sont très variés. Ils sont généralement définis par les finalités qu'on prétend atteindre. Un des objectifs principaux du projet SISBAT est d'effectuer des calculs stochastiques, pour lesquels il est question d'évaluer l'intégrité d'une structure vis-à-vis d'un critère de défaillance associé à un événement sismique. Ainsi, les signaux doivent être suffisamment pénalisants pour bien encadrer la limite de résistance de la structure. Dans cette perspective, le *BRGM* s'est focalisé sur la recherche de scénarios ou d'événements sismiques uniquement dans les zones d'aléa modéré, moyen et forts en se référant à la carte d'aléa sismique en France. Chacune de ces zones est définie par une gamme d'accélération maximale et les sites qui leurs sont associés sont choisis dans la partie haute de la gamme. Ainsi, la zone d'aléa modéré est représentée par la ville de Nice, la zone d'aléa moyen est représentée par la ville de Lourdes et la zone d'aléa fort est représentée par la ville du Moule en Guadeloupe (cf. tableau III.7). On note qu'afin de couvrir la gamme spectrale du spectre de l'aléa uniforme correspondant au site choisi pour

représenter la zone d'aléa fort, deux scénarios distincts ont été considérés correspondant respectivement à un séisme lointain (i.e. spectre riche en basses fréquences) et un séisme proche (i.e. spectre riche en hautes fréquences).

**Tableau III. 7 :** Coordonnées géographiques et caractéristiques sismiques des sites choisis pour l'aléa modéré, moyen et fort [153]

Site	Longitude	Latitude	Aléa	Gamme d'accélération
Nice	7,24°E	43,71°N	modéré	$1,1 \text{ m/s}^2 \leq a_g < 1,6 \text{ m/s}^2$
Lourdes	0,05°O	43,09°N	moyen	$1,6 \text{ m/s}^2 \leq a_g < 3,0 \text{ m/s}^2$
Le Moule	61,39°O	16,37°N	fort	$a_g > 3,0 \text{ m/s}^2$

En résumé, quatre scénarios distincts sont considérés dans notre travail, correspondant respectivement à un aléa modéré, un aléa moyen, un aléa fort lointain et un aléa fort proche, pour lesquels le BRGM a fourni respectivement 15, 14, 13 et 11 signaux. Pour chaque scénario, 3 à 4 signaux représentent un type de sol A, B, C, ou D selon la classification de l'Eurocode 8. La sélection des accélérogrammes a été effectuée dans la base de données européenne de mouvements forts, ainsi que dans des bases de données mondiales, dont la classification était basée sur l'évaluation de plusieurs indicateurs du mouvement sismique tels que le PGA, le PGV, l'intensité d'Arias et la durée de la phase forte. De plus, une hiérarchisation des signaux a été effectuée en se basant sur la moyenne des différents indicateurs du mouvement sismique.



**Figure III. 23 :** Quatre signaux sismiques fournis par le BRGM respectivement pour une zone d'aléa modéré, une zone aléa moyen, une zone d'aléa fort-proche et une zone d'aléa fort-lointain

Les spectres des signaux fournis ont été comparés par le BRGM aux spectres de l'Eurocode 8, représentatifs des différentes zones d'aléa sismique et des différents types de sol, qui a conclu que le contenu fréquentiel est cohérent, sauf pour quelques signaux pour lesquels on a constaté des divergences acceptables dans le sens où ils sont plus pénalisants que ceux de la réglementation. La figure III.23 illustre une sélection de signaux fournis par le BRGM pour

les quatre scénarios considérés dans ce travail. Les directions  $\overrightarrow{OX}$  et  $\overrightarrow{OY}$  qui apparaissent sur les légendes de la figure III.23 désignent les deux directions horizontales définies par rapport à la source d'enregistrement du signal sismique, et elles n'ont aucun lien avec l'orientation de la structure.

### III.5.4 Sélection des signaux sismiques

Les signaux fournis par le *BRGM* correspondent à un niveau de référence du *PGA*. Ainsi, pour observer l'endommagement dans la structure, ce *PGA* de référence est amplifié lors des simulations numériques. De plus, afin de réduire le temps de calcul, nous considérons uniquement la portion du signal susceptible de provoquer la majeure partie de l'endommagement accumulé dans la structure au cours de la durée totale de l'évènement sismique. L'étendue de cette portion est délimitée par les bornes  $t_i$  et  $t_f$  données dans la figure III.24. Ces bornes correspondent aux instants auxquels l'accélération atteint un seuil de 5% de l'accélération maximale, en partant respectivement du début et la fin du signal sismique.

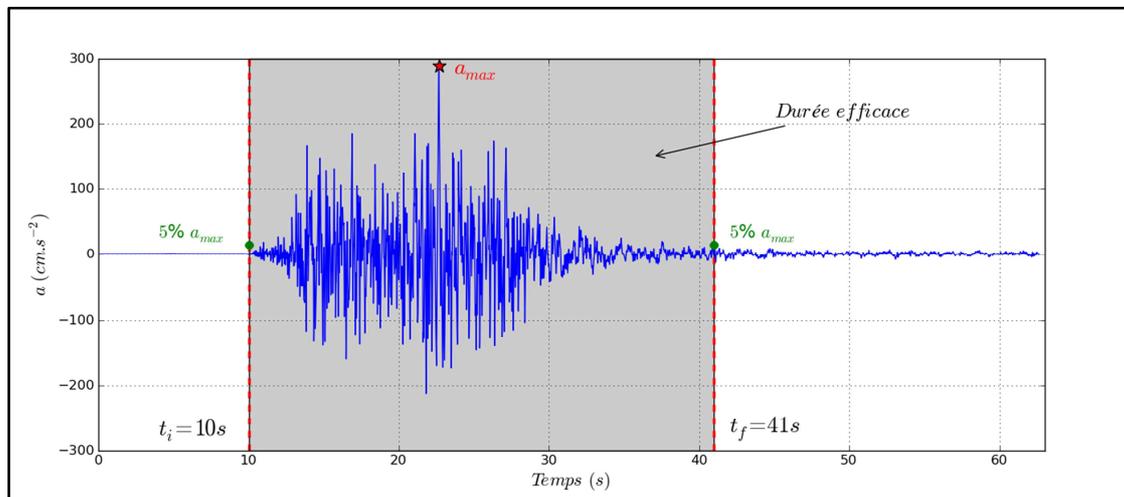


Figure III. 24 : Définition de la durée efficace du signal sismique

Avant d'entamer les calculs stochastiques, il est indispensable de déterminer l'amplification de référence qui correspond à l'atteinte du critère de défaillance de la toiture présenté précédemment dans le paragraphe III.5.2. Ce critère de défaillance est purement déterministe, et il ne faut pas le confondre avec la notion de critère de défaillance rencontré dans le cadre des analyses de fiabilité. Autrement dit, le paramètre  $D$  défini par l'équation (III.8) caractérise la réponse de la structure permettant de mesurer l'endommagement cumulé dans la structure lors du passage d'un séisme. Dans ce travail, nous nous proposons de nous focaliser sur l'évaluation de la résistance de la toiture à ossature bois installée dans une zone d'aléa fort. Dans cette perspective, huit signaux lointains associés au site de la ville du Moule en Guadeloupe ont été sélectionnés (cf. figure III.23). Pour chacun d'eux, le niveau d'amplification du signal de référence, nécessaire pour atteindre le critère de défaillance de la toiture, a été obtenu par interpolation linéaire et en effectuant plusieurs simulations du modèle numérique pour différentes amplifications.

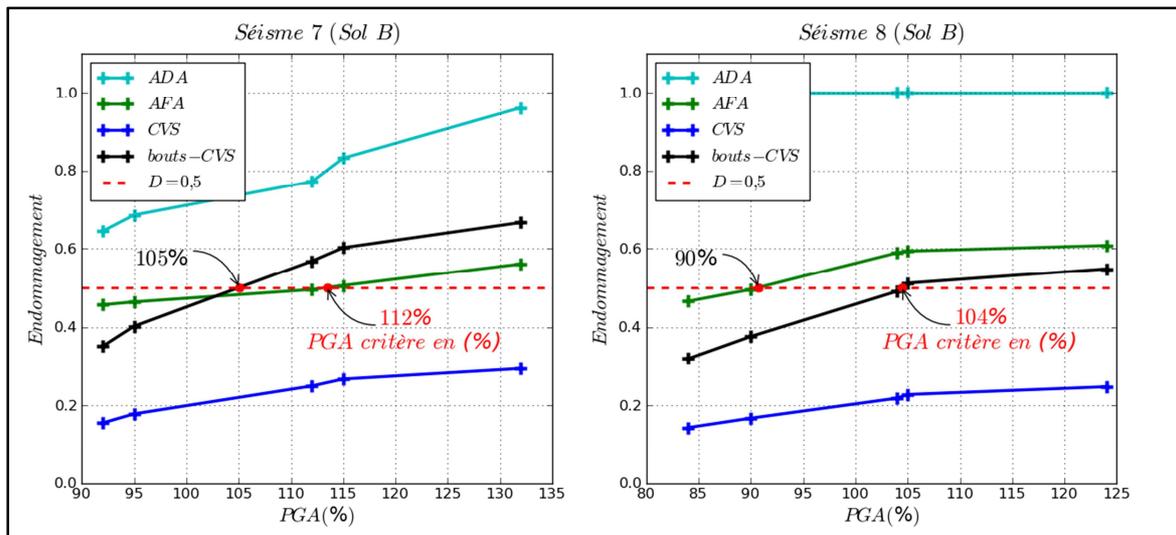
Les résultats concernant les huit signaux représentatifs d'un scénario lointain de l'aléa sismique sont donnés dans le tableau III.8. A première vue, on constate que les signaux qui ont le *PGA* spectre le plus élevé ne sont pas forcément ceux les plus nocifs sur la structure. En effet, pour le signal sismique 6, qui vient en deuxième position si on classe les signaux sismiques par ordre décroissant de *PGA* du spectre, la défaillance de la structure se manifeste

lorsque le *PGA* du spectre est amplifié de 220%. Cette valeur est la plus élevée pour le panel de signaux sismiques représentatifs du scénario lointain. Le *PGA* critère (amplification du *PGA* du spectre) est plus informatif, dans la mesure où il tient compte de plusieurs indicateurs de nocivité du signal sismique à la fois, tandis que le *PGA* spectre ne représente que l'accélération maximale du signal sismique.

**Tableau III. 8 :** Amplifications correspondant à l'atteinte du critère de ruine pour les signaux de la ville du Moule en Guadeloupe – scénario lointain

Séisme	1	2	3	4	5	6	7	8
Référence BRGM	4655ya	4653ya	8659ya	9623ya	9622ya	9588ya	9618ya	9592ya
Sol	B	C	C	C	D	A	B	B
Durée efficace (s)	[10,41]	[10,52]	[15,60]	[14,70]	[17,106]	[15,68]	[14,70]	[15,75]
PGA spectre	6,19	6,15	5,91	6,20	6,03	6,62	7,26	6,49
Amplification	106%	111%	166%	147%	135%	220%	112%	104%
PGA critère	6,56	6,82	9,81	9,11	8,14	14,56	8,13	6,74

On note que la durée d'une simulation par éléments finis dépend de la nocivité et de la durée du signal sismique. En terme de temps machine, elle est comprise entre 12 et 28 heures. On signale aussi que l'amplification du *PGA* du spectre, prise comme référence pour définir la défaillance de la structure, correspond au maximum des amplifications enregistrées pour atteindre les deux critères de ruine (i.e. critère *AFA* et critère bouts des *CVS*), comme le montre la figure III.25. Pour tous les signaux sismiques, on constate que l'amplification du *PGA* du spectre induisant la défaillance de la structure est supérieure à 100%. Cela signifie que la structure, telle qu'elle est conçue, résiste à l'aléa sismique associé à la ville du Moule.



**Figure III. 25 :** Illustration de la détermination de l'amplification correspondant au critère de défaillance de la structure pour les signaux sismiques 7 et 8 (cf. tableau. III.8)

Cette constatation ne doit pas être généralisée puisqu'elle n'est vraie que pour le critère de ruine adopté dans ce travail. En effet, dans certains cas, notamment comme pour le séisme 8 (cf. figure III.25), ce critère est conservatif dans la mesure où si on considère que la défaillance de la structure est induite par l'atteinte de l'endommagement cumulé d'un niveau de 0,5 seulement dans les assemblages situés sur les éléments d'anti-flambement, on constate que la structure, telle qu'elle est conçue, ne résiste pas à l'évènement sismique, puisque la défaillance est enregistrée pour une amplification de 90% du *PGA* du spectre. On constate de plus, que la réalisation des deux conditions de défaillance (i.e. critère *AFA* et critère bouts *CVS*) ne suit pas la même chronologie pour les différents signaux sismiques. En effet, si nous

nous référerons encore à la figure III.25, on constate, dans le cas du séisme 7, que l'endommagement cumulé dans les assemblages situés sur les extrémités des éléments de contreventement atteint un niveau critique pour une amplification du signal sismique inférieure à l'amplification de réalisation du critère de défaillance de la structure, tandis que c'est le contraire pour le séisme 8.

En effectuant des analyses statistiques sur les résultats obtenus pour le panel des signaux sismiques représentant le scénario lointain, nous obtenons, une moyenne de 8,73 et un coefficient de variation de 28% pour le *PGA* induisant la défaillance de la structure. Ces résultats fournissent des informations sur l'effet de l'aléa, non intrinsèque du signal sismique, qui est lié plutôt au type du sol sur lequel la structure est installée. En ce point, nous signalons que ces résultats ne peuvent pas être généralisés dans le sens où ils sont relatifs au type de la structure étudiée et au critère de défaillance adopté. Si nous nous référons uniquement à la valeur moyenne, nous constatons que la structure, telle qu'elle est conçue, résiste à l'aléa sismique de la ville du Moule. Cependant, il faut signaler que cette valeur particulière, est entachée d'une forte incertitude qu'il faut prendre en compte dans la conception de la structure.

### III.6 Analyses stochastiques d'une toiture à ossature bois

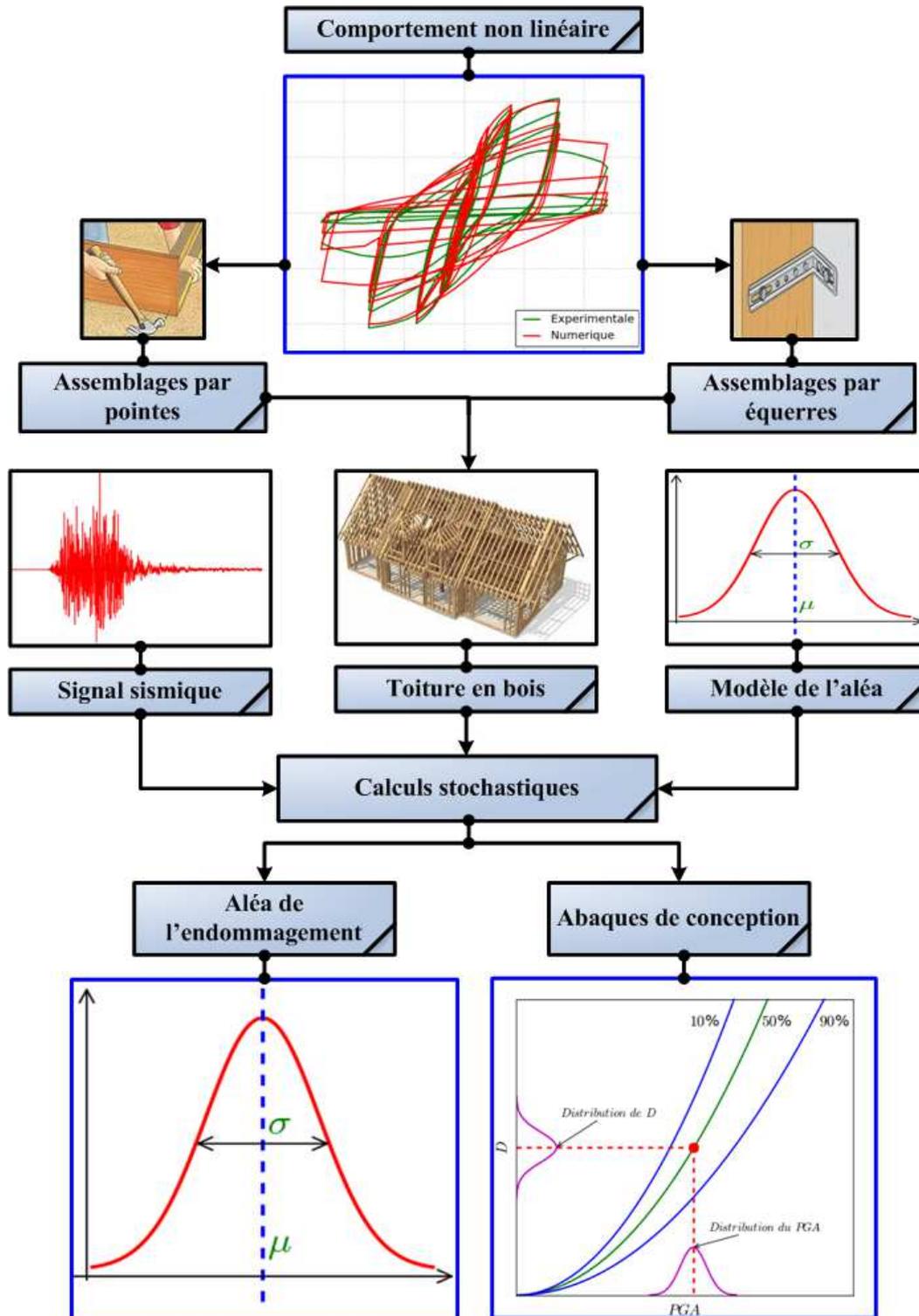
#### III.6.1 Stratégie d'évaluation de l'effet de l'aléa

Un des objectifs principaux de notre travail est d'évaluer l'effet des incertitudes sur la résistance d'une structure réelle correspondant à une toiture à ossature bois. La réponse mécanique est représentée par le paramètre  $D$  mesurant l'endommagement cumulé dans la structure durant un événement sismique. Nous rappelons que nous ne modélisons que l'aléa associé aux paramètres représentant le comportement non linéaire des assemblages bois réalisés par organes métalliques, en particulier les assemblages réalisés par des équerres ou par des pointes métalliques, étant donné que le comportement des assemblages réalisés par plaques métalliques embouties est supposé élastique linéaire. L'unique paramètre incertain considéré dans chaque assemblage est la force au pic  $F_1$  de la loi de comportement avec hystérésis (cf. figure III.2). Ce choix est dicté par les résultats de l'analyse de sensibilité basée sur la décomposition de la variance et sur l'évaluation des indices de sensibilité de Sobol présentée dans le paragraphe III.3.1. Le nombre très élevé d'assemblages constituant la toiture est lié directement au nombre de paramètres incertains dans les calculs stochastiques, induisant ainsi le problème de dimension stochastique élevée. Pour le résoudre, nous appliquons la stratégie développée dans le chapitre précédent.

La finalité du calcul stochastique est d'une part d'évaluer l'aléa induit par les paramètres incertains sur l'indicateur d'endommagement  $D$ , et le calcul des moments statistiques et la construction de la densité de probabilité, et d'autre part de proposer une formulation explicite de la relation entre l'indicateur de d'endommagement  $D$  et le *PGA* du signal sismique. A un niveau donné d'amplification du *PGA* du signal sismique, on peut construire la densité de probabilité de la réponse mécanique de la structure sous forme d'histogramme. On peut ainsi déterminer la courbe d'évolution de l'endommagement accumulé durant un événement sismique et pour différents niveaux de confiance. Réciproquement, pour un niveau d'endommagement cible de conception, on peut construire la distribution de probabilité du *PGA* pour différents signaux sismiques, à partir de laquelle le *PGA* de référence peut être déterminé pour un seuil de confiance donné.

La motivation du choix d'une telle stratégie, pour l'évaluation de l'effet de l'aléa sur la résistance d'une structure vis-à-vis d'un événement sismique, vient du fait qu'elle est

informative et qu'elle peut être utilisée pour justifier les solutions technologiques dans la conception parasismique. L'organigramme de cette stratégie est donné sur la figure III.26.



**Figure III. 26 :** Stratégie d'évaluation de l'aléa associé à l'endommagement cumulé dans une toiture à ossature bois sous l'action sismique

### III.6.2 Stratégie de mise en œuvre du calcul stochastique

D'un point de vue pratique, la méthodologie proposée pour l'évaluation de l'effet de l'aléa sur le comportement de la toiture est basée sur une stratégie de calcul en parallèle. Autrement dit,

les simulations du modèle éléments finis représentant la toiture, pour les jeux de variables aléatoires exigés par le calcul stochastique, ne sont pas effectuées d'une façon séquentielle mais simultanément. En effet, nous avons à disposition un cluster comprenant 27 nœuds, chacun d'eux équipé de deux microprocesseurs. Ainsi, nous pouvons lancer  $27 \times 2$  simulations du modèle éléments finis en même temps. Cependant, l'exploitation de ces moyens de calcul n'est pas triviale, car il faut développer une procédure permettant la gestion automatique des opérations numériques. Pour cette raison, nous avons développé une procédure numérique en langage *Python*, qui permet à la fois, de communiquer entre un *PC* maître et le cluster, de générer les fichiers de données et de commandes, et de les affecter sur les nœuds du cluster, et de récupérer les résultats de simulations pour le post-traitement. Cette procédure est explicitée schématiquement par l'organigramme de la figure III.27.

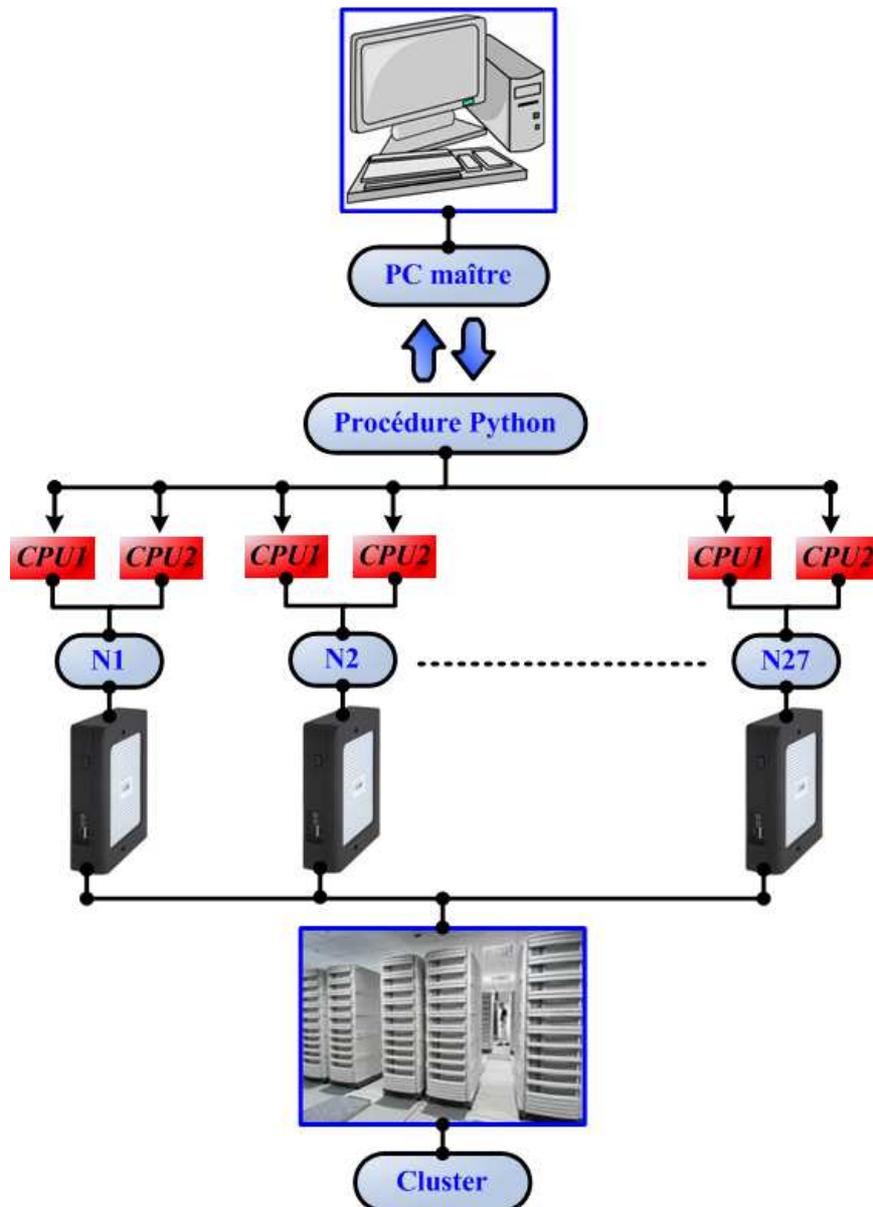


Figure III. 27 : Organigramme de la mise en œuvre de la stratégie du calcul stochastique

### III.6.3 Analyse de l'effet de l'aléa des propriétés du matériau bois

Avant d'appliquer la stratégie présentée ci-dessus pour évaluer l'aléa associé à l'endommagement dans la toiture, nous nous proposons dans ce paragraphe d'étudier l'effet

de l'aléa associé aux propriétés du matériau bois par rapport à celui de l'aléa associé aux propriétés mécaniques des assemblages constituant la toiture. Dans cette perspective, nous nous restreignons à une étude de sensibilité qualitative, visant la hiérarchisation des paramètres incertains, en utilisant une méthode de criblage. La désignation et la localisation des paramètres incertains dans la structure sont données dans le tableau III.9. Ils sont modélisés par des variables aléatoires lognormales indépendantes, dont les caractéristiques statistiques sont données dans le tableau III.10.

**Tableau III. 9 :** Désignation et localisation des paramètres incertains dans la toiture à ossature bois- étude de l'effet de l'aléa des propriétés du matériau bois

<b>Désignation</b>	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
<b>Paramètre</b>	$FI_{\text{clous } 31/90}$	$FI_{\text{clous } 31/90}$	$FI_{\text{clous } 28/70}$	$FI_{\text{clous } 28/70}$	$FI_{\text{clous } 28/70}$	$FI_{\text{équerrres}}$
<b>Localisation</b>	clous AFA	clous ADA	clous CVS	clous liteau	clous lisses	Equerres
<b>Désignation</b>	$X_7$	$X_8$	$X_9$	$X_{10}$	$X_{11}$	$X_{12}$
<b>Paramètre</b>	$E_{\text{bois } C18}$	$E_{\text{bois } C18}$				
<b>Localisation</b>	AFA	CVS1 gauche	CVS2 gauche	CVS3 gauche	CVS4 gauche	CVS5 gauche
<b>Désignation</b>	$X_{13}$	$X_{14}$	$X_{15}$	$X_{16}$	$X_{17}$	$X_{18}$
<b>Paramètre</b>	$E_{\text{bois } C18}$	$E_{\text{bois } C18}$				
<b>Localisation</b>	CVS1 droite	CVS2 droite	CVS3 droite	CVS4 droite	CVS5 droite	lisses
<b>Désignation</b>	$X_{19}$	$X_{20}$	$X_{21}$	$X_{22}$	$X_{23}$	$X_{24}$
<b>Paramètre</b>	$E_{\text{bois } C18}$	$E_{\text{bois } C18}$	$E_{\text{bois } C24}$	$E_{\text{bois } C24}$	$E_{\text{bois } C24}$	$E_{\text{bois } C24}$
<b>Localisation</b>	ADA	entretoises	Ent1+Arba1	Ent2+Arba2	Ent3+Arba3	Ent4+Arba4
<b>Désignation</b>	$X_{25}$	$X_{26}$	$X_{27}$	$X_{28}$	$X_{29}$	$X_{30}$
<b>Paramètre</b>	$E_{\text{bois } C24}$	$E_{\text{bois } C24}$				
<b>Localisation</b>	Ent5+Arba5	Ent6+Arba6	Ent7+Arba7	Ent8+Arba8	Ent9+Arba9	Ent10+Arba10
<b>Désignation</b>	$X_{31}$	$X_{32}$	$X_{33}$	$X_{34}$	$X_{35}$	$X_{36}$
<b>Paramètre</b>	$E_{\text{bois } C24}$	$E_{\text{bois } C24}$				
<b>Localisation</b>	Ent11+Arba11	Ent12+Arba12	Ent13+Arba13	Ent14+Arba14	Ent15+Arba15	Ent16+Arba16
<b>Désignation</b>	$X_{37}$	$X_{38}$	$X_{39}$	$X_{40}$	$X_{41}$	$X_{42}$
<b>Paramètre</b>	$E_{\text{bois } C24}$	$E_{\text{bois } C24}$				
<b>Localisation</b>	Ent17+Arba17	Ent18+Arba18	Ent19+Arba19	Ent20+Arba20	Fiches longues	Fiches courtes

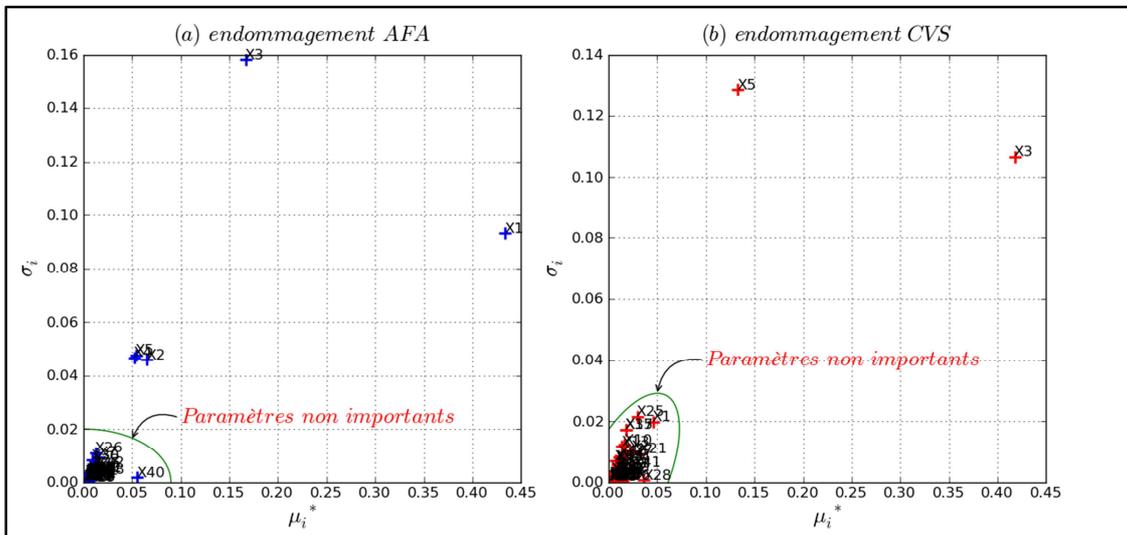
Les distributions des paramètres incertains ont été déterminées par des analyses statistiques. Pour le matériau bois, les caractéristiques statistiques du module de Young sont estimées en se basant sur un échantillon de 1181 spécimens, composés de 603 éléments en épicéas et 578 éléments en sapins. Pour le comportement non linéaire des assemblages, les caractéristiques statistiques de la force au pic  $F_1$  sont déterminées à partir des résultats de 15 essais cycliques dans le cas des assemblages réalisés par vis ou par pointes, et de 5 essais cycliques et 2 essais monotones dans le cas des assemblages réalisés par des équerrres [149, 150].

**Tableau III. 10 :** Caractéristiques statistiques des paramètres incertains- étude de l'effet de l'aléa des propriétés du matériau bois

	$E_{\text{bois } C18}$	$E_{\text{bois } C24}$	$FI_{\text{clous } 31/90}$	$FI_{\text{clous } 28/70}$	$FI_{\text{équerrres}}$
<b>Distribution</b>	lognormale	lognormale	lognormale	lognormale	lognormale
<b><math>\mu</math></b>	9	11	2500	2250	4100
<b>CV(%)</b>	22	22	12	12	12

Afin de déterminer les paramètres incertains les plus influents sur l'intégrité de la structure, nous utilisons la méthode de Morris. Les estimations des caractéristiques statistiques des effets élémentaires utilisés pour hiérarchiser les 42 paramètres incertains sont obtenues par

deux trajectoires (i.e. deux plans d'expérimentation), soit 86 appels au modèle éléments finis. Les figures III.28a et III.28b illustrent la hiérarchisation des paramètres incertains par rapport à la mesure de l'endommagement dans les liaisons des éléments d'anti-flambement et de contreventement.



**Figure III. 28 :** Hiérarchisation des paramètres incertains vis-à-vis de l'endommagement dans les assemblages situés dans les éléments d'anti-flambement et de contreventement

On constate que seules les variables aléatoires  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$  et  $X_5$ , représentant l'aléa de la force au pic  $F_1$ , associées aux assemblages par clous situés respectivement sur les éléments d'anti-flambement (AFA), les éléments d'anti-devers d'appuis (ADA), les éléments de contreventement (CVS), le liteau, et les lisses, sont les plus influents sur la variabilité des deux critères de ruine de la structure, tandis que, l'aléa associé aux propriétés du matériau bois a un effet non significatif. Par conséquent, ces derniers seront considérés dans la suite de notre travail comme des paramètres déterministes en les fixant à leurs valeurs moyennes respectives, et nous ne modélisons que l'aléa associé au comportement non linéaire des assemblages.

### III.6.4 Analyse de l'effet de l'aléa des propriétés des assemblages

#### III.6.4.1 Modélisation des sources d'aléa

A ce niveau de l'étude, les assemblages par pointes situés sur les éléments d'anti-flambement (AFA), les éléments de contreventement (CVS), les éléments d'anti-devers d'appuis (ADA), les lisses, les liteaux et les assemblages par équerres représentant les liaisons entre les chaînages en béton et les fermettes en bois, sont traités séparément dans le calcul stochastique et non pas par groupe comme nous l'avons fait dans l'étude d'importance par criblage traitée dans le paragraphe précédent. Autrement dit, l'aléa est introduit dans chaque assemblage à travers la force au pic  $F_1$  de la loi avec hystérésis utilisée dans la modélisation par éléments finis. On note tout de même que les assemblages par pointes situés dans les liteaux sont représentés par une seule variable aléatoire, et ainsi que pour les assemblages situés sur les lisses. Ce choix n'est pas arbitraire dans la mesure où les résultats des essais sur des maquettes de toitures à ossature bois, ont montré que ces assemblages résistent lors de l'application de la sollicitation sismique. En effet, pour les différentes configurations testées [160], la rupture a toujours eu lieu dans les assemblages entre les éléments d'anti-flambement et les éléments d'anti-devers d'appuis, ou les assemblages entre les éléments d'anti-

flambement et les fermettes. Le tableau III.11 donne l'affectation de l'aléa dans les assemblages constituant la toiture à ossature bois.

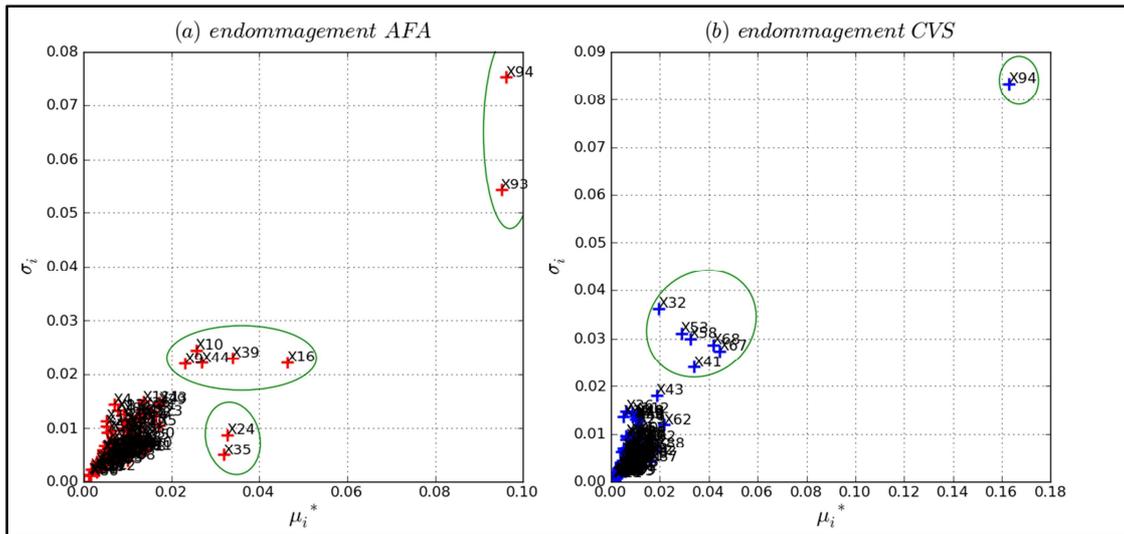
**Tableau III. 11** : Affectation de l'aléa dans les assemblages de la toiture à ossature bois

<i>Organe d'assemblage</i>	<i>Assemblage</i>	<i>Nombre</i>	<i>Modélisation</i>
<i>Pointes</i> $\phi 3,1 \times 90$	AFA	38	1 assemblage $\rightarrow$ 1 v.a.
	ADA	6	1 assemblage $\rightarrow$ 1 v.a.
<i>Pointes</i> $\phi 2,8 \times 70$	CVS	48	1 assemblage $\rightarrow$ 1 v.a.
	Liteaux	680	680 assemblage $\rightarrow$ 1 v.a.
	Lisses	268	268 assemblage $\rightarrow$ 1 v.a.
<i>Equerres</i>	Chaînages	40	1 assemblage $\rightarrow$ 1 v.a.
			Total = 134 v.a.

Dans cette étude nous avons 134 paramètres incertains modélisés par des lois lognormales indépendantes dont les caractéristiques statistiques dépendent du type d'organe métallique utilisé pour l'assemblage des éléments bois. Elles ont les mêmes valeurs que dans l'étude menée dans le paragraphe III.6.2 (cf. tableau III.10). Si nous considérons les 134 paramètres incertains dans le calcul stochastique proprement dit, il en résulte un temps de calcul conséquent que nous sommes incapables de gérer, compte tenu des moyens de calcul numérique à notre disposition. En effet, malgré l'efficacité de la méthode de réduction de la dimension qui est peu sensible à la dimension stochastique, puisque le nombre d'appels au modèle éléments finis croît d'une façon polynômiale vis-à-vis du nombre de paramètres incertains, elle ne peut pas être appliquée directement. Afin de pallier ce problème, nous nous proposons d'utiliser la stratégie développée dans le chapitre II (cf. §.II.2 et §.II.5) pour déterminer la dimension stochastique efficace du problème. Nous rappelons que cette stratégie préconise l'utilisation de la méthode de criblage de Morris pour hiérarchiser à faible coût les paramètres incertains par ordre d'importance. Cette hiérarchisation est basée sur le calcul des statistiques des effets élémentaires. Ainsi, seuls les paramètres les plus importants sont considérés dans la procédure de calcul stochastique.

#### III.6.4.2 Identification de la dimension stochastique efficace

Afin de déterminer la dimension stochastique efficace du problème, nous appliquons la méthode de criblage de Morris. Dans ce contexte, le calcul des statistiques (i.e. moyenne et écart-type) des effets élémentaires associés aux 134 paramètres incertains est basé sur 6 trajectoires ou plans d'expériences de type *OAT* (One At time) probabilisés. Chaque plan d'expérience est constitué de 135 points. De plus, les coordonnées de ces points sont échantillonnées dans l'espace aléatoire standard uniforme dont chacune des directions est discrétisée en 4 niveaux. Ainsi, le nombre total d'appels au modèle éléments finis est 810. Les figures III.29a et III.29b illustrent respectivement la hiérarchisation des paramètres incertains en considérant les deux critères d'endommagement de la toiture (cf. §.III.5.2).



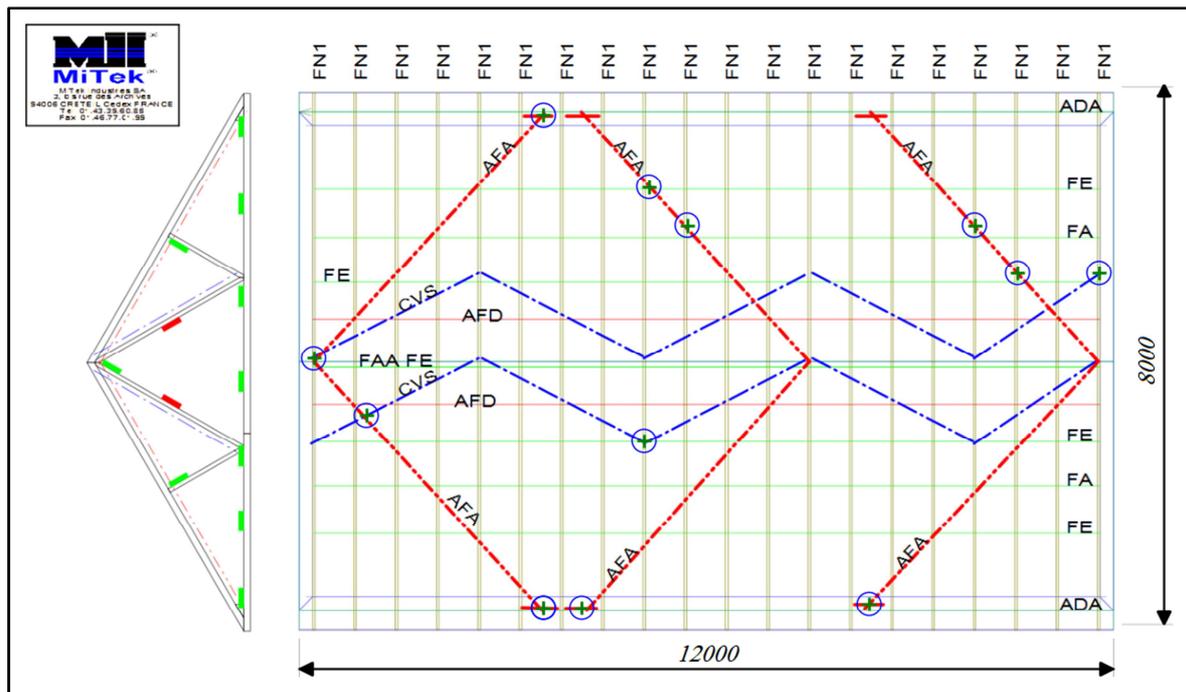
**Figure III. 29 :** Hierarchisation des paramètres incertains du modèle basée sur les mesures  $\mu_i^*$  et  $\sigma_i$  pour l'endommagement enregistré au niveau des éléments (a) d'anti-flambement AFA, (b) de contreventement CVS

Les méthodes de criblage permettent une évaluation qualitative de la sensibilité d'un modèle vis-à-vis de ses paramètres incertains. Elles servent plus à identifier les paramètres dont l'influence n'est pas significative sur la réponse du modèle qui sont les paramètres situés au voisinage de l'origine des graphes définis en abscisses et en ordonnées respectivement par la moyenne standardisée  $\mu_i^*$  et l'écart-type  $\sigma_i$  des effets élémentaires. L'analyse des figures III.29a et III.29b nous a permis de déterminer les paramètres incertains les plus influents sur la résistance de la toiture à ossature bois. Ils sont donnés dans le tableau III.12.

**Tableau III. 12 :** Identification des paramètres importants sur la résistance de la toiture

<b>N° VA</b>	<b>X<sub>9</sub></b>	<b>X<sub>10</sub></b>	<b>X<sub>16</sub></b>	<b>X<sub>24</sub></b>	<b>X<sub>32</sub></b>
<b>Assemblage</b>	AFA	AFA	AFA	AFA	AFA
<b>Type</b>	Clous $\phi 3,1 \times 90$				
<b>N° VA</b>	<b>X<sub>35</sub></b>	<b>X<sub>39</sub></b>	<b>X<sub>41</sub></b>	<b>X<sub>44</sub></b>	<b>X<sub>53</sub></b>
<b>Assemblage</b>	AFA	AFA	ADA	ADA	CVS
<b>Type</b>	Clous $\phi 3,1 \times 90$	Clous $\phi 2,8 \times 70$			
<b>N° VA</b>	<b>X<sub>58</sub></b>	<b>X<sub>67</sub></b>	<b>X<sub>68</sub></b>	<b>X<sub>93</sub></b>	<b>X<sub>94</sub></b>
<b>Assemblage</b>	CVS	CVS	CVS	liteaux	lisses
<b>Type</b>	Clous $\phi 2,8 \times 70$	Clous $\phi 2,8 \times 70$	Clous $\phi 2,8 \times 70$	Clous $\phi 2,1 \times 50$	Clous $\phi 2,8 \times 70$

Du point de vue de la répartition de l'endommagement dans la structure, ces résultats, sont en accord avec les observations réalisées durant les essais sur maquettes de la toiture à ossature bois [160]. En effet, on constate que les assemblages situés sur les éléments d'anti-flambement et les éléments de contreventement sont les plus sollicités par l'action sismique. De plus, l'aléa associé au comportement non linéaire de ces assemblages a un effet significatif sur la variabilité des deux indicateurs de l'endommagement. Cela est plus clair sur la figure III.30 qui illustre la localisation des assemblages les plus importants pour la résistance de la toiture. Ainsi, la dimension stochastique est passée de 134 à 15 variables aléatoires. Nous précisons que les simulations sont effectuées avec le signal sismique désigné par séisme 1 dans le tableau III.8.



**Figure III. 30 :** Localisation des assemblages les plus importants pour la résistance de la toiture à ossature bois

Dans la procédure de calcul stochastique, nous ne modéliserons que ces 15 paramètres importants par des variables aléatoires. Elles auront les distributions et les caractéristiques statistiques présentées dans le tableau III.10. Ainsi, le nombre d'appels au modèle éléments finis sera considérablement réduit.

### III.6.4.3 Quantification de l'aléa associé aux indicateurs d'endommagement

Après avoir déterminé la dimension stochastique efficace du problème, nous nous proposons maintenant de quantifier la variabilité des indicateurs d'endommagement, induite par l'aléa associé au comportement des assemblages. Dans cette perspective, nous avons appliqué la méthode de réduction de la dimension à l'ordre  $s = 1$  pour calculer les caractéristiques statistiques des indicateurs d'endommagement dans les assemblages situés sur les éléments d'anti-flambement et ceux situés sur les éléments de contreventement. Le choix d'une réduction d'ordre 1 est justifié par le fait qu'on a montré, notamment dans les applications traitées dans le chapitre précédent dans le contexte d'analyse de tendance centrale, que cet ordre garantit une bonne précision sur les estimations des deux premiers moments statistiques. Les intégrales élémentaires de la méthode de réduction de la dimension, notamment des intégrales unidimensionnelles, sont calculées par quadrature de Gauss-Hermite d'ordre 5. Ce choix est basé sur une étude de convergence concernant les estimations des deux premiers moments statistiques en faisant varier le nombre de points d'intégration de 3 à 5. Par conséquent, le nombre total d'appels au modèle éléments finis est 76. Le tableau III.13 donne les résultats obtenus pour les huit signaux sismiques représentant le scénario lointain de la zone d'aléa fort associée à la ville du Moule en Guadeloupe. On note que, pour chaque signal sismique, les résultats sont obtenus pour le *PGA* critère fourni dans le tableau III.8.

**Tableau III. 13 :** Estimations des deux premiers moments statistiques des deux indicateurs d'endommagement de la toiture pour les signaux de la ville du Moule en Guadeloupe – scénario lointain

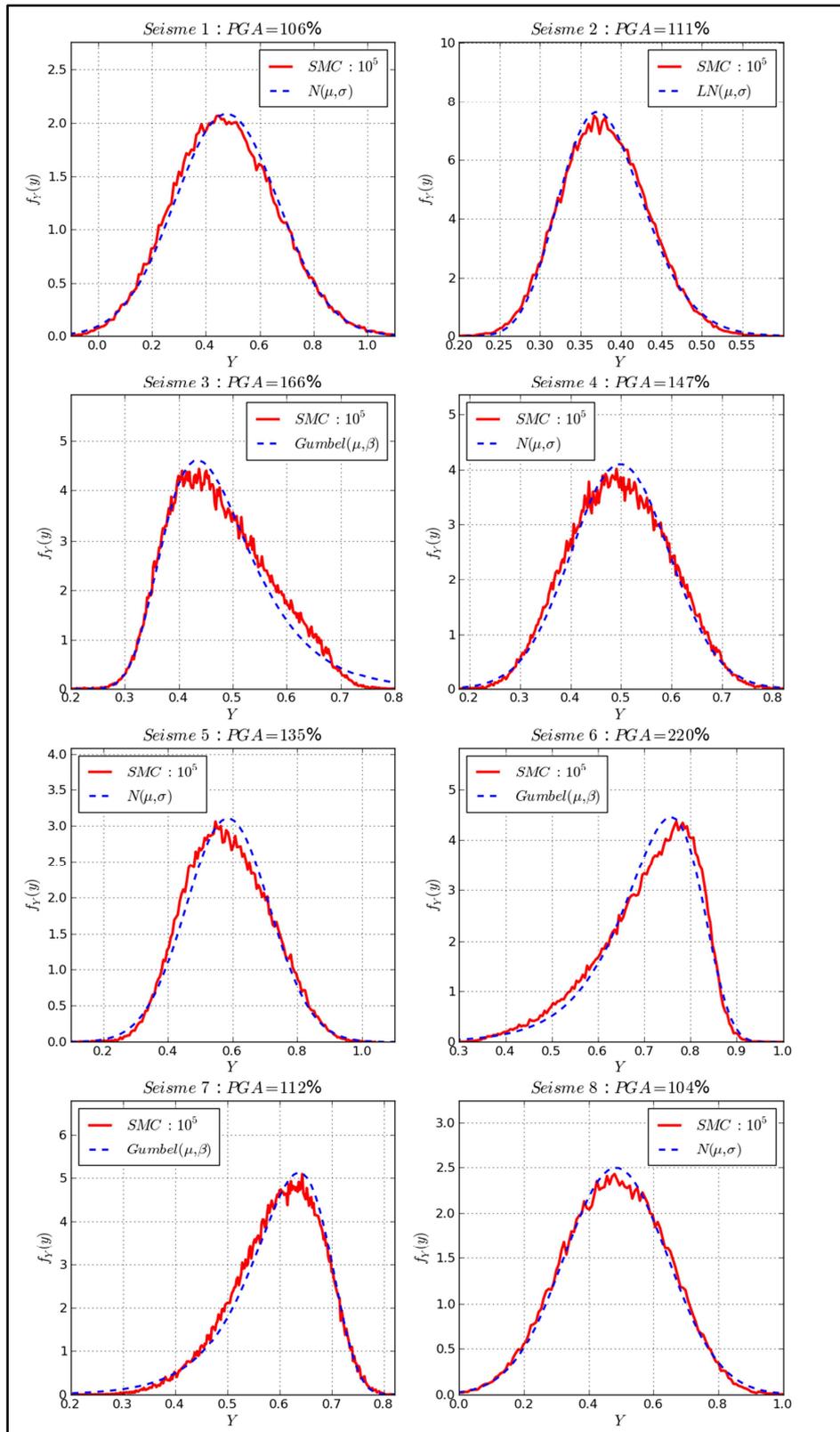
<i>Séisme</i>	<i>1</i>		<i>2</i>		<i>3</i>		<i>4</i>	
<i>Critère</i>	<i>AFA</i>	<i>b-CVS</i>	<i>AFA</i>	<i>b-CVS</i>	<i>AFA</i>	<i>b-CVS</i>	<i>AFA</i>	<i>b-CVS</i>
$\mu_Y$	0,525	0,477	0,474	0,381	0,674	0,478	0,542	0,498
$\sigma_Y$	0,044	0,191	0,020	0,053	0,023	0,091	0,022	0,097
<i>CV (%)</i>	8,4	40	4,2	14	3,4	19	4	19,5
<i>Séisme</i>	<i>5</i>		<i>6</i>		<i>7</i>		<i>8</i>	
<i>Critère</i>	<i>AFA</i>	<i>b-CVS</i>	<i>AFA</i>	<i>b-CVS</i>	<i>AFA</i>	<i>b-CVS</i>	<i>AFA</i>	<i>b-CVS</i>
$\mu_Y$	0,556	0,585	0,495	0,709	0,501	0,597	0,536	0,484
$\sigma_Y$	0,033	0,129	0,030	0,106	0,016	0,084	0,056	0,159
<i>CV (%)</i>	6	22	6	15	2,7	14,1	10,5	32,8

Nous constatons que les résultats obtenus sont en bon accord avec le calcul déterministe effectués dans le paragraphe III.5.4. En effet, si nous nous référons aux valeurs moyennes des indicateurs d'endommagement, nous constatons que la réalisation des deux critères de défaillance pour chaque signal sismique garde la même chronologie que lors du calcul déterministe. De plus, on constate que l'effet de l'aléa associé à la force au pic  $F_1$  n'est pas significatif sur la variabilité de l'endommagement accumulé dans les assemblages situés sur les éléments d'anti-flambement, puisqu'il ne dépasse pas 11%. Ceci peut être expliqué par le fait que, parmi tous les groupes d'assemblages constituant la toiture bois, ces derniers atteignent en premier le critère de ruine (i.e.  $D = 0,5$ ) durant la période efficace du signal sismique. En effet, pour la plupart des signaux sismiques, le *PGA* induisant la défaillance de la structure, correspond au *PGA* de réalisation du deuxième critère de défaillance (endommagement des assemblages situés aux extrémités des éléments de contreventement) qui est supérieur au *PGA* de réalisation du première critère de défaillance (endommagement des assemblages situés dans les éléments d'anti-flambement). Par conséquent, les assemblages situés dans les éléments d'anti-flambement sont les plus exposés à l'endommagement, et l'intensité du signal sismique prend le dessus sur l'effet de l'aléa associé au comportement non linéaire des assemblages. Cette constatation est d'autant plus claire pour les signaux sismiques 3 et 7 puisque les coefficients de variations sont les plus faibles. Ce résultat est important dans la mesure où, si nous choisissons de concevoir la structure vis-à-vis à ce critère de défaillance uniquement, il serait possible de fixer les paramètres incertains à leurs valeurs moyennes respectives.

Concernant le deuxième indicateur de résistance, représenté par l'endommagement accumulé dans l'ensemble des assemblages situés sur les extrémités des éléments de contreventement, nous constatons que sa variabilité est significative puisqu'elle dépasse 14%. Ceci est très clair pour les signaux sismiques 1 et 8 pour lesquels les coefficients de variation atteignent respectivement 40% et 32%. Il faut signaler que ces deux signaux sont les plus nocifs, puisque les *PGA* (i.e. les amplifications du *PGA* du spectre) induisant la défaillance sont les plus bas (cf. tableau III.8). Ces résultats sont importants dans la mesure où ils fournissent des informations supplémentaires sur l'endommagement accumulé dans la structure. Cependant, cette observation ne peut pas être généralisée à tous les signaux sismiques, dans le but de trouver un lien entre la criticité de l'évènement sismique et la variabilité des indicateurs de résistance de la structure. Autrement dit, le classement des signaux sismiques par ordre croissant de *PGA* induisant la défaillance de la structure ne correspond pas à celui effectué en se référant à la valeur du coefficient de variation des indicateurs de résistance de la structure. Néanmoins, chercher à établir une correspondance entre le *PGA* de défaillance et la variabilité des indicateurs de résistance, nécessite peut-être l'introduction d'autres éléments de

comparaison tels que la variabilité intrinsèque du signal sismique. Cependant, il faut signaler, à ce niveau de l'étude, l'apport du calcul stochastique par rapport au calcul déterministe. En effet, outre le fait que les résultats du calcul stochastique offrent un outil d'aide à la décision pour les concepteurs, ils peuvent être utilisés comme moyen de classification des signaux sismiques pour un site donné. Dans ce contexte, la nocivité d'un événement sismique ne sera pas évaluée uniquement par le niveau de *PGA* et le niveau d'endommagement qu'il induit dans la structure, mais aussi par le niveau de l'aléa sur cet endommagement.

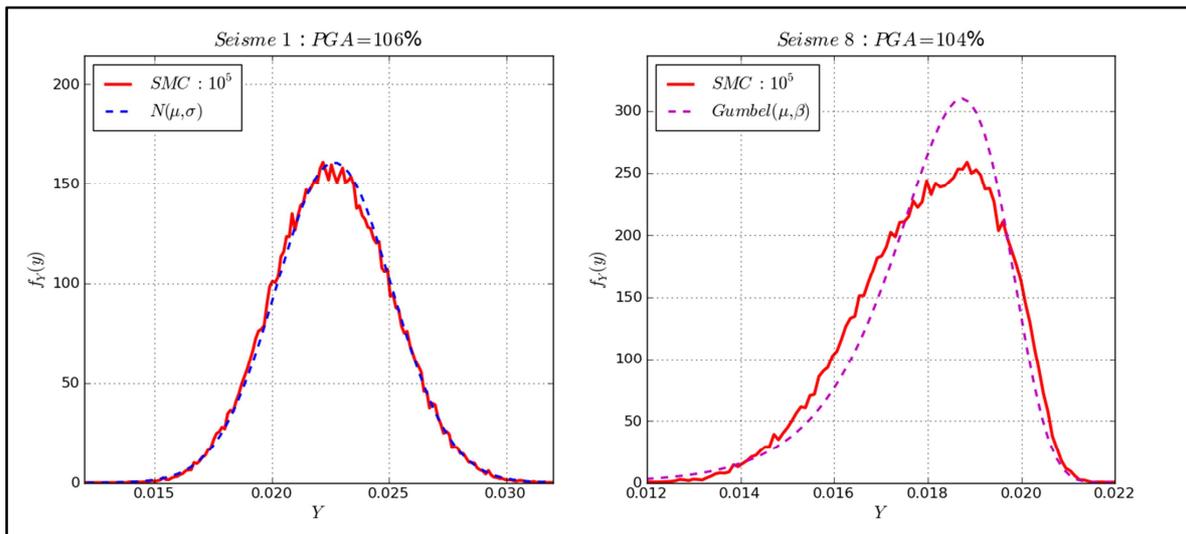
L'expérience a montré, notamment lorsque nous procédons à des analyses stochastiques, que la façon la plus convenable pour caractériser une grandeur aléatoire est de construire sa densité de probabilité. Ceci est possible en faisant appel à la méthode de décomposition de la dimension comme nous l'avons vu dans le deuxième chapitre de ce manuscrit. En effet, à partir des 76 réponses mécaniques, utilisées pour l'estimation des deux premiers moments statistiques, nous pouvons, par projection sur une base de polynômes de Lagrange, construire une représentation explicite ou un méta-modèle pour les deux indicateurs de résistance de la structure. Sur ce méta-modèle, des simulations de Monte-Carlo sont effectuées pour construire la densité de probabilité. Cette procédure ne nécessite pas d'évaluations supplémentaires du modèle éléments finis ; de plus elle n'est pas coûteuse en temps de calcul puisqu'on manipule des expressions analytiques. Bien que cette procédure ait été appliquée pour les deux indicateurs utilisés pour l'évaluation de l'intégrité de la toiture, nous nous restreignons dans cette étude à présenter les résultats concernant l'endommagement accumulé dans les assemblages situés dans les extrémités des éléments de contreventement, étant donné sa forte variabilité. La figure III.31 illustre les évolutions des densités de probabilité du deuxième indicateur de résistance pour les huit signaux sismiques de la ville du Moule. On constate que la densité de probabilité fournit des informations supplémentaires sur l'aléa de la réponse de la structure, puisqu'elle permet de visualiser l'endroit où la masse de probabilité est concentrée. En effet, pour les signaux sismiques 3, 6 et 7, on constate que la majeure partie de la masse de probabilité est concentrée au voisinage des valeurs extrêmes. Dans ces cas, le concepteur doit se focaliser sur ces régions pour cerner l'événement de défaillance, plutôt que de cibler le voisinage de la valeur moyenne. Comme nous pouvons l'observer sur la figure III.31, nous avons approximé les densités de probabilité par des lois normale, lognormale et Gumbel en se basant sur les moments statistiques. Nous constatons que ces lois usuelles sont en bon accord avec les résultats donnés par les simulations de Monte-Carlo. Ces approximations, assez précises, sont faciles à exploiter, et ne demandent pas de connaissances mathématiques approfondies. Nous notons tout de même que les densités de probabilités obtenues par les simulations de Monte-Carlo peuvent ne pas être assez précises aux extrémités des distributions. Ceci ne doit pas être considéré comme une limitation, dans le sens où, dans les analyses de tendance centrale, ces régions n'ont pas d'effet significatif sur la précision des résultats. Par contre, la précision sur les queues des distributions est essentielle dans les analyses de fiabilité. Cette difficulté peut être surmontée par l'utilisation d'une décomposition de la dimension d'ordre 2 pour la construction d'une représentation explicite des indicateurs de résistance de la structure au lieu d'une décomposition d'ordre 1. Cependant, cette alternative n'est pas convenable en terme de nombre d'appels au modèle éléments finis.



**Figure III. 31 :** Densité de probabilité du deuxième indicateur d'endommagement pour les signaux sismiques de la ville du Moule – scénario lointain

En plus de l'endommagement accumulé au niveau des assemblages, les ingénieurs s'intéressent également au déplacement maximal global de la toiture pour évaluer sa résistance pendant l'évènement sismique. En effet, cet indicateur de résistance est plus

accessible car en réalité l'endommagement ne peut pas être directement mesuré, et qu'on se restreint le plus souvent à mesurer les déplacements de points particuliers de la structure. Dans ce contexte, nous nous intéressons à identifier l'aléa sur le déplacement maximal du point central de la toiture induit par la dispersion au niveau du comportement des assemblages. Les figures III.32a et III.32b illustrent, respectivement pour les signaux sismiques 1 et 8, l'évolution de sa densité de probabilité construite par  $10^5$  simulations de Monte-Carlo sur le méta-modèle obtenu par projection sur les polynômes de Lagrange. Nous rappelons que les calculs stochastiques ont été effectués pour les *PGA* induisant la défaillance de la structure.



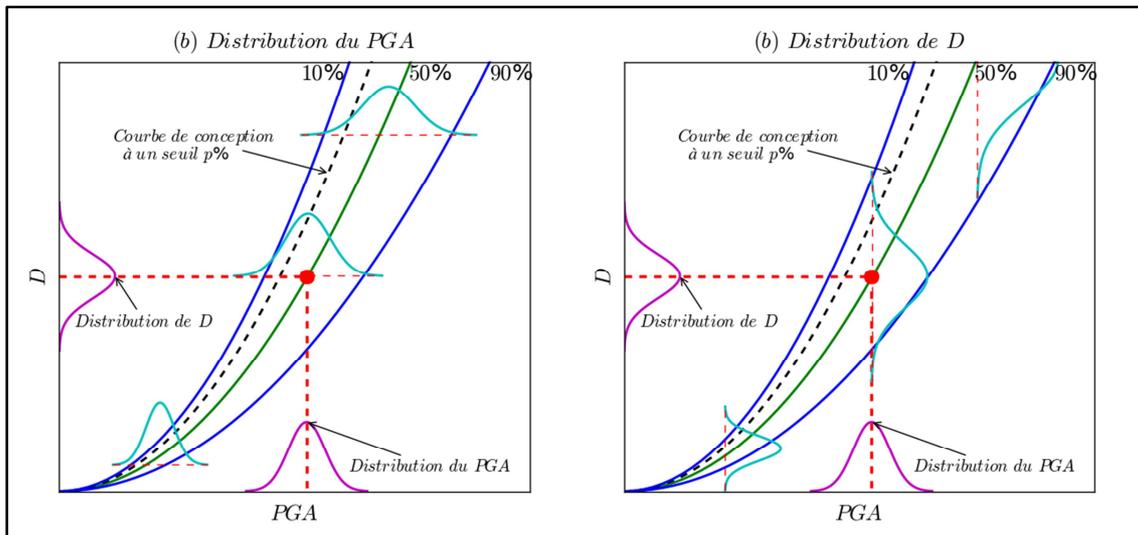
**Figure III. 32 :** Evolutions de la densité de probabilité du déplacement maximal de la structure pour les signaux sismiques 1 et 8

L'effet de l'aléa associé au comportement non linéaire des assemblages constituant la toiture n'est pas significatif sur la variabilité du déplacement maximal, puisque le coefficient de variation ne dépasse pas 11%. Ceci peut être expliqué par le fait que nous n'avons conservé, dans le calcul stochastique, que l'aléa le plus important sur la résistance des assemblages situés sur les éléments d'anti-flambement et les éléments de contreventement, qui n'a pas forcément le même effet sur la variabilité du déplacement maximal. Dans le cas du signal sismique 1, nous constatons que la masse de probabilité est concentrée autour de la valeur médiane, tandis que pour le signal sismique 8, elle est plutôt concentrée à la queue représentant les valeurs maximales de la distribution. Afin, de faciliter l'exploitation des densités de la probabilité, nous les avons approximé par des distributions usuelles. Ainsi, pour le signal sismique 1, nous avons pu représenter correctement le déplacement maximal, par une distribution Gaussienne. Cependant, pour le signal sismique 8, une distribution de Gumbel, qui est la plus adaptée parmi les distributions usuelles, ne permet de représenter que les deux queues de la densité de probabilité obtenue par simulations de Monte-Carlo.

#### III.6.4.4 Construction des courbes probabilistes d'endommagement

Le calcul stochastique mené dans le paragraphe précédent nous a fourni des informations supplémentaires sur l'endommagement accumulé dans la toiture sollicitée par une action sismique. Dans le but de valoriser les résultats obtenus, mais surtout de les rendre plus accessibles aux ingénieurs, nous proposons dans ce paragraphe une procédure de construction de courbes probabilistes d'évolution de l'endommagement en fonction du *PGA* du signal sismique. Ces dernières sont obtenues à partir d'un post-traitement des résultats du modèle éléments finis. Elles permettent, d'une part de construire la distribution du *PGA* permettant de

garantir un niveau donné d'endommagement, et, d'autre part, de construire la distribution de l'endommagement pour un  $PGA$  cible. A partir de ces distributions, on peut déterminer, pour un seuil de confiance donné, la courbe d'évolution des indicateurs d'endommagement dans la structure en fonction du  $PGA$ . Cette procédure est résumée sur la figure III.33.

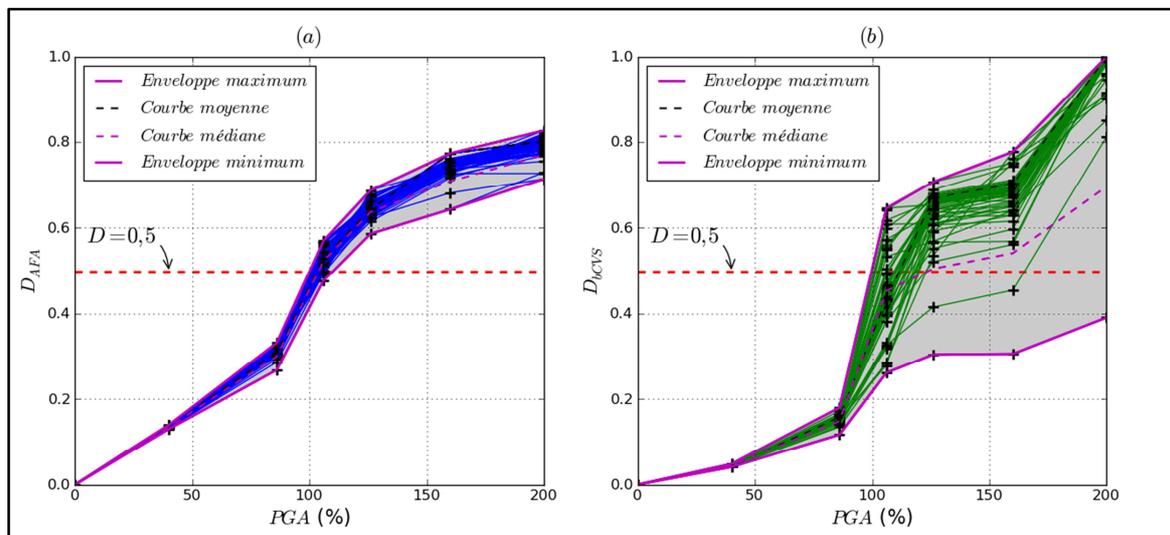


**Figure III. 33 :** Illustration schématique de la procédure de construction des courbes probabilistes d'évolution de l'endommagement

L'idée adoptée pour construire les courbes d'évolution de l'endommagement en fonction du  $PGA$  du signal sismique consiste à utiliser les couples de points ( $PGA, D$ ) correspondant aux réponses du modèle éléments finis associées aux différentes combinaisons des paramètres incertains générées dans l'espace aléatoire lors du calcul des moments statistiques. Ainsi, le calcul des moments statistiques est effectué pour plusieurs niveaux distincts de  $PGA$ , choisis en se basant sur l'allure des courbes déterministes d'endommagement (i.e. courbes utilisées dans la détermination de  $PGA$  de défaillance cf. figure III.25), et en se focalisant sur le voisinage du  $PGA$  induisant la défaillance de la structure. Nous disposons pour chaque niveau de  $PGA$ , de 76 valeurs pour chacun des deux indicateurs d'endommagement de la structure. Cette procédure est appliquée pour les signaux sismiques 1 et 8, puisqu'ils sont les plus endommageants pour la structure d'une part, et parce que leur dispersion est la plus significative pour les deux indicateurs d'endommagement d'autre part.

Les figures III.34a et III.34b illustrent, dans le cas du séisme 1, l'évolution des courbes probabilistes d'endommagement associées aux deux critères de défaillance de la structure. Nous constatons, pour un  $PGA$  donné, que l'effet de l'aléa qui entache le comportement non linéaire des assemblages élémentaires est plus significatif sur l'endommagement cumulé dans les assemblages situés sur les éléments de contreventement que celui enregistré dans les assemblages situés sur les éléments d'anti-flambement. De plus, nous constatons que les courbes d'évolution de l'endommagement se croisent par moments, tandis que nous nous attendons à ce qu'elles restent parallèles lorsque le  $PGA$  du signal sismique augmente. Ceci est plus visible lorsque le  $PGA$  du signal sismique dépasse le  $PGA$  induisant la réalisation des deux critères de défaillance de la structure. En effet, au-delà de ce niveau d'amplification du signal sismique, d'autres assemblages prennent le relais pour encaisser l'effet du chargement, ce qui induit une redistribution de l'endommagement sur l'ensemble de la structure. Ceci est justifié par les résultats présentés sur les figures III.34a et III.34b : nous constatons qu'une fois que l'endommagement dans les assemblages situés sur les éléments d'anti-flambement a atteint un niveau critique, l'endommagement dans les assemblages situés dans les éléments de contreventement enregistre une augmentation brusque.

Si nous comparons, pour le même jeu de paramètres aléatoires, les courbes d'évolution des deux indicateurs d'endommagement, nous observons que la réalisation des deux critères de défaillance conserve la même chronologie que lors du calcul déterministe (cf. §.III.5.4). Autrement dit, pour les 76 réalisations des paramètres aléatoires d'entrée, la défaillance des assemblages situés sur les éléments d'anti-flambement est enregistrée avant celle des assemblages situés sur les éléments de contreventement. Lorsque le  $PGA$  augmente, les réalisations des deux indicateurs d'endommagement sont plutôt concentrées au voisinage de la courbe enveloppe maximum et non pas autour de la courbe médiane. En effet, lorsque l'amplification du signal sismique augmente, l'endommagement résultant dans les assemblages est de plus en plus significatif et tend vers une valeur asymptotique qui est en principe égale à l'unité. De plus, dans cette situation, l'effet du signal sismique prend le dessus sur l'aléa dans le comportement des assemblages. Ceci est justifié par le fait que la réponse de la structure est presque invariable pour la plupart des jeux des paramètres incertains.

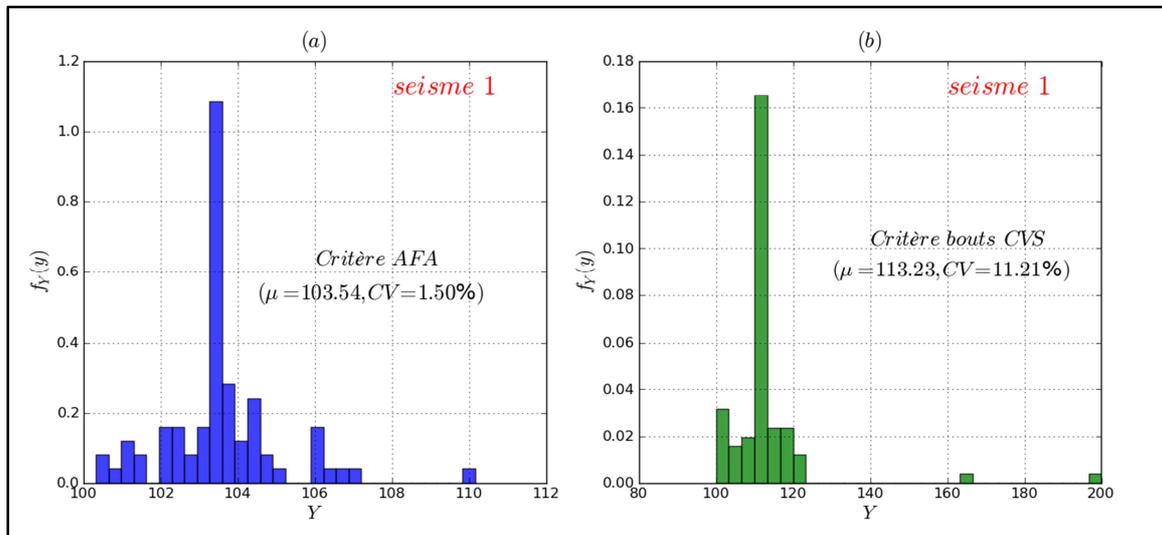


**Figure III. 34 :** Courbes probabilistes d'endommagement associées aux deux critères de défaillance de la structure – signal sismique 1

L'avantage majeur des courbes probabilistes d'évolution de l'endommagement, réside dans le fait que pour un  $PGA$  donné, nous pouvons construire la distribution des indicateurs d'endommagement, et vice-versa. En effet, les figures III.35a et III.35b illustrent la distribution du  $PGA$  du signal sismique, lorsque les deux indicateurs de défaillance de la structure atteignent un seuil  $D = 0,5$ . Bien que le nombre de réalisations soit trop limité pour pouvoir construire la distribution réelle, nous obtenons tout de même des informations supplémentaires concernant le  $PGA$  induisant la défaillance de la structure, notamment, sa moyenne et son coefficient de variation. En se référant à ces deux grandeurs, nous constatons que la dispersion du  $PGA$  pour le deuxième indicateur de défaillance est plus élevée que celle enregistrée pour le premier indicateur de défaillance. Cependant, cette dispersion n'est pas significative, puisque le coefficient de variation ne dépasse pas 12%. Nous constatons, de plus, que toutes les réalisations du  $PGA$  induisant la réalisation des deux critères de défaillance sont supérieures à 100%. Ceci veut dire que la structure telle quelle est conçue résiste à l'évènement sismique représenté par le signal sismique 1 et pour le scénario de défaillance adopté.

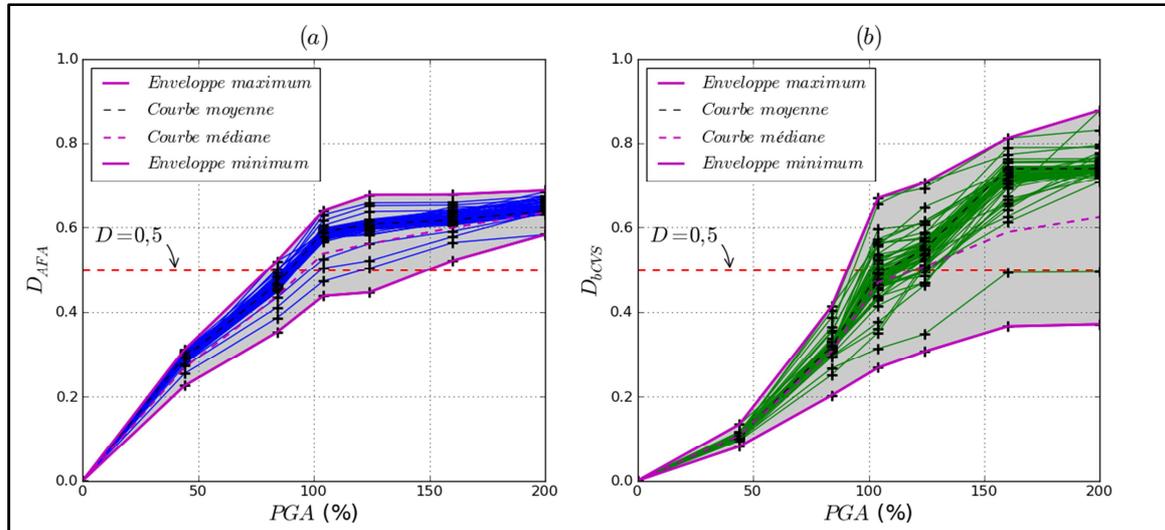
Le problème du nombre limité de réalisations utilisées dans la construction des distributions du  $PGA$  présentées sur les figures III.35a et III.35b, peut être surmonté si nous arrivons, pour chaque niveau de  $PGA$ , à construire avec précision la distribution de l'endommagement. En

effet, une fois ces distributions obtenues et approximées par des distributions usuelles, nous pouvons générer le nombre désiré de points. Cependant, plus nous augmentons le *PGA* du signal sismique, plus la densité de probabilité de l'endommagement est concentrée dans les queues de la distribution et il est difficile d'identifier une loi usuelle qui peut la représenter.



**Figure III. 35 :** Distributions du *PGA* du signal sismique pour un niveau d'endommagement seuil  $D = 0,5$  – signal sismique 1

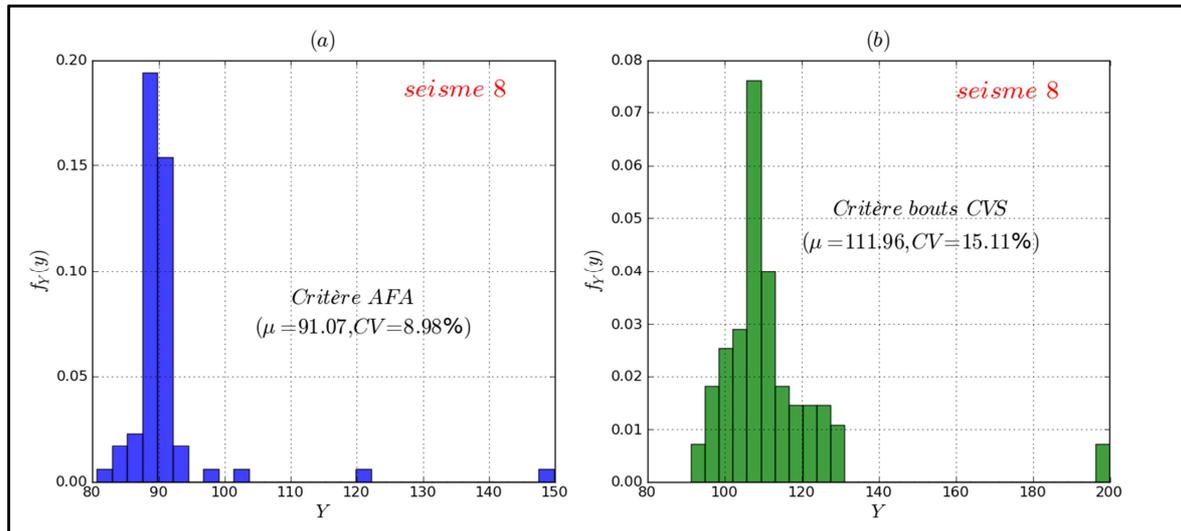
Nous nous intéressons maintenant à l'évènement sismique représenté par le signal 8. Les figures III.36a et III.36b illustrent l'évolution des courbes probabilistes d'endommagement associées aux deux critères de défaillance.



**Figure III. 36 :** Courbes probabilistes d'endommagement associées aux deux critères de défaillance de la structure – signal sismique 8

Comme pour le signal sismique 1, nous constatons que, pour un *PGA* donné, la dispersion de l'endommagement est plus significative dans les assemblages situés dans les éléments d'anti-flambement que celle de l'endommagement dans les assemblages situés dans les éléments de contreventement. Toutefois, nous constatons que la majeure partie des réalisations des courbes d'endommagement est concentrée au voisinage de la courbe médiane, contrairement au premier cas traité pour lequel ces dernières sont concentrées au voisinage de la courbe enveloppe maximum (sachant que dans les deux cas nous partons des mêmes réalisations des

paramètres incertains d'entrée). Cette constatation est précieuse car elle peut valider l'hypothèse de dimensionner la structure en se référant à la courbe médiane. Cependant, il est important de signaler que, si nous considérons un niveau d'endommagement seuil  $D = 0,5$ , nous constatons que plusieurs réalisations des courbes probabilistes d'endommagement donnent un  $PGA$  de défaillance inférieur à 100% (cf. figures III.37a et III.37b). Ainsi, si nous choisissons de dimensionner la structure vis-à-vis de la courbe d'endommagement médiane, le niveau de fiabilité serait trop faible pour des applications réelles.



**Figure III. 37 :** Distributions du  $PGA$  du signal sismique pour un niveau d'endommagement seuil  $D = 0,5$  – signal sismique 8

Dans ce qui a précédé, nous avons montré l'importance des informations que peuvent fournir les courbes probabilistes d'endommagement, non seulement dans l'évaluation de la résistance de la structure, mais aussi comme aide pour faire les choix adéquats lors de la conception de la structure. En plus des distributions du  $PGA$  pour un niveau d'endommagement seuil, et de la distribution des indicateurs de défaillance pour un  $PGA$  cible, nous pouvons construire pour un seuil de confiance préalablement fixé, la courbe d'évolution de l'endommagement en fonction de l'amplification d'un signal sismique associé à un site donné.

### III.7 Conclusion

Dans ce chapitre, nous nous sommes focalisés sur l'analyse stochastique de structures bois constituées d'assemblages par connecteurs métalliques. Nous avons analysé le comportement de structures bois de différentes échelles allant d'un simple assemblage constitué de deux éléments en bois connectés par une plaque métallique à dents, jusqu'à la toiture à ossature bois d'une habitation individuelle, en passant par une structure en treillis représentant une ferme en bois constituée de plusieurs assemblages. De plus, nous avons traité les trois aspects du calcul stochastique que sont l'analyse de sensibilité basée sur la décomposition de la variance, l'analyse de tendance centrale visant la quantification de l'aléa de la réponse d'un modèle, et l'analyse de fiabilité visant l'évaluation de l'intégrité d'une structure vis-à-vis d'un critère de défaillance. Nous avons mis en œuvre plusieurs approches de couplage mécano-probabiliste, notamment des approches non intrusives telles que le développement en chaos polynômial et la décomposition de la dimension.

Dans un premier temps, nous avons effectué des analyses de sensibilité sur un assemblage élémentaire basées sur le calcul des indices de sensibilité de Sobol au moyen du post-traitement des coefficients d'un développement en chaos polynômial. Ces analyses nous ont

permis de déterminer les paramètres incertains de la loi de comportement avec hystérésis contribuant le plus dans la variabilité de la résistance de l'assemblage et de l'énergie dissipée. Lorsque l'endommagement est faible, nous avons montré que l'aléa associé au paramètre  $K_0$  de la loi de comportement, correspondant à la rigidité initiale, est le plus influant sur la variabilité de la capacité résistante de l'assemblage et le cumul de l'énergie dissipée. Mais, au fur et à mesure que l'endommagement accumulé dans l'assemblage devient significatif, nous avons observé une redistribution des rôles puisque la variabilité de la capacité résistante de l'assemblage est due principalement à l'aléa associé au paramètre  $F_1$  de la loi de comportement correspondant à la force au pic. Cette étude a permis de conclure que la variabilité du comportement non linéaire de l'assemblage est due principalement à l'aléa de ce paramètre. Dans un deuxième temps, nous avons étudié la fiabilité d'une fermette en bois sollicitée par une action sismique pour laquelle on a montré que la probabilité de défaillance des assemblages entre les fiches et les entrants est la plus significative.

La particularité du développement effectué dans ce chapitre est la résolution du problème de dimension stochastique élevée pour laquelle nous avons adopté une stratégie en deux étapes permettant dans un premier temps de déterminer la dimension stochastique efficace du problème par une méthode de criblage et d'effectuer dans un deuxième temps une analyse stochastique proprement dite par la méthode de réduction de la dimension. Cette stratégie, qui permet d'optimiser le temps de calculs en terme de nombre d'appels au modèle éléments finis, comme nous l'avons montré dans le deuxième chapitre de ce manuscrit, a été appliquée pour l'analyse stochastique d'une toiture à ossature bois installée sur un site ayant un aléa sismique fort. L'action sismique est modélisée par des accélérogrammes naturels correspondant à des mouvements sismiques sur le site du Moule en Guadeloupe et l'évaluation du comportement de la toiture est basée sur une analyse chronologique (i.e. un calcul dynamique non linéaire). L'intégrité de la structure est évaluée par deux indicateurs de résistance mesurant l'endommagement accumulé dans les assemblages situés sur les éléments d'anti-flambement et les extrémités des éléments de contreventement. La défaillance est enregistrée lorsque ces deux indicateurs de résistance atteignent une valeur supérieure ou égale à 0,5. Nous avons montré, en nous basant sur des analyses déterministes du comportement vis-à-vis des signaux sismiques représentatifs de la ville du Moule, que la dispersion sur le *PGA* induisant la défaillance est significative et, qu'il faut en tenir compte dans la conception de la structure.

Grâce à une étude préliminaire basée sur la hiérarchisation des paramètres incertains par la méthode de criblage de Morris, nous avons montré que l'effet de l'aléa associé aux paramètres intervenant dans la modélisation du comportement non linéaire des assemblages est plus significatif que l'effet de la variabilité dans le module d'Young des éléments en bois. Compte tenu de ce résultat, qui nous a permis de fixer les modules des éléments en bois à leurs valeurs moyennes et par conséquent de réduire le nombre de paramètres incertains, nous avons appliqué la stratégie de résolution du problème de dimension stochastique élevée pour évaluer la variabilité de la réponse mécanique de la structure. Dans ce contexte, nous avons pu classer les sources d'aléa potentielles au niveau des assemblages de la toiture, et nous avons caractérisé d'une manière adéquate la variabilité qu'elles induisent sur les indicateurs d'endommagement. Pour ce dernier point, nous avons calculé les deux premiers moments statistiques. Les résultats obtenus ont montré l'importance du calcul stochastique et de l'intégration des sources de dispersion dans le calcul mécanique. En effet, nous avons montré que l'aléa associé à un indicateur d'endommagement varie d'un événement sismique à un autre. La dispersion du niveau d'endommagement accumulé dans les assemblages situés aux extrémités des éléments de contreventement peut atteindre 40%, tandis qu'elle ne dépasse pas 11% pour les assemblages situés sur les éléments d'anti-flambement. Ce résultat est très important dans la mesure où il peut consolider soit l'hypothèse de fixer les paramètres

incertains à leurs valeurs moyennes si la conception de la structure doit s'effectuer vis-à-vis du premier critère de défaillance, soit la prise en compte de l'effet de l'aléa dans le comportement des assemblages si le deuxième critère de défaillance est adopté.

En plus de l'estimation de la moyenne et de l'écart-type des indicateurs d'endommagement, nous avons pu construire leurs densités de probabilité en effectuant des simulations de Monte-Carlo sur un méta-modèle basé sur les polynômes de Lagrange. Ceci nous a permis d'obtenir des informations supplémentaires sur la variabilité de la réponse mécanique de la structure, puisqu'elles donnent la répartition de la masse de probabilité. En effet, pour certains événements sismiques, la variabilité de l'endommagement dans les assemblages situés dans les extrémités des éléments de contreventement suit des distributions extrêmes pour lesquelles la masse de probabilité est localisée dans l'une des queues de la distribution, plutôt qu'autour de la valeur moyenne. De plus, ces distributions ont été approximées par des lois usuelles, pour faciliter leur exploitation.

Par ailleurs, nous avons développé une procédure de post-traitement des résultats du calcul stochastique afin de rendre leur exploitation plus abordable et plus concrète. Cette procédure, qui consiste à effectuer des calculs stochastiques pour différents niveaux d'amplification du signal sismique, nous a permis de construire des courbes probabilistes d'endommagement. A partir de ces courbes, qui peuvent être assimilées à des *abaques de conception*, nous avons pu construire la distribution du *PGA*, sous la forme d'un histogramme, pour un niveau d'endommagement cible, et inversement. De plus, pour un seuil de confiance préalablement établi dans la phase de conception, nous pouvons construire les courbes d'évolution des deux indicateurs de défaillance en fonction de l'amplification du signal sismique. Nous signalons aussi, qu'à partir de ces courbes probabilistes d'endommagement, nous pouvons obtenir les *courbes de fragilité* [161] couramment utilisées dans l'analyse de la vulnérabilité des structures vis-à-vis d'un événement sismique. Elles représentent l'évolution de la probabilité de défaillance de la structure en fonction d'un des indicateurs (*PGA*, *PGV*, intensité d'Arias) de nocivité du signal sismique. Les résultats du calcul stochastique fournissent des informations supplémentaires sur l'endommagement cumulé dans la structure et constituent un outil d'aide à la décision dans la phase de conception de la structure. Plus généralement, le calcul stochastique permet de caractériser la criticité de l'aléa sismique associé à un site donné, notamment, en ajoutant d'autres paramètres d'évaluation, tels que les moments statistiques et la distribution du *PGA* induisant la défaillance de la structure.



---

## **Conclusion et perspectives**

---

De nos jours, les méthodes de calcul stochastique rencontrent de plus en plus de succès dans plusieurs disciplines tels que le génie mécanique et le génie civil. Ces méthodes permettent de quantifier d'une manière rationnelle l'effet des incertitudes qui peuvent entacher certains paramètres d'un modèle représentant un phénomène physique donné, en les intégrant dans le calcul mécanique. De plus, leur application peut avoir différentes finalités telles que l'analyse de tendance centrale, l'analyse de sensibilité, ou encore l'analyse de fiabilité vis-à-vis d'un scénario de défaillance. Cependant, certaines limites sont apparues. D'une part, il n'y pas une méthode *universelle* qui permette de traiter les trois finalités du calcul stochastique. D'autre part, si on écarte les simulations de Monte-Carlo et leurs variantes dont la convergence est lente, qui s'avèrent ainsi inapplicables pour les modèles exigeants en temps de calcul, la plupart des méthodes de calcul stochastique exhibent le problème de *la dimension stochastique élevée*. Ce problème peut être défini par l'augmentation exponentielle du nombre d'appels au modèle lorsque le nombre de paramètres incertains croît. Dans les applications pratiques, cette difficulté est prépondérante. En effet, on manque souvent d'informations sur le comportement des paramètres d'un modèle, et on est contraint de les modéliser tous par des variables aléatoires dans le calcul stochastique. De plus, la modélisation de certaines sources d'aléa nécessite l'emploi de champs aléatoires comme pour le module de Young, ou de processus aléatoires comme pour un chargement cyclique, qui nécessitent à leur tour une décomposition en variables aléatoires, dont le nombre peut être très élevé. Le problème de la dimension stochastique élevée s'impose donc dans ce travail, puisque l'objectif est d'évaluer et de quantifier l'effet de l'aléa associé au comportement des différents assemblages constituant une toiture à ossature bois sur son intégrité vis-à-vis d'un évènement sismique. En effet, même si l'aléa dans le comportement d'un assemblage élémentaire est représenté par une seule variable aléatoire, le nombre de paramètres incertains à l'échelle de la toiture avoisine les quelques centaines. Ce nombre affecte l'efficacité de la plupart des méthodes de calcul stochastique existantes, d'autant que l'évaluation du comportement de la structure est basée sur une analyse dynamique non linéaire, elle-même exigeante en temps de calcul.

Afin de pallier cette difficulté, la première étape de ce travail était consacrée à une revue bibliographique des méthodes de calcul stochastique disponibles, en nous intéressant aux trois finalités que peut avoir un calcul stochastique. La précision et l'efficacité de ces méthodes ont été évaluées en se basant sur des applications illustratives. Nous avons montré qu'à l'exception des simulations de Monte-Carlo dont la convergence est lente, l'efficacité de la plupart des méthodes de calcul stochastique est affectée lorsque le nombre de paramètres incertains est élevé. Des versions améliorées, telles que les tirages asymptotiques et les subset simulations, permettent d'améliorer la convergence des simulations de Monte-Carlo ; cependant, elles ne sont applicables que pour des analyses de fiabilité. Nous nous sommes ensuite focalisés sur les méthodes d'analyse de sensibilité, en particulier, les méthodes de criblage, pour lesquelles nous avons montré que les mesures de sensibilité qualitatives sont pertinentes, dans le sens où elles sont en accord avec les mesures de sensibilité globale basées sur la décomposition de la variance tels que les indices de Sobol. Elles permettent, par conséquent, de hiérarchiser les paramètres incertains d'un modèle, sans pour autant quantifier la contribution exacte (effet individuel et effet d'interaction) de chacun sur la variabilité de sa réponse. Sur ce point, nous avons montré que la méthode MSDP est la plus stable, tandis que la méthode de Morris offre le meilleur compromis entre précision et efficacité.

En nous basant sur ce constat, nous avons décidé de poser le problème de dimension stochastique élevée d'une façon plus intuitive : au lieu de chercher à développer une approche permettant d'aborder directement les différentes finalités du calcul stochastique tout en prenant en compte l'aléa associé aux différents paramètres incertains, il est plus pertinent de chercher, dans un premier temps, un moyen de réduire le nombre de paramètres incertains, et

de ne considérer dans un deuxième temps que les paramètres les plus importants dans le calcul stochastique proprement dit. Cette idée est soutenue également par le fait que, dans les problèmes d'ingénierie, l'expérience a montré que le plus souvent, parmi l'ensemble des paramètres incertains considérés dans le calcul stochastique, seul un nombre limité contribue efficacement à l'aléa observé sur la réponse du modèle. Ceci nous a amené à distinguer la dimension stochastique nominale de la dimension stochastique efficace. Ainsi, nous avons proposé une approche en deux étapes pour résoudre le problème de dimension stochastique élevée.

La première étape consiste à identifier les paramètres importants, en se basant sur une hiérarchisation des paramètres du modèle par la méthode de Morris. Dans le cas où les statistiques des effets élémentaires ne suffisent pas pour distinguer les paramètres importants, nous avons proposé un nouveau critère de sélection basé sur une mesure de sensibilité normalisée  $G_i^*$ . Une fois les paramètres importants identifiés, ils sont modélisés par des variables aléatoires et les autres paramètres moins influents sont fixés à leurs valeurs moyennes respectives.

La deuxième étape consiste à effectuer le calcul stochastique proprement dit. Dans ce contexte, nous avons retenu le développement en chaos polynômial et la méthode de décomposition de la dimension pour la construction d'un méta-modèle. Cette idée est motivée par le fait que le modèle éléments finis représentant le comportement de la toiture à ossature bois vis-à-vis d'une sollicitation sismique est coûteux en temps de calcul, et il est plus judicieux de le remplacer par un modèle explicite pour gagner en efficacité dans les analyses stochastiques. Nous avons montré que les deux méthodes nous permettent d'aborder les trois finalités que peut avoir le calcul stochastique. De plus, nous avons montré que la méthode de décomposition de la dimension offre le meilleur compromis entre efficacité et précision pour effectuer des analyses de tendance centrale ou des analyses de fiabilité, tandis que le développement en chaos polynômial est meilleur pour l'estimation des indices de sensibilité de Sobol.

L'approche en deux étapes proposée, est validée à l'aide d'une application traitant le comportement d'une structure à cinq niveaux sollicitée par des charges latérales, et pour laquelle nous avons considéré 21 paramètres incertains. Nous avons montré que cette approche est efficace, puisqu'elle permet d'économiser environ 92% du temps de calcul si la méthode de décomposition de la dimension est appliquée directement pour le calcul des moments statistiques de la réponse du modèle, sans passer par l'étape préliminaire d'identification des paramètres importants.

Le troisième chapitre de ce manuscrit était consacré à l'analyse stochastique des structures bois à différentes échelles. Dans un premier temps, des analyses de sensibilité globale, basées sur la décomposition de la variance, sont effectuées sur un assemblage bois élémentaire sollicité par un déplacement cyclique, afin de déterminer la contribution de l'aléa associé aux paramètres de la loi anisotrope avec hystérésis sur son comportement. L'estimation des indices de Sobol, basée sur un post-traitement sur les coefficients d'un développement en chaos polynômial obtenus par quadrature de Smolyak, a montré que l'aléa associé à la rigidité initiale  $K_0$  de l'assemblage est le plus contributif dans la variabilité de l'énergie dissipée durant le chargement, tandis que la variabilité de la capacité résistante de l'assemblage est due en grande partie à la force au pic  $F_1$ . Ce constat nous a amené à ne modéliser, à l'échelle d'un assemblage élémentaire, que le paramètre  $F_1$  par une variable aléatoire dans les analyses stochastiques ultérieures.

Dans un deuxième temps, nous avons évalué la fiabilité d'une ferme sollicitée par une action sismique représentant le tremblement de terre de Kobe 1995 (Japon). Nous avons

montré que la défaillance de la fermette est engendrée par la rupture des assemblages situés entre les fiches et l'entrait.

Enfin, nous nous sommes focalisés sur le problème posé dans le cadre du projet SISBAT, qui est l'analyse stochastique d'une toiture à ossature bois d'une habitation individuelle installée sur un site d'aléa sismique fort (Le Moule en Guadeloupe). L'intégrité de la structure est évaluée à travers deux indicateurs de résistance représentant l'évolution de l'endommagement cumulé dans les assemblages situés sur les éléments d'anti-flambement et ceux situés sur les éléments de contreventement. La défaillance de la structure est enregistrée lorsque ces deux indicateurs atteignent simultanément un seuil critique, préalablement établi. Une étude préliminaire basée sur la méthode de Morris a montré que l'aléa associé au comportement non linéaire des assemblages constituant la toiture est plus influant que l'aléa associé au module de Young des éléments bois. Ce constat nous a permis de réduire la dimension stochastique dans l'analyse stochastique de la toiture. Malgré cela, nous n'avons pu éviter le problème de dimension stochastique élevée, puisque le nombre de paramètres incertains dépasse 300.

Pour y remédier, nous avons fait appel à l'approche de calcul stochastique en deux étapes que nous avons développée dans le deuxième chapitre. Dans la première étape du calcul stochastique, nous avons identifié et localisé dans la toiture les assemblages élémentaires dont l'aléa contribue significativement à la variabilité de la réponse de la structure. Dans la deuxième étape, nous avons utilisé la méthode de décomposition de la dimension pour l'estimation des deux premiers moments statistiques des deux indicateurs de résistance de la toiture. Nous avons montré, pour le panel de signaux sismiques fournis par le BRGM, que l'effet de l'aléa est plus significatif sur l'endommagement accumulé dans les assemblages situés dans les éléments de contreventement. Ce constat est plus visible pour les signaux sismiques les plus nocifs sur la structure (ceux dont l'amplification du *PGA* de référence, permettant la réalisation des deux critères de défaillance, est la plus basse), puisque le coefficient de variation atteint 40%. Afin de mieux caractériser l'aléa associé à l'indicateur d'endommagement dans les assemblages situés sur les éléments de contreventement, nous avons construit la densité de probabilité, en effectuant des simulations de Monte-Carlo sur le méta-modèle obtenu par la méthode de décomposition de la dimension en conjonction avec les polynômes d'interpolation de Lagrange. Ceci nous a permis de visualiser les zones de concentration de la masse de probabilité. En effet, pour certains signaux, cette dernière est localisée dans l'une des queues de la distribution, plutôt qu'autour de la valeur moyenne. De plus, nous avons pu approximer les densités de probabilité par des distributions usuelles telles que les lois Gaussienne, lognormale et Gumbel. Nous avons proposé une méthodologie basée sur le post-traitement des résultats du calcul stochastique pour différents niveaux de *PGA* du signal sismique, pour construire des courbes probabilistes d'évolution des deux indicateurs de résistance de la structure. Ces courbes probabilistes, qui peuvent être assimilées à des *abaques de conception*, permettent de construire la distribution, sous forme d'histogramme, du *PGA* pour un niveau seuil d'endommagement et vice-versa. Nous avons montré, pour les événements sismiques pour lesquels nous avons enregistré une forte variabilité sur la réponse de la structure, que la dispersion du *PGA* pour un endommagement critique de 50% n'est pas significative, puisque le coefficient de variation ne dépasse pas, dans le pire des cas, 16%.

Dans ce travail, nous nous sommes attaqués au problème de la dimension stochastique élevée, qui compte parmi les problèmes les plus importants quand on procède à des analyses de propagation d'incertitudes sur des problèmes réels. Le niveau de complexité est amplifié lorsque le problème mécanique porte sur l'évaluation de l'intégrité des structures bois sous sollicitation sismiques. Nous avons apporté des solutions assurant un bon compromis entre l'efficacité, la précision et la simplicité. Malgré ces avancées, certaines ambiguïtés persistent, qui nous mettent en face de plusieurs défis, et les perspectives de ce travail sont variées. En

effet, nous nous sommes focalisés uniquement sur l'analyse de l'effet de l'aléa associé au comportement des assemblages qui constituent la structure. Cependant, des travaux [156] ont montré que l'action sismique exhibe un aléa significatif, dont l'origine peut être intrinsèque à l'action sismique elle-même et/ou au site sur lequel est installée la structure. Sur ce point, Sommerville [157] a montré que la variabilité de l'action sismique enregistrée sur plusieurs types de sols peut être plus significative que la variabilité associée à plusieurs signaux sismiques enregistrés sur un même site. Dans ce contexte, la variabilité associée au type de sol est prise en compte dans les analyses effectuées dans ce travail dans la mesure où les huit événements sismiques traités représentent différents types de sol selon la classification de l'Eurocode8. Cependant, ce nombre est relativement limité pour que la modélisation de cette source d'aléa soit pertinente. La variabilité intrinsèque de l'événement sismique peut être abordée en utilisant les processus aléatoires. Cependant, ce type de modèles probabilistes nécessite une décomposition en variables aléatoires ce qui engendre une amplification de la dimension stochastique, et ainsi nous nous retrouvons encore en face du problème de la dimension stochastique élevée. Par conséquent, l'approche développée dans ce travail doit être encore améliorée, en particulier sa première étape qui consiste à déterminer la dimension stochastique efficace. Dans ce contexte, les effets élémentaires sur lesquels est basée la méthode de Morris peuvent être estimés par la méthode proposée par Pradlwarter [156] utilisant les simulations de Monte-Carlo et la théorie des chaînes de Markov. Elle consiste à calculer d'une façon globale le vecteur gradient, à partir duquel on peut construire un estimateur des mesures d'importance associées aux paramètres du modèle. La variance de cet estimateur est indépendante de la dimension stochastique ce qui peut nous faire gagner en efficacité dans le cas où le nombre de paramètres incertains du modèle est élevé. Cette méthode a fait ses preuves pour l'estimation des mesures d'importance dans les analyses de sensibilité [157], et du gradient dans les problèmes d'optimisation [158]. De plus, ces mesures d'importances, qui permettent de mesurer seulement les effets principaux, peuvent être améliorées en utilisant la théorie des graphes [67] pour pouvoir quantifier les effets d'interactions de différents ordres.

Dans le cadre du projet SISBAT, le travail mené dans cette thèse sera prolongé, d'une part pour étudier d'autres sites représentant différents aléas sismiques en France métropolitaine et en territoires d'outre mer, notamment, le scénario fort proche (Le Moule), le scénario moyen représenté par la ville de Lourdes, et le scénario modéré représenté par la ville de Nice. D'autre part, l'approche de calcul stochastique proposée dans ce travail sera appliquée pour l'analyse de l'intégrité d'une toiture à comble habitable, et d'une maison individuelle ayant des murs en maçonnerie.



# Bibliographie

- [1] R. Gupta, M. Vatovec, T. H. Miller. *Metal-Plate-Connected Wood Joints: A Literature Review*. Tech. rep., Oregon State University., 1996.
- [2] A. Ayoub. *Seismic analysis of wood building structures.*, Engineering Structures., 29: 213-223, 2007.
- [3] E. Frühwald-Hansson. *Analysis of structural failures in timber structures: Typical causes for failure and failure modes.*, 33: 2978-2982, 2011.
- [4] Eurocode 5. *Design of Timber Structures.*, Comité Européen de Normalisation (CEN)., Brussels, Belgium 2005.
- [5] Eurocode 8. *Design of Structures for Earthquake Resistance.*, Comité Européen de Normalisation (CEN)., Brussels, Belgium, 2004.
- [6] PS92. *Règles de construction parasismique*. AFNOR, NF P 06-013 edn, 1992.
- [7] SISBAT. *Analyse probabiliste de la vulnérabilité sismique de l'habitat avec toiture en charpentes industrialisées en bois - Couplage modélisation probabiliste /expérimentation*. ANR RiskNat 2008, projet ref. ANR-08-RISK-008, 2009-2013.
- [8] P. Sommerville. *Reducing uncertainty in strong motion predictions*. In *Proc. World Conf. on Earthquake Engineering*. New Zealand Society for Earthquake Engineering, Upper Hutt, New Zealand., 2000.
- [9] B. Sudret., A. Der Kiureghian. *Stochastic finite element methods and reliability, A state-of-the-art-report*. Report No. UCB/SEMM-2000/08, Department of Civil & Engineering University of California, Berkeley, 2000.
- [10] M. Wand., M. Jones. *Kernel smoothing*. Chapman and Hall, 1995.
- [11] M. R. McKay., R. J. Beckman, and W. J. Conover. *A comparison of three methods for selecting values of input variables in the analysis of output from a computer code*. Technometrics 2, 239-245, 1979.
- [12] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*. SIAM, Philadelphia, PA, USA, 1992.
- [13] H. Baldeweck. *Méthodes des éléments finis stochastiques – Application à la géotechnique et à la mécanique de la rupture*. Thèse de doctorat, Université d'Evry-Val d'Essonne , France, 1999.
- [14] E.S. Pearson., M. Tukey. *Distributions whose first four moments are known*. Biometrika, 6: 126-132, 1965.
- [15] N.L. Johnson. *Systems of frequency curves generated by methods of translation*. Biometrika, 36: 149-176, 1949.
- [16] A. Stroud. *Approximate calculation of multiple integrals*. Prentice-Hall, Englewood Cliffs, NJ, 1971.

## Bibliographie

- [17] S. Smolyak. *Quadrature and interpolation formulas for tensor products of certain classes of functions*. Sov Math Dokl, 4: 240-243, 1963.
- [18] D. Xiu., J.S. Hesthaven. *High-order collocation methods for differential equations with random inputs*. SIAM J. Sci. Comput, 27:1118-1139, 2005.
- [19] I. Babuska., F. Nobile., R. Tempone. *A stochastic collocation method for elliptic partial differential equations with random input data*. SIAM, J. Numer. Anal, 45 (3): 1005-1034, 2007.
- [20] J. Baroth., L. Bodé., Ph. Bressollette., M. Fogli. *SFE method using Hermite polynomials: An approach for solving nonlinear mechanical problems with uncertain parameters*. Comput Methods Appl Mech Engrg, 195:6479-6501, 2006.
- [21] J. Baroth., Ph. Bressollette., C. Chauvière., M. Fogli. *An efficient SFE method using Lagrange polynomials: Application to nonlinear problems with uncertain parameters*. Comput Methods Appl Mech Engrg, 196:4419-4429, 2007.
- [22] Ph. Bressollette., M. Fogli., C. Chauvière. *A stochastic collocation method for large classes of mechanical problems with uncertain parameters*. Prob Eng Mech, 25:255-270, 2010.
- [23] H. Riahi., Ph. Bressollette., A. Chateaneuf. *Random fatigue crack growth in mixed mode by stochastic collocation method*. Eng Frac Mech. 77(16): 3292-3309, 2010.
- [24] H.O. Madsen., S. Krenk., N.C. Lind. *Methods of structural safety*. Prentice-Hall, Englewood Cliffs, New Jersey, 1986.
- [25] N. Metropolis., S. Ulam. *The Monte Carlo method*. J Am Stat Assoc. 44: 335-341, 1949.
- [26] G.I. Schuëller., R. Stix. *A critical appraisal of methods to determine failure probabilities*. Struct Safety. 4 (4): 293-309, 1987.
- [27] A.M. Hasofer., N.C. Lind. *Exact and invariant second moment code format*. J Eng Mech, ASCE. 100 (1): 111-121, 1974.
- [28] K. Breitung. *Asymptotic approximation for multinormal integrals*. J Eng Mech, ASCE. 110 (3): 357-366, 1984.
- [29] A. Der Kiureghian., H.Z. Lin., S.J. Hwang. *Second order reliability approximations*. J Eng Mech, ASCE. 113 (8): 1208-1225, 1987.
- [30] L. Tvedt. *Two second-order approximations of failure probability*. Technical note, Det Norske Veritas RDIV/20-004-83, Oslo, Norway, September 1984.
- [31] A. Der Kiureghian., T. Dakessian. *Multiple design points in first and second-order reliability*. Structural Safety, 20 (1): 37-49, 1998.
- [32] S.K. Au., J.L. Beck. *Importance sampling in high dimensions*. Structural Safety, 25: 139-163, 2003.
- [33] A. Harbitz. *An efficient sampling method for probability of failure calculation*. Structural Safety, 3 (2): 109-115, 1986.
- [34] F. Grooteman. *Adaptative radial-based importance sampling method for structural reliability*. Structural Safety, 30 (6): 533-542, 2008.
- [35] F. Grooteman. *An adaptative directional importance sampling method for structural reliability*. Structural Safety, 26 (2): 134-141, 2011.

## Bibliographie

- [36] C. Bucher. *Adaptive sampling-an iterative fast Monte Carlo procedure*. Structural Safety, 5 (2): 119-126, 1988.
- [37] C. Bucher. *Asymptotic sampling for high-dimensional reliability analysis*. Prob Eng Mech, 24: 504-510, 2009.
- [38] K.W. Breitung. *Asymptotic approximations for multinormal integrals*. Journal of Engineering Mechanics, 110 (3): 357-366, 1984.
- [39] M.T. Sichani., S.R.K. Nielsen., C. Bucher. *Applications of asymptotic sampling on high dimensional structural dynamic problems*. Structural Safety, 33: 305-316, 2011.
- [40] S.K. Au., J.L. Beck. *Estimation of small failure probability in high dimensions by subset simulations*. Prob Eng Mech, 16: 263-277, 2001.
- [41] N. Metropolis., A.W. Rosenbluth., M.N. Rosenbluth., A.H. Teller. *Equations of state calculations by fast computing machines*. J Chem Phys, 21 (6): 1087-1092, 1953.
- [42] W.K. Hastings. *Monte Carlo sampling methods using Markov chains and their applications*. Biometrika, 57:97-109, 1970.
- [43] S.K. Au., J.L. Beck. *Subset simulations and its application to seismic risk based on dynamic analysis*. J Eng Mech, 129 (8): 901-917, 2003.
- [44] R. Rackwitz., B. Fiessler. *Structural reliability under combined load sequences*. Comput Struct. 9: 489-494, 1978.
- [45] C.G. Bucher., U. Bourgund. *A fast efficient response surface approach for structural reliability problems*. Structural Safety. 7: 57-66, 1990.
- [46] M.R. Rajashekhar., B.R. Ellingwood. *A new look at the response surface approach for reliability analysis*. Structural Safety. 12 (3): 205-220, 1993.
- [47] L. Faravelli L. *Surface-response approach for reliability analysis*. J Eng Mech, ASCE, 115 (12): 2763-81, 1989.
- [48] S.M. Wong., R.E. Hobbs., C. Onof. *An adaptive response surface method for reliability analysis of structures with multiple loading sequences*. Structural Safety, 27: 287-308, 2005.
- [49] S.H. Kim., S.W. Na. *Response surface method using vector projected sampling points*. Structural Safety, 19 (1): 3-19, 1997.
- [50] M.D. Morris. *Factorial sampling plans for preliminary computational experiments*. Technometrics, 33: 161-174, 1991.
- [51] I. Sobol. *Sensitivity estimates for nonlinear mathematical models*. Mathematical Modeling & Computational Experiments, 1: 407-414, 1993.
- [52] A. Saltelli, K. Chan, EM Schott. *Sensitivity analysis*. Wiley Series in Probability and Statistics. 2000.
- [53] T. Homma, A. Saltelli. *Importance measures in global sensitivity analysis of model output*. Reliab. Eng. Syst. Saf, 52 (1): 1-17, 1996.
- [54] A. Saltelli. *Making best use of model evaluations to compute sensitivity indices*. Comput. Phys. Comm, 145: 280-297, 2002.
- [55] B. Sudret. *Global sensitivity analysis using polynomial chaos expansions*. Reliab. Eng. Syst. Saf, 93: 964-979, 2008.
- [56] J.H. Halton. *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*. Numer. Math, 2: 84-90, 1960.

## Bibliographie

- [57] H. Dai., W. Wang. *Application of low-discrepancy sampling method in structural reliability analysis*. Struct. Saf, 31: 55-64, 2009.
- [58] R.I. Cukier., C.M. Fortune., K.E. Schuler., A.G. Petschek., J.H. Schaibly. *Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, I theory*. J. Chem. Phys, 59: 3873-3878, 1973.
- [59] R.I. Cukier., J.H. Schaibly., K.E. Schuler. *Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, III analysis of the approximations*. J. Chem. Phys, 63: 1140-1149, 1975.
- [60] R.I. Cukier., C.M. Levine., K.E. Schuler. *Nonlinear sensitivity analysis of multiparameter model systems*. J. Comput. Phys, 26: 1-41, 1978.
- [61] A. Saltelli., R. Bolado. *An alternative way to compute Fourier amplitude sensitivity test (FAST)*. Comput. Stat. Data. Anal, 26: 445-460, 1998.
- [62] M. Koda., G.J. McRay., J.H. Seinfeld. *Automatic sensitivity analysis of kinetic mechanisms*. Int. J. Chem. Kinetics, 11: 427-444, 1979.
- [63] A. Saltelli., S. Tarantola., K. Chan. *A quantitative model-independent method for global sensitivity analysis of model output*. Technometrics, 41 (1): 39-56, 1999.
- [64] T. Ishigami, T. Homma. *An importance quantification technique in uncertainty analysis for computer models*. In: Proceedings of the ISUMA'90, first international symposium on uncertainty modeling and analysis, University of Maryland, 398-403, 1990.
- [65] H.C. Van Houwelingen., H.C. Boshuizen., M. Capannesi. *Sensitivity analysis of state-transition models: How to deal with a large number of inputs*. Computers in Biology and Medicine, 41: 838-842, 2011.
- [66] F. Campolongo., J. Cariboni., A. Saltelli. *An effective screening design for sensitivity analysis of large models*. Environmental Modeling & Software, 22: 1509-1518, 2007.
- [67] F. Campolongo., R.D. Braddock. *The use of graph theory in the sensitivity analysis of the model output: a second order screening method*. Reliab. Eng. Syst. Saf, 64: 1-12, 1999.
- [68] R.A. Cropp., R.D. Braddock. *The new Morris Method: an efficient second-order screening method*. Reliab. Eng. Syst. Saf, 78: 77-83, 2002.
- [69] I. Sobol., A. Gresham. *On an alternative global sensitivity estimators*. In: Proceedings of SAMO, Belgirate, 40-42, 1995.
- [70] S. Kucherenko., M. Rodriguez-Fernandez., C. Pantelides., N. Shah. *Monte Carlo evaluation of derivative-based sensitivity measures*. Reliab. Eng. Syst. Saf, 94: 1135-1148, 2009.
- [71] I. Sobol., S. Kucherenko. *Derivative based global sensitivity measures and their link with global sensitivity indices*. Mathematics and Computers in Simulation, 79: 3009-3017, 2009.
- [72] F. Campolongo., J. Cariboni., A. Saltelli. *An effective screening design for sensitivity analysis of large models*. Environmental Modeling & Software, 22: 1509-1518, 2007.
- [73] Nataf A. *Determination des distributions dont les marges sont données*. Comptes Rendus de l'Académie des Sciences, 225: 42-43, 1962.
- [74] Rosenblatt M. *Remarks on a multiformation*. Ann Math Statist, 23: 470-472, 1952.
- [75] R.G. Ghanem, S.D. Spanos. *Stochastic finite elements: a spectral approach*. Berlin: Springer; 1991.

## Bibliographie

- [76] H. Rabitz, Ö.F. Aliş. *General foundations of high-dimensional model representations*. Journal of Mathematical Chemistry. 25: 197-233, 1999.
- [77] R.E. Caflisch, W. Morokoff, A.B. Owen. *Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension*. Journal of Mathematical Chemistry. 25: 197-233, 1999.
- [78] S. Paskov, J. Traub. *Faster valuation of financial derivatives*. J. Portfolio Management. 22: 113-120, 1995.
- [79] I.H. Sloan, H. Wozniakowski. *When are quasi-Monte Carlo algorithms efficient for high dimensional integration?* J. complexity. 14: 1-31, 1998.
- [80] A.B. Owen. *The dimension distribution and quadrature test function*. Stat. Sinica. 13: 1-17, 2003.
- [81] D. Asotsky, E. Myshetskaya, I. Sobol. *The average dimension of multidimensional function for quasi-Monte Carlo estimates of an integral*. Comput Math Math Phys. 46: 2061-2067, 2006.
- [82] S. Kucherenko, B. Feil, N. Shah, W. Mauntz. *The identification of model effective dimensions using global sensitivity analysis*. Reliab. Eng. Syst. Saf, 96: 440-449, 2011.
- [83] N. Wiener. *The homogeneous chaos*. American Journal of Mathematics. 60(4): 897-936, 1938.
- [84] R.H. Cameron, W.T. Martin. *The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals*. Annals of Mathematics. 48(2): 385-392, 1947.
- [85] D. Xiu, GE Karniadakis. *The Weiner-Askey polynomial chaos for stochastic differential equations*. SIAM J Sci Comput. 24(2): 619-644, 2002.
- [86] R.L. Iman, W.J. Conover. *A distribution-free approach to including rank correlation among input variables*. Communications in Statistics Part B. Simulation and Computation. 11(3): 311-334, 1982.
- [87] A.M.J. Olsson, G.E. Sandberg. *On Latin hypercube sampling for stochastic finite element analysis*. J. Eng. Mech. 128(1): 121-125, 2002.
- [88] A.B. Owen. *Controlling correlations in Latin hypercube samples*. Journal of the American Statistical Association. 89(428): 1517-1522, 1994.
- [89] M. Stein. *Large sample properties of simulations using Latin hypercube sampling*. Technometrics. 29(2): 143-151, 1987.
- [90] K. Žiha. *Descriptive sampling in structural safety*. Structural Safety. 17(1): 33-41, 1995.
- [91] A.M.J. Olsson, G.E. Sandberg, O. Dahlblom. *On Latin hypercube sampling for structural reliability analysis*. Structural Safety. 25: 47-68, 2003.
- [92] A. B. Owen. (1992). *A central limit theorem for Latin hypercube sampling*. J. Royal Stat. Soc, Series B. 54: 541-551, 1992.
- [93] D. Ghiocel, R. Ghanem. *Stochastic finite element analysis of seismic soil-structure interaction*. J. Eng. Mech. 128: 66-77, 2002.
- [94] C.J. Sallaberry, J.C. Helton, S.C. Hora. *Extension of Latin hypercube samples with variables*. Reliab. Eng. Syst. Saf, 93: 1047-1059, 2008.

## Bibliographie

- [95] R. Schürer. *A comparison between (quasi-) Monte Carlo and cubature rule based methods for solving high-dimensional integration problems*. Mathematics and Computers in Simulation, 62: 509-517, 2003.
- [96] W.J. Morokoff, R.E. Caflish. *Quasi-Monte Carlo integration*. Journal of Computational Physics, 122: 218-230, 1995.
- [97] I. Sobol. *On the distribution of points in cube and the approximate evaluation of integrals*. Comp Math Math Phys, (Engl. Transl), 7(4): 86-112, 1967.
- [98] H. Faure. *Discrépance de suites associées à un système de numération (en dimension  $s$ )*. Acta Arithmetica, 41: 337-351, 1982.
- [99] I. Sobol. *On quasi-Monte Carlo integrations*. Mathematics and Computers in Simulation, 47: 103-112, 1998.
- [100] B. Sudret, G. Blatman, M. Berveiller. *Quasi-random numbers in stochastic finite element analysis-application to global sensitivity analysis*. In proc. 10<sup>th</sup> Int. Conf. On Applications of stat. and Prob. in Civil Engineering (ICASP10), Tokyo (2007).
- [101] M. Abramowitz. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. 9<sup>th</sup> ed. New York: Dover Publications; 1972.
- [102] E.D. Fichtl, A.K. Prinja. *The stochastic collocation method for radiation transport in random media*. J Quant Spect Rad Trans. 112: 646-659, 2011.
- [103] M. Liu, Z. Gao, J.S. Hesthaven. *Adaptive sparse grid algorithms with applications to electromagnetic scattering under uncertainty*. Appl Num Math. 61: 24-37, 2011.
- [104] E. Novak, K. Ritter. *Simple cubature formulas with high polynomial exactness*. Constructive approximation. 15: 499-522, 1999.
- [105] L. Plaskota, W. Wasilkowski. *Smolyak's algorithm for integration and  $L_1$ -approximation of multivariate functions with bounded mixed derivatives of second order*. Numerical Algorithms. 36 (3): 229-246, 2005.
- [106] I. Mysovskih. *Proof of the minimality of the number of nodes in the cubature formula for a hypersphere*. USSR Comput. Math. Math. Phys. 6: 15-27, 1966.
- [107] D. Xiu. *Numerical integration formulas of degree two*. Applied Numerical Mathematics. 58: 1515-1520, 2008.
- [108] S.S. Isukapalli. *Uncertainty analysis of transport-transformation models*. PhD thesis, The State University of New Jersey. 1999.
- [109] J. An, A. Owen. *Quasi-regression*. J. Complexity. 17(4): 588-607, 2001.
- [110] M. Berveiller. *Éléments finis stochastiques : approches intrusive et non intrusive pour des analyses de fiabilité*. Thèse de l'université Blaise Pascal, France 171 pages, 2005.
- [111] Å Björck. *Numerical methods for least squares problems*. Philadelphia, PA: SIAM press, 1996.
- [112] A. Owen. *Detecting near linearity in high dimensions*. Technical report, Stanford University, Department of Statistics, 1998.
- [113] D. Lewandowski, R.M. Cooke, R.J.D. Tabbens. *Sample-based estimation of correlation ratio with polynomial approximation*. ACM Transactions on Modeling and Computer Simulation. 18(1):1-16, 2007.

## Bibliographie

- [114] J. Oakley, A. O'Hagan. *Probabilistic sensitivity analysis of complex models: a Bayesian approach*. Journal of the Royal Statistical Society. 66: 751-769, 2004.
- [115] R. Ihaka, R. Gentleman. *R: A language for data analysis and graphics*. Journal of Computational and Graphical Statistics. 5(3): 299-314, 1996.
- [116] Y. Yu. *Coupling a stochastic finite element solver with ANSYS and visualization of the results*. Technical report Matr.-Nr. : 2663324, Institute of scientific Computing. Technische Universität Carolo-Wilhelmina zu Braunschweig, 2003.
- [117] S. Rahman. *Stochastic sensitivity analysis by dimensional decomposition and score functions*. Prob. Eng. Mech. 24: 278-287, 2009.
- [118] B. Sudret. *Global sensitivity analysis using polynomial chaos expansions*. Reliab Eng Sys Saf. 93: 964-979, 2008.
- [119] W. Hoeffding. *A class of statistics with asymptotically normal distributions*. Ann. Math. Stat. 19: 293-325, 1948.
- [120] B. Efron, C. Stein. *The jackknife estimate of variance*. Annals of Statistics. 9: 586-596, 1981.
- [121] G. Li, S.W. Wang, H. Rabitz. *Practical approaches to construct RS-HDMR component functions*. Journal of Physical Chemistry. 106: 8722-8733, 2002.
- [122] S. Rahman, H. Xu. *A univariate dimension-reduction method for multidimensional integration in stochastic mechanics*. Prob. Eng. Mech. 19: 393-408, 2004.
- [123] H. Xu, S. Rahman. *Decomposition methods for structural reliability analysis*. Prob. Eng. Mech. 20: 239-250, 2005.
- [124] I. Sobol. *Theorems and examples on high dimensional model representation*. Reliab Eng Sys Saf. 79: 187-193, 2003.
- [125] H. Xu, S. Rahman. *A generalized dimension-reduction method for multidimensional integration in stochastic mechanics*. International Journal for Numerical Methods in Engineering. 61: 1992-2019, 2004.
- [126] S. Rahman. *Global sensitivity analysis by polynomial dimension decomposition*. Reliab Eng Sys Saf. 96: 825-837, 2011.
- [127] K. Entacher. *Quasi-Monte Carlo methods for numerical integration of multivariate Haar Series*. BIT. 4(37): 845-860, 1997.
- [128] D. Wei, S. Rahman. *Structural reliability analysis by univariate decomposition and numerical integration*. Prob. Eng. Mech. 22: 27-38, 2007.
- [129] D. Wei, S. Rahman. *A multi-point univariate decomposition method for structural reliability analysis*. International Journal of Pressure Vessels and Piping. 87: 220-229, 2010.
- [130] M. Hohenbichler, S. Gollwitzer, W. Kruse, R. Rackwitz. *New light on first and second order reliability methods*. Structural Safety. 4: 267-284, 1987.
- [131] S. Rahman. *A polynomial dimensional decomposition for stochastic computing*. International Journal for Numerical Methods in Engineering. 76: 2091-2116, 2008.
- [132] M. Lemaire, A. Chateauneuf, J.C. Mitteau. *Fiabilité des structures*. Hermès-Lavoisier, Paris, 2006.
- [133] A. Der Kiureghian, M. De Stefano. *Efficient algorithm for second-order reliability analysis*. J Eng Mech ASCE, 117 (12): 2904-2923, 1987.

## Bibliographie

- [134] P.L. Liu, A. Der Kiureghian. *Optimization algorithms for structural reliability*. Struct Safety. 9: 161-177, 1991.
- [135] G. Blatman, B. Sudret. *Efficient computation of global sensitivity indices using sparse polynomial chaos expansions*. Reliab. Eng. Syst. Saf. 95: 1216-1229, 2010.
- [136] R.O. Foschi. *Analysis of wood diaphragms and trusses. Part one: Diaphragms*. Can. J. of Civil Engrg., NRC, 4 (3): 345-352, 1977.
- [137] M. Yasumura, S. Yasui. *Pseudodynamic tests and earthquake response analysis of timber structures I: plywood-sheathed conventional wooden walls with opening*. J.Wood Sci. 52: 63-68, 2006.
- [138] N. Richard, M. Yasumura, L. Davenne. *Prediction of seismic behaviour of wood-framed shear walls with openings by pseudodynamic test and FE model*. J. Wood Sci. 49: 145-151, 2003.
- [139] J. Humbert. *Caractérisation du comportement des structures bois à assemblages métalliques sous sollicitations sismiques : Essais expérimentaux, Modélisation numérique, Analyse de sensibilité*. Thèse, Grenoble, France 2010.
- [140] SCIBO, Bordeaux. Rapport d'essais N°F-R/86/05/226/06/124. 2006.
- [141] EN 1075. *Timber structures. Test methods. Joints made with punched metal plate fasteners*. Comité Européen de Normalisation (CEN), Brussels, Belgium, 2000.
- [142] D. Rosowsky. *Reliability-based seismic design of wood shear walls*. J Struct Eng ASCE, 128 (11): 1439-1453, 2002.
- [143] J.W. Van de Lindt, M.A. Walz. *Development and application of wood shear wall reliability model*. J Struct Eng ASCE, 129 (3): 405-413, 2003.
- [144] R. Gupta, K. G. Gebremedhin. *Resistance distributions of a metal-plate connected wood truss*. Forest Products Journal, 42(7/8): 11-16, 1992.
- [145] R. Gupta, K. G. Gebremedhin, J. R. Cooke. *Analysis of metal-plate connected wood trusses with semi-rigid joints*. American Society of Agricultural Engineers (ASAE), 35(3): 1011-1018, 1992.
- [146] M. Hansson and P. Ellegaard. *Monte Carlo simulation and reliability analysis of roof trusses with punched metal plate fasteners*. Tech. Rep. 36-14-2, CIB-W18, Colorado, USA, 2003.
- [147] X. Song and F. Lam. *Reliability analysis of metal plate connected wood truss assemblies concerning buckling failure*. In Proc. of the 11<sup>th</sup> World Conference on Timber Engineering (WCTE). Presentation 437, 2010.
- [148] H. Dohi, T. Inaba, T. Sato, T. Ninomiya, K. Okuta, H. Akagi. *Seismic response analysis of the NTT Kobe Ekimae building during the Hyogo-ken Nanbu earthquake of 17<sup>th</sup> January 1995*. The 10<sup>th</sup> Japan Earthquake Symposium, 1803-1808.
- [149] A. Chateauneuf. *Projet ANR SISBAT – Modèles probabilistes des assemblages*. juillet 2011.
- [150] A. Chateauneuf. *Projet ANR SISBAT – Modèles probabilistes des matériaux*. juin 2010.
- [151] CSTB. *Rapport d'essais N° EEM 09 26019348/A concernant des essais mécaniques sur équerres E5 et E14 (Projet ANR-08-RIKNAT-08-02/SISBAT)*. mars 2010.

## Bibliographie

- [152] CSTB. *Rapport d'essais N° EEM 09 26019348/B concernant des essais mécaniques sur des pointes et des vis (Projet ANR-08-RIKNAT-08-02/SISBAT)*. mars 2010.
- [153] D. Bertil, J. Rey, M. Belvaux. *Projet ANR SISBAT – Modélisation de l'action sismique*. Rapport final. BRGM/RP-58886-FR, 132 pages. 2010.
- [154] K. Toki. *Lessons learned from the Kobe 1995 earthquake - Low frequency but high impact event*. In Y. W. N. Shiraishi, M. Shinozuka (ed.), *Structural Safety and Reliability*, vol. 1, p. 61-67. A.A Balkema, Rotterdam, Netherlands. 1998.
- [155] P. Sommerville. *Reducing uncertainty in strong motion predictions*. In Proc. World. Conf. on Earthquake Engineering. New Zealand Society for Earthquake Engineering, Upper Hutt, New Zealand. 2000.
- [156] H.J. Pradlwarter. *Relative importance of uncertain structural parameters, part I: algorithm*. *Comput Mech.* 40: 627-635, 2007.
- [157] M.F. Pellissetti, H.J. Pradlwarter, G.I. Schuëller. *Relative importance of uncertain structural parameters, part II: applications*. *Comput Mech.* 40: 637-649, 2007.
- [158] E. Patelli, H.J. Pradlwarter. *Monte Carlo gradient estimation in high dimensions*. *International Journal for Numerical Methods in Engineering.* 81: 172-188, 2010.
- [159] L. Davenne. *Projet ANR SISBAT – Modélisation de toitures combles perdues testées au FCBA*. mars 2012.
- [160] CSTB. *Rapport d'essais N° 403/11/194 essais sismiques de toitures (Projet ANR-08-RIKNAT-08-02/SISBAT)*. mais 2012.
- [161] G.C. Morano, R. Greco, E. Morrone. *Analytical evaluation of essential facilities fragility curves by using a stochastic approach*. *Engineering Structures.*, 33: 191-201, 2011.



## Résumé

Le problème de la dimension stochastique élevée est récurrent dans les analyses probabilistes des structures. Il correspond à l'augmentation exponentielle du nombre d'évaluations du modèle mécanique lorsque le nombre de paramètres incertains est élevé. Afin de pallier cette difficulté, nous avons proposé dans cette thèse, une approche à deux étapes. La première consiste à déterminer la dimension stochastique efficace, en se basant sur une hiérarchisation des paramètres incertains en utilisant les méthodes de criblage. Une fois les paramètres prépondérants sur la variabilité de la réponse du modèle identifiés, ils sont modélisés par des variables aléatoires et le reste des paramètres est fixé à leurs valeurs moyennes respectives, dans le calcul stochastique proprement dit. Cette tâche fut la deuxième étape de l'approche proposée, dans laquelle la méthode de décomposition de la dimension est utilisée pour caractériser l'aléa de la réponse du modèle, par l'estimation des moments statistiques et la construction de la densité de probabilité. Cette approche permet d'économiser jusqu'à 90% du temps de calcul demandé par les méthodes de calcul stochastique classiques. Elle est ensuite utilisée dans l'évaluation de l'intégrité d'une toiture à ossature bois d'une habitation individuelle installée sur un site d'aléa sismique fort. Dans ce contexte, l'analyse du comportement de la structure est basée sur un modèle éléments finis, dans lequel les assemblages en bois sont modélisés par une loi anisotrope avec hystérésis et l'action sismique est représentée par huit accélérogrammes naturels fournis par le BRGM. Ces accélérogrammes permettent de représenter différents types de sols selon en se référant à la classification de l'Eurocode 8. La défaillance de la toiture est définie par l'atteinte de l'endommagement, enregistré dans les assemblages situés sur les éléments de contreventement et les éléments d'anti-flambement, d'un niveau critique fixé à l'aide des résultats des essais. Des analyses déterministes du modèle du modèle éléments finis ont montré que la toiture résiste à l'aléa sismique de la ville du Moule en Guadeloupe. Les analyses probabilistes ont montré que parmi les 134 variables aléatoires représentant l'aléa dans le comportement non linéaire des assemblages, 15 seulement contribuent effectivement à la variabilité de la réponse mécanique ce qui a permis de réduire la dimension stochastique dans le calcul des moments statistiques. En s'appuyant sur les estimations de la moyenne et de l'écart-type on a montré que la variabilité de l'endommagement dans les assemblages situés dans les éléments de contreventement est plus importante que celle de l'endommagement sur les assemblages situés sur les éléments d'anti-flambement. De plus, elle est plus significative pour les signaux les plus nocifs sur la structure.

## Abstract

The problem of the curse of dimensionality is frequently encountered in practical applications. It can be defined as the significant increase of the number of mechanical model calls with the number of uncertain parameters. To overcome this difficulty, a two-steps stochastic approach has been developed in this work. The first step of this approach consists in calculating the stochastic effective dimension by the means of Morris screening method. Once the most significant uncertain parameters on the variability of the mechanical responses are identified, they are modeled as random variables and the remaining parameters are fixed to their respective mean values. This allows us to reduce significantly the stochastic dimension of the problem in the second step of the approach where the decomposition method is used to estimate the statistical characteristics of the mechanical responses. The efficiency and the accuracy of this approach are evaluated through an academic problem dealing with the assessment of the integrity of a three-span five-story frame structure subjected to horizontal loads. We have demonstrate that we can reduce about 90% of the computation time required by the classical stochastic methods. Then, the proposed approach is used to the analysis of the integrity of timber roofs under seismic loading. The behaviour of this structure is described through a finite element model where the timber joints are modeled by anisotropic hysteresis law, and the seismic action is represented by eight real earthquake ground motion records. These accelerograms provided by the French institution involved in geosciences BRGM allow us to take into account different soil types according to the classification provided by the european design code dealing with seismic events Eurocode 8. The failure of timber roofs is reached when the damage levels in the timbers joints localized on the buckling and bracing members reach the critical value. It is shown, through a deterministic analysis, that the structure resists the seismic hazard representing the city of Le Moule in Guadeloupe. The stochastic analysis has shown that, among the 134 random variables representing the uncertainty in the nonlinear behaviour of the timber joints, only 15 have a significant effect on the variability of the structural response, which allow us to reduce the stochastic dimension in the computation of the statistical moments. According to the estimates of the mean and the standard deviation, we have shown that the variability of bracing members damage is greater than the variability of buckling members damage. Moreover, the variability of the bracing members damage is more significant for the earthquake ground motion records having the lowest collapse PGA.