



HAL
open science

Réduction du coût numérique des méthodes d'apprentissage actif pour l'analyse de fiabilité

Morgane Menz

► **To cite this version:**

Morgane Menz. Réduction du coût numérique des méthodes d'apprentissage actif pour l'analyse de fiabilité. Sciences de l'ingénieur [physics]. UNIVERSITE DE TOULOUSE, 2020. Français. NNT : . tel-03132391

HAL Id: tel-03132391

<https://hal.science/tel-03132391>

Submitted on 5 Feb 2021

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.



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

soutenue le *04/12/2020* par :

MORGANE MENZ

**Réduction du coût numérique des méthodes d'apprentissage
actif pour l'analyse de fiabilité**

JURY

PR. N. GAYTON	SIGMA, Clermont-Ferrand	Rapporteur
DR. C. HELBERT	École Centrale Lyon	Rapporteur
DR. S. DA VEIGA	Safran Tech, Paris-Saclay	Examineur
PR. P. LAFON	Université de Technologie de Troyes	Examineur
PR. M. SALAÜN	ISAE-SUPAERO, Toulouse	Examineur
DR. C. GOGU	Université Toulouse III	Directeur de thèse
PR. J. MORIO	ONERA, Toulouse	Co-directeur de thèse
PR. N. BARTOLI	ONERA, Toulouse	Co-directrice de thèse
DR. S. DUBREUIL	ONERA, Toulouse	Co-encadrant

École doctorale et spécialité :

ED 467 Aéronautique Astronautique : Génie Mécanique

Unité de Recherche :

Institut Clément Ader (UMR 5312)

Directeur(s) de Thèse :

Christian Gogu et Jérôme Morio

Rapporteurs :

Nicolas Gayton et Céline Helbert

Résumé

Les analyses de fiabilité sur des problèmes d'ingénierie complexes impliquent souvent des temps de calcul très élevés et requièrent l'utilisation de méthodes numériques avancées pour l'estimation d'une probabilité de défaillance. Les approches d'apprentissage actif de processus gaussiens pour la fiabilité sont un moyen de fortement réduire les temps de calcul en diminuant le nombre d'appels au code de simulation numérique coûteux tout en conservant une précision d'estimation suffisante de la probabilité de défaillance. Cette thèse s'est concentrée sur l'amélioration de ces techniques autour de deux axes : la caractérisation des erreurs d'estimation et leur prise en compte dans l'algorithme d'apprentissage ainsi que le couplage entre apprentissage actif et les approches de réduction de modèles.

Les approches d'apprentissage actif de processus gaussiens pour la fiabilité consistent à construire une approximation par processus gaussien de la fonction de performance. L'approximation est construite en se basant sur une fonction d'apprentissage de manière à la rendre précise au voisinage de l'état limite de défaillance. L'estimation de la probabilité de défaillance est alors typiquement obtenue en classant une population d'échantillons évalués à partir du métamodèle construit. Ainsi, on peut identifier deux sources d'incertitude influant sur l'estimateur de la probabilité de défaillance, qui sont respectivement l'approximation par métamodèle et la variabilité de l'échantillonnage. Dans ces travaux de thèse, nous proposons une méthodologie pour quantifier la sensibilité de l'estimateur de la probabilité de défaillance à ces deux sources d'incertitude. Cette analyse est ensuite intégrée à l'algorithme d'apprentissage actif afin de focaliser l'enrichissement sur la réduction de la source majeure d'incertitude. Elle permet également de quantifier et contrôler l'erreur globale associée à l'estimation de la probabilité de défaillance et ainsi de définir un critère d'arrêt de l'algorithme. L'approche est développée aussi bien dans le cas d'un échantillonnage par Monte Carlo que par échantillonnage préférentiel pour l'estimation de probabilités d'évènements rares. La méthode a été validée sur différentes applications et a montré un potentiel de réduction du nombre de points d'apprentissage.

Une autre voie de réduction du coût numérique des analyses de fiabilité étudiée dans cette thèse concerne les techniques de réduction de modèles qui constituent une autre approche prometteuse pour diminuer les temps de résolution de certains modèles numériques. Ainsi, afin de réduire encore davantage le coût numérique des analyses de fiabilité, nous proposons un couplage entre les approches d'apprentissage actif et la réduction de modèle de type base réduite. Plus précisément, nous avons mis en place une méthodologie adaptative permettant de choisir, par le biais d'un critère de couplage, si une solution réduite peut être utilisée à la place du modèle numérique complexe pour évaluer la fonction de performance. L'utilisation du couplage sur différentes applications

(mécaniques et thermiques) montre un fort potentiel de réduction des temps de calcul tout en assurant des approximations raisonnables des probabilités de défaillance.

Mots clés : modélisation par processus gaussien, classification, apprentissage actif, probabilité de défaillance, réduction de modèles, base réduite, analyse de fiabilité.

Abstract

Running a reliability analysis on engineering problems involving complex numerical models can be computationally very expensive, requiring advanced simulation methods to reduce the overall numerical cost. Gaussian process based active learning approaches for reliability analysis are thus a way to strongly decrease the computational time by reducing the number of calls to the expensive numerical model while maintaining a sufficiently accurate estimation of the probability of failure. The work of this PhD thesis is focused on two improvement paths for these methods : the characterization of the estimation errors and their inclusion in the learning algorithm as well as the coupling of active learning and model order reduction approaches.

Gaussian process based active learning methods for reliability analysis involve a learning phase that consists in building a Gaussian process approximation of the performance function. The resulting approximation, obtained using a learning function, is built such as to be accurate in the vicinity of the limit state of the failure domain. Then, the failure probability estimation can typically be obtained by a classification of a sample population evaluated on the final surrogate model. Hence, two sources of uncertainty that are impacting the failure probability estimator can be identified. These are respectively related to the metamodel approximation and the sampling variability. In this PhD thesis project, we propose a methodology to quantify the sensitivity of the failure probability estimator to these two uncertainty sources. This analysis is then integrated into an active learning algorithm in order to focus the enrichment on the reduction of the major source of uncertainty. It also allows to quantify and control the global error associated to the probability estimation and thus to define a stopping criterion for the algorithm. The approach is developed for both a Monte Carlo sampling and an Importance Sampling to tackle rare events probability estimation problems. The method has been validated on several applications and has shown a potential to reduce the number of learning points.

Another way to reduce the numerical cost of reliability analyses that is studied in this thesis is related to model order reduction techniques, which are another promising way to reduce the numerical model resolution time. In order to further reduce the numerical cost of reliability analyses, we propose a coupling between active learning approaches and model order reduction, specifically the reduced basis approach. More precisely, we set an adaptive methodology that allows to choose, on the basis of a coupling criterion, if a reduced solution can be used instead of the whole numerical model to evaluate the performance function. The use of the coupling on different applications (mechanical and thermal) shows great potential of computational time reduction while ensuring reasonable approximations of failure probabilities.

Keywords : Gaussian process modeling, classification, active learning, probability

of failure, model order reduction, reduced basis, reliability analysis.

Remerciements

Je souhaite tout d'abord remercier mes directeurs et encadrants de thèse, Christian Gogu, Jérôme Morio, Nathalie Bartoli et Sylvain Dubreuil, pour m'avoir encadrée et guidée durant ces trois années. Merci à vous quatre d'avoir été présents pour moi, j'ai beaucoup appris de vos nombreux conseils et idées et ça a été un réel plaisir de travailler avec vous.

J'adresse également mes remerciements à Nicolas Gayton et Céline Helbert d'avoir accepté de rapporter ma thèse et à Michel Salaün, Pascal Lafon et Sébastien Da Veiga d'avoir accepté d'examiner ma thèse et pour leurs nombreuses questions pertinentes lors de la soutenance et des pré-rapports.

Je remercie également les membres de l'ICA et de l'ONERA que j'ai côtoyé pour leur accueil et grâce à qui je garde de supers souvenirs de ces trois années. Je pense plus particulièrement aux autres doctorants et post-doctorants qui sont devenus entre quelques pauses cafés, séances de torture au fitness et parties de tarot, de vrais copains. Ainsi je dis un grand merci en particulier à Rémy, Gabriel, Gaspard et Maxime, ainsi qu'à Pierre, Alessandro, Jean-Emmanuel, Tiphaine, Margot, Vincent, Marco, Vincenzo et Marie.

Évidemment, je remercie aussi ma famille et mes amis pour leur soutien.

Table des matières

Table des matières	vii
Liste des figures	ix
Liste des tableaux	xiii
Liste des symboles	xv
Liste des acronymes	xvii
Introduction	1
1 État de l’art	9
1.1 Modèles pour la réduction du coût numérique	10
1.2 Analyse de fiabilité	33
2 Apprentissage actif basé sur l’étude de la variance de l’estimateur de la probabilité de défaillance	65
2.1 Introduction	67
2.2 Mesure de la sensibilité de la probabilité de défaillance aux incertitudes du GP et à l’estimation par une simulation de Monte-Carlo	68
2.3 Approche basée sur les mesures de sensibilité aux incertitudes	76
2.4 Applications	79
2.5 Amélioration avec l’Échantillonnage préférentiel, ou <i>Importance Sampling</i>	88
2.6 Conclusion	97

3 Apprentissage actif combiné à la modélisation en base réduite	99
3.1 Introduction	101
3.2 Approche combinant apprentissage actif et modélisation en base réduite	102
3.3 Applications	113
3.4 Conclusion	140
 Conclusion générale et perspectives	 143
 Références	 161

Liste des figures

1.1	Illustration d'un modèle de GP pour un Plan d'expériences, ou <i>Design of Experiments</i> (DoE) initial de 5 points	14
1.2	Covariance entre l'ensemble des points du domaine et les points $x^* = 0.5$ et $x^* = 0.7$	16
1.3	Exemple de trajectoires générées à partir d'un GP	17
1.4	Exemple de génération d'une trajectoire de $\mathcal{G}_n(x)$ à partir de l'Eq. (1.21)	19
1.5	Exemple de représentation de représentation de Karhunen-Loève (KL) du noyau de covariance d'un processus gaussien, ou <i>Gaussian process</i> (GP)	22
1.6	Schéma des différentes étapes de la méthode de quantification des incertitudes.	34
1.7	Exemple d'application de la méthode FORM.	37
1.8	Approximation linéaire de l'état limite par la méthode FORM dans le cas de deux modes de défaillance distincts.	38
1.9	Classification d'une population de Monte-Carlo.	40
1.10	Méthode de Monte-Carlo pour une probabilité faible. Illustration de l'inefficacité de l'inefficacité de la méthode puisqu'aucun points n'est trouvé dans le domaine de défaillance.	41
1.11	Illustration des lignes de contour de la densité des entrées (lignes grises) et des lignes de contour de la densité optimale de l'Échantillonnage préférentiel, ou <i>Importance Sampling</i> (IS) (lignes en couleur).	43
1.12	Illustration de l'échantillonnage préférentiel pour une probabilité faible.	44
1.13	Estimation par noyaux de la densité optimale en utilisant la méthode NAIS pour un état limite complexe.	48

1.14	Exemple d'apprentissage actif pour le cas 4 branches	49
1.15	Initialisation de <i>Active learning reliability method combining Kriging and Monte Carlo Simulation</i> (AK-MCS).	55
1.16	Exemple d'enrichissement du GP par AK-MCS à une étape de l'apprentissage.	56
1.17	Résultat final de AK-MCS à la fin de l'apprentissage.	57
1.18	Diagramme de l'algorithme AK-MCS.	58
1.19	Évolution de la probabilité normalisée pour les critères d'apprentissage <i>EFF</i> et <i>U</i> sur le cas 4 branches.	59
2.1	Évolution de l'estimation de la probabilité de défaillance en fonction du nombre d'itérations au cours d'une exécution de l'algorithme AK-MCS+ <i>EFF</i> ($\hat{P}_f^{MC} = 4.46 \times 10^{-3}(1.6\%)$)	75
2.2	Évolution des estimations des variances $V_{\hat{\mathbf{x}}}$ et $V_{\mathcal{G}_n}$ en fonction du nombre d'itérations au cours d'une exécution de l'algorithme AK-MCS+ <i>EFF</i>	75
2.3	Diagramme de l'algorithme d'apprentissage actif Apprentissage actif de GP basé sur la variance, ou <i>Variance based Active GP learning</i> (Vb-AGP) + Simulations de Monte-Carlo, ou <i>Monte Carlo Simulations</i> (MCS)	79
2.4	Estimation de $COV_{\mathcal{G}_n} = \frac{\sqrt{V_{\mathcal{G}_n}}}{\hat{P}_f}$ pour un nombre de points d'intégration n_q croissant	81
2.5	Comparaison de deux DoE : un résultant d'une exécution de la méthode AK-MCS+ <i>EFF</i> et un autre obtenu par la méthode proposée pour les mêmes DoE et population de Monte-Carlo (MC) initiaux avec $COV_{max} = 3\%$	84
2.6	Évolution des logarithmes des variances $V_{\hat{\mathbf{x}}}$ et $V_{\mathcal{G}_n}$ estimées en fonction du nombre d'itérations au cours de l'algorithme Vb-AGP	85
2.7	Histogramme des estimations \hat{P}_f sur 100 répétitions de Vb-AGP + MCS sur le cas test système en série à 4 branches	85
2.8	Schéma de l'oscillateur non-linéaire à un degré de liberté	86
2.9	Histogramme des estimations \hat{P}_f sur 100 répétitions de Vb-AGP + MCS sur l'exemple de l'oscillateur non linéaire	88
2.10	Diagramme de l'algorithme d'apprentissage actif Vb-AGP + IS	93
2.11	Illustration de la méthode Vb-AGP + IS (<i>Non parametric Adaptive Importance Sampling</i> (NAIS)) sur l'exemple du système série à 4 branches	94
2.12	Histogramme des estimations \hat{P}_f sur 100 répétitions de Vb-AGP + IS sur le cas test système en série à 4 branches - cas probabilité faible	95
2.13	Histogramme des estimations \hat{P}_f sur 50 répétitions de Vb-AGP + IS sur l'exemple de l'oscillateur non linéaire - cas probabilité faible	97

3.1	Résultat final de AK-MCS à la fin de l'apprentissage	103
3.2	Schéma de l'algorithme générique de couplage	104
3.3	Schéma de la chambre de combustion d'une fusée refroidie par régénération.	114
3.4	Maillage Éléments Finis (EF) de la paroi de la chambre de combustion et conditions aux limites du problème thermique.	115
3.5	Champ de température calculé au point moyen.	116
3.6	Corrélation entre les résidus et la vraie erreur pour le problème ther- mique. Trois tailles de base réduite sont considérées (5, 10 et 20).	117
3.7	Trois premiers modes de température obtenus en utilisant la procédure de construction à la volée au cours de AK-MCS pour le problème ther- mique.	118
3.8	Précision (étoiles rouges) de P_f et accélération (points bleus) de la mé- thode proposée en fonction du seuil ϵ pour le résidu non préconditionné pour différents DoE et populations de MC générés par l'algorithme AK- MCS pour le problème thermique.	119
3.9	Précision (étoiles rouges) de P_f et accélération (points bleus) de la mé- thode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le préconditionneur du point moyen pour différents DoE et populations de MC générés par l'algorithme AK-MCS pour le problème thermique.	120
3.10	Précision (étoiles rouges) de P_f et accélération (points bleus) de la mé- thode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le préconditionneur du point le plus proche pour différents DoE et po- pulations de MC générés par l'algorithme AK-MCS pour le problème thermique.	120
3.11	Conditions aux limites et chargement imposés à la plaque stratifiée.	127
3.12	Maillage EF de la plaque stratifiée trouée.	128
3.13	Champ de déplacement du stratifié nominal.	129
3.14	Corrélation entre les résidus et la vraie erreur pour le problème lié à la plaque stratifiée trouée. Trois tailles de base réduite sont considérées (5, 10 et 20).	131
3.15	Quatre premiers modes de déplacement selon l'axe Z obtenus en utilisant la procédure de construction à la volée au sein de l'algorithme AK-MCS pour le problème de la plaque stratifiée trouée.	132
3.16	Précision (étoiles rouges) de P_f et accélération (points bleus) de la mé- thode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le préconditionneur au point moyen P_{mean} pour différents DoE et popu- lations de MC générés par l'algorithme AK-MCS pour le problème de la plaque stratifiée trouée.	132

3.17	Précision (étoiles rouges) de P_f et accélération (points bleus) de la méthode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le préconditionneur calculé au point le plus proche $\mathbf{P}_{nearest}$ pour différents DoE et populations de MC générés par l'algorithme AK-MCS pour le problème de la plaque stratifiée trouée.	133
3.18	Intervalles de confiance obtenus par l'approche Vb-AGP + IS + BR ($COV_{max} = 0.03$) de la précision de P_f en fonction du seuil ϵ sur le résidu préconditionné pour le préconditionneur calculé au point moyen \mathbf{P}_{mean} pour le problème thermique.	138
3.19	Histogramme des estimations \hat{P}_f sur 70 répétitions de Vb-AGP + IS + BR (\mathbf{P}_{mean}) avec $\epsilon = 10^{-3}$ et $COV_{max} = 0.03$ sur le cas test thermique .	139

Liste des tableaux

1.1	Tableau récapitulatif des éléments des méthodes EGRA et AK-MCS . . .	55
2.1	Système en série à 4 branches— résultats moyens sur 100 répétitions - paramètres de la méthode adaptative : $n_{MC}^{init} = 5 \times 10^4$, $n_{DoE}^{init} = 16$, $COV_{max} = 0.03$	85
2.2	Oscillateur non-linéaire à un degré de liberté - Distributions des para- mètres du problème	87
2.3	Oscillateur non linéaire ($\mu_{F_1} = 1$, $\sigma_{F_1} = 0.2$) — Résultats moyens sur 100 répétitions - Paramètres de l'algorithme : $n_{MC}^{init} = 10^4$, $n_{DoE}^{init} = 12$, $COV_{max} = 0.03$	87
2.4	Système en série à 4 branches - cas probabilité faible — résultats moyens sur 100 répétitions - paramètres de la méthode adaptative : $n_{MC}^{init} =$ 1×10^4 , $n_{DoE}^{init} = 12$, $COV_{max} = 0.03$	95
2.5	Oscillateur non linéaire ($\mu_{F_1} = 0.6$, $\sigma_{F_1} = 0.1$) — Résultats moyens sur 50 répétitions - Paramètres de l'algorithme : $n_{MC}^{init} = 10^4$, $n_{DoE}^{init} = 12$, $COV_{max} = 0.03$	96
3.1	Lois de probabilité des paramètres du problème thermique.	114
3.2	Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le résidu non préconditionné pour le problème thermique.	121
3.3	Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point moyen pour le problème thermique. . .	121

3.4	Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point le plus proche pour le problème thermique.	121
3.5	Résultats de la 2 ^{ème} stratégie pour le co-krigeage	125
3.6	Résultats de la 3 ^{ème} stratégie pour le co-krigeage	126
3.7	Paramètres de la loi de probabilité gamma utilisée pour modéliser les épaisseurs h des plis.	127
3.8	Paramètres de la loi de probabilité uniforme utilisée pour modéliser les angles d'orientation des fibres.	128
3.9	Constantes élastiques d'un pli.	129
3.10	Contraintes ultimes d'un pli.	129
3.11	Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point moyen pour le problème de la plaque stratifiée trouée.	134
3.12	Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point le plus proche pour le problème de la plaque stratifiée trouée.	134
3.13	Temps CPU (en secondes) des opérations pour l'analyse de fiabilité du problème de thermique en grande dimension.	135
3.14	Temps CPU (en secondes) des opérations pour l'analyse de fiabilité du problème mécanique en grande dimension.	136
3.15	Résultats de l'analyse de sensibilité fiabiliste.	137
3.16	Résultats moyens de AK-MCS et Vb-AGP + IS + BR (\mathbf{P}_{mean}) avec $\epsilon = 10^{-3}$ pour le problème thermique - paramètres de la méthode adaptative : $n_{MC}^{init} = 10^4$, $n_{DoE}^{init} = 14$, $COV_{max} = 0.03$	138

Liste des symboles

- \mathcal{D}_f Domaine de défaillance. 34, 35, 47, 49, 51, 61, 70
- $\hat{\mathcal{D}}_f$ Approximation du domaine de défaillance basée sur le GP. 49, 50, 61
- \mathcal{D}_s Domaine de sûreté. 34, 61
- \mathcal{G}_n Processus Gaussien conditionné. ix, 13, 16–19, 50–53, 55, 67, 68, 70, 71, 73, 74, 76, 77, 80, 88–90, 92, 103, 108
- \mathcal{G} Processus Gaussien non conditionné. 12, 13, 17, 19–21

Liste des acronymes

- AK-MCS** *Active learning reliability method combining Kriging and Monte Carlo Simulation.* x–xiv, 53–58, 62–64, 74, 75, 83–87, 94–96, 102, 103, 106, 107, 111, 113, 115, 118–122, 128, 130, 132–139
- CE** Entropie Croisée, ou *Cross-Entropy*. 44, 46, 145
- COV** Coefficient de variation. 39, 56, 57, 68, 73, 74, 77, 78, 82, 83, 86, 87, 91, 94–98, 115, 121, 129, 137, 139, 145
- ddl** Degrés de liberté. 26–28, 101, 106, 119, 128, 134, 135, 141
- DoE** Plan d’expériences, ou *Design of Experiments*. ix–xii, 11, 13–15, 17, 18, 24, 31, 49, 50, 54–56, 62, 64, 76, 78, 83, 84, 86–88, 92, 93, 95, 96, 102, 103, 106, 108, 118–121, 123–125, 130, 132, 133, 137, 140
- EF** Éléments Finis. xi, 25–28, 30, 31, 33, 38, 41, 64, 101, 102, 105, 107, 111, 114–116, 128, 135, 141
- EOLE** *Expansion optimal linear estimation*. 21, 80
- GP** processus gaussien, ou *Gaussian process*. vii, ix, x, xv, 12–19, 21–24, 32, 35, 50–53, 55, 56, 62–64, 66–80, 82, 83, 88–94, 101–103, 108, 112, 122–126, 135, 137, 143–146
- GPR** Régression par processus gaussiens, ou *Gaussian process Regression*. 12, 23, 24
- IC** Intervalle de confiance. 13

- IS** Échantillonnage préférentiel, ou *Importance Sampling*. vii, ix, x, xii, xiv, 41–44, 46, 55, 64, 66, 68, 88–98, 136–139, 143–145
- KL** représentation de Karhunen-Loève. ix, 19–22, 28, 80–82, 98
- LHS** Échantillonnage par hypercube latin, ou *Latin Hypercube Sampling*. 76
- MC** Monte-Carlo. x–xii, 39–41, 50, 51, 53–56, 60, 61, 67–74, 76–78, 80, 82–84, 86, 88, 91, 92, 98, 115, 116, 119–121, 124, 125, 129, 132, 133, 137, 138, 143
- MCS** Simulations de Monte-Carlo, ou *Monte Carlo Simulations*. x, xiii, xiv, 39, 41, 43, 45, 55, 61, 68, 71, 75, 76, 79, 82–90, 92, 94–96, 98, 115, 121, 122, 124, 125, 128, 129, 134, 136, 137, 145
- NAIS** *Non parametric Adaptive Importance Sampling*. x, 46–48, 91–94, 96–98, 144
- PGD** Décomposition Propre Généralisée, ou *Proper Generalized Decomposition*. 28
- POD** Décomposition Orthogonale aux valeurs Propres, ou *Proper Orthogonal Decomposition*. 27–29
- RBDO** Dimensionnement optimal sous contrainte de fiabilité, ou *Reliability Based Design Optimisation*. 145, 146
- SVD** Décomposition en Valeurs Singulières, ou *Singular Value Decomposition*. 29
- SVM** Machines à vecteurs de support ou, *Support Vector Machine*. 11, 50
- Vb-AGP** Apprentissage actif de GP basé sur la variance, ou *Variance based Active GP learning*. x, xii, xiv, 76, 78, 79, 83–88, 90–98, 135–139, 141, 143–145

Introduction

Cadre

Ce travail de thèse s'est déroulé entre l'Université Toulouse 3 Paul Sabatier (UT3) et l'Office National d'Etudes et de Recherches Aérospatiales (ONERA) au sein de l'école doctorale aéronautique et astronautique (EDAA). Le travail de recherche s'est effectué dans l'unité Modélisation des Systèmes et Microsystèmes Mécaniques (MS2M) de l'Institut Clément Ader (ICA) et l'unité Méthodes Multidisciplinaires, Concepts Intégrés (M2CI) du département Traitement de l'Information et Systèmes (DTIS) de l'ONERA.

Les développements effectués durant la thèse s'inscrivent dans le cadre du projet de recherche ReBReD porté par Christian Gogu (ICA) et financé par l'Agence Nationale de la Recherche.

Contexte et formulation du problème

Les problèmes d'ingénierie sont sujets à de nombreuses incertitudes sur la modélisation (mathématique ou numérique) du problème ou sur les valeurs des paramètres physiques (incertitudes des propriétés matériaux, variabilités sur les charges appliquées, variabilités géométriques, etc.). Ces incertitudes peuvent alors affecter le comportement ou les performances du système et mener à une défaillance. Ainsi, les différentes sources d'incertitude d'un système doivent être identifiées et quantifiées pour assurer la fiabilité de celui-ci. Les incertitudes des paramètres d'entrée sont souvent prises en compte par des méthodes d'analyses probabilistes, dont l'utilisation a pris son essor en ingénierie mécanique depuis les années 70. Ces approches reposent sur une modélisation

probabiliste des incertitudes, en représentant les paramètres d'entrée incertains par des variables aléatoires.

Une fois les incertitudes quantifiées et modélisées, l'analyse de la fiabilité d'une structure passe par la définition d'un état limite (par exemple un déplacement excessif ou un dépassement de contrainte) à partir duquel le système est considéré défaillant. Au cours des dernières décennies, de nombreux développements ont concerné des outils mathématiques pour la modélisation, l'amélioration du calcul parallèle et les ressources informatiques disponibles. Ces progrès permettent aujourd'hui de simuler de manière plus précise le comportement de nombreux phénomènes à l'aide de modèles numériques dont la complexité va en grandissant. Ces modèles reposent, par exemple, sur la résolution numérique d'équations aux dérivées partielles par les méthodes des éléments finis ou des volumes finis.

Ainsi, le comportement de nombreux systèmes peut être simulé de façon suffisamment précise de façon à pouvoir étudier numériquement leur fiabilité. Pour cela, les incertitudes sur les paramètres d'entrée sont propagées à travers le modèle pour pouvoir analyser l'influence de ces incertitudes sur la variabilité de la sortie et estimer la probabilité de défaillance de la structure considérée.

Au cours des dernières décennies, différentes méthodes d'analyse de fiabilité ont été développées pour évaluer les risques liés aux diverses incertitudes sur les paramètres d'un système dans le cas où l'état limite de défaillance n'est pas connu de façon explicite, c.-à-d. lorsqu'on ne dispose pas d'une expression analytique de l'état limite. Ces méthodes peuvent être classées en deux catégories : les approches d'approximation analytique et les approches statistiques.

Dans le premier cas, il s'agit essentiellement des méthodes FORM (*First Order*) et SORM (*Second Order Reliability Method*) [LEMAIRE et al., 2009] qui sont basées sur un développement de Taylor (respectivement à l'ordre 1 et 2) de la fonction de performance du système autour du point de défaillance le plus probable, obtenu par une optimisation. La méthode SORM est plus largement applicable que la méthode FORM, qui repose sur une hypothèse de linéarité de la surface d'état limite, mais nécessite le calcul des dérivées premières et secondes de la fonction de performance et est donc plus coûteuse. Ces méthodes ont néanmoins un coût de calcul relativement raisonnable (souvent de l'ordre de la centaine d'évaluations). Malgré tout, les hypothèses sur lesquelles reposent l'approximation de l'état limite de défaillance par ces méthodes sont restrictives et l'erreur commise si ces hypothèses sont fausses n'est ni contrôlable ni vérifiable.

Dans le second cas, il s'agit essentiellement des méthodes d'échantillonnage. En général, ces méthodes requièrent un nombre important de simulations ($> 10^3$ simulations) dans l'optique d'effectuer une analyse statistique (méthode de Monte-Carlo) et ainsi évaluer la probabilité de défaillance. Dans ce cadre, la probabilité de défaillance

obtenue est une estimation statistique qui a donc une variabilité liée aux tirages aléatoires que l'on peut caractériser par sa variance ou par son coefficient de variation. De nombreuses méthodes ayant pour objectif de réduire la variance des estimateurs de la probabilité de défaillance obtenus par échantillonnage ont ainsi été développées. On peut notamment citer l'échantillonnage préférentiel [MELCHERS, 1989] et l'échantillonnage directionnel [DITLEVSEN et al., 1988] parmi les précurseurs de telles méthodes, ayant donné lieu à un riche panel de méthodes dérivées. Malgré cela, dû à l'augmentation continue de la taille et de la complexité des modèles numériques, les coûts de calculs n'ont jamais cessé de représenter un obstacle majeur à l'utilisation pratique des méthodes d'analyse de fiabilité. Ainsi ce problème de coût numérique a également été abordé sous deux nouveaux angles par la suite, visant à diminuer le plus possible le temps de calcul lié aux appels des modèles numériquement coûteux : les modèles de substitution et la réduction de modèle.

Une première approche de réduction de coût consiste à considérer un modèle de substitution de la sortie du code de calcul. Ce modèle de substitution est construit à partir d'une base de données et permet de prédire à moindre coût la réponse pour d'autres paramètres d'entrée indépendamment du modèle physique. Historiquement, les premières approches pour les analyses de fiabilité se sont appuyées sur l'utilisation de surfaces de réponse polynomiales. Une des problématiques liées à l'utilisation de modèles de substitution globaux, tels que les surfaces de réponse polynomiales, est qu'ils sont construits de manière à être précis sur l'ensemble du domaine de variation des paramètres. Cela est en revanche inutile en vue de la réalisation d'analyses de fiabilité, où la précision n'est requise qu'à proximité de l'état limite de défaillance. Des nouvelles méthodes ont ainsi été développées. Elles consistent à choisir judicieusement la base de données de manière à être précis du point de vue de la classification plutôt que du point de vue de la régression sur l'ensemble du domaine. Ainsi, la probabilité de défaillance peut être estimée de façon précise en évaluant le modèle de substitution sur une population générée par une méthode d'échantillonnage. Les méthodes d'apprentissage actif de modèles de processus gaussiens [ECHARD et al., 2011; BICHON et al., 2008] pour la fiabilité, développées au cours des dix dernières années, ont pour objectif de construire la base de données pour obtenir une approximation suffisamment précise de l'état limite avec un nombre réduit d'évaluations du modèle numérique. Ces approches sont des méthodes adaptatives qui reposent sur un critère d'enrichissement, utilisé pour enrichir judicieusement la base de données, et sur deux critères pour l'arrêt de l'algorithme : d'une part un critère d'arrêt d'apprentissage, qui permet de maîtriser la précision de l'approximation de l'état limite, et d'autre part un critère sur le coefficient de variation de l'estimation par échantillonnage, qui permet de contrôler que le nombre d'échantillons est suffisant. Or, la précision locale de l'approximation de l'état limite de défaillance n'est qu'un des facteurs affectant la variance globale de

l'estimation de la probabilité de défaillance. Ainsi, la stratégie qui consiste à imposer un critère très précis sur l'approximation de l'état limite de défaillance n'est pas forcément la mieux adaptée pour contrôler la variabilité de l'estimation de la probabilité de défaillance et peut mener à un apprentissage trop conservatif. Dans ce manuscrit, nous nous intéressons à la caractérisation des erreurs d'estimation induites à la fois par l'approximation par processus gaussiens et par l'échantillonnage. On note également, que l'évaluation de chaque point d'apprentissage nécessite de résoudre le modèle numérique. Il peut être intéressant de considérer des modèles de fidélité moins élevée avec des coûts de calcul plus faibles pour évaluer certains points d'apprentissage, ce qui est l'objet de notre second axe de travail.

Un second type d'approche visant à réduire les coûts associés aux modèles numériques considère la réduction de systèmes d'équations obtenus par la méthode des éléments finis. Cela consiste donc à chercher une solution dans un sous-espace engendré par les modes dominants du problème, obtenus grâce à des solutions connues du modèle numérique. C'est le principe des méthodes de réduction de modèles appelées en base réduite [KERSCHEN et al., 2005], qui ont pour objectif de réduire le coût de résolution du problème en le projetant sur un sous-espace de dimension notablement réduite par rapport à l'espace initial. Cet espace réduit est généralement construit afin de minimiser sur l'ensemble du domaine l'erreur d'approximation sur la solution, ou encore plus précisément l'erreur sur une quantité d'intérêt qui dépend de la solution à l'aide de méthodes dites *goal-oriented*. Néanmoins, ces méthodes ont pour objectif de minimiser l'erreur sur l'ensemble du domaine et peuvent donc potentiellement mener à évaluer des solutions dans des régions où la défaillance est très improbable. Afin de prendre en compte la distribution des entrées dans le processus de construction de la base réduite, l'article [GOGU et al., 2016] a proposé une approche de construction à la volée de la base réduite sur les points d'une population échantillonnée par Monte-Carlo ou par échantillonnage préférentiel en contrôlant l'erreur sur la solution réduite. Cette méthode permet de réduire les points d'apprentissage de la base réduite aux points localisés dans le domaine dans lequel la réponse peut préférentiellement prendre ses valeurs. Or en fiabilité, il est uniquement nécessaire d'être précis du point de vue de la classification. Pour cela, il suffit de connaître précisément l'état limite et non pas les valeurs exactes de la réponse sur l'ensemble des points échantillonnés. Ainsi, il serait judicieux de choisir les candidats pour l'enrichissement de la base réduite parmi des points améliorant la connaissance de l'état limite. Il serait donc intéressant d'utiliser les points déterminés par le critère d'enrichissement d'une méthode d'apprentissage actif pour les analyses de fiabilité comme candidat pour la construction de l'espace réduit. On disposerait alors de solutions de fidélités différentes (évaluées par le modèle éléments finis ou le modèle réduit) pour les points d'enrichissement d'un processus gaussien, ce qui permettrait de réduire davantage le coût d'apprentissage de celui-ci.

Objectifs et plan de la thèse

D'une part le nombre de points d'apprentissage ajoutés au cours de la phase d'apprentissage du processus gaussiens peut potentiellement être réduit en se basant sur l'erreur introduite par l'incertitude du métamodèle sur l'estimation de la probabilité de défaillance. D'autre part, une réduction significative des temps de calcul liés à l'évaluation des points d'apprentissage peut être rendue possible via une synergie entre les approches d'analyse de fiabilité par apprentissage actif et la modélisation en base réduite.

Ce travail de thèse se place dans le cadre de l'estimation de défaillances faibles en considérant un seul scénario de défaillance avec un nombre modéré de variables et son objet est alors de :

1. proposer des estimateurs des contributions de l'approximation par processus gaussien et de l'intégration numérique des méthodes d'échantillonnage (Monte Carlo et échantillonnage préférentiel) à la variabilité de l'estimateur de la probabilité de défaillance. Pour cela l'analyse de l'influence des deux sources d'incertitude peut être étudiée comme un problème d'analyse de sensibilité basé sur la décomposition de la variance. Une stratégie d'enrichissement basée sur la quantification de la variabilité de l'estimateur liée aux incertitudes pourra ainsi être construite à partir de ces estimateurs.
2. développer le couplage entre les approches d'apprentissage actif et de réduction de modèles afin de rendre possible une synergie permettant de réduire les temps de calcul. Le développement de cette nouvelle méthode permettra ainsi des analyses de fiabilité à fidélité adaptative, ajustant la fidélité du modèle aux besoins de l'analyse de fiabilité.

Ce manuscrit de thèse se décompose en trois chapitres. Le premier chapitre correspond un l'état de l'art de ces problématiques et les deux chapitres suivants décrivent les contributions de la thèse :

Le [chapitre 1](#) a pour objectif de dresser un état de l'art des différents métamodèles et des différents outils et méthodes numériques utilisés dans cette thèse. Dans un premier temps, la modélisation par processus gaussiens pour prédire la réponse d'un simulateur ainsi que les méthodes de réduction d'ordre de modèles sont introduites. Dans un second temps, ce chapitre présente un ensemble de méthodes d'estimation de probabilités, classiques et dédiées à l'estimation de probabilités d'évènements rares.

Dans le [chapitre 2](#), nous nous intéresserons à l'analyse de la contribution des incertitudes liées à l'approximation par processus gaussien et à l'échantillonnage à la variabilité de l'estimateur de la probabilité de défaillance en montrant que cette problématique s'écrit comme un problème d'analyse de sensibilité. En se basant sur cette

analyse, nous proposerons un apprentissage ciblant la source majeure d'incertitude et s'arrêtant à l'aide du coefficient de variation total de l'estimation. Les performances de cette stratégie seront étudiées sur différents exemples.

Dans le [chapitre 3](#), des stratégies basées sur l'utilisation de solutions réduites pour l'amélioration de l'efficacité numérique des approches d'apprentissage actif pour certains problèmes d'analyse de fiabilité seront présentées. L'hybridation entre les deux approches est rendue possible grâce à une stratégie de construction adaptative des bases réduites basée sur un critère de couplage. Deux critères seront proposés et leurs performances seront étudiés sur des cas d'applications thermique et mécanique. Finalement un exemple combinant les deux contributions est développé dans ce chapitre.

Une conclusion et des perspectives viendront dans ce document.

Contributions

Certains des travaux effectués durant cette thèse ont fait l'objet de publications dans des journaux et de communications lors de conférences.

Publication

- Menz, M., Gogu, C., Dubreuil, S., Bartoli, N., Morio, J., 2020. Adaptive coupling of reduced basis modeling and Kriging based active learning methods for reliability analyses. *Reliability Engineering System Safety* 196, 106771. <https://doi.org/10.1016/j.ress.2019.106771> [MENZ et al., 2020]

Publication soumise

- Menz, M., Dubreuil, S., , Morio, J., Bartoli, N., Gogu, C., Chiron, M., Variance based sensitivity analysis for Monte Carlo and importance sampling reliability assessment with Gaussian processes, soumis à *Structural Safety* le 11/10/20.

Communications à des conférences sur résumé

- Menz, M., Gogu, C., Dubreuil, S., Bartoli, N., Morio, J., “Towards reliability analysis by adaptive sampling with multiple fidelity levels by the use of reduced basis methods”, 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, Jun 2019, Crete, Greece
- Menz, M., Dubreuil, S., , Morio, J., Bartoli, N., Gogu, C., “Variance analysis of AK-MCS failure probability estimator and algorithmic application”, 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference, 1 - 6 November 2020, Venice, Italy

Communications à des conférences sur article

- Menz, M., Gogu, C., Dubreuil, S., Bartoli, N., Morio, J., “Vers des analyses de fiabilité à fidélité adaptative par l'utilisation de modèles en base réduite”, 10èmes journées Fiabilité des Matériaux et des Structures, Bordeaux, 27-28 Mars 2018
- Menz, M., Gogu, C., Dubreuil, S., Bartoli, N., Morio, J., “Couplage adaptatif des méthodes d'analyses de fiabilité par apprentissage actif basées sur les métamodèles de Krigeage et de la modélisation en base réduite”, 24ème Congrès Français de Mécanique, 26-30 Août 2019, Brest, France

CHAPITRE 1

État de l'art

Objectifs du chapitre

- Introduire les différents modèles et outils pour la réduction du coût numérique utilisés dans la thèse.
- Présenter un panel des différentes méthodes numériques d'estimation de probabilité d'évènements rares ou non.

Sommaire

1.1	Modèles pour la réduction du coût numérique	10
1.1.1	Introduction	10
1.1.2	Régression par processus gaussiens	11
1.1.3	Méthodes de réduction de modèles	24
1.1.4	Conclusion	32
1.2	Analyse de fiabilité	33
1.2.1	Introduction	33
1.2.2	Méthodes FORM/SORM	35
1.2.3	Méthodes par échantillonnage	39
1.2.4	Analyse de fiabilité basée sur l'utilisation de métamodèles . .	49
1.2.5	Conclusion	62

1.1 Modèles pour la réduction du coût numérique

1.1.1 Introduction

Dans le cadre de la quantification d'incertitudes ou de l'optimisation d'un système, les études nécessitent d'effectuer un nombre important de simulations correspondant aux variations des paramètres de ce système. Pour un jeu de n_u paramètres $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^{n_u}$, la réponse d'une structure à une sollicitation est obtenue grâce à un code numérique $\mathcal{M}^u(\cdot)$ calculant (par exemple par éléments finis) le champ de réponse \mathbf{u} :

$$\mathcal{M}^u : \left\{ \begin{array}{l} \mathcal{X} \subseteq \mathbb{R}^{n_u} \longrightarrow V \subseteq \mathbb{R}^{n_{dd}} \\ \mathbf{x} \longmapsto \mathbf{u} = \mathcal{M}^u(\mathbf{x}) \end{array} \right. \quad (1.1)$$

Au cours d'une étude paramétrique on ne s'intéresse généralement pas à la réponse \mathbf{u} elle-même mais plutôt à une variable d'intérêt y qui permet de caractériser les performances du système et qui dépend de \mathbf{u} . Finalement, on considère le calcul paramétrique défini par le code de calcul $\mathcal{M}(\mathbf{x}) = \mathcal{M}(\mathbf{u}(\mathbf{x}), \mathbf{x})$ et un jeu de paramètres d'entrée \mathbf{x} dont la sortie est la variable scalaire y :

$$\mathcal{M} : \left\{ \begin{array}{l} \mathcal{X} \longrightarrow \mathbb{R} \\ \mathbf{x} \longmapsto y = \mathcal{M}(\mathbf{x}) \\ \qquad \qquad \qquad = \mathcal{M}(\mathbf{u}(\mathbf{x}), \mathbf{x}) \end{array} \right. \quad (1.2)$$

Différentes études sont effectuées dans le cadre du calcul paramétrique, par exemple l'analyse de la sensibilité aux paramètres \mathbf{x} , la détection ou la prévision de défaillances de systèmes, l'optimisation des paramètres de conception d'une structure ou la caractérisation de la sortie y du modèle. Ces analyses requièrent souvent un nombre important, plusieurs milliers, d'appels au code numérique et leur coût peut devenir rédhibitoire pour des modèles $\mathcal{M}(\cdot)$ avec des temps de calcul élevés. Il est alors nécessaire de remplacer les modèles précis et coûteux par des modèles donnant une approximation de leur(s) réponse(s) avec un coût d'exécution fortement réduit.

On a alors deux approches pour construire un modèle approchant la sortie du code de calcul :

- Une première approche consiste à construire un modèle mathématique $\widehat{\mathcal{M}}(\mathbf{x})$ de la réponse $y = \mathcal{M}(\mathbf{x})$ du code de calcul, appelé méta-modèle ou modèle de substitution. Ces modèles permettent de prédire la réponse en se basant sur les simulations pour d'autres paramètres d'entrée du code de calcul et ce à moindre coût.
- Une autre approche consiste à construire un modèle réduit du modèle numérique, qui tient compte de la physique du problème par des méthodes de réduction de

modèles, pour obtenir une approximation $\mathbf{u}_r(\mathbf{x})$ du champ de réponse $\mathbf{u}(\mathbf{x})$. Ainsi, une estimation $\hat{y} = \mathcal{M}(\mathbf{u}_r(\mathbf{x}), \mathbf{x})$ de la quantité y peut être obtenue en utilisant l'approximation de $\mathbf{u}(\mathbf{x})$. En effet, le coût numérique est généralement lié au calcul du champ $\mathbf{u}(\mathbf{x})$ à partir duquel la quantité y est obtenue pour un coût qui est en comparaison négligeable.

Dans un premier temps, cette section introduit la notion de méta-modélisation et présente plus précisément la modélisation par processus gaussien. Dans un second temps, le principe de la réduction de modèles est présenté, ainsi que quelques unes des méthodes les plus utilisées.

1.1.2 Régression par processus gaussiens

1.1.2.1 Introduction

La méta-modélisation consiste à construire une approximation par un modèle mathématique $\widehat{\mathcal{M}}(\mathbf{x})$, appelé méta-modèle ou modèle de substitution, de la sortie scalaire $y \in \mathbb{R}$ d'un simulateur numérique $\mathcal{M}(\mathbf{x})$ en fonction de ses entrées $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^{n_u}$, à partir d'une base de données composée de n jeux d'entrées $\mathbf{x}_{doe} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, appelés *Plan d'expériences*, ou *Design of Experiments (DoE)*, et des sorties correspondantes du simulateur $\mathbf{y}_{doe} = \mathcal{M}(\mathbf{x}_{doe})$. Il s'agit alors d'estimer les paramètres $\boldsymbol{\theta} \in \mathbb{R}^{n_u}$ du méta-modèle $\widehat{\mathcal{M}}(\mathbf{x}) = \widehat{\mathcal{M}}(\mathbf{x}, \boldsymbol{\theta})$ par apprentissage statistique [VAPNIK, 1998; HASTIE et al., 2009] à partir des observations du DoE afin d'obtenir une précision optimale (dans un sens à définir) de l'approximation de la sortie du simulateur sur l'espace des entrées \mathcal{X} .

Dans la littérature, les méthodes consistant à apprendre une fonction de prédiction à partir d'observations sont plus généralement connues sous le nom d'apprentissage supervisé. Parmi les méta-modèles classiques utilisés pour les codes de calcul, on trouve les méta-modèles polynômiaux [MYERS et al., 2012], les modèles de fonctions à base radiale (RBF) [BUHMANN, 2000], le chaos polynomial [GHANEM et SPANOS, 1991], les réseaux de neurones [HOPFIELD, 1982; HAYKIN, 1994; SCHMIDHUBER, 2015], les *Machines à vecteurs de support* ou, *Support Vector Machine (SVM)* [VAPNIK, 1998; BOURINET et al., 2011; SMOLA et SCHÖLKOPF, 2004; HURTADO JORGE E. et ALVAREZ DIEGO A., 2003] et les modèles de régression par processus gaussien [RASMUSSEN et WILLIAMS, 2006; JONES et al., 1998]. L'ensemble de ces métamodèles ne sera pas décrit dans ce manuscrit par souci de concision mais pour plus de détails, on pourra se référer à [HASTIE et al., 2009; MURPHY, 2012].

La nature probabiliste des processus gaussiens étant particulièrement intéressante pour nos travaux, nous nous sommes concentrés sur cette approche dans le manuscrit. En effet, ils présentent l'avantage par rapport aux autres méta-modèles de fournir intrinsèquement une estimation de l'incertitude a priori de la prédiction puisque la

variance du processus gaussien conditionné peut servir d'indicateur d'erreur. Cet indicateur trouve son intérêt notamment dans le cadre de l'enrichissement adaptatif des données d'apprentissage.

1.1.2.2 Généralités sur la modélisation par processus gaussiens

La Régression par processus gaussiens, ou *Gaussian process Regression* (GPR) [RASMUSSEN et WILLIAMS, 2006; FORRESTER et al., 2008], aussi dénommée Krigeage, d'abord introduite par Daniel Krige [KRIGE, 1951] en géostatistique et formalisée plus tard par Georges Matheron [MATHERON, 1963], est une méthode de prédiction bayésienne. Cette méthode consiste à prédire la réponse d'une fonction par une GPR et son utilisation pour la prédiction des résultats de codes de calcul a été introduite dans [SACKS et al., 1989].

1.1.2.3 Formulation

Un processus gaussien, ou *Gaussian process* (GP) $\mathcal{G}(\mathbf{x})$, $\mathbf{x} \in \mathcal{X}$, est une collection infinie de variables aléatoires définies sur le même espace de probabilité $(\Omega_P, \mathcal{A}, \mathbb{P})$, c.-à-d. un processus stochastique, dont chaque sous-ensemble fini suit une loi normale multi-variée.

La distribution d'un GP \mathcal{G} est entièrement caractérisée par sa moyenne $\mu(\mathbf{x})$ et un noyau (ou fonction de covariance) $k(\cdot, \cdot)$:

$$\mu(\mathbf{x}) = \mathbb{E}[\mathcal{G}(\mathbf{x})], \quad \mathbf{x} \in \mathcal{X} \quad (1.3)$$

$$k(\mathbf{x}, \mathbf{x}') = \text{CoVar}(\mathcal{G}(\mathbf{x}), \mathcal{G}(\mathbf{x}')), \quad (\mathbf{x}, \mathbf{x}') \in \mathcal{X}^2 \quad (1.4)$$

où $\text{CoVar}(\cdot, \cdot)$ désigne la covariance entre deux variables aléatoires.

Par la suite, cette distribution est notée :

$$\mathcal{G} \sim GP(\mu(\cdot), k(\cdot, \cdot)) \quad (1.5)$$

La distribution a priori du GP \mathcal{G} peut être définie comme suit :

$$\mathcal{G}(\mathbf{x}) = \mu(\mathbf{x}) + Z(\mathbf{x}) \quad (1.6)$$

où

- $\mu(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}$ avec $\mathbf{f}(\mathbf{x})$ le vecteur des fonctions d'une base à définir et $\boldsymbol{\beta}$ le vecteur contenant les coefficients de régression associés.
- $Z(\mathbf{x})$ un GP stationnaire centré et de variance σ_Z^2 , vérifiant que le noyau du GP

s'écrit :

$$k(\mathbf{x}, \mathbf{x}') = \text{CoVar}(\mathcal{G}(\mathbf{x}), \mathcal{G}(\mathbf{x}')) = \sigma_Z^2 r_\theta(\mathbf{x}, \mathbf{x}')$$

où $r_\theta(\mathbf{x}, \mathbf{x}')$ est une fonction de corrélation définie par un ensemble d'hyperparamètres $\boldsymbol{\theta}$.

La distribution a posteriori \mathcal{G}_n du GP sur le domaine \mathcal{X} correspond à la distribution a priori du GP \mathcal{G} conditionné aux n observations $\{\mathbf{x}_{doe} = (\mathbf{x}_1, \dots, \mathbf{x}_n), \mathbf{y}_{doe} = \mathcal{M}(\mathbf{x}_{doe})\} \in \mathcal{X}^n \times \mathbb{R}^n$, c.-à-d. aux réponses du modèle $\mathcal{M}(\mathbf{x})$ évalué aux points d'un DoE \mathbf{x}_{doe} .

La distribution a priori du GP non conditionné \mathcal{G} est supposée gaussienne, la distribution a posteriori est donc gaussienne. On a alors :

$$\mathcal{G}_n = \mathcal{G} | \mathbf{x}_{doe}, \mathbf{y} \sim GP(\mu_n(\cdot), k_n(\cdot, \cdot)) \quad (1.7)$$

où la moyenne $\mu_n(\mathbf{x})$, la variance $\sigma_n^2(\mathbf{x})$ et la covariance $k_n(\mathbf{x}, \mathbf{x}')$ de la prédiction de la réponse $G(\mathbf{x})$ par le GP conditionné \mathcal{G}_n sont données par les expressions :

$$\mu_n(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + \mathbf{k}(\mathbf{x})^T \mathbf{C}^{-1} (\mathbf{y}_{doe} - \mathbf{F} \boldsymbol{\beta}) \quad (1.8)$$

$$\sigma_n^2(\mathbf{x}) = \mathbf{k}(\mathbf{x}, \mathbf{x}) - \begin{pmatrix} \mathbf{k}(\mathbf{x})^T & \mathbf{f}(\mathbf{x})^T \end{pmatrix} \begin{pmatrix} \mathbf{C} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{k}(\mathbf{x}) \\ \mathbf{f}(\mathbf{x}) \end{pmatrix} \quad (1.9)$$

$$k_n(\mathbf{x}, \mathbf{x}') = \mathbf{k}(\mathbf{x}, \mathbf{x}') - \begin{pmatrix} \mathbf{k}(\mathbf{x})^T & \mathbf{f}(\mathbf{x})^T \end{pmatrix} \begin{pmatrix} \mathbf{C} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{k}(\mathbf{x}') \\ \mathbf{f}(\mathbf{x}') \end{pmatrix} \quad (1.10)$$

où $\mathbf{k}(\mathbf{x}) = (k(\mathbf{x}, \mathbf{x}_1), \dots, k(\mathbf{x}, \mathbf{x}_n))^T$, \mathbf{F} est une matrice dont les lignes i valent $\mathbf{f}(\mathbf{x}_i)^T$, $\mathbf{C} := (k(\mathbf{x}_i, \mathbf{x}_j))_{i,j}$ est la matrice de covariance entre les observations.

La Figure 1.1 représente une fonction $G(\mathbf{x})$ et son interpolation par un GP, construit à partir d'un DoE de 5 points, dont on a représenté la moyenne $\mu_n(\mathbf{x})$ et l'Intervalle de confiance (IC) à 99% (en prenant $\mu_n(\mathbf{x}) \pm 3\sigma_n(\mathbf{x})$) sur le domaine $\mathcal{X} = [0, 1]$.

1.1.2.4 Estimation des hyperparamètres

À noyau fixé $k(\cdot, \cdot)$, les hyperparamètres $\boldsymbol{\theta}$, σ_Z et $\boldsymbol{\beta}$ doivent être estimés pour approcher la réponse en tout point du domaine. Il existe de nombreuses techniques permettant d'obtenir les valeurs des hyperparamètres, par exemple ils peuvent être obtenus par maximum de vraisemblance [JONES, 2001], par maximum a posteriori [FUGLSTAD et al., 2019] ou par validation-croisée [RASMUSSEN et WILLIAMS, 2006].

Lorsque les hyperparamètres sont remplacés par les estimateurs obtenus par maximum de vraisemblance, les estimateurs de σ_Z et $\boldsymbol{\beta}$ ont des expressions analytiques

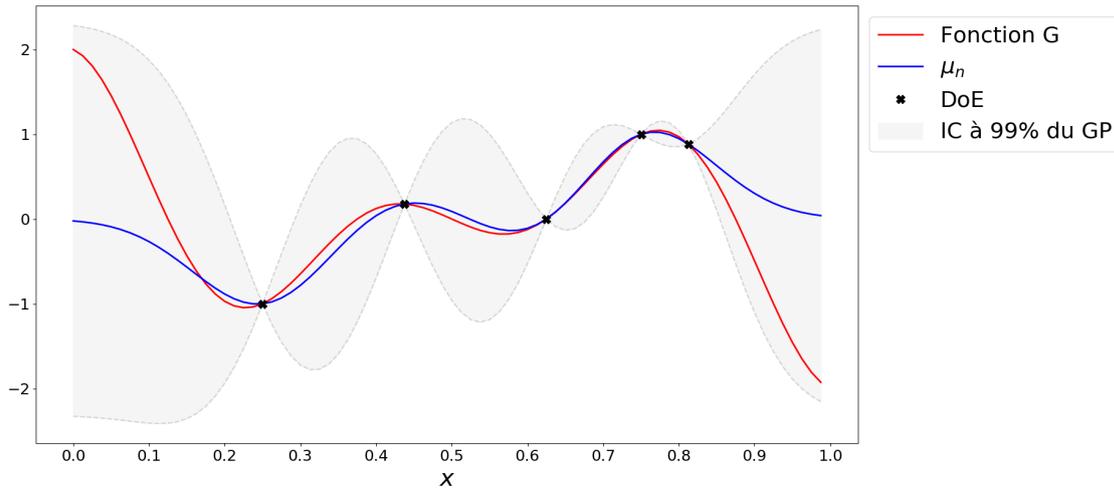


FIGURE 1.1 – Illustration d'un modèle de GP pour un DoE initial de 5 points

alors que celui de θ nécessite de résoudre numériquement un problème d'optimisation non linéaire. Les estimateurs de σ_Z et β sont donnés respectivement par les équations (1.11) et (1.12) :

$$\hat{\beta} = (\mathbf{F}^T \mathbf{C}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{C}^{-1} \mathbf{y}_{doe} \quad (1.11)$$

$$\hat{\sigma}_Z^2 = \frac{1}{n} (\mathbf{y}_{doe} - \hat{\beta} \mathbf{F})^T \mathbf{C}^{-1} (\mathbf{y}_{doe} - \hat{\beta} \mathbf{F}) \quad (1.12)$$

1.1.2.5 Noyaux de covariance

Les noyaux de covariance permettent de modéliser la corrélation entre les valeurs du GP en deux points \mathbf{x} et \mathbf{x}' . Un noyau de covariance $k(\mathbf{x}, \mathbf{x}')$ sur l'ensemble \mathcal{X} est un noyau défini positif, c.-à-d. une fonction satisfaisant :

- $\forall (\mathbf{x}, \mathbf{x}') \in \mathcal{X}^2, k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}', \mathbf{x})$ (fonction symétrique)
- $\forall N \in \mathbb{N}, (\mathbf{x}_1, \dots, \mathbf{x}_N) \in \mathcal{X}^N, (a_1, \dots, a_N) \in \mathbb{R}^N,$

$$\sum_{i=1}^N \sum_{j=1}^N a_i a_j k(\mathbf{x}_i, \mathbf{x}_j) \geq 0$$

De nombreux modèles de noyaux existent pour définir une fonction de corrélation.

Dans le cadre de la modélisation par GP, les noyaux les plus couramment utilisés sont les noyaux gaussiens, de Matérn et exponentiel. Ces noyaux ont la particularité d'être des noyaux stationnaires, c'est-à-dire qu'ils ne dépendent que de la distance entre deux points $\mathbf{d} = |\mathbf{x} - \mathbf{x}'|$.

Le noyau le plus utilisé est certainement le noyau gaussien, dont la fonction de corrélation est donnée par :

$$r^{\text{SE}}(\mathbf{d}) = \exp \left[-\frac{\mathbf{d}^2}{\boldsymbol{\theta}^2} \right] \quad (1.13)$$

Cette fonction de corrélation est indéfiniment dérivable, ce qui implique que les GPs construits à partir de ce noyau sont très lisses.

Pour de nombreuses applications numériques, il a été montré que l'utilisation du noyau gaussien abouti à des matrices de covariance mal conditionnées [ZIMMERMANN, 2015]. En effet, le conditionnement dépend du nombre de points et de la distance maximale entre ces points. Ainsi, les matrices de covariance construites en se basant sur ce noyau sont particulièrement mal conditionnées quand les points d'apprentissage sont fortement corrélés [KOSTINSKI et KOIVUNEN, 2000].

Parmi la classe Matérn, ce sont les noyaux Matérn $3/2$ et Matérn $5/2$ qui sont fréquemment considérés. Leurs fonctions de corrélation sont données par :

$$r^{\text{Mat}^{3/2}}(\mathbf{d}) = \left(1 + \frac{\sqrt{3}\mathbf{d}}{\boldsymbol{\theta}} \right) \exp \left[-\frac{\sqrt{3}\mathbf{d}}{\boldsymbol{\theta}} \right] \quad (1.14)$$

$$r^{\text{Mat}^{5/2}}(\mathbf{d}) = \left(1 + \frac{\sqrt{5}\mathbf{d}}{\boldsymbol{\theta}} + \frac{5\mathbf{d}^2}{3\boldsymbol{\theta}^2} \right) \exp \left[-\frac{\sqrt{5}\mathbf{d}}{\boldsymbol{\theta}} \right] \quad (1.15)$$

L'expression de la fonction de corrélation du noyau exponentiel est la suivante :

$$r^{\text{exp}}(\mathbf{d}) = \exp \left[-\frac{\mathbf{d}}{\boldsymbol{\theta}} \right] \quad (1.16)$$

Il s'agit du noyau le moins régulier comparé à ceux précédemment définis.

La Figure 1.2 représente l'évolution sur l'ensemble du domaine $[0, 1]$ de la covariance $k_n(\cdot, x^*)$ du GP conditionné de l'exemple précédent, construit avec un noyau de covariance gaussien, pour $x^* = 0.5$ et $x^* = 0.7$.

Sur la figure, on peut observer que la covariance est plus forte au voisinage du point x^* . On peut également voir que la covariance est nulle aux points du DoE de construction, qui correspondent aux 5 points du DoE de la Figure 1.1, car il s'agit de points sans bruit (déterministes). On a donc aux points pour lesquels la distance $d = |x - x^*|$ est faible et qui sont éloignés des points du DoE la corrélation la plus importante.

Dans la section suivante, nous allons présenter les techniques de génération de trajectoires de GP. Ces techniques nous seront utiles pour caractériser certaines quantités d'intérêt dépendantes d'un GP à partir des réalisations (ou trajectoires) de celui-ci.

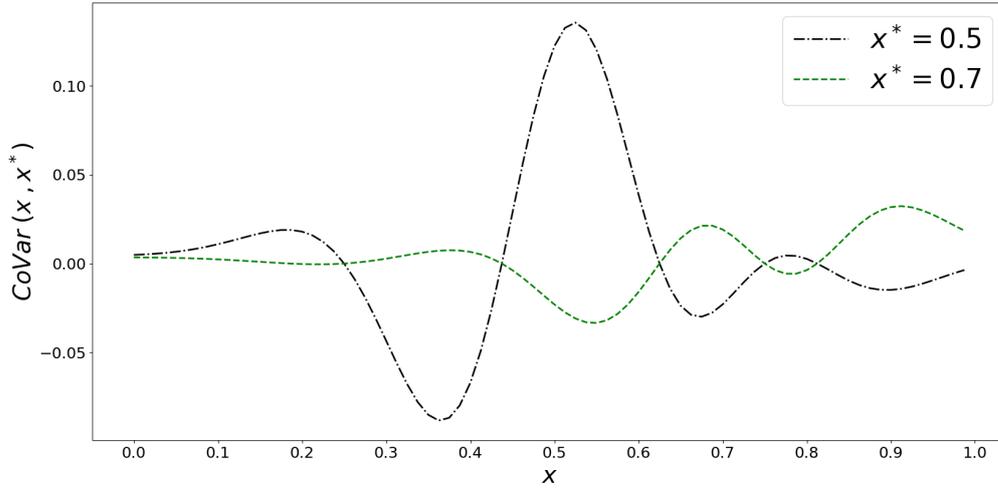


FIGURE 1.2 – Covariance entre l'ensemble des points du domaine et les points $x^* = 0.5$ et $x^* = 0.7$

1.1.2.6 Génération de trajectoires de processus gaussiens

Une trajectoire d'un GP $\mathcal{G}_n \sim GP(\mu_n(\cdot), k_n(\cdot, \cdot))$ simulée sur un ensemble de n_{data} points $\tilde{\mathbf{x}} = (\mathbf{x}_i)_{i=1, \dots, n_{data}} \in \mathcal{X}^{n_{data}}$ correspond à une réalisation de $\mathcal{G}_n(\tilde{\mathbf{x}}) \sim GP(\boldsymbol{\mu}_n(\tilde{\mathbf{x}}), \mathbf{C}_n(\tilde{\mathbf{x}}))$, avec $\mathbf{C}_n(\tilde{\mathbf{x}}) := (k_n(\mathbf{x}_i, \mathbf{x}_j))_{i,j}$ la matrice de covariance de taille $n_{data} \times n_{data}$ conditionnée par $(\mathbf{x}_{doe}, \mathcal{M}(\mathbf{x}_{doe}))$ calculée aux points $\tilde{\mathbf{x}}$. Elle est obtenue suivant l'expression suivante :

$$\mathcal{G}_n(\tilde{\mathbf{x}}) = \boldsymbol{\mu}_n(\tilde{\mathbf{x}}) + \mathbf{L}_n(\tilde{\mathbf{x}})\boldsymbol{\xi} \quad (1.17)$$

avec $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}_{n_{data}}, \mathbf{I}_{n_{data}, n_{data}})$ et $\mathbf{L}_n(\tilde{\mathbf{x}}) \in \mathbb{R}^{n_{data} \times n_{data}}$ obtenue par décomposition de Cholesky de $\mathbf{C}_n(\tilde{\mathbf{x}})$ tel que :

$$\mathbf{C}_n(\tilde{\mathbf{x}}) = \mathbf{L}_n(\tilde{\mathbf{x}})\mathbf{L}_n(\tilde{\mathbf{x}})^T \quad (1.18)$$

La procédure de génération des trajectoires d'un GP conditionné \mathcal{G}_n est donnée par l'Algorithme 1.

Algorithm 1 Procédure d'évaluation de n_t trajectoires d'un GP

Require: $\tilde{\mathbf{x}}, \mathcal{G}_n(\mathbf{x}), n_t$

- 1: Construction de $\mathbf{C}_n(\tilde{\mathbf{x}}) := (k_n(\mathbf{x}_i, \mathbf{x}_j))_{i,j}$
 - 2: Décomposition de Cholesky de $\mathbf{C}_n(\tilde{\mathbf{x}}) = \mathbf{L}_n(\tilde{\mathbf{x}})\mathbf{L}_n(\tilde{\mathbf{x}})^T$
 - 3: Génération d'une matrice $\boldsymbol{\xi}_{n_t}$ de taille $n_{data} \times n_t$ dont les colonnes sont échantillonnées selon la distribution $\mathcal{N}(\mathbf{0}_{n_{data}}, \mathbf{I}_{n_{data}, n_{data}})$
 - 4: Évaluation des n_t trajectoires : $g_n(\tilde{\mathbf{x}}) = \boldsymbol{\mu}_n(\tilde{\mathbf{x}}) + \mathbf{L}_n(\tilde{\mathbf{x}})\boldsymbol{\xi}_{n_t}$
-

La Figure 1.3 représente 3 trajectoires simulées à partir du GP de la Figure 1.1 à $n_{data} = 100$ points du domaine $[0, 1]$.

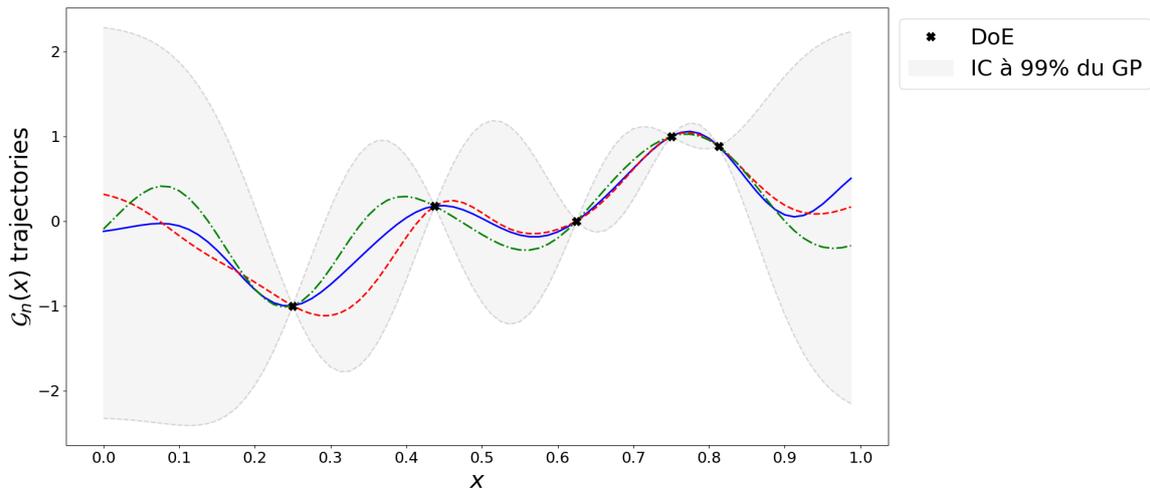


FIGURE 1.3 – Exemple de trajectoires générées à partir d'un GP

Cependant pour des jeux de données de taille n_{data} importante et une dimension n_u importante, le calcul de réalisations de $\mathcal{G}_n(\mathbf{x})$ représente un coût numérique très élevé et peut être sujet à des problèmes numériques tels que des matrices de covariance mal conditionnées.

De plus, la variance de la prédiction du GP \mathcal{G}_n est proche de zéro aux points voisins des points du DoE ce qui mène quasi systématiquement à des matrices de covariance conditionnées $\mathbf{C}_n(\tilde{\mathbf{x}})$ singulières, rendant ainsi impossible le calcul de la décomposition de Cholesky et donc le calcul des réalisations du GP par cette méthode. Ce problème numérique peut être évité en simulant un GP non conditionné [RASMUSSEN et WILLIAMS, 2006; VILLEMONTAIX et al., 2009; LE GRATIET et al., 2014], puis en conditionnant les réalisations comme suit :

Soit $\tilde{\mathcal{G}}(\mathbf{x})$ un GP centré ayant la même fonction de covariance $k(\mathbf{x}, \mathbf{x}') = \sigma_Z^2 r(\mathbf{x}, \mathbf{x}')$ que $\mathcal{G}(\mathbf{x})$:

$$\tilde{\mathcal{G}}(\mathbf{x}) \sim GP(0, \sigma_Z^2 r(\mathbf{x}, \mathbf{x}')) \quad (1.19)$$

où les $\boldsymbol{\theta}$ et σ_Z sont les hyperparamètres estimés par maximum de vraisemblance pour \mathcal{G}_n à partir des observations $(\mathbf{x}_{doe}, \mathcal{M}(\mathbf{x}_{doe}))$.

On note $\tilde{\mu}(\mathbf{x})$ sa moyenne prédite en conditionnant l'hyperparamètre $\boldsymbol{\beta}$ aux variables aléatoires $\tilde{\mathcal{G}}(\mathbf{x}_{doe})$:

$$\tilde{\mu}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \tilde{\boldsymbol{\beta}} + \mathbf{k}(\mathbf{x})^T \mathbf{C}^{-1} (\tilde{\mathcal{G}}(\mathbf{x}_{doe}) - \mathbf{F} \tilde{\boldsymbol{\beta}}) \quad (1.20)$$

avec $\tilde{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{C}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{C}^{-1} \tilde{\mathcal{G}}(\mathbf{x}_{doe})$ l'estimateur par maximum de vraisemblance des hyperparamètres $\boldsymbol{\beta}$ conditionnés aux points $(\mathbf{x}_{doe}, \tilde{\mathcal{G}}(\mathbf{x}_{doe}))$.

Définissons ensuite $\tilde{\mathcal{G}}_n(\mathbf{x})$:

$$\tilde{\mathcal{G}}_n(\mathbf{x}) = \mu_n(\mathbf{x}) - \tilde{\mu}(\mathbf{x}) + \tilde{\mathcal{G}}(\mathbf{x}) \quad (1.21)$$

avec $\mu_n(\mathbf{x})$ la moyenne prédite par $\mathcal{G}_n(\mathbf{x})$ aux points \mathbf{x} .

On note que $\tilde{\mathcal{G}}_n(\mathbf{x})$ suit la même loi de probabilité que $\mathcal{G}_n(\mathbf{x})$, c.-à-d. $\tilde{\mathcal{G}}_n(\mathbf{x})$ est un GP de même moyenne et de même variance que $\mathcal{G}_n(\mathbf{x})$, conditionnellement aux observations $(\mathbf{x}_{doe}, \mathbf{y}_{doe})$. On a en d'autres mots :

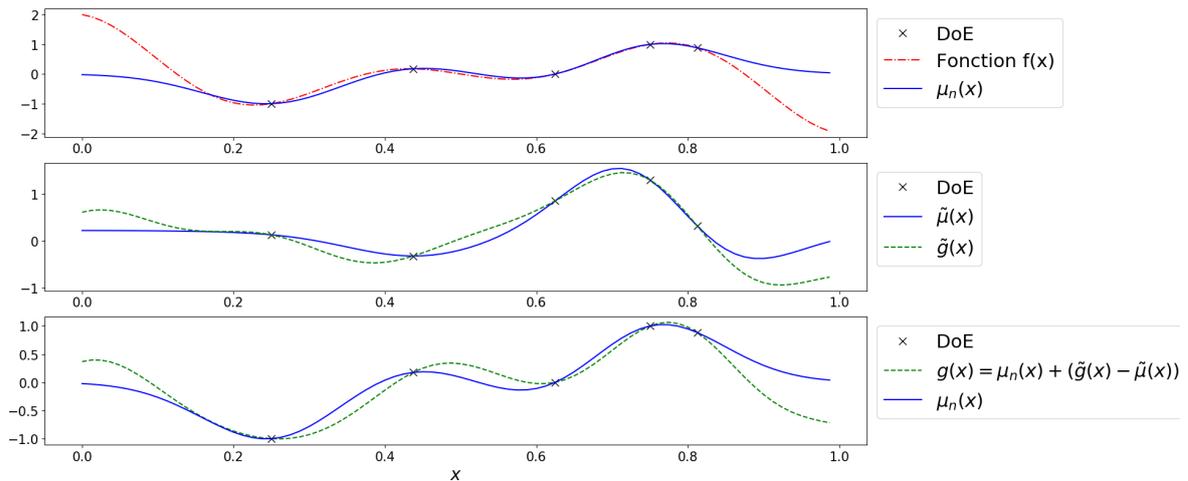
$$\tilde{\mathcal{G}}_n(\mathbf{x}) \stackrel{\mathcal{L}}{=} \mathcal{G}_n(\mathbf{x}) \quad (1.22)$$

En résumé, les différentes étapes de la méthode présentée ci-dessus pour obtenir une réalisation $g(\mathbf{x})$ du GP $\mathcal{G}_n(x)$ sont les suivantes :

1. Construction du GP $\mathcal{G}_n(x)$ conditionné aux observations $(\mathbf{x}_{doe}, \mathcal{M}(\mathbf{x}_{doe}))$, c.-à-d. estimation des hyperparamètres $\boldsymbol{\theta}$, σ_z et β par maximum de vraisemblance.
2. Construction du GP $\tilde{\mathcal{G}}(\mathbf{x}) \sim GP(0, \sigma_z^2 r(\mathbf{x}, \mathbf{x}'))$ en utilisant les estimations obtenues pour $\boldsymbol{\theta}$ et σ_z dans l'étape 1.
3. Calcul de la moyenne $\mu_n(\mathbf{x})$ de la prédiction par $\mathcal{G}_n(x)$.
4. Simulation d'une réalisation \tilde{g} de $\tilde{\mathcal{G}}$ aux points $(\mathbf{x}, \mathbf{x}_{doe})$
5. Calcul de $\tilde{\mu}(\mathbf{x})$ pour la réalisation $\tilde{g}(\mathbf{x}_{doe})$ à partir de l'Eq. (1.20).
6. Calcul de $g(\mathbf{x}) = \mu_n(\mathbf{x}) - \tilde{\mu}(\mathbf{x}) + \tilde{g}(\mathbf{x})$.

Ainsi, la simulation d'une trajectoire de $\mathcal{G}_n(x)$ peut être obtenue en ajoutant l'erreur de prédiction $\tilde{\mathcal{G}}(\mathbf{x}) - \tilde{\mu}(\mathbf{x})$ de $\tilde{\mathcal{G}}(\mathbf{x})$ à sa moyenne $\mu_n(\mathbf{x})$. Cette méthode permet de simuler le GP non conditionné $\tilde{\mathcal{G}}(\mathbf{x})$ à la place de $\mathcal{G}_n(\mathbf{x})$. Contrairement au GP conditionné, les valeurs de la variance du GP non conditionné aux points proches du DoE ne sont pas proches de zéro. Cela permet d'avoir un meilleur conditionnement de la matrice de covariance et d'éviter ainsi les problèmes numériques qui en résultent.

La Figure 1.4 illustre l'obtention d'une réalisation $g(x)$ du GP $\mathcal{G}_n(x)$ en utilisant cette procédure sur l'exemple utilisé précédemment dans cette section. Sur la Figure du haut, on a représenté la fonction interpolée ainsi que la moyenne $\mu_n(\mathbf{x})$ du GP \mathcal{G}_n conditionné aux points du DoE, dont on souhaite simuler des trajectoires. Sur la Figure centrale, on a tracé une réalisation $\tilde{g}(x)$ du GP $\tilde{\mathcal{G}}(x)$ et une réalisation de la moyenne $\tilde{\mu}(\mathbf{x})$ calculée à partir de la réalisation aux points du DoE $\tilde{g}(\mathbf{x}_{doe})$. Finalement, sur la dernière Figure (figure du bas), la réalisation $g(x)$ du GP \mathcal{G}_n obtenue à partir de la réalisation $\tilde{g}(x)$ du GP $\tilde{\mathcal{G}}(x)$ et de la formule Eq. (1.21) est représentée.


 FIGURE 1.4 – Exemple de génération d'une trajectoire de $\mathcal{G}_n(x)$ à partir de l'Eq. (1.21)

De plus, l'utilisation de cette méthode permet d'utiliser des représentations de champs aléatoires comme l'expansion de [représentation de Karhunen-Loève](#) pour générer des trajectoires du GP non conditionné $\tilde{\mathcal{G}}(\mathbf{x})$. Cette technique est très intéressante car elle permet de résoudre les problèmes numériques liées à la simulation de trajectoires sur des grands ensembles de données (mauvais conditionnement des matrices de covariance et coûts de calcul très importants). La représentation de [représentation de Karhunen-Loève de GPs](#) est présentée dans la section suivante.

1.1.2.7 Représentation de Karhunen-Loève d'un processus gaussien

L'expansion de KL d'un GP $\mathcal{G}(\mathbf{x})$ est basée sur une décomposition spectrale de sa fonction d'autocovariance ou noyau. L'expansion de KL d'un GP $\mathcal{G} \sim GP(\mu(\cdot), k(\cdot, \cdot))$ s'écrit [[KARHUNEN, 1947](#); [LOEVE, 1948](#)] :

$$\mathcal{G}(\mathbf{x}, \omega) = \mu(\mathbf{x}) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \phi_i(\mathbf{x}) \xi_i(\omega) \quad (1.23)$$

où les $\xi_i(\omega) : \Omega_P \rightarrow \mathbb{R}$ sont des variables aléatoires normales centrées réduites indépendantes et $\lambda_i \geq 0$, $\phi_i : \mathcal{X} \rightarrow \mathbb{R}$ les valeurs et fonctions propres normalisées du noyau.

En effet d'après le théorème de Mercer, les λ_i sont positives et le noyau k admet la représentation :

$$k(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \phi_i(\mathbf{x}) \phi_i(\mathbf{x}') \quad (1.24)$$

Les fonctions propres et les valeurs propres associées sont obtenues en résolvant

l'équation intégrale suivante :

$$\lambda_i \phi_i(\mathbf{x}) = \int_{\mathcal{X}} k(\mathbf{x}, \mathbf{x}') \phi_i(\mathbf{x}') d\mathbf{x}' \quad (1.25)$$

En pratique, il existe un nombre infini de fonctions $\phi_i(\mathbf{x})$.

L'expansion de KL peut être approchée en tronquant l'expansion au $M^{\text{ième}}$ terme après avoir classé les valeurs propres λ_i et les fonctions propres $\phi_i(\mathbf{x})$ associées par ordre décroissant des λ_i :

$$\hat{\mathcal{G}}(\mathbf{x}, \omega) = \mu(\mathbf{x}) + \sum_{i=1}^M \sqrt{\lambda_i} \phi_i(\mathbf{x}) \xi_i(\omega) \quad (1.26)$$

Pour $M > 0$ fixé, l'approximation $\hat{\mathcal{G}}(\mathbf{x}, \omega)$ est optimal par rapport à l'erreur quadratique moyenne globale $e_{\mathcal{G}}$:

$$e_{\mathcal{G}} = \int_{\mathcal{X}} \mathbb{E} \left[(\mathcal{G}(\mathbf{x}, \omega) - \hat{\mathcal{G}}(\mathbf{x}, \omega))^2 \right] d\mathbf{x} \quad (1.27)$$

Et la variance de $\hat{\mathcal{G}}(\mathbf{x}, \omega)$ est donnée par :

$$\text{Var} [\hat{\mathcal{G}}(\mathbf{x}, \omega)] = \sum_{i=1}^M \lambda_i \phi_i^2(\mathbf{x}) \quad (1.28)$$

Les fonctions et valeurs propres peuvent alors être estimées numériquement par la procédure de Nyström ou par des méthodes de Galerkin comme présenté dans [BETZ et al., 2014]. Dans cette thèse, la méthode de Nyström a été utilisée et est détaillée ci-dessous.

Dans la méthode de Nyström, l'intégrale de l'Eq. (1.25) est estimée par intégration numérique. On a ainsi :

$$\hat{\lambda}_i \hat{\phi}_i(\mathbf{x}) = \sum_{j=1}^{n_q} w_j k(\mathbf{x}, \mathbf{x}_j) \hat{\phi}_i(\mathbf{x}_j) \quad (1.29)$$

avec $(\mathbf{x}_j)_{j=1, \dots, n_q} \in \mathcal{X}^{n_q}$ un ensemble de n_q points d'intégration et w_j les poids d'intégration associés à chaque \mathbf{x}_j . Ainsi dans la méthode de Nyström, l'équation (1.29) est résolue aux points d'intégration, ce qui équivaut à résoudre le système matriciel suivant :

$$\mathbf{C} \mathbf{W} \mathbf{y}_i = \hat{\lambda}_i \mathbf{y}_i \quad (1.30)$$

avec \mathbf{C} la matrice (symétrique définie positive) de covariance $n_q \times n_q$ calculée aux points d'intégration, $\mathbf{W} = \text{Diag}(w_1, \dots, w_{n_q})$ et \mathbf{y}_i un vecteur de taille n_q tel que $y_{i,j} = \hat{\phi}_i(\mathbf{x}_j)$.

Le problème de l'Eq. (1.30) peut se réécrire sous la forme :

$$\mathbf{B}\mathbf{y}_i^* = \hat{\lambda}_i \mathbf{y}_i^* \quad (1.31)$$

avec :

$$\mathbf{B} = \mathbf{W}^{\frac{1}{2}} \mathbf{C} \mathbf{W}^{\frac{1}{2}} \quad (1.32)$$

\mathbf{B} est une matrice symétrique définie positive, ses valeurs propres $\hat{\lambda}_i$ sont donc positives et ses vecteurs propres \mathbf{y}_i^* sont orthogonaux. On peut en déduire les vecteurs propres $\mathbf{y}_i = \mathbf{W}^{-\frac{1}{2}} \mathbf{y}_i^*$. L'approximation suivante des fonctions propres $\phi_i(\mathbf{x})$ est ainsi donnée par :

$$\hat{\phi}_i(\mathbf{x}) = \frac{1}{\hat{\lambda}_i} \sum_{j=1}^{n_q} \sqrt{w_j} \mathbf{y}_{i,j}^* k(\mathbf{x}, \mathbf{x}_j) \quad (1.33)$$

Finalement, on peut approcher le GP \mathcal{G} par :

$$\hat{\mathcal{G}}(\mathbf{x}, \omega) \approx \mu(\mathbf{x}) + \sum_{i=1}^M \frac{1}{\sqrt{\hat{\lambda}_i}} \sum_{j=1}^{n_q} k(\mathbf{x}, \mathbf{x}_j) \sqrt{w_j} \mathbf{y}_{i,j}^* \xi_i(\omega) \quad (1.34)$$

Un cas particulier de la méthode de Nyström consiste à utiliser une distribution uniforme des points d'intégration $(\mathbf{x}_j)_{j=\llbracket 1, n_q \rrbracket}$ sur le domaine \mathcal{X} . Cette procédure est dérivée de la méthode *Expansion optimal linear estimation (EOLE)* développée dans [LI et DER KIUREGHIAN, 1993], qui est une méthode de développement en série de champs aléatoires.

Dans le cadre de cette méthode, les n_q points d'intégration $(\mathbf{x}_j)_{j=\llbracket 1, n_q \rrbracket}$ sont soit des échantillons générés aléatoirement selon la distribution uniforme sur \mathcal{X} , soit des noeuds équidistants d'une grille structurée. Dans ces deux cas, les poids d'intégrations w_j sont tous égaux $w_j = w$, $\forall j \in \{1, \dots, n_q\}$ et valent $w = \frac{|\mathcal{X}|}{n_q}$ où $|\mathcal{X}| = \int_{\mathcal{X}} d\mathbf{x}$.

Notons que dans le cas d'une grille structurée, la méthode utilisée est en fait une intégration par la formule des rectangles. D'autres formules de quadrature d'ordre supérieur peuvent également être utilisées pour cette intégration en adaptant alors les valeurs prises par les poids d'intégration.

Remarque

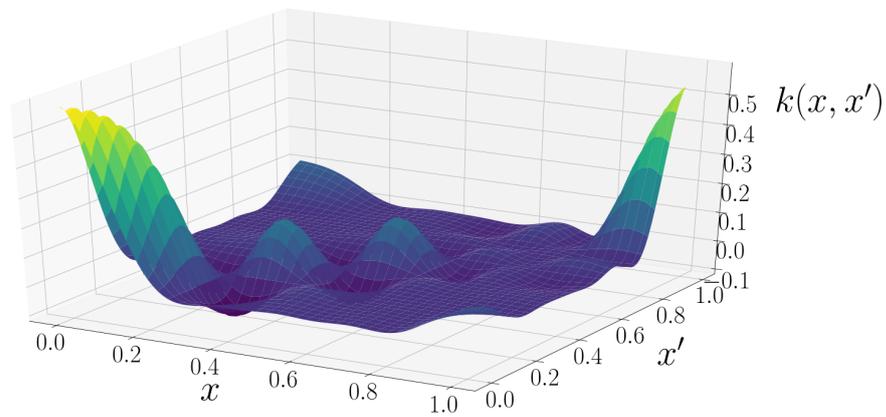
- Dans le cadre de la Section 1.1.2.6, une représentation de KL du GP non conditionné $\tilde{\mathcal{G}}(\mathbf{x})$ peut être utilisée pour simuler ses trajectoires. Dans ce cas on a la représentation tronquée de KL suivante du GP $\tilde{\mathcal{G}}(\mathbf{x})$:

$$\tilde{\mathcal{G}}(\mathbf{x}) \approx \sum_{i=1}^{n_q} \frac{1}{\sqrt{\hat{\lambda}_i}} \sum_{j=1}^{n_q} k(\mathbf{x}, \mathbf{x}_j) \sqrt{w_j} \mathbf{y}_{i,j}^* \xi_i \quad (1.35)$$

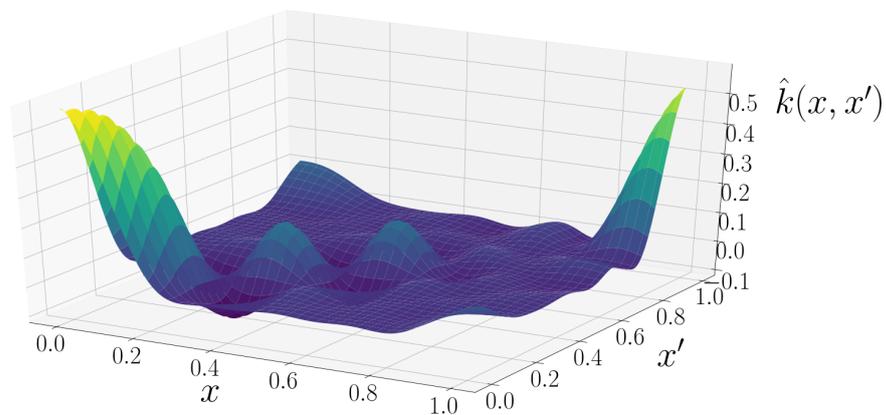
et l'approximation de $\tilde{\mu}(\mathbf{x})$ définie par l'Eq. (1.20) pour une trajectoire ξ de $\tilde{\mathcal{G}}(\mathbf{x})$ est calculée pour :

$$\tilde{\mathcal{G}}(\mathbf{x}_{doe}) = \sum_{i=1}^{n_q} \frac{1}{\sqrt{\hat{\lambda}_i}} \sum_{j=1}^{n_q} k(\mathbf{x}_{doe}, \mathbf{x}_j) \sqrt{w_j} y_{i,j}^* \xi_i \quad (1.36)$$

La Figure 1.5 représente le noyau de covariance $k(x, x') = \sigma_Z^2 r(x, x')$ conditionné par $(\mathbf{x}_{doe}, \mathbf{y}_{doe})$ sur l'exemple précédent de la fonction $G(\mathbf{x})$ et les valeurs de sa représentation de **KL**, estimée grâce à la méthode de Nyström sur 13 points d'intégration, sur le domaine \mathcal{X}^2 . On observe sur cette Figure que l'expansion de **KL** du **GP** calculée pour 13 points d'intégration donne une approximation correcte du noyau de covariance.



(a) Noyau de covariance $k(x, x')$ sur \mathcal{X}^2



(b) Approximation du noyau de covariance par une représentation de **KL** sur \mathcal{X}^2

FIGURE 1.5 – Exemple de représentation de **KL** du noyau de covariance d'un **GP**

Pour résumer, dans cette section et les sections précédentes, nous avons présenté la **GPR** et les différents outils numériques permettant de construire un modèle de **GP** puis d'en obtenir des réalisations.

Dans la section suivante, une extension des modèles de **GP**, le modèle de co-Krigeage, est présentée. Ce dernier utilise des **GPs** pour modéliser une quantité pour laquelle on dispose des données de fidélités différentes.

1.1.2.8 Modèle pour la multi-fidélité : co-krigeage

Un code de calcul peut être exécuté avec différents niveaux de complexité ou différents niveaux de précision d'un modèle. Le co-krigeage permet de combiner une grande quantité de données basse fidélité d'un modèle avec peu de données haute fidélité pour améliorer la précision du métamodèle du modèle haute fidélité, qui correspond généralement au modèle le plus coûteux à simuler. Un premier modèle de co-krigeage a été proposé par [KENNEDY et O'HAGAN, 2000] pour des analyses à plusieurs niveaux de fidélité distincts. La présentation du co-krigeage dans [FORRESTER et al., 2008] donne davantage de détails sur l'estimation des paramètres du modèle. Des approches bayésiennes pour l'estimation des hyperparamètres θ de ce modèle ont également été proposées [LE GRATIET, 2013; QIAN et WU, 2008] car elles permettent une meilleure estimation de la variance de la prédiction.

Dans cette section, nous présenterons uniquement les formules pour la prédiction par co-Krigeage d'un problème pour lequel on dispose de deux modèles à deux niveaux de fidélité distincts : $G_b(\mathbf{x})$ et $G_h(\mathbf{x})$.

Le co-krigeage est une forme **GPR** qui permet de corrélérer plusieurs ensembles de données $\mathbf{x}_{doe} = (\mathbf{x}_b, \mathbf{x}_h)$, qui sont des données basse fidélité $\{\mathbf{x}_b, y_b = G_b(\mathbf{x}_b)\}$ et haute fidélité $\{\mathbf{x}_h, y_h = G_h(\mathbf{x}_h)\}$. Ce modèle est basé sur des **GPs** \mathcal{G}_h et \mathcal{G}_b approchant respectivement les réponses y_h et y_b .

Le modèle auto-régressif de [KENNEDY et O'HAGAN, 2000] impose que :

$$CoVar(\mathcal{G}_h(\mathbf{x}'), \mathcal{G}_b(\mathbf{x}) | \mathcal{G}_b(\mathbf{x}')) = 0, \quad \forall \mathbf{x} \neq \mathbf{x}' \quad (1.37)$$

Ce qui implique que lorsque la réponse haute fidélité est connue en un point, la réponse basse fidélité n'apporte aucune information supplémentaire.

La réponse haute-fidélité peut s'écrire :

$$\mathcal{G}_h(\mathbf{x}) = \rho \mathcal{G}_b(\mathbf{x}) + \delta(\mathbf{x}) \quad (1.38)$$

avec ρ un facteur d'échelle constant et $\delta(\mathbf{x})$ la différence entre \mathcal{G}_h et $\rho \mathcal{G}_b$.

La prédiction de la sortie du modèle haute fidélité peut alors être obtenue en conditionnant le **GP** à l'ensemble des observations $\{\mathbf{x}_{doe}, \mathbf{y}_{doe} = (\mathbf{y}_b \quad \mathbf{y}_h)\}$, aux hyperpara-

mètres $\boldsymbol{\beta}$, σ_Z , $\boldsymbol{\theta}$ de chaque GP et à ρ :

$$\mathcal{G}_{h,n} = \mathcal{G}_h | \mathbf{x}_{doe}, \mathbf{y}, (\boldsymbol{\beta}_h, \boldsymbol{\beta}_b, \rho), (\sigma_{Z,h}^2, \sigma_{Z,b}^2), (\boldsymbol{\theta}_h, \boldsymbol{\theta}_b) \quad (1.39)$$

La moyenne $\mu_{h,n}(\mathbf{x})$ et la variance $\sigma_{h,n}^2(\mathbf{x})$ de la prédiction de la réponse par le GP conditionné $\mathcal{G}_{h,n}$ sont données par les expressions :

$$\mu_{h,n}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + \mathbf{k}(\mathbf{x})^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{H} \boldsymbol{\beta}) \quad (1.40)$$

$$\sigma_{h,n}^2(\mathbf{x}) = \rho^2 \sigma_{Z,b}^2 + \sigma_{Z,h}^2 - \mathbf{k}(\mathbf{x})^T \mathbf{V}^{-1} \mathbf{k}(\mathbf{x}) \quad (1.41)$$

où $\boldsymbol{\beta} = (\boldsymbol{\beta}_b \ \boldsymbol{\beta}_h)$, $\mathbf{f}(\mathbf{x})^T = (\rho \mathbf{f}_b(\mathbf{x})^T \ \mathbf{f}_h(\mathbf{x})^T)$, $\mathbf{k}(\mathbf{x})^T = \text{CoVar}(\mathcal{G}_{n,h}(\mathbf{x}), \mathbf{y})$, \mathbf{H} est une matrice définie comme suit :

$$\mathbf{H} = \begin{pmatrix} \mathbf{F}_b(\mathbf{x}_b) & \mathbf{0} \\ \rho \mathbf{F}_b(\mathbf{x}_h) & \mathbf{F}_h(\mathbf{x}_h) \end{pmatrix} \quad (1.42)$$

et \mathbf{V} est la matrice de covariance aux points du DoE (haute et basse fidélités) :

$$\mathbf{V} = \begin{pmatrix} \sigma_{Z,b}^2 r_b(\mathbf{x}_b) & \rho \sigma_{Z,b}^2 r_b(\mathbf{x}_b, \mathbf{x}_h) \\ \rho \sigma_{Z,b}^2 r_b(\mathbf{x}_b, \mathbf{x}_h) & \rho \sigma_{Z,b}^2 r_b(\mathbf{x}_h) + \sigma_{Z,h}^2 r_h(\mathbf{x}_h) \end{pmatrix} \quad (1.43)$$

1.1.2.9 Bilan

Dans cette section, nous avons présenté des métamodèles basés sur les GPs. Dans un premier temps, la GPR a été introduite ainsi que différents outils numériques pour la génération de réalisations d'un GP. Dans un second temps, une extension des modèles basés sur les GPs pour la multi-fidélité, le co-Krigeage, a été présenté.

Ces métamodèles permettent de prédire la sortie du code de calcul avec un coût de calcul moindre car les modèles construits par apprentissage statistique donnent une approximation à partir d'expressions analytiques.

Dans la suite, une approche différente de modélisation consistant à construire un modèle réduit du modèle numérique initial, gardant ainsi des informations sur la physique du problème, est présentée.

1.1.3 Méthodes de réduction de modèles

1.1.3.1 Introduction

Nous nous intéressons maintenant aux méthodes de réduction du modèle numérique, qui ont l'avantage par rapport aux métamodèles de tenir compte de la physique du

problème considéré et aussi d'approcher un champ de solutions d'un modèle et pas seulement un scalaire.

Les approches de réduction de modèles reposent sur l'idée qu'un phénomène physique, même complexe, peut être décrit par ses modes dominants et que la solution d'un problème modélisant le comportement d'un système peut donc être obtenue dans le sous-espace engendré par ces modes. Ces approches sont notamment utilisées dans le cadre de la réduction de systèmes d'équations obtenus par la méthode des éléments finis, dont la résolution peut atteindre des temps de calcul très élevés pour des systèmes complexes.

Dans cette section, la modélisation d'un problème de mécanique par la méthode des éléments finis et le principe de la réduction de ces modèles sont introduits. Un tour d'horizon, non exhaustif, des différentes approches de réduction de modèles est ensuite effectué puis une description plus précise des méthodes pour la réduction de modèles linéaires paramétriques est proposée.

1.1.3.2 Formulation du problème

La méthode des **EF** permet de décrire numériquement le comportement physique d'un système régi par une équation d'équilibre et défini par un ensemble de paramètres géométriques et physiques. Dans le cadre de la mécanique des milieux continus, elle permet de calculer numériquement la déformation des solides et l'écoulement des fluides. Ici on s'intéresse uniquement à la déformation des solides.

On considère le domaine d'étude $\Omega \subset \mathbb{R}^d$, $d = 1, 2$ ou 3 , de frontière $\Gamma = \partial\Omega$. L'état de la structure étudiée est défini par les champs de déplacements, de déformations et de contraintes, définis sur Ω et sur le domaine de temps $[0, T]$.

La formulation la plus générale permettant de décrire le comportement physique d'une structure, qui correspond au cas de la dynamique avec amortissement, est donnée par le système suivant :

$$\left\{ \begin{array}{ll} \operatorname{div} \sigma + f_v = \rho \ddot{u} + \mu \dot{u} & \text{sur } \Omega \text{ équation d'équilibre local} \\ \sigma = D\epsilon & \text{sur } \Omega \text{ loi de comportement} \\ \epsilon = \frac{1}{2} (\operatorname{grad}(u) + \operatorname{grad}^T(u)) = \mathcal{L}u & \text{sur } \Omega \text{ relation cinématique} \end{array} \right. \quad (1.44)$$

auquel s'ajoute les conditions aux limites suivantes :

$$\left\{ \begin{array}{ll} u = u_D & \text{sur } \Gamma_D \text{ condition limite (Dirichlet)} \\ \sigma \cdot n = \sigma_N & \text{sur } \Gamma_N \text{ condition limite (Neumann)} \\ \dot{u}(t=0) = \dot{u}_0 & \text{sur } \Omega \text{ condition initiale} \\ u(t=0) = u_0 & \text{sur } \Omega \text{ condition initiale} \end{array} \right. \quad (1.45)$$

avec ρ la masse volumique, μ le facteur d'amortissement visqueux, f_v la force volumique, ϵ le tenseur des déformations, σ le tenseur des contraintes, D la loi de comportement (relation de Hooke), $\Gamma_D \subset \Gamma$ et $\Gamma_N \subset \Gamma$.

Notons que cette formulation correspond au cas le plus général. Ainsi, si une structure est sans amortissement alors $\mu = 0$ ou si le problème est statique, c.-à-d. qu'il ne dépend pas du temps, $\ddot{u} = \dot{u} = 0$.

Ce problème mis sous forme faible (formulation en déplacements) est alors donné par :

$$\int_{\Omega} \rho v \ddot{u} + \int_{\Omega} \mu v \dot{u} + \int_{\Omega} \mathcal{L}v : D\mathcal{L}u - \int_{\Omega} f_v v - \int_{\Gamma_N} \sigma_N v = 0, \quad \forall v \text{ tel que } v = u_D \text{ sur } \Gamma_D \quad (1.46)$$

La méthode des EF se propose de discrétiser le problème Eq. (1.46) afin d'obtenir un système matriciel à résoudre.

Ainsi, après discrétisation par la méthode des EF des équations d'équilibre et l'application des conditions aux limites, le système matriciel suivant est obtenu :

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}(\mathbf{x})\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{F}(t) \quad (1.47)$$

avec \mathbf{M} la matrice de masse, \mathbf{C} la matrice d'amortissement, \mathbf{K} la matrice de rigidité, $\mathbf{u}(t)$ le vecteur des déplacements, \mathbf{F} le vecteur des efforts extérieurs.

Dans le cadre de cette thèse, on s'intéresse plus précisément à des problèmes statiques linéaires, c.-à-d. indépendants du temps. La modélisation EF correspondante est donnée par le système linéaire suivant :

$$\mathbf{K}\mathbf{u} = \mathbf{F} \quad (1.48)$$

Ce système est en fait paramétré par l'ensemble des paramètres physiques et géométriques $\mathbf{x} \in \mathcal{X}$ du modèle étudié. Finalement, le système matriciel étudié est le suivant :

$$\mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \quad (1.49)$$

La méthode des EF revient à étudier le comportement d'une structure dans un espace V de dimension n_{ddl} , le nombre de Degrés de liberté (ddl) de la discrétisation du problème. A l'heure actuelle les modèles EF donnent les résultats les plus précis pour l'étude du comportement des systèmes physiques notamment en comparaison aux méthodes analytiques utilisées avant l'avènement des EF. Cependant, la résolution de systèmes tel que le système (1.47) pour un grand nombre de ddl nécessite des temps de calcul importants.

Les méthodes de réduction de modèles consistent à construire une base optimale en

un sens à définir $\Psi \in \mathbb{R}^{n_{ddl} \times n_r}$ afin de se ramener dans un sous-espace V_r de dimension n_r plus faible que la dimension de l'espace de départ V . Une telle base est constituée de n_r vecteurs Ψ_i de dimension n_{ddl} , la dimension d'une solution du problème EF, et est généralement choisie orthonormée. Ces méthodes permettent alors d'obtenir une approximation \mathbf{u}_r du vecteur de déplacements \mathbf{u} en l'exprimant dans la base Ψ :

$$\begin{aligned} \mathbf{u}_r(\mathbf{x}) &= \sum_{i=1}^{n_r} \alpha_i(\mathbf{x}) \Psi_i \\ &= \Psi \boldsymbol{\alpha}(\mathbf{x}) \end{aligned} \tag{1.50}$$

Les coefficients $\boldsymbol{\alpha}(\mathbf{x})$ du vecteur déplacement $\mathbf{u}(\mathbf{x})$ dans la base Ψ sont obtenus en résolvant le problème de l'équation (1.47) projeté dans cette base :

$$\mathbf{K}_r(\mathbf{x}) \boldsymbol{\alpha}(\mathbf{x}) = \Psi^T \mathbf{F}(\mathbf{x}) \tag{1.51}$$

avec $\mathbf{K}_r(\mathbf{x}) = \Psi^T \mathbf{K}(\mathbf{x}) \Psi$ la matrice de rigidité réduite.

L'obtention d'une solution par le modèle réduit passe alors par la résolution du système matriciel (1.51) de taille $n_r \ll n_{ddl}$ et est donc numériquement moins coûteuse que le calcul de la solution EF.

Notons que la décomposition (1.50) n'est pas unique vis à vis du choix de Ψ . Les sections suivantes décrivent différentes approches de construction de la base Ψ .

1.1.3.3 Tour d'horizon des méthodes de réduction majeures

De nombreuses méthodes de réduction de modèles existent dans la littérature que l'on peut classer en différentes catégories : optimalité de l'ordre de réduction, capacité à traiter des problèmes non-linéaires, capacité à conserver des paramètres de conception dans le modèle réduit.

Parmi les méthodes existantes, on peut citer les méthodes de sous-structuration telles que la méthode de Guyan [GUYAN, 1965], la méthode de Craig-Bampton [CRAIG et BAMPTON, 1968; FRISWELL et al., 1995; RIXEN, 2004] ou la méthode des modes de composants [ROCHES et al., 2010], qui s'appuient sur une condensation des ddl pour réduire la taille du modèle.

D'autres approches existent permettant de construire de manière optimale par rapport à un critère d'erreur la base réduite, dont la *Décomposition Orthogonale aux valeurs Propres*, ou *Proper Orthogonal Decomposition* (POD), les approches de type algorithme glouton, les approches goal-oriented ou des méthodes de constructions adaptatives, qui seront décrites en détail dans la section suivante. Les méthodes citées précédemment sont des méthodes a posteriori, se basant sur des observations du problème étudié. Il existe également des méthodes de réduction a priori, qui ne nécessitent aucune information sur la solution pour construire la base réduite. C'est le cas de la

Décomposition Propre Généralisée, ou *Proper Generalized Decomposition* (PGD) [CHINESTA et al., 2011; LADEVÈZE, 2012] introduite par [LADEVÈZE, 1985], qui consiste à construire itérativement la base réduite en partant directement des équations d'équilibre. Néanmoins, les méthodes de construction a priori sont bien plus intrusives que les approches a posteriori. Des travaux récents [ZOU et al., 2018; LEON et al., 2019] ont ainsi proposé des méthodes de PGD non intrusives afin de faciliter leur utilisation pour des codes de calcul industriels.

Pour notre étude, le caractère non intrusif des méthodes a posteriori nous a poussés à nous intéresser uniquement à ces dernières.

1.1.3.4 Méthodes optimales pour la réduction de modèles linéaires paramétriques

Dans les paragraphes suivants, les méthodes optimales de type a posteriori de construction de la base réduite Ψ pour des systèmes linéaires paramétrés du type (1.49) sont présentées.

1.1.3.4.1 Décomposition Orthogonale aux valeurs Propres

La *Décomposition Orthogonale aux valeurs Propres*, ou *Proper Orthogonal Decomposition* (POD) est une technique pour extraire une base de modes propres orthogonaux dominants, permettant d'accéder à un modèle réduit représentatif du comportement d'un système. Elle a d'abord été introduite en mécanique des fluides et adaptée pour la turbulence par [LUMLEY, 1967]. Cette méthode est aujourd'hui présente dans divers domaines d'applications sous le nom de Décomposition de *représentation de Karhunen-Loève* (KL) ou Analyse en Composantes Principales (Principal Component Analysis). En effet, elle a d'abord été développée pour caractériser à moindre coût le comportement de systèmes complexes mais elle est aussi utilisée dans les domaines des statistiques, du traitement du signal, de la compression de données ou encore en contrôle optimal.

La méthode POD repose sur la représentation de la forme (1.50) de la solution $\mathbf{u}(\mathbf{x})$ d'un modèle EF. Pour construire l'approximation de \mathbf{u} , on dispose d'un ensemble de solutions $\mathcal{U} = (\mathbf{u}(\mathbf{x}_1), \dots, \mathbf{u}(\mathbf{x}_{n_v}))$ d'un modèle EF à n_{ddl} ddl pour n_v jeux de paramètres $\mathbf{x}_1, \dots, \mathbf{x}_{n_v}$, où $(\mathbf{x}_i)_{i \in \llbracket 1, n_v \rrbracket} \in \mathbb{R}^{n_u}$. Cependant la décomposition (1.50) n'est pas unique et est dépendante de la forme des fonctions de base. Dans le cadre de la POD, ces fonctions de base sont appelées modes POD et sont construites de manière à minimiser l'erreur de reconstruction au sens des moindres carrés :

$$\Psi = \arg \min \sum_{i=1}^{n_v} \left\| \mathbf{u}(\mathbf{x}_i) - \sum_{k=1}^{n_r} \langle \mathbf{u}(\mathbf{x}_i), \Psi_k \rangle \Psi_k \right\|_2^2 \quad (1.52)$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire entre deux fonctions et $n_r \leq n_v$ la taille retenue pour la base Ψ .

L'ensemble des observations de \mathbf{u} aux points $\mathbf{x}_1, \dots, \mathbf{x}_{n_v}$ est rangé dans la matrice

$$\mathbf{A} = [\mathbf{u}(\mathbf{x}_1), \dots, \mathbf{u}(\mathbf{x}_{n_v})] \quad (1.53)$$

de $\mathbb{R}^{n_{ddl} \times n_v}$, appelée matrice des snapshots.

Dans la littérature, on trouve plusieurs méthodes pour calculer ces modes mais la méthode la plus utilisée est certainement la *Décomposition en Valeurs Singulières, ou Singular Value Decomposition* (SVD), dont le principe est expliquée dans le paragraphe suivant.

La *Décomposition en Valeurs Singulières, ou Singular Value Decomposition* (SVD) de la matrice des Snapshots \mathbf{A} correspond à sa factorisation sous la forme :

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (1.54)$$

où $\mathbf{U} \in \mathbb{R}^{n_{ddl} \times n_{ddl}}$ et $\mathbf{V} \in \mathbb{R}^{n_v \times n_v}$ sont des matrices orthogonales, appelées vecteurs singuliers gauches et droits respectivement, et $\mathbf{\Sigma} \in \mathbb{R}^{n_{ddl} \times n_v}$ est une matrice diagonale contenant $r = \min(n_{ddl}, n_v)$ éléments diagonaux $(\sigma_i)_{i=1, \dots, r}$, appelés valeurs singulières de \mathbf{A} , tels que $\sigma_1 \geq \dots \geq \sigma_r \geq 0$. Dans le cadre de notre étude, on se trouve généralement dans le cas où $n_v \ll n_{ddl}$, donc $r = n_v$.

En pratique, l'obtention des valeurs singulières et des vecteurs singuliers associés par un calcul direct est fastidieuse. Ils sont donc calculés en résolvant le problème aux valeurs propres équivalent associé à la matrice carrée :

$$\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T, \quad \text{où } \mathbf{\Lambda} = \mathbf{\Sigma}^T \mathbf{\Sigma} \quad (1.55)$$

Ainsi les valeurs propres λ_i de ce problème sont telles que $\sigma_i = \sqrt{\lambda_i}$ et les vecteurs propres correspondent aux vecteurs singuliers droits \mathbf{V} . On peut ainsi retrouver les vecteurs singuliers gauches à l'aide de la relation $\mathbf{U} = \mathbf{\Sigma}^{-1} \mathbf{A} \mathbf{V}$.

La méthode de calcul des modes *POD* utilisant la *SVD* est la méthode des snapshots. Dans le cadre de cette méthode, les modes *POD* correspondent aux vecteurs singuliers gauche. Ainsi l'approximation de \mathbf{u} dans le sous-espace de taille $n_r \ll n_{ddl}$ est obtenue en retenant les $n_r \leq r$ premières valeurs singulières non nulles de la matrice $\mathbf{\Sigma}$ et les fonctions propres associées déterminées en résolvant le problème de l'Eq. (1.55).

1.1.3.4.2 Approche *offline/online* : algorithme glouton

Une stratégie *offline/online* peut être adoptée pour exploiter la base réduite. Celle-ci consiste à construire l'espace réduit au cours d'une phase *offline* afin de pouvoir projeter le système sur cet espace réduit et calculer rapidement les solutions réduites

lors de la phase *online*.

Une approche naturelle pour construire l'espace réduit lors de la phase *offline* consiste à enrichir itérativement la base réduite par des vecteurs afin de minimiser l'erreur d'approximation sur le domaine \mathcal{X} . Dans cette optique, il est possible d'utiliser une méthode comme l'algorithme glouton. Cette méthode permet de contrôler l'erreur du modèle réduit tout en assurant un bon compromis entre la précision et la taille du modèle réduit [VEROY et al., 2003; ZAHM et NOUY, 2016].

Ainsi à chaque itération de l'algorithme glouton, on recherche le paramètre \mathbf{x}^* pour lequel l'estimation de l'erreur $\Delta(\mathbf{x})$ est maximale :

$$\mathbf{x}^* = \arg \sup_{\mathbf{x} \in \mathcal{X}} \Delta(\mathbf{x}) \quad (1.56)$$

La mesure utilisée pour estimer l'erreur $\Delta(\mathbf{x})$ peut prendre plusieurs formes précisées ci-après.

La solution EF est alors évaluée pour le jeu de paramètre \mathbf{x}^* et est utilisée pour enrichir la base réduite Ψ .

L'utilisation de l'erreur $e(\mathbf{x}) = \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_V$ pour le problème d'optimisation donné par l'Eq. (1.56) nécessite de résoudre le problème EF complet à chaque itération de l'optimisation. Or on cherche à réduire au minimum le nombre de résolutions du système EF complet.

En pratique, ce sont donc d'autres mesures de l'erreur qui sont utilisées telles qu'une borne a posteriori de l'erreur $\Delta_r(\mathbf{x})$:

$$\|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_V \leq \Delta_r(\mathbf{x}) \quad (1.57)$$

où $\Delta_r(\mathbf{x})$ est donnée par exemple par l'expression :

$$\Delta_r(\mathbf{x}) = \frac{\|\mathbf{F}(\mathbf{x}) - \mathbf{K}(\mathbf{x})\mathbf{u}_r(\mathbf{x})\|_{V'}}{\gamma(\mathbf{x})^{1/2}} \quad (1.58)$$

où V' est l'espace dual de V et $\gamma(\mathbf{x}) > 0$ vérifie :

$$\gamma(\mathbf{x}) \leq \frac{\|\mathbf{K}(\mathbf{x})\mathbf{v}\|_{V'}}{\|\mathbf{v}\|_{V'}}, \quad \forall \mathbf{v} \in V \quad (1.59)$$

Dans de nombreux cas, on s'intéresse plutôt à une quantité d'intérêt $\mathbf{y}(\mathbf{x}) = \mathcal{M}(\mathbf{u}(\mathbf{x}), \mathbf{x}) \in Y$ dépendante de $\mathbf{u}(\mathbf{x})$ et pas à la solution $\mathbf{u}(\mathbf{x})$ elle-même. On souhaite alors contrôler l'erreur commise en approchant $\mathbf{y}(\mathbf{x})$ par $\mathbf{y}_r(\mathbf{x}) = \mathcal{M}(\mathbf{u}_r(\mathbf{x}), \mathbf{x})$. Les approches goal-oriented sont des méthodes populaires pour construire la base réduite [ZAHM et al., 2016] dans ce cadre. Celles-ci consistent alors à sélectionner les vecteurs de la base réduite par un algorithme glouton en contrôlant l'erreur $\|\mathbf{y}(\mathbf{x}) - \mathbf{y}_r(\mathbf{x})\|_Y$. Elles sont

toutefois limitées à $\mathcal{M}(\cdot)$ linéaire.

Ces approches *offline/online* sont coûteuses lors de la construction *offline* afin de garantir une erreur faible dans tout le domaine. L'étape *online* est elle au contraire très rapide car on applique la projection sur la base trouvée durant la phase *offline*.

1.1.3.4.3 Méthode de construction de base à la volée

Cette méthode a été développée pour l'utilisation de solution en base réduite pour la construction de surfaces de réponse [GOGU et PASSIEUX, 2013].

Elle consiste à construire adaptativement une base au fil de l'évaluation des points d'un DoE en utilisant un critère sur un estimateur d'erreur pour choisir si le point courant du DoE doit être évalué par le modèle numérique pour enrichir ou non la base réduite. En comparaison aux méthodes *offline/online*, la base réduite est à la fois exploitée et enrichie au cours de la phase *online*. En pratique, les points du DoE correspondent alors aux points d'un DoE généré pour construire une surface de réponse ou aux points d'apprentissage sélectionnés par une méthode d'apprentissage actif pour l'analyse de fiabilité ou pour une optimisation. Notons ainsi qu'en fonction de la position des points du DoE, la base réduite obtenue peut permettre de calculer des solutions réduites très précises dans certaines régions du domaine et très mauvaises dans d'autres régions. Contrairement, aux méthodes *offline/online*, l'erreur n'est pas contrôlée sur l'ensemble du domaine mais seulement dans les zones d'intérêt pour l'exploitation des solutions.

L'estimateur d'erreur proposé dans [GOGU et PASSIEUX, 2013] est basé sur le résidu de la solution en base réduite, sur lequel on reviendra dans la suite du manuscrit (voir Section 3.2.3).

La méthodologie débute par la génération d'un DoE sur lequel on veut construire une surface de réponse. Après une phase d'initialisation, l'évaluation successive des points du DoE est ensuite effectuée en utilisant la solution en base réduite si l'estimateur d'erreur basé sur le résidu respecte un certain seuil, sinon la solution du modèle EF est calculée. Dans ce dernier cas, la solution \mathbf{u}_i , correspondant à la i -ème solution EF calculée, est utilisée pour enrichir la base réduite Ψ après avoir été orthonormalisée par le procédé de Gram-Schmidt en suivant les équations Eq. (1.60) et Eq. (1.61) :

$$\Psi_i = \mathbf{u}_i - \sum_{k=1}^{i-1} \langle \mathbf{u}_i, \Psi_k \rangle \Psi_k \quad (1.60)$$

$$\Psi_i = \frac{\Psi_i}{\|\Psi_i\|_2} \quad (1.61)$$

avec $\langle \cdot, \cdot \rangle$ le produit scalaire dans L^2 .

1.1.3.5 Bilan

Parmi les différentes techniques de réduction de modèles linéaires vues dans cette section, la méthode de construction à la volée d'une base POD se démarque dans le cadre de ces travaux. En comparaison aux stratégies *offline/online* qui consistent à construire une fois pour toute une base réduite précise sur l'ensemble du domaine, cette approche ne comporte pas de phase *offline* coûteuse et permet d'enrichir la base réduite tout au long de l'exploitation des solutions afin d'améliorer les approximations dans les zones d'intérêt lorsque cela est nécessaire. Ainsi, le caractère adaptatif de la construction de la base est particulièrement adéquate pour les travaux effectués dans la thèse car dans le [chapitre 3](#) on s'intéressera notamment à l'utilisation de solutions réduites dans un algorithme d'apprentissage actif pour l'étude de la fiabilité. De plus, elle a l'avantage d'être simple à mettre en oeuvre et d'offrir des estimateurs d'erreur relativement peu coûteux à évaluer.

1.1.4 Conclusion

Dans cette section, plusieurs modèles pour la réduction du coût numérique ont été introduits.

D'une part la modélisation par GP a été présentée, qui permet de construire une fonction approchant les sorties scalaires d'un simulateur complexe afin d'en obtenir une prédiction à moindre coût. De plus, la structure de ces modèles leur donne l'avantage de fournir une prédiction probabiliste et donc l'incertitude sur l'approximation effectuée.

D'autre part, différentes méthodes de réduction pour les modèles linéaires paramétriques ont été présentées. Parmi ces méthodes, la méthode de construction à la volée d'une base POD se démarque dans le cadre de ces travaux de thèse. En effet, la méthode itérative de construction de la base réduite est particulièrement adaptée à l'évaluation successive d'une base d'apprentissage d'un métamodèle.

Ces deux approches peuvent être utilisées pour remplacer des codes numériques complexes précis et coûteux par des modèles donnant une approximation de leur(s) réponse(s) avec un coût numérique fortement réduit. Elles sont donc particulièrement utiles pour effectuer des analyses nécessitant de nombreux appels au modèle numérique.

Dans la suite, on s'intéresse à l'analyse de fiabilité. Dans un premier temps, les différentes méthodes numériques classiques d'analyses de fiabilité seront présentées puis les analyses de fiabilité basées sur l'approximation par un métamodèle de la fonction de performance relative au critère de défaillance seront introduites.

1.2 Analyse de fiabilité

1.2.1 Introduction

1.2.1.1 Prise en compte des incertitudes

Les problèmes d'ingénierie sont sujets à de nombreuses incertitudes, qui peuvent être classées selon leur origine :

- incertitudes sur la modélisation (mathématiques ou numérique) du système ou des paramètres ;
- incertitudes sur les valeurs des paramètres physiques (incertitudes des propriétés matériaux, variabilités sur les charges appliquées, variabilités géométriques).

Dans le cadre de cette thèse, on considère que la modélisation par **EF** des systèmes étudiés est suffisamment précise pour que l'incertitude associée puisse être négligée. En revanche on considère le système est affectée par des incertitudes sur les paramètres physiques.

Les approches probabilistes sont un moyen de prendre en compte les incertitudes affectant les paramètres physiques \boldsymbol{x} d'une structure modélisée numériquement par un modèle déterministe $\mathcal{M}(\boldsymbol{x})$ (cf Eq. (1.2)) dans le cadre d'une étude paramétrique.

La Figure 1.6 décrit le fonctionnement général d'une étude liée à la quantification des incertitudes en mécanique [SUDRET, 2007].

- *Étape 1* : La première étape consiste à construire un modèle numérique $\mathcal{M}(\cdot)$ permettant de simuler le comportement du système que l'on souhaite étudier et les sorties du modèle que l'on souhaite obtenir.
- *Étape 2* : La deuxième étape consiste à identifier et modéliser les différentes incertitudes sur les paramètres d'entrées du modèle. Dans le cadre de la modélisation probabiliste des incertitudes, les paramètres d'entrée sont des variables aléatoires \boldsymbol{X} suivant des lois de probabilité (gaussienne, Weibull, uniforme, etc.) qui seront à définir.
- *Étape 3* : La troisième étape consiste à propager les incertitudes portant sur les paramètres d'entrée \boldsymbol{X} en utilisant le modèle numérique. La sortie y du modèle est alors également une variable aléatoire $Y = \mathcal{M}(\boldsymbol{X})$, qu'il sera intéressant de caractériser afin d'effectuer les analyses permettant d'évaluer l'impact des incertitudes (analyse de fiabilité, analyse de sensibilité, etc.). Celle-ci peut être caractérisée en simulant le modèle numérique pour des réalisations des paramètres d'entrée aléatoires \boldsymbol{X} pour obtenir des réalisations de Y .

- *Étape 4* : La quatrième étape consiste à effectuer une analyse de sensibilité afin d'étudier la contribution des entrées à la variabilité des sorties en utilisant les résultats de la propagation d'incertitudes effectuée à l'*étape 3*.

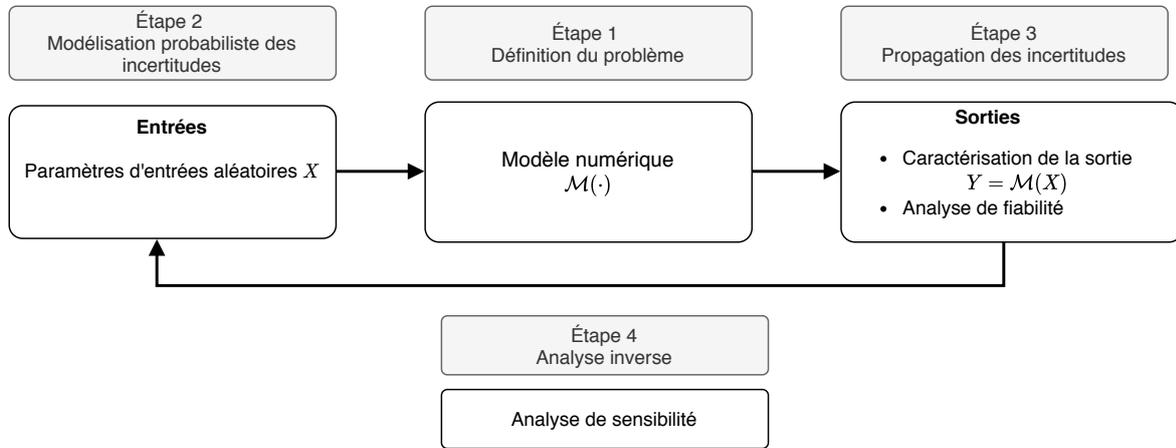


FIGURE 1.6 – Schéma des différentes étapes de la méthode de quantification des incertitudes.

1.2.1.2 Formulation du problème de fiabilité

Considérons un problème à n_u paramètres incertains $(x^1, \dots, x^{n_u}) \in \mathcal{X}$. Ces paramètres sont modélisés par un vecteur aléatoire absolument continu \mathbf{X} de variables aléatoires X^k , $k = 1, \dots, n_u$ caractérisé par une distribution jointe de densité de probabilité $f_{\mathbf{X}} : \mathcal{X} \subseteq \mathbb{R}^{n_u} \rightarrow \mathbb{R}^+$.

La défaillance d'une structure est caractérisée à l'aide d'une fonction de performance $G(\cdot)$, encore appelée fonction de défaillance, qui associe la performance d'un système à un jeu de paramètres \mathbf{x} . En fiabilité, la fonction de performance est généralement formulée pour vérifier que la valeur de la sortie du modèle $\mathcal{M}(\mathbf{x})$ ne dépasse pas le seuil $S \in \mathbb{R}$ à partir duquel le système est considéré défaillant :

$$G : \begin{cases} \mathcal{X} & \longrightarrow \mathbb{R} \\ \mathbf{x} & \longmapsto S - \mathcal{M}(\mathbf{x}) \end{cases} \quad (1.62)$$

Par convention, le domaine de défaillance \mathcal{D}_f correspond à une valeur négative de la fonction de performance, il est donc donné par :

$$\mathcal{D}_f = \{\mathbf{x} \in \mathcal{X}, G(\mathbf{x}) \leq 0\} \quad (1.63)$$

Le domaine de sûreté \mathcal{D}_s correspond à une valeur strictement positive de la fonction de performance. La frontière entre ces deux domaines ($G(\mathbf{x}) = 0$) est appelée état limite.

L'objectif d'une analyse de fiabilité est d'estimer la probabilité de défaillance d'un système en tenant compte des incertitudes sur ses paramètres.

La probabilité de défaillance P_f est définie par :

$$P_f = \mathbb{P} [G(\mathbf{X}) \leq 0] \quad (1.64)$$

$$= \int_{\mathcal{D}_f} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (1.65)$$

$$= \int_{\mathcal{X}} \mathbb{1}_{G(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (1.66)$$

$$= \mathbb{E}_{f_{\mathbf{X}}} [\mathbb{1}_{G(\mathbf{X}) \leq 0}] \quad (1.67)$$

En pratique, la probabilité de défaillance est évaluée à l'aide de méthodes numériques. Dans la suite, nous présenterons les méthodes basées sur une approximation de l'état limite par un hyperplan ou une surface quadratique, puis les méthodes classiques d'échantillonnage pour l'intégration de l'intégrale donnée par l'Eq. (1.66) et finalement nous introduirons les méthodes basées sur l'apprentissage de l'état limite par un métamodèle de GP, qui sont plus générales que les approximations du premier ou second ordre.

1.2.2 Méthodes FORM/SORM

Les méthodes de fiabilité du premier et du second ordre respectivement FORM, pour *First Order Reliability Method*, et SORM, pour *Second Order Reliability Method* [LEMAIRE et al., 2009; DITLEVSEN et MADSEN, 1996], sont des méthodes d'approximation de l'état limite. Elles reposent sur l'utilisation d'une approximation linéaire (FORM) ou quadratique (SORM) de la fonction d'état limite dans l'espace normal standard \mathcal{U} (l'espace normal standard désigne l'espace où toutes les variables aléatoires suivent des lois normales centrées réduites indépendantes) pour évaluer la probabilité de défaillance. Ces méthodes reposent néanmoins sur certaines hypothèses, qui en limitent les cas d'application, mais donnent des résultats satisfaisants quand ces hypothèses sont vérifiées.

Une transformation isoprobabiliste $T(\cdot)$ de Nataf [NATAF, 1962] ou de Rosenblatt [ROSENBLATT, 1952] est effectuée sur les variables d'entrées \mathbf{X} pour se ramener à des variables aléatoires gaussiennes centrées réduites indépendantes \mathbf{U} . La fonction de performance dans l'espace normal standard est alors notée $H(\mathbf{u})$ et vérifie :

$$\mathbb{P} [G(\mathbf{X}) \leq 0] = \mathbb{P} [H(\mathbf{U}) \leq 0] \quad \text{avec } \mathbf{U} = T(\mathbf{X}) \quad (1.68)$$

La probabilité de défaillance s'écrit alors :

$$P_f = \int_{\mathcal{D}_{f_u}} \phi_{n_u}(\mathbf{u}) d\mathbf{u} \quad (1.69)$$

où $\phi_{n_u}(\cdot)$ est la densité de probabilité jointe de la loi normale centrée réduite de dimension n_u et $\mathcal{D}_{f_u} = \{\mathbf{u} \in \mathcal{U}, H(\mathbf{u}) \leq 0\}$ est le domaine de défaillance dans l'espace standard.

Le point de défaillance le plus probable, ou encore appelé point de conception, est le point appartenant à l'état limite où la densité de probabilité des variables aléatoires est maximale. Ainsi le point de conception, noté P^* de coordonnées u^* , correspond au point du domaine de défaillance le plus proche de l'origine de l'espace normal standard et est obtenu en résolvant un problème d'optimisation.

L'indice de fiabilité d'Hasofer Lind [HASOFER, 1974] β correspond à la distance entre l'origine de l'espace standard et le point P^* :

$$\beta = \|u^*\| \quad \text{avec } u^* = \arg \min_{u \in \mathcal{D}_{f_u}} \|u\| \quad (1.70)$$

Si la fonction d'état limite est linéaire dans l'espace normal standard, alors on peut montrer qu'on a :

$$P_f = \Phi(-\beta) \quad (1.71)$$

où $\Phi(\cdot)$ est la fonction de répartition de la loi normale centrée réduite.

La méthode FORM consiste alors à se ramener dans l'espace standard pour pouvoir approcher la surface d'état limite $H(\mathbf{u}) = 0$ par un hyperplan tangent au point de conception P^* vérifiant :

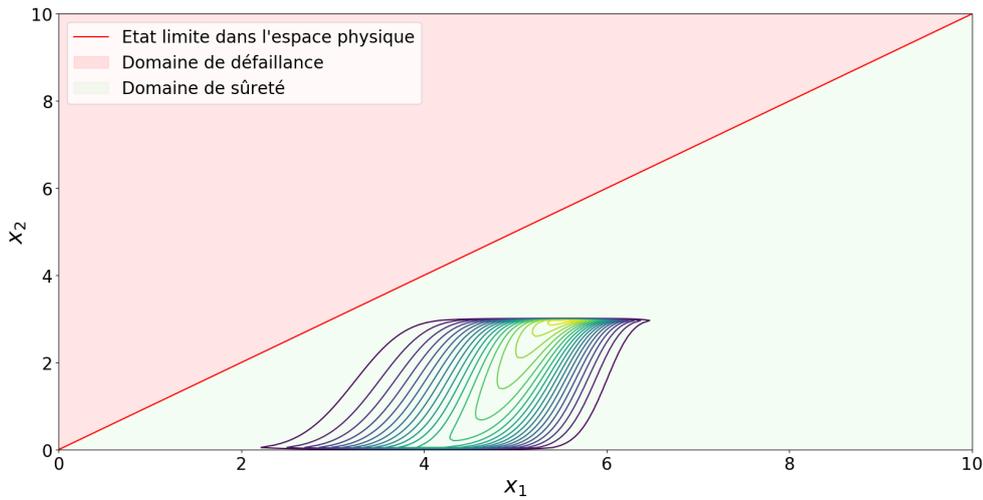
$$\hat{H}(\mathbf{u}) = -\boldsymbol{\alpha}\mathbf{u} + \beta = 0 \quad (1.72)$$

où les $\alpha_i = \frac{u_i^*}{\beta}$, $i = 1, \dots, n_u$ sont les cosinus directeurs du vecteur OP^* (où O désigne l'origine de l'espace standard). L'hyperplan est orthogonal au vecteur OP^* et l'approximation de la probabilité de défaillance \hat{P}_f^u obtenue en utilisant l'approximation par l'hyperplan de l'état limite est donnée par :

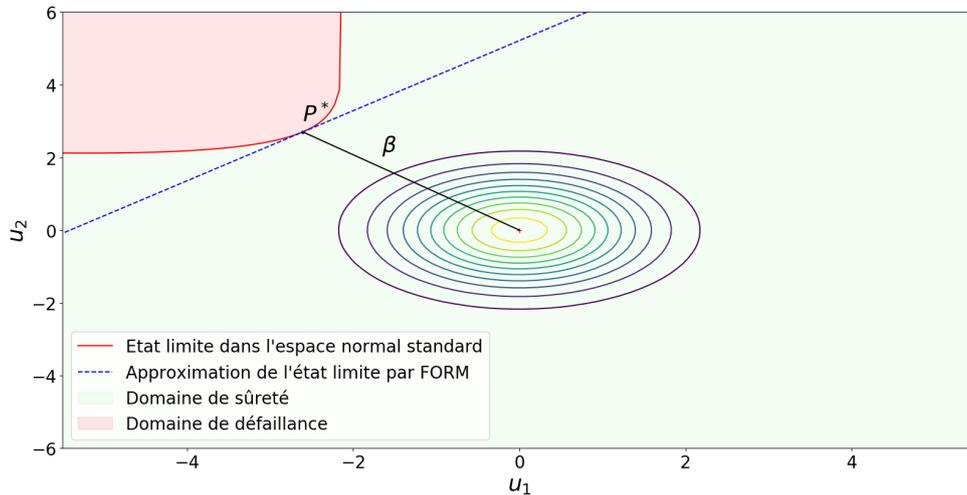
$$\begin{aligned} \hat{P}_f^u &= \mathbb{P} [\hat{H}(\mathbf{U}) \leq 0] \\ &= \mathbb{P} [\boldsymbol{\alpha}\mathbf{U} \leq -\beta] \\ &= \Phi(-\beta) \end{aligned} \quad (1.73)$$

La Figure 1.7 illustre la méthode FORM sur un exemple en deux dimensions. L'état limite et la distribution des entrées \mathbf{X} sont représentés dans l'espace physique \mathcal{X} sur la Figure 1.7a. Sur la Figure 1.7b l'état limite dans l'espace normal standard \mathcal{U} , son approximation par un hyperplan et la distribution des variables \mathbf{U} sont représentés.

Lorsque la fonction d'état limite n'est pas linéaire, une approximation de second ordre peut être utilisée. La méthode SORM [LEMAIRE et al., 2009; DITLEVSEN et MADSEN, 1996] est alors utilisée pour approcher l'état limite par une surface qua-



(a) Etat limite dans l'espace physique.



(b) Approximation linéaire de l'état limite dans l'espace normal standard par la méthode FORM.

FIGURE 1.7 – Exemple d'application de la méthode FORM.

dratique pour déterminer une approximation de la probabilité de défaillance. Cette approximation est valable asymptotiquement et vaut :

$$\hat{P}_f^u = \Phi(-\beta) \prod_{i=1}^{n_u-1} \frac{1}{\sqrt{1 + \beta \kappa_i}} \quad (1.74)$$

où les $(\kappa_i)_{i=1, \dots, n_u-1}$ sont les courbures principales de la surface d'état limite.

Avantages et inconvénients

Ces méthodes sont très populaires pour leur coût de calcul relativement raisonnable lié à la méthode utilisée pour résoudre l'optimisation de l'Eq. (1.70). Cependant, elles ont certaines limitations qui ne leur permettent pas de traiter correctement certains

problèmes de défaillance. En effet, l'approximation de la probabilité de défaillance par ces méthodes n'est exacte que si l'état limite est linéaire ou quadratique dans l'espace standard et l'erreur engendrée par une mauvaise approximation de l'état limite n'est ni connue ni contrôlable. Notons que la méthode FORM est plus largement applicable que la méthode FORM. Cependant, cette méthode nécessite le calcul des courbures principales, donc des dérivées premières et secondes de la fonction de performance, dont le calcul peut être très élevé lorsque la dimension stochastique augmente et que la fonction de performance est évaluée en utilisant un code de calculs [EF](#) de la structure. De plus, si les hypothèses sur lesquelles reposent l'approximation de l'état limite de défaillance par ces méthodes sont fausses, l'erreur commise n'est ni contrôlable ni vérifiable.

De plus ces approches ne sont adaptées que pour un seul mode de défaillance. En effet, ces méthodes ne permettent de déterminer qu'un unique point de conception et ne peuvent pas approcher correctement l'état limite lorsqu'il existe plusieurs zones de défaillance. Un exemple d'application de la méthode FORM sur un cas comportant deux zones de défaillance distinctes et de même importance est illustré sur la Figure 1.8. En effet, sans connaissance a priori du nombre de zones de défaillance, il s'agit d'un cas où l'application de la méthode FORM mène à une approximation erronée de l'état limite et donc de la probabilité de défaillance, d'autant plus que sur cet exemple les zones de défaillance sont de même importance. Afin de déterminer plusieurs points de conception, l'algorithme "multi-FORM" a été proposé par [[DER KIUREGHIAN et DAKESSIAN, 1998](#)]. La méthode proposée consiste à répéter l'algorithme FORM avec une fonction de performance modifiée qui permet d'orienter la recherche d'un prochain point de conception hors de la zone où se trouvent le(s) point(s) de conception trouvé(s).

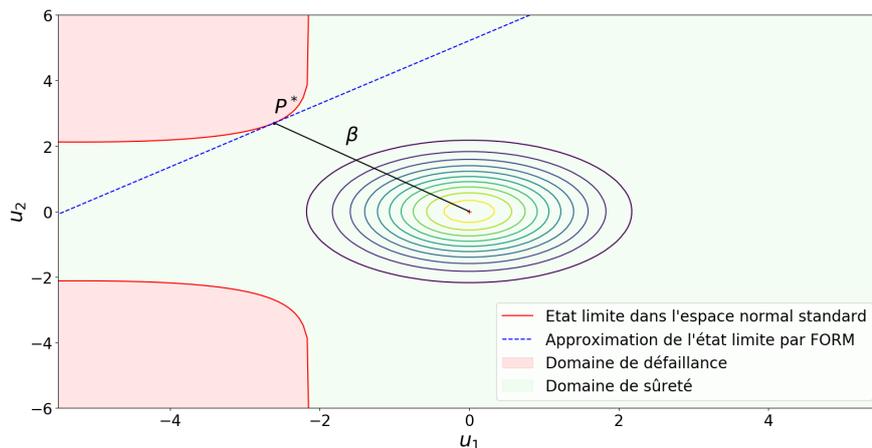


FIGURE 1.8 – Approximation linéaire de l'état limite par la méthode FORM dans le cas de deux modes de défaillance distincts.

Les méthodes par échantillonnage présentées par la suite, ne souffrant pas des

mêmes limitations, permettent d'aborder les calculs de probabilité de défaillance dans un cadre plus général.

1.2.3 Méthodes par échantillonnage

Dans cette partie, nous allons introduire différentes techniques d'intégration de l'intégrale donnée par l'Eq. (1.66) basées sur l'échantillonnage d'une population [MORIO et al., 2014]. Tout d'abord, la méthode de MC, adaptée à l'analyse de domaines de défaillance complexes, sera introduite. Le principe de l'échantillonnage préférentiel pour l'estimation d'évènements rares sera ensuite présenté ainsi que les différentes méthodes numériques permettant d'utiliser cette technique.

1.2.3.1 Méthode de Monte Carlo

Plusieurs méthodes d'échantillonnage peuvent s'appliquer pour approcher l'intégrale donnée par l'Eq. (1.66). Une des méthodes les plus simples et la plus générale est la méthode de Monte-Carlo (MC) puisqu'elle fonctionne en théorie pour tout type de problème de fiabilité. Les MCS permettent d'estimer la probabilité de défaillance en s'appuyant sur un tirage aléatoire pour intégrer numériquement l'intégrale de l'Eq. (1.66). En effet, la probabilité de défaillance s'exprime comme une espérance (voir Eq. (1.67)) :

$$P_f = \mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} \right] \quad (1.75)$$

Cette définition de la probabilité de défaillance montre que le problème revient à estimer une moyenne. L'estimation de la probabilité de défaillance est alors obtenue par la classification des échantillons du tirage aléatoire basée sur les valeurs prises par la fonction de performance en ces points.

Étant donnée un tirage indépendant et identiquement distribué (i.i.d) d'échantillons $\mathbf{X}_1, \dots, \mathbf{X}_{n_{MC}}$ de taille n_{MC} de distribution $f_{\mathbf{X}}$, la probabilité de défaillance est alors estimée par :

$$\hat{P}_f^{MC} = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{G(\mathbf{x}) \leq 0}(\mathbf{X}_i) \quad (1.76)$$

La loi forte des grands nombres précise que \hat{P}_f^{MC} converge presque sûrement vers $\mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} \right] = P_f$.

La variance de l'estimateur \hat{P}_f^{MC} est donnée par :

$$Var \left(\hat{P}_f^{MC} \right) = \frac{Var \left(\mathbb{1}_{G(\mathbf{X}_1) \leq 0}(\mathbf{X}_1) \right)}{n_{MC}} = \frac{P_f(1 - P_f)}{n_{MC}} \quad (1.77)$$

En pratique, le Coefficient de variation (COV) est généralement utilisé pour quan-

tifier l'erreur de cet estimateur. Son estimation est obtenue en utilisant l'estimation empirique de la variance et est donnée par :

$$\widehat{COV}_{\hat{P}_f^{MC}} = \sqrt{\frac{(1 - \hat{P}_f^{MC})}{n_{MC} \hat{P}_f^{MC}}} \quad (1.78)$$

L'exemple classique du système en séries à 4 branches, qui correspond à un cas en deux dimensions à zones de défaillances multiples, sera utilisé comme exemple fil rouge dans la suite de la Section 1.2 pour illustrer les différentes méthodes sur le cas de défaillances multimodales. Pour cet exemple, on considère deux variables aléatoires X_1 et X_2 suivant des lois normales standards et une fonction de performance donnée par :

$$y = \min_{x_1, x_2} \left\{ \begin{array}{l} 3 + 0.1(x_1 + x_2)^2 - \frac{(x_1 + x_2)}{\sqrt{(2)}}; \\ 3 + 0.1(x_1 + x_2)^2 + \frac{(x_1 + x_2)}{\sqrt{(2)}}; \\ (x_1 - x_2) + \frac{6}{\sqrt{2}}; \\ (x_2 - x_1) + \frac{6}{\sqrt{2}} \end{array} \right\} \quad (1.79)$$

La méthode de MC est illustrée sur cet exemple sur la Figure 1.9.

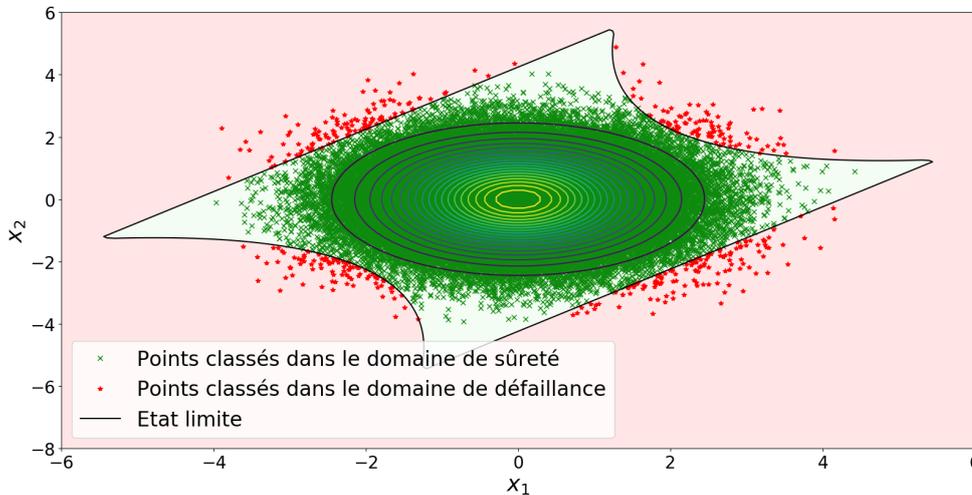


FIGURE 1.9 – Classification d'une population de Monte-Carlo.

Avantages et inconvénients La méthode de Monte-Carlo étant la plus générale possible, elle est adaptée en théorie pour tout type de problème de fiabilité, c.-à-d. pour des domaines de défaillance très complexes. En pratique, elle a cependant certaines limitations quand l'évaluation des échantillons représente un coût de calcul important. En effet, l'inconvénient principal de la méthode de Monte-Carlo est le nombre important

de simulations nécessaires pour obtenir un coefficient de variation de MC raisonnable ($< 10\%$), et ce d'autant plus que la probabilité recherchée est faible comme le montre l'Eq. (1.78).

En effet, pour obtenir une probabilité de défaillance de 10^{-m} avec un coefficient de variation de 10% , on déduit de l'Eq. (1.78) que 10^{m+2} simulations sont nécessaires. Pour estimer des probabilités de défaillance faibles ($< 10^{-3}$), le coût de calcul de cette méthode devient donc prohibitif lorsque le modèle numérique est coûteux en temps de calcul (par exemple pour un code EF). Sur la Figure 1.10, ce problème est illustré sur un exemple où la probabilité de défaillance est estimée à 0 par MC.

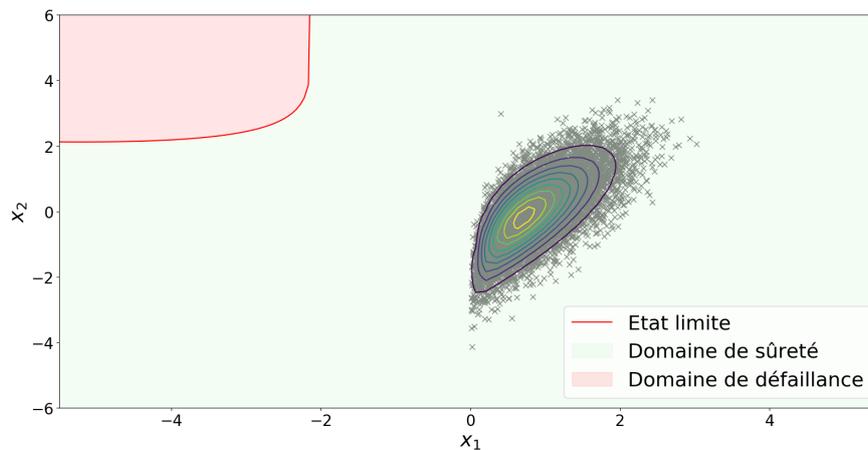


FIGURE 1.10 – Méthode de Monte-Carlo pour une probabilité faible. Illustration de l'inefficacité de la méthode puisque aucun points n'est trouvé dans le domaine de défaillance.

Afin de diminuer ces temps de calcul des méthodes telles que les méthodes d'Échantillonnage préférentiel, ou *Importance Sampling* [MELCHERS, 1989] ou les Subset Simulations (SS) [AU et BECK, 2001a] ont été proposées.

1.2.3.2 Échantillonnage préférentiel

1.2.3.2.1 Principe de l'échantillonnage préférentiel

Dans la section précédente, la méthode classique de MC a été présentée et nous avons vu que malgré sa capacité théorique à traiter tout type de problème son efficacité est fortement réduite lorsqu'il s'agit de l'estimation de probabilité d'évènements rares. La raison pour laquelle les MCS ne sont pas efficaces pour estimer des probabilités faibles est que la majorité des points d'un échantillonnage standard sont localisés dans la région où la densité de probabilité est la plus élevée. Cette zone correspond au voisinage du point moyen de la distribution des entrées, alors que les échantillons menant à la défaillance se trouvent typiquement dans la queue de la distribution. Les méthodes Échantillonnage préférentiel, ou *Importance Sampling* (IS) utilise une densité auxiliaire

f_{aux} qui permet de concentrer l'échantillonnage dans la zone d'intérêt. Le principe des méthodes **IS** est d'utiliser une densité de probabilité auxiliaire f_{aux} permettant d'échantillonner dans les zones d'importance pour l'estimation de la probabilité de défaillance.

En effet, la description naturelle de P_f comme une espérance par rapport à $\mathbf{X} \sim f_{\mathbf{X}}$, donnée par l'Eq. (1.67), peut se réécrire comme une espérance par rapport à la densité f_{aux} :

$$\begin{aligned}
 P_f &= \mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} \right] \\
 &= \int_{\mathcal{X}} \mathbb{1}_{G(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\
 &= \int_{\mathcal{X}} \mathbb{1}_{G(\mathbf{x}) \leq 0} \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{aux}(\mathbf{x})} f_{aux}(\mathbf{x}) d\mathbf{x} \\
 &= \mathbb{E}_{f_{aux}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} \frac{f_{\mathbf{X}}(\mathbf{X})}{f_{aux}(\mathbf{X})} \right]
 \end{aligned} \tag{1.80}$$

La probabilité de défaillance peut donc être estimée en se basant sur des échantillons générés selon la nouvelle densité f_{aux} qu'il va falloir déterminer. Or, cette densité étant biaisée, on ne peut pas appliquer directement l'estimation sur les simulations obtenues pour un échantillonnage selon cette densité. En effet, pour estimer P_f les simulations sont pondérées comme le montre l'expression de l'Eq. (1.80).

Ainsi considérant un tirage indépendant et identiquement distribué (i.i.d) d'échantillons $\mathbf{X}_1, \dots, \mathbf{X}_{n_{IS}}$ de taille n_{IS} de distribution f_{aux} , l'estimation de la probabilité de défaillance par **IS** est donnée par :

$$\hat{P}_f^{IS} = \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w(\mathbf{X}_i) \mathbb{1}_{G(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) \tag{1.81}$$

avec $w(\mathbf{X}_i) = \frac{f_{\mathbf{X}}(\mathbf{X}_i)}{f_{aux}(\mathbf{X}_i)}$ les poids des échantillons \mathbf{X}_i .

La variance de cet estimateur est alors donnée par :

$$\begin{aligned}
 Var(\hat{P}_f^{IS}) &= \frac{Var(w(\mathbf{X}_1) \mathbb{1}_{G(\mathbf{X}_1) \leq 0}(\mathbf{X}_1))}{n_{IS}} \\
 &= \frac{1}{n_{IS}} \left(\mathbb{E}_{f_{aux}} \left[\left(w(\mathbf{X}_1) \mathbb{1}_{G(\mathbf{X}_1) \leq 0}(\mathbf{X}_1) \right)^2 \right] - P_f^2 \right)
 \end{aligned} \tag{1.82}$$

Ainsi, si elle est bien choisie la densité auxiliaire f_{aux} peut permettre de réduire considérablement la variance de l'estimateur et améliorer sa convergence. La meilleure densité auxiliaire f_{aux} possible et notée f_{opt} correspond naturellement à la densité permettant d'obtenir une variance nulle de l'estimateur. D'après l'Eq. (1.82), son expression est donnée par :

$$f_{opt}(\mathbf{x}) = \frac{\mathbb{1}_{G(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x})}{P_f} \quad (1.83)$$

En reprenant l'exemple de la Figure 1.10 illustrant les limites de la méthode MCS pour les probabilités faibles, la densité optimale de l'IS est représentée sur la Figure 1.11. L'échantillonnage selon cette densité permet d'avoir davantage de points défaillants pour une taille de population plus faible.

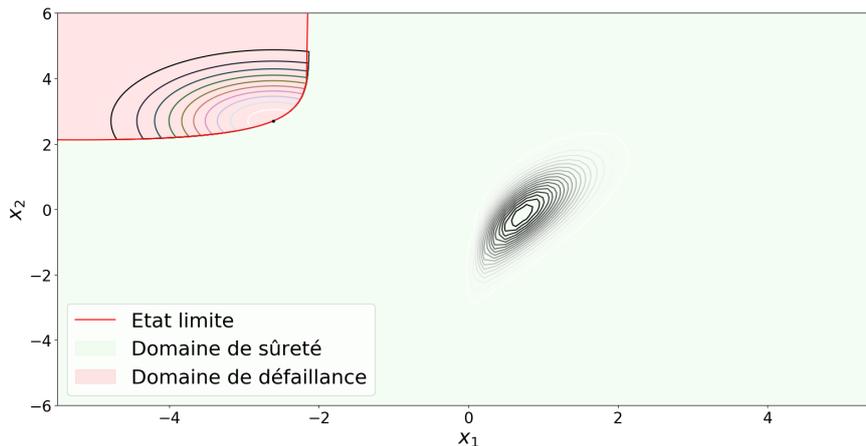


FIGURE 1.11 – Illustration des lignes de contour de la densité des entrées (lignes grises) et des lignes de contour de la densité optimale de l'IS (lignes en couleur).

En pratique l'expression théorique de la densité optimale f_{opt} donnée par l'Eq. (1.83) ne peut être utilisée directement puisqu'elle dépend de la quantité P_f que l'on cherche à déterminer. Une difficulté des méthodes d'IS est alors de déterminer une densité f_{aux} correcte, aussi proche que possible de la densité optimale f_{opt} . De nombreuses méthodes ont été développées pour estimer la densité optimale [SCHUËLLER et STIX, 1987; MELCHERS, 1989; ZHANG, 1996; RUBINSTEIN, 1999; DE BOER et al., 2005; MORIO, 2012; PAPAIOANNOU et al., 2016]. Par exemple, le point de défaillance le plus probable peut être utilisé comme centre de la densité optimale [MELCHERS, 1989; SCHUËLLER et STIX, 1987]. Or dans le cas où il existe plusieurs domaines disjoints dans l'espace physique ou standard, plusieurs points de conception doivent être déterminés ce qui peut être en pratique difficilement réalisable. De nombreuses méthodes adaptatives d'apprentissage de la densité optimale de d'IS ont été proposées. Ces méthodes ont pour objectif d'approcher itérativement la densité optimale en estimant la densité de manière paramétrique [DE BOER et al., 2005; RUBINSTEIN, 1999] ou de manière non paramétrique [MORIO, 2012; ZHANG, 1996; PAPAIOANNOU et al., 2016] à l'aide d'échantillons bien choisis.

Un exemple d'échantillonnage utilisant l'approximation de la densité optimale obtenue par une méthode d'IS adaptatif est illustrée sur la Figure 1.12, où la population

est concentrée autour de l'état limite.

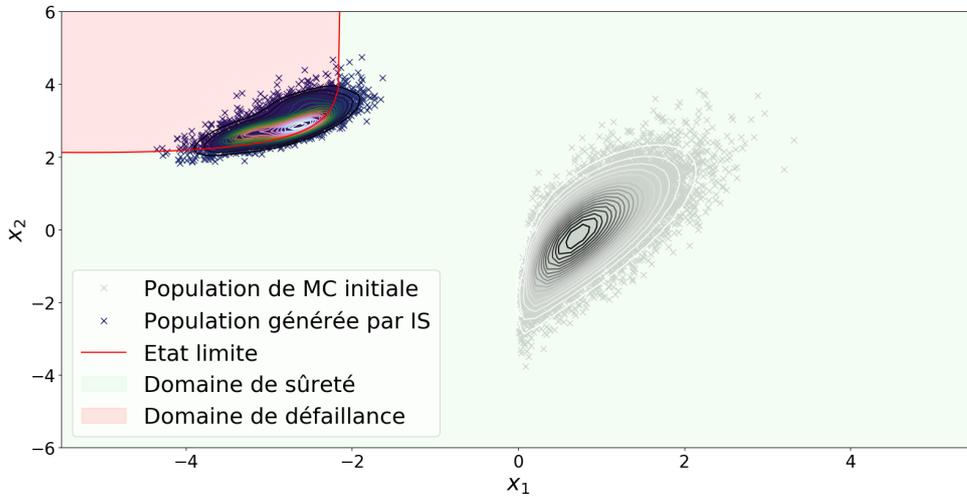


FIGURE 1.12 – Illustration de l'échantillonnage préférentiel pour une probabilité faible.

1.2.3.2.2 Échantillonnage préférentiel adaptatif

Dans cette section, deux procédures classiques d'apprentissage itératif de la densité optimale d'IS sont présentées.

- **Méthode d'entropie croisée**

Une manière d'estimer la densité auxiliaire est de considérer qu'elle appartient à une famille de distributions paramétriques, dont il faut alors déterminer les paramètres pour approcher au mieux la densité optimale [RUBINSTEIN, 1999, 1997]. Supposons alors que la densité auxiliaire f_{aux} appartient à la famille paramétrique suivante :

$$f_{aux} \in \{f(\cdot, \mathbf{p}), \mathbf{p} \in \mathbb{R}^{n_p}\} \quad (1.84)$$

où $\mathbf{p} = (p_1, \dots, p_{n_p})$ est un vecteur de n_p paramètres réels. L'objectif est alors de déterminer les paramètres \mathbf{p} pour que f_{aux} soit la meilleure approximation possible de la densité optimale f_{opt} .

La méthode d'Entropie Croisée, ou *Cross-Entropy* (CE) [RUBINSTEIN et KROESE, 2004; DE BOER et al., 2005] permet d'estimer la densité auxiliaire f_{aux} en se basant sur la divergence de Kullback-Leibler, ou encore entropie relative, [KULLBACK et LEIBLER, 1951], qui permet de mesurer la dissimilarité entre deux densités.

La divergence de Kullback-Leibler entre f_{aux} et f_{opt} est donnée par :

$$D_{KL}(f_{aux}(\cdot, \mathbf{p}), f_{opt}) = \int_{\mathcal{X}} f_{opt}(\mathbf{x}) \ln \left(\frac{f_{opt}(\mathbf{x})}{f_{aux}(\mathbf{x}, \mathbf{p})} \right) d\mathbf{x} \quad (1.85)$$

$$= \int_{\mathcal{X}} f_{opt}(\mathbf{x}) \ln(f_{opt}(\mathbf{x})) d\mathbf{x} - \int_{\mathcal{X}} f_{opt}(\mathbf{x}) \ln(f_{aux}(\mathbf{x}, \mathbf{p})) d\mathbf{x} \quad (1.86)$$

Les paramètres optimaux \mathbf{p}^* de la densité auxiliaire correspondent à la solution du problème d'optimisation :

$$\mathbf{p}^* = \arg \min_{\mathbf{p}} D_{KL}(f_{aux}(\cdot, \mathbf{p}), f_{opt}) \quad (1.87)$$

$$= \arg \max_{\mathbf{p}} \int_{\mathcal{X}} f_{opt}(\mathbf{x}) \ln(f_{aux}(\mathbf{x}, \mathbf{p})) d\mathbf{x} \quad (1.88)$$

En pratique le problème d'optimisation formulé par l'Eq. (1.88) ne peut pas être résolu tel quel car il dépend de la densité auxiliaire optimale f_{opt} que l'on cherche à déterminer. Les paramètres optimaux \mathbf{p}^* peuvent être évalués en utilisant le problème équivalent [RUBINSTEIN et KROESE, 2004] suivant :

$$\mathbf{p}^* = \arg \max_{\mathbf{p}} \int_{\mathcal{X}} \mathbb{1}_{G(\mathbf{x}) \leq 0}(\mathbf{x}) \ln(f_{aux}(\mathbf{x}, \mathbf{p})) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (1.89)$$

$$= \arg \max_{\mathbf{p}} \mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} \ln(f_{aux}(\mathbf{X}, \mathbf{p})) \right] \quad (1.90)$$

En considérant que l'on veut échantillonner selon une densité mieux adaptée f_{aux} plutôt qu'en utilisant la densité initiale $f_{\mathbf{X}}$, on a :

$$\mathbf{p}^* = \arg \max_{\mathbf{p}} \mathbb{E}_{f_{aux}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} \frac{f_{\mathbf{X}}(\mathbf{X})}{f_{aux}(\mathbf{X}, \mathbf{p})} \ln(f_{aux}(\mathbf{X}, \mathbf{p})) \right] \quad (1.91)$$

Ce problème peut alors être résolu de manière adaptative à partir de seuils intermédiaires $\gamma_0 > \gamma_1 > \dots > \gamma_k > \dots \geq 0$, où le seuil γ_k correspond au ρ -quantile de l'échantillon $(G(\mathbf{X}_i^k))_{i=1, \dots, n_{IS}}$ obtenu en évaluant la fonction de performance sur n_{IS} échantillons distribués selon $f_{aux}(\mathbf{X}, \mathbf{p}^{(k-1)})$, densité auxiliaire estimée à l'itération $(k-1)$.

Les paramètres $\mathbf{p}^{(k)}$ de la densité auxiliaire f_{aux} à une itération (k) sont alors estimés en maximisant une estimation par MCS de l'Eq. (1.91) :

$$\mathbf{p}^k = \arg \max_{\mathbf{p}} \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} \mathbb{1}_{G(\mathbf{X}_i) \leq \gamma_k} \frac{f_{\mathbf{X}}(\mathbf{X}_i)}{f_{aux}(\mathbf{X}_i, \mathbf{p}^{(k-1)})} \ln(f_{aux}(\mathbf{X}_i, \mathbf{p})) \quad (1.92)$$

Le critère d'arrêt de l'algorithme correspond à une valeur de seuil intermédiaire $\gamma_k \leq 0$.

Finalement, la probabilité est estimée en utilisant la formule donnée par l'Eq. (1.81)

avec la densité auxiliaire optimale estimée $\widehat{f_{aux}}(\cdot, \mathbf{p}^{(k^*)})$ estimée à la dernière étape (k^*) de l'algorithme d'IS adaptatif, ou *Adaptive Importance Sampling*, basée sur la CE et noté AIS-CE :

$$\hat{P}_f^{AIS-CE} = \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} \mathbb{1}_{G(\mathbf{X}_i^{k^*}) \leq 0} \frac{f_{\mathbf{X}}(\mathbf{X}_i^{k^*})}{\widehat{f_{aux}}(\mathbf{X}_i^{k^*}, \mathbf{p}^{(k^*)})} \quad (1.93)$$

Avantages et inconvénients

L'inconvénient principal de la méthode CE est qu'elle repose sur l'hypothèse que la densité auxiliaire optimale appartient à une famille paramétrique. Cette limitation ne permet pas de traiter correctement des problèmes complexes avec des domaines de défaillance multimodaux. Afin de résoudre ce problème, l'utilisation des modèles de mélange gaussien avec l'AIS-CE [KURTZ et SONG, 2013] a été proposée. Cependant, pour des problèmes de grande dimension ces méthodes sont très peu efficaces. Ainsi, certains travaux récents proposent d'utiliser d'autres modèles de mélange afin de résoudre ce problème [WANG et SONG, 2016].

L'avantage de la méthode de CE par rapport aux méthodes non-paramétriques utilisant des noyaux, décrites dans la section suivante, est qu'elle n'est pas soumise aux difficultés de paramétrage des noyaux tels que le réglage de la largeur de bande.

• *Non parametric Adaptive Importance Sampling*

L'objectif de l'échantillonnage préférentiel non paramétrique est d'approcher la densité auxiliaire optimale f_{opt} de l'IS sans connaissance a priori du modèle, c.-à-d. sans hypothèse sur la nature de la distribution. Une estimation non paramétrique de la densité, plutôt que l'estimation des paramètres d'une loi paramétrique choisie a priori, est davantage en mesure d'apprendre une loi complexe dans le cas de domaine de défaillance non régulier. L'utilisation pratique de l'estimation par noyau requiert d'une part le choix d'un noyau k et d'autre part un paramètre de lissage h à estimer.

La méthode NAIS [MORIO, 2012] repose sur l'estimation par noyaux pondérés de la densité de probabilité optimale f_{opt} [ZHANG, 1996], donnée par l'Eq. (1.83) à partir de la densité de probabilité initiale $f_{\mathbf{X}}$. Cette méthode consiste à apprendre la densité auxiliaire optimale de l'IS de façon adaptative comme pour la méthode CE mais en estimant chaque densité intermédiaire par un estimateur à noyau à la place d'une loi paramétrique.

L'algorithme NAIS permet ainsi d'adapter itérativement la densité auxiliaire estimée par noyau, en considérant des seuils intermédiaires γ_k . La méthode introduit un paramètre $\rho \in]0, 1]$ permettant de définir chaque seuil intermédiaire γ_k comme le ρ^e -quantile de l'ensemble des réponses de la fonction de performance $(G(\mathbf{X}_i^k))_{i=1, \dots, n_{IS}}$ évaluées pour n_{IS} échantillons générés à partir de la dernière densité auxiliaire estimée par noyau.

A l'itération k de l'algorithme, la densité auxiliaire mise à jour est donnée par :

$$\widehat{f_{opt}}^{(k+1)}(\mathbf{x}) = \frac{1}{kn_{IS}\det(B)I_k} \sum_{j=1}^k \sum_{i=1}^{n_{IS}} w_i^j K_{n_u} \left(B^{-1} (\mathbf{x} - \mathbf{X}_i^j) \right) \quad (1.94)$$

où les poids w_i^j sont donnés par :

$$w_i^j = \frac{\mathbb{1}_{G(\mathbf{X}_i^j) \leq \gamma_k} f_{\mathbf{X}}(\mathbf{X}_i^j)}{\widehat{f_{aux}}^{opt,j}(\mathbf{X}_i^j)} \quad (1.95)$$

I_k vaut :

$$I_k = \frac{1}{kn_{IS}} \sum_{j=1}^k \sum_{i=1}^{n_{IS}} w_i^j \quad (1.96)$$

K_{n_u} est un noyau gaussien de dimension n_u et B est une matrice de covariance diagonale.

A la dernière étape k^* de l'algorithme **NAIS**, la probabilité est estimée avec la dernière densité auxiliaire optimale estimée $\widehat{f_{opt}}^{(k^*)}$:

$$\hat{P}_f^{NAIS} = \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} \frac{\mathbb{1}_{G(\mathbf{X}_i^{k^*}) \leq 0} f_{\mathbf{X}}(\mathbf{X}_i^{k^*})}{\widehat{f_{opt}}^{(k^*)}(\mathbf{X}_i^{k^*})} \quad (1.97)$$

L'algorithme de la méthode **NAIS** est donné par l'Algorithme 2.

Finalement, la probabilité est estimée en utilisant la formule de l'Eq. (1.81) avec la densité auxiliaire optimale estimée $\widehat{f_{opt}}^{(k^*)}$ estimée à la dernière étape k^* de l'algorithme **NAIS** :

$$\hat{P}_f^{NAIS} = \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} \frac{\mathbb{1}_{G(\mathbf{X}_i^{k^*}) \leq 0} f_{\mathbf{X}}(\mathbf{X}_i^{k^*})}{\widehat{f_{opt}}^{(k^*)}(\mathbf{X}_i^{k^*})} \quad (1.98)$$

où les $\mathbf{X}_i^{k^*}$ sont tirés i.i.d. selon $\widehat{f_{opt}}^{(k^*)}$.

Avantages et inconvénients de la méthode

L'avantage majeur de la méthode **NAIS** est qu'elle est bien adaptée pour des domaines de défaillance \mathcal{D}_f complexes (c.-à-d. dans le cas de zones de défaillance multiples). L'efficacité de la méthode **NAIS** sur ce type de problème est illustrée sur la Figure 1.13 où les 4 zones de défaillances non connexes ont été détectées. De plus, l'erreur sur l'estimation peut être contrôlée à l'aide de l'estimation de la variance en se basant sur l'Eq. (1.82).

Cependant, l'algorithme **NAIS** n'est efficace que pour des entrées dont la dimension est relativement faible ($n_u < 10$). En effet, la méthode étant basée sur l'estimation par noyau, son efficacité diminue lorsque la dimension augmente. De plus, l'estimation du

Algorithm 2 Non parametric Adaptive Importance Sampling (NAIS)

Require: $\rho, n_{IS}, f_{\mathbf{X}}, G$

- 1: $k \leftarrow 0$
 - 2: $\tilde{\mathbf{X}}^0 \leftarrow$ population de taille n_{IS} générée suivant $f_{\mathbf{X}}$
 - 3: $\mathbf{Y}^0 \leftarrow G(\tilde{\mathbf{X}}^0)$
 - 4: $S^* \leftarrow \rho^e$ quantile de \mathbf{Y}^0
 - 5: $\gamma_0 \leftarrow \max(S^*, 0)$
 - 6: $\mathbf{w}^0 \leftarrow \mathbb{1}_{\mathbf{Y}^0 \leq \gamma_0}$
 - 7: $I_0 \leftarrow \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i^0$
 - 8: $\widehat{f_{aux}^{opt}}^{(1)}(\mathbf{x}) \leftarrow \frac{1}{n_{IS} \det(B) I_0} \sum_{i=1}^{n_{IS}} w_i^0 K_{n_u}(B^{-1}(\mathbf{x} - \mathbf{X}_i^0))$
 - 9: **while** $\gamma_k > 0$ **do**
 - 10: $k \leftarrow k + 1$
 - 11: $\tilde{\mathbf{X}}^k \leftarrow$ population de taille n_{IS} générée suivant $\widehat{f_{aux}^{opt}}^{(k)}$
 - 12: $\mathbf{Y}^k \leftarrow G(\tilde{\mathbf{X}}^k)$
 - 13: $S^* \leftarrow \rho^e$ quantile de \mathbf{Y}^k
 - 14: $\gamma_k \leftarrow \max(S^*, 0)$
 - 15: $w_i^j \leftarrow \frac{\mathbb{1}_{\mathbf{Y}_i^j \leq \gamma_k} f_{\mathbf{X}}(\mathbf{X}_i^j)}{\widehat{f_{aux}^{opt}}^{(j)}(\mathbf{X}_i^j)}$, for $i, j \in [1, n_{IS}] \times [1, k]$
 - 16: $I_k \leftarrow \frac{1}{kn_{IS}} \sum_{j=1}^k \sum_{i=1}^{n_{IS}} w_i^j$
 - 17: $\widehat{f_{aux}^{opt}}^{(k+1)}(\mathbf{x}) \leftarrow \frac{1}{kn_{IS} \det(B) I_k} \sum_{j=1}^k \sum_{i=1}^{n_{IS}} w_i^j K_{n_u}(B^{-1}(\mathbf{x} - \mathbf{X}_i^j))$
 - 18: **end while**
 - 19: $\hat{P}_f^{NAIS} \leftarrow \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} \frac{\mathbb{1}_{\mathbf{Y}^k \leq 0} f_{\mathbf{X}}(\mathbf{X}_i^k)}{\widehat{f_{aux}^{opt}}^{(k)}(\mathbf{X}_i^k)}$
-

paramètre de lissage optimal pour l'estimation par noyau peut être complexe et peut donc affecter la qualité de l'estimation par NAIS.

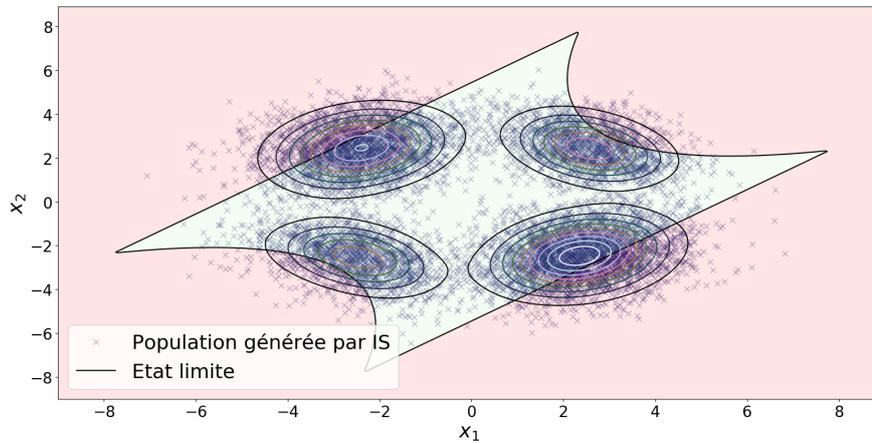


FIGURE 1.13 – Estimation par noyaux de la densité optimale en utilisant la méthode NAIS pour un état limite complexe.

1.2.3.3 Bilan

Dans cette section, différentes méthodes d'échantillonnage ont été introduites pour résoudre des problèmes de fiabilité dans le cas de domaines de défaillance complexes (c.-à-d. non réguliers ou multimodaux) et pour l'estimation de probabilités faibles. D'autres méthodes basées sur l'échantillonnage telles que l'échantillonnage directionnel [DITLEVSEN et al., 1988] ou les Subset Simulation [AU et BECK, 2001b] existent mais ne sont pas considérées dans le cadre de cette thèse par souci de concision.

1.2.4 Analyse de fiabilité basée sur l'utilisation de métamodèles

Afin de réduire le coût de l'analyse de fiabilité, un métamodèle (cf Section 1.1.1) $\hat{G}(\mathbf{x})$ de la fonction de performance peut être construit pour obtenir une approximation $\hat{\mathcal{D}}_f$ du domaine de défaillance \mathcal{D}_f . Dans cette optique, le choix du DoE sur lequel se base la construction du métamodèle est primordial car il doit permettre d'obtenir une approximation suffisamment précise de l'état limite mais contenir un nombre minimal de points afin que le coût numérique de la construction du métamodèle ne soit pas trop élevé.

Les méthodes d'analyse de fiabilité par apprentissage actif, ou d'échantillonnage adaptatif, permettent d'enrichir le métamodèle en apprenant de façon précise l'état limite là où il affecte le plus l'estimation de la probabilité de défaillance. La Figure 1.14 illustre le principe de ces méthodes sur l'exemple du cas 4 branches en remarquant que des points ont été ajoutés dans les quatre zones d'intérêts, c.-à-d. la grande majorité des points se trouvent au voisinage de l'état limite.

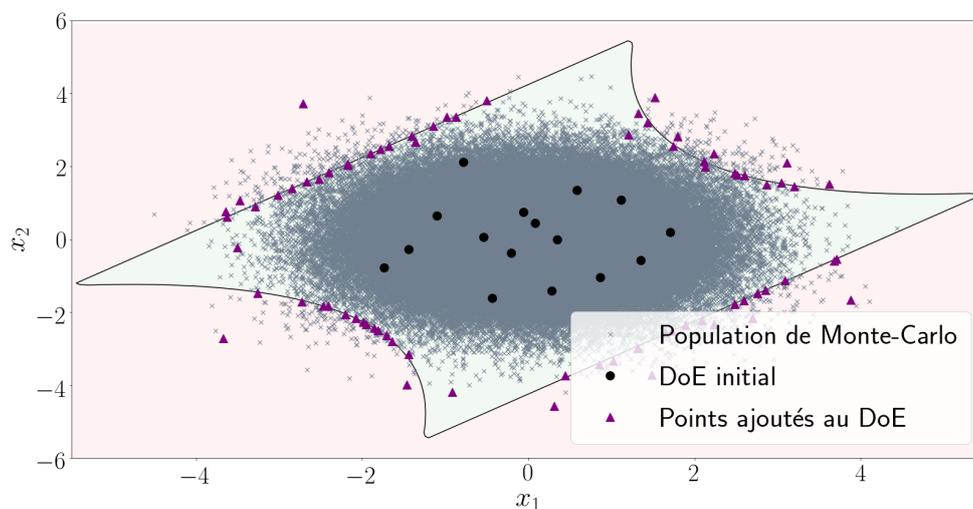


FIGURE 1.14 – Exemple d'apprentissage actif pour le cas 4 branches

Ces méthodes reposent principalement sur le choix de 4 éléments :

- le type de métamodèle.
- la méthode d'échantillonnage.
- le critère d'enrichissement utilisé pour enrichir le DoE afin d'obtenir une approximation précise de l'état limite.
- le critère d'arrêt de l'algorithme, qui est imposé pour déterminer quand l'apprentissage du métamodèle est suffisant pour obtenir une classification précise des échantillons.

Il existe de nombreuses méthodes d'analyse de fiabilité par apprentissage actif, notamment pour les SVM [BASUDHAR et MISSOUM, 2013; BOURINET et al., 2011], les GPs [VAZQUEZ et BECT, 2009; ECHARD et al., 2013; BICHON et al., 2008, 2011; FAURIAT et GAYTON, 2014; LI et al., 2012; BECT et al., 2012; DUBOURG et al., 2013] et le Krigeage combiné au chaos polynomial [SCHÖBI R. et al., 2017].

Considérons un métamodèle $\hat{G}(\mathbf{x})$ de la fonction de performance. Dans ce cas, l'approximation du domaine de défaillance obtenue en utilisant le métamodèle est donnée par :

$$\hat{\mathcal{D}}_f = \{\mathbf{x} \in \mathcal{X}, \hat{G}(\mathbf{x}) \leq 0\} \quad (1.99)$$

Finalement, une approximation de la probabilité de défaillance peut être estimée à moindre coût en remplaçant les appels à la fonction de performance coûteuse $G(\mathbf{x})$ par des évaluations du métamodèle $\hat{G}(\mathbf{x})$:

$$P_f \approx \mathbb{P} [\hat{G}(\mathbf{X}) \leq 0] \quad (1.100)$$

$$= \mathbb{E}_{f_{\mathbf{X}}} [\mathbb{1}_{\hat{G}(\mathbf{X}) \leq 0}] \quad (1.101)$$

$$= \int_{\mathcal{X}} \mathbb{1}_{\hat{G}(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (1.102)$$

$$= \int_{\hat{\mathcal{D}}_f} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (1.103)$$

Dans cette thèse, le type de métamodèle utilisé est un GP conditionné construit sur n observations noté \mathcal{G}_n . Dans la prochaine section, les différentes manières possibles d'estimer l'intégrale donnée par l'Eq. (1.103) pour un métamodèle \mathcal{G}_n seront présentées.

1.2.4.1 Estimation de la probabilité de défaillance

L'estimateur de la probabilité de défaillance calculée sur une population de MC $\tilde{\mathbf{X}} = (\mathbf{X}_i)_{i=1, \dots, n_{MC}}$ avec \mathbf{X}_i i.i.d de même distribution que \mathbf{X} est alors donné par :

$$\hat{P}_f = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mathcal{G}_n(\mathbf{x}_i) \leq 0}(\mathbf{X}_i) \quad (1.104)$$

En pratique, la moyenne de la prédiction du GP est souvent utilisée pour estimer la probabilité de défaillance car on considère que les GPs construits par apprentissage actif sont généralement très précis du point de vue de la classification. Dans ce cas, le domaine de défaillance \mathcal{D}_f est approché par :

$$\hat{\mathcal{D}}_f^0 = \{\mathbf{x} \in \mathcal{X}, \mu_n(\mathbf{x}) \leq 0\} \quad (1.105)$$

et la probabilité de défaillance est estimée par :

$$\begin{aligned} \hat{P}_f^0 &= \mathbb{P} [\mu_n(\mathbf{X}) \leq 0] \\ &= \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mu_n(\mathbf{X}) \leq 0}(\mathbf{X}_i) \end{aligned} \quad (1.106)$$

Cependant, si le métamodèle construit n'est pas suffisamment précis son utilisation peut introduire un biais dans l'estimateur de la probabilité de défaillance.

La structure d'incertitude du GP offre une prédiction probabiliste. La probabilité de défaillance s'exprime alors comme :

$$\begin{aligned} P_f &\approx \mathbb{E}_{\mathcal{G}_n} \left[\mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{\mathcal{G}_n(\mathbf{X}) \leq 0} \right] \right] \\ &= \int_{\Omega_P} \int_{\mathcal{X}} \mathbb{1}_{g_n(\mathbf{x}) \leq 0} f_{\mathbf{X}}(\mathbf{x}) f_{\mathcal{G}_n}(g_n) d\mathbf{x} dg_n \end{aligned} \quad (1.107)$$

Lorsque la méthode de MC est utilisée pour approcher $\mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{\mathcal{G}_n(\mathbf{X}) \leq 0} \right]$, on se ramène à :

$$\begin{aligned} P_f &\approx \mathbb{E}_{\mathcal{G}_n} \left[\mathbb{E}_{f_{\mathbf{X}}} \left[\mathbb{1}_{\mathcal{G}_n(\mathbf{X}) \leq 0} \right] \right] \\ &\approx \mathbb{E}_{\mathcal{G}_n} \left[\hat{P}_f \right] \end{aligned} \quad (1.108)$$

En un point \mathbf{x} , on peut déterminer la probabilité que ce point appartienne au domaine de défaillance en évaluant la fonction de classification probabiliste $p(\mathbf{x})$ définie par :

$$p(\mathbf{x}) = \mathbb{P} [\mathcal{G}_n(\mathbf{x}) \leq 0] = \Phi \left(-\frac{\mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})} \right) \quad (1.109)$$

Notons que $\mathbb{1}_{\mathcal{G}_n(\mathbf{x}) \leq 0}(\mathbf{x})$ est une variable aléatoire suivant une loi de Bernoulli $\mathcal{B}(p(\mathbf{x}))$ de paramètre $p(\mathbf{x})$, qui correspond à la probabilité que \mathbf{x} appartienne au domaine de défaillance considérant le GP \mathcal{G}_n . L'espérance de \hat{P}_f définie par l'Eq. (1.104) est donnée par :

$$\begin{aligned}
 \mathbb{E}_{\mathcal{G}_n} [\hat{P}_f] &= \mathbb{E}_{\mathcal{G}_n} \left[\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mathcal{G}_n(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) \right] \\
 &= \mathbb{E}_{\mathcal{G}_n} \left[\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathcal{B}(p(\mathbf{X}_i)) \right] \\
 &= \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{E}_{\mathcal{G}_n} [\mathcal{B}(p(\mathbf{X}_i))] \\
 &= \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} p(\mathbf{X}_i)
 \end{aligned} \tag{1.110}$$

Les travaux [PICHENY et al., 2010; PICHENY, 2009; BECT et al., 2012] proposent un estimateur de la probabilité de défaillance utilisant toute l'information du GP en calculant pour chaque point la probabilité qu'il appartienne au domaine de défaillance :

$$\hat{P}_f^p = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} p(\mathbf{X}_i) \tag{1.111}$$

L'utilisation de la fonction de répartition permet de prendre en compte l'incertitude de la prédiction, et donc de la classification, en prenant une valeur comprise entre 0 et 1 si la variance est élevée ou si la moyenne de la prédiction est proche de 0. La fonction de répartition prend des valeurs proches de 0 aux points où la variance de la prédiction est faible et la moyenne bien supérieure à 0 et proches de 1 aux points où la variance de la prédiction est faible et la moyenne bien inférieure à 0.

Dans la prochaine section, nous nous intéresserons aux processus adaptatifs d'apprentissage du GP basés sur des fonctions d'apprentissage. Rappelons que ces méthodes reposent principalement sur les 4 éléments suivants : un choix de métamodèle, une technique d'échantillonnage et les critères d'enrichissement et d'arrêt qui guident l'apprentissage du modèle. Le métamodèle considéré ici étant un GP, un état de l'art des critères d'enrichissement classiques et des critères d'arrêt pour l'apprentissage de ce type de métamodèle est proposé dans la section suivante.

1.2.4.2 Analyses de fiabilité basées sur l'apprentissage actif d'un GP

1.2.4.2.1 Critères d'enrichissement

Dans la littérature de nombreuses fonctions d'apprentissage d'un GP ont été proposées dans le cadre des analyses de fiabilité, parmi les plus connues on trouve les fonctions d'apprentissage introduites dans [RANJAN et al., 2008; BICHON et al., 2008; ECHARD et al., 2011; PICHENY et al., 2010; BECT et al., 2012; CHEVALIER et al., 2014]. Cette section se limite aux fonctions U [ECHARD et al., 2011] et EFF [BICHON et al., 2008] qui sont simples à implémenter et dont l'évaluation est rapide en comparaison aux critères basés sur une intégration.

Fonction U

La fonction U a été proposée dans le cadre de la méthode d'apprentissage actif *Active learning reliability method combining Kriging and Monte Carlo Simulation* (AK-MCS) par [ECHARD et al., 2011].

La fonction U à un point de l'apprentissage \mathbf{x} est donnée par :

$$U(\mathbf{x}) = \frac{|\mu_n(\mathbf{x})|}{\sigma_n(\mathbf{x})} \quad (1.112)$$

La fonction U quantifie la distance en terme d'écart type entre la moyenne de la prédiction par GP et l'état limite. Elle est évaluée sur une population de MC et la fonction de performance $G(\mathbf{x})$ est calculée au point de l'échantillonnage qui minimise U . Dans AK-MCS, le critère de fin d'apprentissage est défini par :

$$\min_{\mathbf{x}}(U(\mathbf{x})) \geq 2 \quad (1.113)$$

Ce critère correspond à une probabilité inférieure à $\Phi(-2) = 0.023$ de faire une erreur sur le signe de la fonction de performance.

Fonction EFF

La fonction $EFF(\mathbf{x})$ pour *Expected Feasibility Function*, inspirée des travaux sur l'estimation de contour [RANJAN et al., 2008], a été introduite pour la méthode *Efficient Global Reliability Analysis* (EGRA) [BICHON et al., 2008] et peut également être utilisée pour la méthode AK-MCS [ECHARD et al., 2011]. Cette fonction correspond à l'espérance que la vraie valeur de la fonction de performance en un point \mathbf{x} appartienne à l'état limite, c.-à-d. vérifie $G(\mathbf{x}) = 0$.

$$EFF(\mathbf{x}) = \mathbb{E}_{\mathcal{G}_n(\mathbf{x})} \left[\epsilon(\mathbf{x}) - \min_{\mathbf{x}} (\epsilon(\mathbf{x}), \mathcal{G}_n(\mathbf{x})) \right] \quad (1.114)$$

Le paramètre $\epsilon(\mathbf{x})$ définit une région dans le voisinage proche de l'état limite $\pm\epsilon(\mathbf{x})$, dans laquelle on veut concentrer l'apprentissage.

L'expression de la fonction EFF est obtenue en intégrant sur le domaine $\pm\epsilon(\mathbf{x})$:

$$EFF(\mathbf{x}) = \int_{-\epsilon}^{\epsilon} (\epsilon - |g_n|) f_{\mathcal{G}_n(\mathbf{x})} dg_n \quad (1.115)$$

où $f_{\mathcal{G}_n(\mathbf{x})}$ est la distribution du GP \mathcal{G}_n au point \mathbf{x} et g_n est une réalisation de $\mathcal{G}_n(\mathbf{x})$.

Cette intégrale peut être exprimée analytiquement comme suit :

$$\begin{aligned}
 EFF(\mathbf{x}) = & \mu_n(\mathbf{x}) \left[2\Phi\left(-\frac{\mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \Phi\left(-\frac{\epsilon(\mathbf{x}) + \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \Phi\left(\frac{\epsilon(\mathbf{x}) - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \right] \\
 & - \sigma_n(\mathbf{x}) \left[2\phi\left(-\frac{\mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \phi\left(-\frac{\epsilon(\mathbf{x}) + \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \phi\left(\frac{\epsilon(\mathbf{x}) - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \right] \\
 & + \epsilon(\mathbf{x}) \left[\Phi\left(\frac{\epsilon(\mathbf{x}) - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) - \Phi\left(-\frac{\epsilon(\mathbf{x}) + \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \right]
 \end{aligned} \tag{1.116}$$

où $\Phi(\cdot)$ est la fonction de répartition de la loi normale standard et $\phi(\cdot)$ la densité de probabilité de la loi normale standard. Dans les méthodes EGRA et **AK-MCS**+*EFF*, la fonction *EFF* est construite avec $\epsilon(\mathbf{x}) = 2\sigma_n(\mathbf{x})$.

A chaque itération de l'algorithme, le meilleur point pour l'évaluation correspond au point qui maximise la fonction *EFF*. Dans la méthode **AK-MCS**, il est déterminé parmi les points d'une population de **MC** et dans la méthode EGRA il est déterminé à l'aide d'une optimisation.

L'apprentissage est stoppé lorsque le critère suivant est vérifié :

$$\max_{\mathbf{x}}(EFF(\mathbf{x})) \leq 0.001 \tag{1.117}$$

On peut différencier les critères d'enrichissement des différentes méthodes par la fonction d'apprentissage utilisée mais aussi par la procédure suivie pour déterminer le meilleur point en s'appuyant sur le critère choisi. En effet selon la méthode, le meilleur point peut être :

- sélectionné en évaluant la fonction d'apprentissage sur un ensemble déterministe de points d'une population de **MC** ;
- déterminé par une optimisation sur le critère d'apprentissage.

Concernant les deux méthodes citées ci-dessus, la méthode EGRA [**BICHON et al., 2008**] s'appuie sur l'optimisation du critère d'apprentissage *EFF* pour sélectionner le meilleur point à ajouter au **DoE** tandis que la méthode **AK-MCS** [**ECHARD et al., 2011**] sélectionne le meilleur point parmi les échantillons d'une population de **MC**. Le Tableau 1.1 fournit un récapitulatif des différents éléments de ces deux méthodes.

La procédure d'apprentissage de la méthode **AK-MCS**, sur laquelle on s'appuiera plus particulièrement pour certains développements proposés dans ces travaux de thèse, est détaillée ci-après.

Méthode **AK-MCS**

Les différentes étapes de la méthode **AK-MCS** sont décrites ci-dessous et illustrées

Méthode	EGRA	AK-MCS
Fonction(s) d'apprentissage	<i>EFF</i>	<i>EFF</i> ou <i>U</i>
Méthode de sélection du meilleur point	Optimisation sur \mathcal{X}	Sélection déterministe sur une population de MC
Méthode d'échantillonnage	IS multimodal adaptatif	MCS

TABLEAU 1.1 – Tableau récapitulatif des éléments des méthodes EGRA et AK-MCS

sur l'exemple du cas 4 branches :

1. Génération d'une population de **MC** de n_{MC} échantillons.
2. Définition d'un **DoE** constitué de n échantillons (en utilisant la méthode d'échantillonnage Latin Hypercube Sampling (LHS) [MCKAY et al., 1979; MACKAY, 1992]). Évaluation de la fonction de performance $G(\mathbf{x})$ en chacun des n échantillons.

L'initialisation de la méthode AK-MCS est illustrée sur la Figure 1.15, sur laquelle est représentée une population de **MC** initiale et un **DoE** initial généré par LHS.

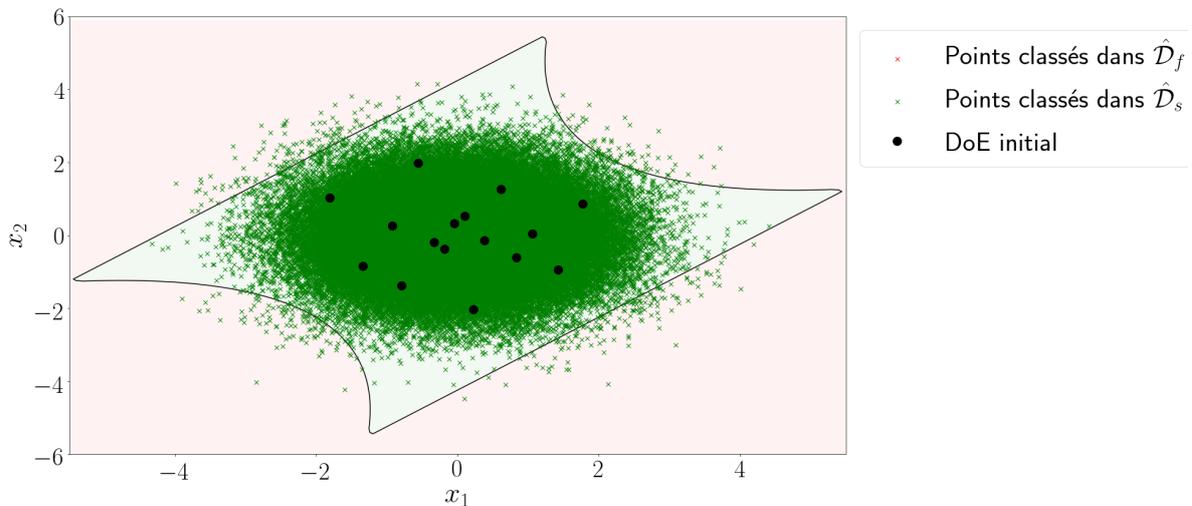


FIGURE 1.15 – Initialisation de AK-MCS.

3. Construction du métamodèle de **GP** \mathcal{G}_n de la fonction de performance $G(\mathbf{x})$ en utilisant les observations du **DoE**.
4. Estimation de la probabilité de défaillance P_f sur la population **MC** en se basant sur la formule suivante :

$$\hat{P}_f^{AK-MCS} = \hat{P}_f^0 = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mu_n(\mathbf{x}) \leq 0}(\mathbf{x}_i) \quad (1.118)$$

5. La fonction d'apprentissage U donnée par l'Eq. (1.112) est évaluée sur toute la population de **MC** pour sélectionner le meilleur point à évaluer \mathbf{x}^* pour enrichir le **GP**. Ce point correspond au point de la population de **MC** qui minimise U , $\mathbf{x}^* = \arg \min U(\mathbf{x})$.

Une étape de l'apprentissage itératif de la méthode **AK-MCS** est illustrée sur la Figure 1.16, où le point x^* symbolisé par l'étoile sur le graphique a été ajouté au **DoE**.

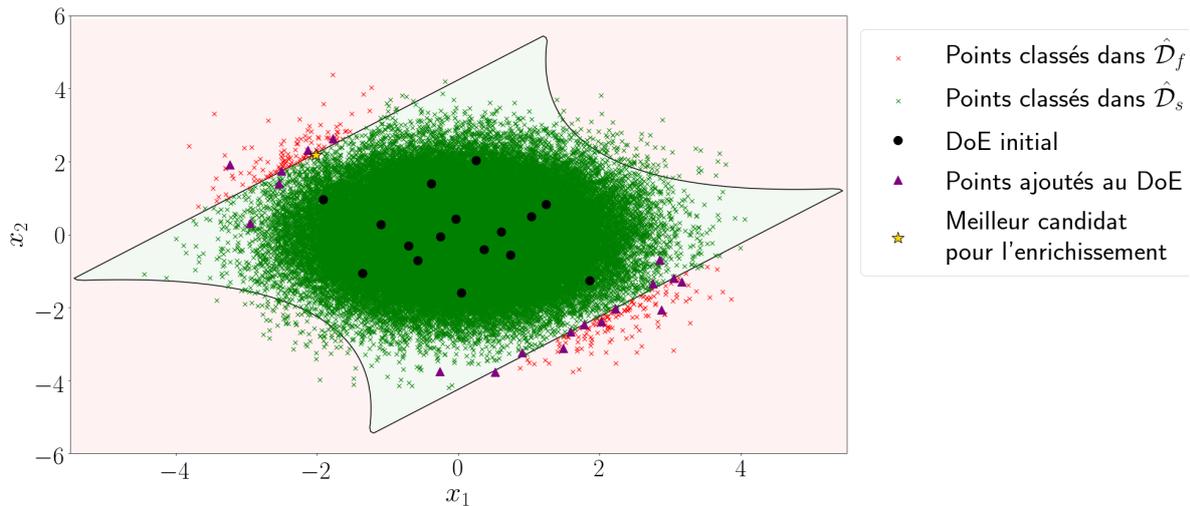


FIGURE 1.16 – Exemple d'enrichissement du **GP** par **AK-MCS** à une étape de l'apprentissage.

6. Si le critère de fin d'apprentissage donné par l'Eq. (1.113) est atteint le métamodèle est considéré comme suffisamment précis pour la population et l'apprentissage est arrêté puis l'algorithme passe à l'étape 8. Sinon il passe à l'étape 7.
7. La fonction de performance $G(\mathbf{x})$ est évaluée au point minimisant la fonction U et le nouveau point x^* est ajouté au **DoE**. Puis l'algorithme retourne à l'étape 3.
8. Le **COV** de **MC** de la probabilité de défaillance estimée \hat{P}_f^{AK-MCS} est calculé en se basant sur l'expression suivante :

$$\widehat{COV}_{MC} = \sqrt{\frac{(1 - \hat{P}_f^{AK-MCS})}{n_{MC} \hat{P}_f^{AK-MCS}}} \quad (1.119)$$

La valeur estimée du **COV** de **MC** \widehat{COV}_{MC} est vérifiée pour s'assurer de la cohérence de la taille de la population de **MC** par rapport à l'ordre de grandeur de la probabilité de défaillance estimée. Si le **COV** est trop élevé, de nouveaux échantillons sont ajoutés à la population de **MC** utilisée et l'algorithme retourne

à l'étape 7. Sinon, si le **COV** est suffisamment faible par rapport à un seuil fixé par l'utilisateur, la dernière estimation de la probabilité de défaillance obtenue est retenue comme estimation finale. Le résultat final de l'algorithme, où l'on a

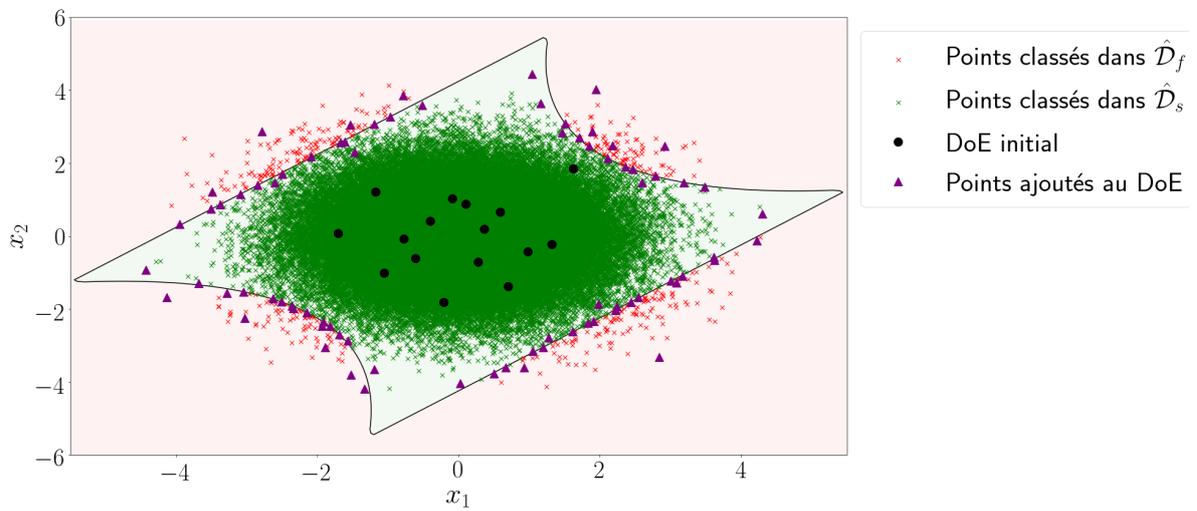


FIGURE 1.17 – Résultat final de AK-MCS à la fin de l'apprentissage.

bien identifié les quatre zones de défaillance, est illustré sur la Figure 1.17.

Les différentes étapes d'AK-MCS sont résumées sur le diagramme de la Figure 1.18.

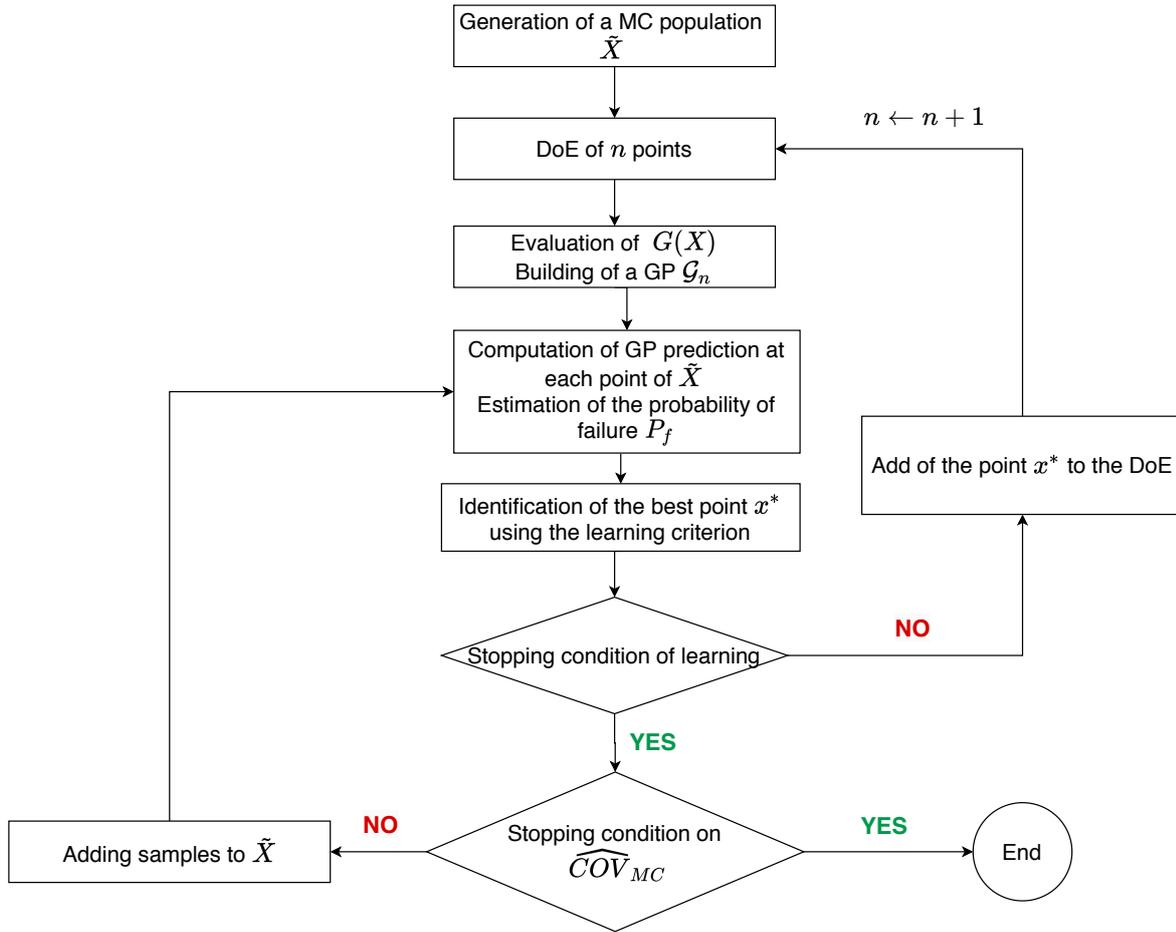


FIGURE 1.18 – Diagramme de l'algorithme AK-MCS.

Remarques

- D'autres méthodes basées sur la fonction d'apprentissage U ont été proposées afin de pouvoir traiter des problèmes plus spécifiques tels que l'estimation d'évènements rares (ou estimation de probabilités faibles) [ECHARD et al., 2013; HUANG et al., 2016; LELIÈVRE et al., 2018; BALESDENT et al., 2013; DUBOURG et al., 2013; LI et al., 2012; CADINI et GIOLETTA, 2016], problèmes à régions de défaillance multiples [CADINI et al., 2014; LV et al., 2015; ZHANG et al., 2019; WANG et SHAFIEEZADEH, 2019b] ou l'obtention de probabilité de défaillance de systèmes [FAURIAT et GAYTON, 2014; GASPARD et al., 2017; BICHON et al., 2011; VAZQUEZ et BECT, 2009; BECT et al., 2012; PERRIN, 2016].
- Les fonctions d'apprentissage ont un comportement différent. Sur la Figure 1.19, l'évolution de la probabilité de défaillance au cours de l'algorithme AK-MCS est donnée pour les deux critères d'apprentissage EFF et U sur le cas 4 branches.

Concernant, la fonction U on peut observer une évolution par paliers de la probabilité de défaillance au cours de la convergence de l'algorithme. Chaque palier correspond à l'apprentissage d'une branche différente de l'état limite, c'est-à-dire que le critère va apprendre précisément une portion de l'état limite avant d'explorer le reste du domaine. Concernant le critère EFF , on peut voir que des valeurs plus dispersées de la probabilité de défaillance sont obtenues au début de l'algorithme ; cette phase correspond à une exploration plus large du domaine. Le critère EFF alterne donc davantage entre l'exploration et l'exploitation que le critère U . On remarque cependant que les deux critères permettent de converger vers la même valeur de \hat{P}_f .

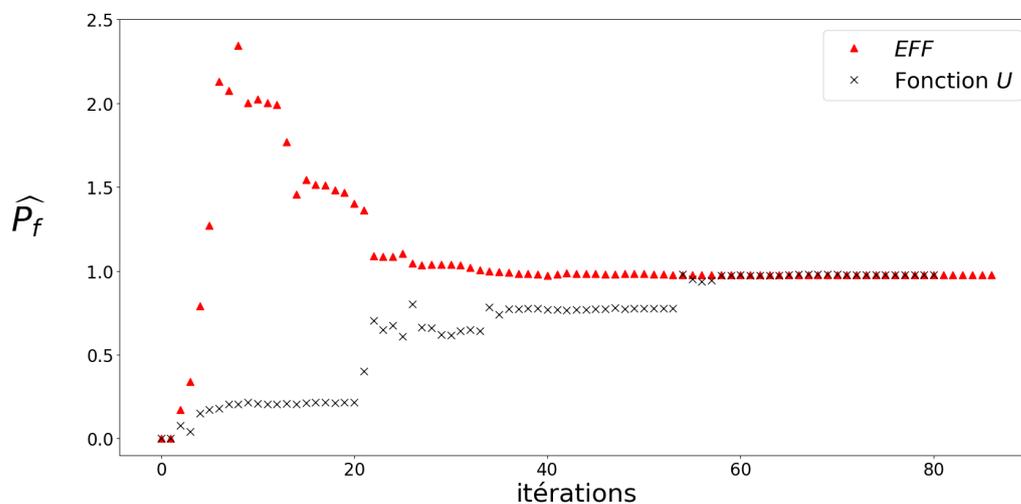


FIGURE 1.19 – Évolution de la probabilité normalisée pour les critères d'apprentissage EFF et U sur le cas 4 branches.

1.2.4.2.2 Autres critères d'arrêt

Dans la partie précédente, les critères d'arrêts d'apprentissage présentés étaient directement liés à la fonction d'apprentissage utilisée mais ces critères sont souvent conservatifs. D'autres critères d'arrêt d'apprentissage ont été développés pour éviter un apprentissage trop conservatif du métamodèle de la fonction de performance tout en vérifiant une certaine précision de la prédiction. C'est notamment le cas dans le domaine de l'optimisation sous contrainte de fiabilité [MOUSTAPHA et SUDRET, 2019; CHAUDHURI et HAFTKA, 2015]. Dans la même optique de réduction du nombre de points d'apprentissage, des critères d'arrêt globaux ont été proposés pour contrôler l'erreur sur l'estimation de la statistique d'intérêt, c.-à-d. ici la probabilité de défaillance. Dans ces travaux, nous nous intéresserons à la variabilité de l'estimateur de la probabilité de défaillance, on choisit donc d'introduire ici les critères construits à partir d'une mesure d'erreur ou de stabilité de l'estimation. Parmi les critères d'arrêt nous pouvons citer :

- Dans [SCHÖBI R. et al., 2017], il a été proposé un critère d'arrêt basé sur la marge de l'état limite \mathbb{M}_f , qui correspond à l'intersection entre les bornes inférieure et supérieure du domaine de défaillance : [DUBOURG et al., 2011] :

$$\mathbb{M}_f = \mathcal{D}_f^+ \cap \mathcal{D}_f^- \quad (1.120)$$

où

$$\mathcal{D}_f^\pm = \{\mathbf{x} \in \mathcal{X}, \mu_n(\mathbf{X}) \pm k\sigma_n(\mathbf{X}) \leq S\} \quad (1.121)$$

Ainsi la phase d'apprentissage est stoppée lorsque la condition suivante est vérifiée pour deux itérations consécutives :

$$\frac{\hat{P}_f^+ - \hat{P}_f^-}{\hat{P}_f^0} \leq \epsilon_{P_f} \quad (1.122)$$

où ϵ_{P_f} permet de mesurer la taille de la marge de l'état limite \mathbb{M}_f et

$$\hat{P}_f^\pm = \mathbb{P} [\mu_n(\mathbf{X}) \pm k\sigma_n(\mathbf{X}) \leq S] \quad (1.123)$$

Dans l'article [SCHÖBI R. et al., 2017], ϵ_{P_f} est fixé à 5% et $k = 2$.

- Dans le cadre de la méthode AK-MC*Si*, [LELIÈVRE et al., 2018] a proposé un nouveau critère d'arrêt basé sur la probabilité d'exactitude de la classification de la population de *MC*, dont le but est notamment d'améliorer l'estimation de probabilités faibles.

Soient C l'évènement "tous les points de la population de *MC* sont classifiés correctement" et C_i l'évènement "le point \mathbf{x}_i est classé correctement". La probabilité de C peut s'exprimer comme suit :

$$\mathbb{P} [C] = \mathbb{P} [C_1 \cap C_2 \cap \dots \cap C_{n_{MC}}] \quad (1.124)$$

On considère l'hypothèse simplificatrice selon laquelle les évènements $C_1, C_2, \dots, C_{n_{MC}}$ sont considérés indépendants pour que le calcul de l'Eq. (1.124) soit faisable pour de grandes populations de *MC*. L'Eq. (1.124) peut alors se réécrire comme suit :

$$\mathbb{P} [C] \approx \prod_{i=1}^{n_{MC}} \mathbb{P} [C_i] \quad (1.125)$$

De plus, la probabilité de chaque évènement C_i est donnée par :

$$\mathbb{P} [C_i] = \Phi \left(\frac{|\mu_n(\mathbf{x}_i)|}{\sigma_n(\mathbf{x}_i)} \right) \quad (1.126)$$

Le critère d'arrêt consiste à vérifier que la probabilité que tous les points du MC soient bien classés soit plus élevée qu'un certain seuil P_{lim} , c.-à-d. que l'algorithme continue tant que :

$$\mathbb{P}[C] \leq P_{lim} \quad (1.127)$$

Ce critère d'arrêt permet de considérer l'ensemble de la population de MC en comparaison au critère classique donné par l'Eq. (1.113) qui ne prend en compte que le pire cas.

- Le critère d'arrêt ESC proposé dans [WANG et SHAFIEEZADEH, 2019a] est basé sur l'erreur relative de l'estimation de la probabilité de défaillance par rapport à la valeur estimée par MCS sur la vraie fonction \hat{P}_f^{MC} :

$$\epsilon = \frac{|\hat{P}_f - \hat{P}_f^{MC}|}{\hat{P}_f^{MC}} = \left| \frac{\hat{n}_f}{n_f} - 1 \right| \quad (1.128)$$

avec n_f le nombre réel de points défailants (c.-à-d. appartenant à \mathcal{D}_f) de la population de Monte-Carlo et \hat{n}_f le nombre de points de la population classés dans \mathcal{D}_f . Le nombre réel n_f de points défailants vérifie :

$$n_f = \hat{n}_f + \hat{S}_s - \hat{S}_f \quad (1.129)$$

avec $\hat{S}_s = \text{card}(\mathcal{D}_f \cap \hat{\mathcal{D}}_f)$ et $\hat{S}_f = \text{card}(\mathcal{D}_f \cap \hat{\mathcal{D}}_s)$, où $\hat{\mathcal{D}}_s$ est l'approximation de \mathcal{D}_s en utilisant le modèle de substitution de la fonction de performance.

Il vient alors :

$$n_f \in [\hat{n}_f - \hat{S}_f^u, \hat{n}_f + \hat{S}_s^u] \quad (1.130)$$

où \hat{S}_f^u et \hat{S}_s^u sont les bornes supérieures de l'intervalle de confiance d'un certain niveau α des estimateurs \hat{S}_f et \hat{S}_s . Il peut être démontré que \hat{S}_f et \hat{S}_s suivent des lois de Poisson binomiales et leurs bornes peuvent donc être calculées en conséquence comme explicité dans l'article [WANG et SHAFIEEZADEH, 2019a].

Finalement, l'erreur maximale est estimée comme suit :

$$\epsilon \leq \max \left(\left| \frac{\hat{n}_f}{\hat{n}_f - \hat{S}_f^u} - 1 \right|, \left| \frac{\hat{n}_f}{\hat{n}_f + \hat{S}_s^u} - 1 \right| \right) = \hat{\epsilon}_{max} \quad (1.131)$$

Le critère d'arrêt ESC consiste à vérifier que :

$$\hat{\epsilon}_{max} \leq \epsilon_l \quad (1.132)$$

où ϵ_l correspond à une valeur limite acceptable pour l'erreur.

- Dans [GASPAR et al., 2015] un critère d'arrêt basé sur la stabilité de l'estimation de la probabilité de défaillance en utilisant le GP est proposé. Ce critère d'arrêt est donné par :

$$\frac{\hat{P}_f^{(i)} - \hat{P}_f^{(1)}}{\hat{P}_f^{(i)}} \leq \gamma, \quad \forall i \in \{2, \dots, n_\gamma\} \quad (1.133)$$

où $\hat{P}_f^{(1)}$ est une valeur de référence utilisée pour détecter la stabilisation de l'estimation de la probabilité de défaillance et $\hat{P}_f^{(i)}$ pour $i = 2, \dots, n_\gamma$ sont les estimations de la probabilité de défaillance aux $n_\gamma - 1$ itérations suivantes. La valeur de référence est initialisée en utilisant la première probabilité estimée puis elle est redéfinie au cours de l'algorithme par la valeur de l'estimation de la probabilité de défaillance lorsque celle-ci ne vérifie pas la condition (1.133). Ce critère dépend des deux paramètres γ , la tolérance sur l'estimation \hat{P}_f , et n_γ , le nombre d'itérations successives de l'apprentissage où la condition (1.133) doit être vérifiée.

Ce critère est plutôt utilisé comme un critère additionnel aux critères proposés dans [ECHARD et al., 2011] pour la méthode AK-MCS et est utile (ou devrait être utilisé) uniquement dans le cas où un fort effet de stabilisation de l'estimation de P_f est observé durant l'apprentissage actif du modèle de GP. En effet, en revenant sur la Figure 1.19 et plus précisément sur l'évolution de l'estimation de P_f au cours de l'algorithme avec le critère U , on peut voir que l'utilisation de ce critère comme unique critère d'arrêt avec un choix de n_γ mal avisé (trop petit) pourrait mener à une estimation finale erronée.

1.2.5 Conclusion

Dans cette section, différentes méthodes d'analyse de fiabilité ont été présentées. Les méthodes FORM/ SORM ont l'avantage d'être peu coûteuses car elles demandent peu d'appels au modèle mais leur efficacité est limitée au cas où l'état limite est relativement simple.

Différentes méthodes d'échantillonnage, permettant de résoudre des problèmes de fiabilité complexes et d'estimer des probabilité d'évènements rares, ont aussi été introduites dans cette section. Ces méthodes ont néanmoins l'inconvénient de nécessiter un nombre important de simulations. Ceci les rend inadaptées pour l'analyse des modèles complexes, pour lesquels les temps de calcul sont très élevés.

Finalement, les méthodes d'analyses de fiabilité basées sur l'utilisation de métamodèles, et plus précisément de métamodèles de GP, ont été introduites. Ces méthodes permettent de traiter des problèmes de fiabilité complexes et sont prometteuses d'un point de vue de la réduction des temps de calcul et ce notamment lorsque le DoE est judicieusement choisi. C'est l'objectif des méthodes basées sur un processus adaptatif

d'apprentissage du GP à partir de fonctions d'apprentissage telles que les méthodes classiques EGRA ou AK-MCS, qui ont été présentées. De plus, la précision de la prédiction de l'état limite étant vérifiée par un critère d'apprentissage, ces techniques adaptatives garantissent que l'erreur sur l'estimation de la probabilité de défaillance est contrôlée. Différents critères d'arrêt de l'apprentissage ont été présentés parmi les critères existants. On peut les classer en deux catégories : d'une part les critères liés à la fonction d'apprentissage, qui sont souvent conservatifs, et d'autre part des critères construits à partir d'une mesure de la stabilité ou de l'erreur de l'estimation de la probabilité de défaillance. En effet, la statistique d'intérêt de ces analyses est l'estimateur de la probabilité de défaillance, il serait donc intéressant d'estimer l'erreur sur la probabilité de défaillance liée à l'approximation par GP pour la contrôler au cours de la procédure d'apprentissage.

Résumé du chapitre

Dans un premier temps, ce chapitre a introduit différentes méthodes de réduction du coût numérique de modèles qui peuvent prendre ou non en compte la physique du modèle. D'une part le principe de la métamodélisation, qui consiste à construire une approximation par une fonction mathématique de la sortie d'un code numérique, a été introduite. Le modèle de *processus gaussien*, ou *Gaussian process* a été présenté plus précisément car sa structure d'incertitude est particulièrement intéressante en fiabilité. En effet, elle lui permet de fournir un indicateur d'erreur utile pour l'enrichissement de la base d'apprentissage dans le cadre de problèmes de classification, qui sont sous-jacents aux études de fiabilité. D'autre part, le principe de la construction de modèles réduits, tenant compte de la physique du problème, a été présenté dans le cadre de la réduction de modèles *Éléments Finis* et un état de l'art plus approfondi des méthodes de réduction de modèles linéaires paramétriques a été dressé.

Dans un second temps, différentes méthodes d'analyses de fiabilité ont été introduites. Tout d'abord, les méthodes FORM/SORM ont été présentés, celles-ci sont très populaires pour leur faible coût de calcul mais uniquement efficaces pour des fonctions d'état limite régulières. Les méthodes d'échantillonnage, plus générales, ont ensuite été introduites. Parmi ces méthodes, nous nous sommes plus particulièrement intéressés aux techniques d'*Échantillonnage préférentiel*, ou *Importance Sampling* qui seront utilisées dans ces travaux pour traiter des probabilités de défaillance faibles.

Cependant, ces méthodes, efficaces d'un point de vue de la réduction de la variance, sont inadaptées lorsque le coût d'évaluation de la fonction de performance est très élevé. L'utilisation d'un métamodèle pour prédire les résultats de la fonction de performance semble être une solution naturelle pour réduire le nombre d'évaluations de la fonction coûteuse. Dans ce cas, le choix du *Plan d'expériences*, ou *Design of Experiments* utilisé pour construire le métamodèle est évidemment primordial. Les méthodes d'analyse de fiabilité basées sur l'apprentissage actif présentées ensuite permettent de choisir le *DoE* de façon adaptative et sont donc particulièrement prometteuses pour la réduction du nombre d'appels à la fonction de performance. Différents critères d'apprentissage et critères d'arrêt de ces algorithmes sont présentés et les différentes étapes de la méthode classique d'apprentissage actif *AK-MCS*, qui sera utilisée dans ces travaux, sont détaillées.

CHAPITRE 2

Étude de la sensibilité de la probabilité de défaillance à l'approximation par processus gaussien et à l'intégration par Monte-Carlo et construction d'un nouveau critère d'apprentissage associé

Objectifs du chapitre

- Analyser l'influence de l'échantillonnage de Monte-Carlo et de l'approximation par processus gaussien sur l'estimateur de la probabilité de défaillance.
- Proposer un nouveau critère d'apprentissage basé sur cette analyse de sensibilité pour les méthodes d'analyse de fiabilité par apprentissage actif.
- Proposer une extension de la méthodologie pour l'estimation de probabilité d'évènements rares.

Sommaire

2.1	Introduction	67
2.2	Mesure de la sensibilité de la probabilité de défaillance aux incertitudes du GP et à l'estimation par une simulation de Monte-Carlo	68
2.2.1	Estimation des contributions à la variance de \hat{P}_f	69
2.2.2	Motivation pour le développement d'une nouvelle approche illustrée sur un cas test	74
2.3	Approche basée sur les mesures de sensibilité aux incertitudes	76
2.3.1	Principe général	76
2.3.2	Algorithme proposé	76
2.4	Applications	79
2.4.1	Méthodologie et mesures de comparaison	79
2.4.2	Système en série à 4 branches	83
2.4.3	Réponse dynamique d'un oscillateur non linéaire	86
2.5	Amélioration avec l'Échantillonnage préférentiel, ou <i>Importance Sampling</i>	88
2.5.1	Estimateur de la probabilité de défaillance dans le cas de l'IS	88
2.5.2	Estimateurs de variance	89
2.5.3	Méthode étendue à l'échantillonnage préférentiel	91
2.5.4	Applications	94
2.6	Conclusion	97

2.1 Introduction

Les approches d'apprentissage actif présentées dans le [chapitre 1](#) sont un moyen de réduire fortement les temps de calcul des analyses de fiabilité en réduisant le nombre d'appels au modèle. Pour cela, ces méthodes s'appuient sur la construction adaptative d'un modèle de GP \mathcal{G}_n , précis du point de vue de la classification, de la fonction de performance coûteuse à évaluer. L'arrêt de l'apprentissage du métamodèle se base généralement sur un critère sur l'erreur de la prédiction au voisinage de l'état limite pour déterminer si celui-ci est suffisamment précis. Finalement, l'estimation de la probabilité de défaillance est obtenue en évaluant une population de MC grâce au métamodèle \mathcal{G}_n ainsi construit.

La probabilité de défaillance estimée sur une population de MC $\tilde{\mathbf{X}} = (\mathbf{X}_i)_{i=1, \dots, n_{MC}}$ avec \mathbf{X}_i i.i.d de même distribution que le vecteur aléatoire \mathbf{X} de dimension n_u est généralement estimée en se basant sur la moyenne de la prédiction du GP :

$$\hat{P}_f^{MC}(\tilde{\mathbf{X}}) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mu_n(\mathbf{X}_i) \leq 0} \quad (2.1)$$

Ainsi, l'erreur finale d'estimation de la probabilité de défaillance obtenue prend uniquement en compte la variabilité due à l'intégration par Monte-Carlo et non l'incertitude de l'approximation par un métamodèle car celui-ci est considéré suffisamment précis. Néanmoins, la méthodologie qui consiste à considérer l'erreur locale de la prédiction au voisinage de l'état limite est souvent trop conservative. Certains travaux ont déjà proposé de prendre en compte la contribution du processus gaussien sur la précision de la probabilité de défaillance en utilisant des bornes de l'estimateur [[SCHÖBI R. et al., 2017](#)] ou une approximation de l'erreur relative [[WANG et SHAFIEEZADEH, 2019a](#)] pour définir un critère d'arrêt d'apprentissage. A partir de là, les axes d'amélioration suivants peuvent être relevés :

- D'une part, il serait intéressant de pouvoir quantifier l'erreur individuelle introduite par chaque source d'incertitude et d'adapter en fonction l'enrichissement, c.-à-d. améliorer le GP ou enrichir la population de MC.
- D'autre part, pour pouvoir valider la précision de l'estimation de la probabilité de défaillance, il serait intéressant de se baser sur la variance totale de l'estimation tenant compte des incertitudes de prédiction et d'intégration.

Dans ces travaux, nous montrons que l'analyse de l'influence de l'échantillonnage du MC et du métamodèle sur l'estimateur de la probabilité de défaillance défini par l'Eq. (2.1) peut s'écrire comme un problème d'analyse de sensibilité basé sur la décomposition de la variance. Dans la [Section 2.2.1](#) nous proposons des estimateurs de

ces indices de sensibilité et de la variance totale. Il est alors possible d'analyser quantitativement quelle source d'incertitude doit être réduite pour améliorer la précision de l'estimation de la probabilité de défaillance à partir de ces estimateurs. De plus, ils permettent de donner une estimation du COV total de l'estimation de la probabilité de défaillance obtenue. Dans la Section 2.3, nous proposons ensuite un nouvel algorithme d'analyse de fiabilité qui intègre cette analyse pour améliorer adaptativement la source d'erreur majeure durant la phase d'apprentissage. Un critère d'arrêt global basé sur la variance totale de l'estimation de la probabilité de défaillance est également proposé pour cet algorithme. Finalement, nous présentons dans la Section 2.5 une extension de cette méthode afin de pouvoir traiter des problèmes d'estimation de probabilité d'évènements rares, via un algorithme basé sur l'IS.

2.2 Mesure de la sensibilité de la probabilité de défaillance aux incertitudes du GP et à l'estimation par une simulation de Monte-Carlo

Dans le cadre des analyses de fiabilité par apprentissage actif, l'estimateur de P_f est en fait une variable aléatoire dont la variance dépend à la fois de l'incertitude de l'intégration par MCS et de l'approximation de la fonction de performance par un modèle de substitution. L'idée est alors de considérer l'estimation par MC donnée par l'Eq. (2.1) mais de remplacer la moyenne du GP par le GP \mathcal{G}_n :

$$\hat{P}_f(\tilde{\mathbf{X}}, \mathcal{G}_n) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mathcal{G}_n(\mathbf{x}_i) \leq 0}(\mathbf{x}_i) = \Gamma(\tilde{\mathbf{X}}, \mathcal{G}_n) \quad (2.2)$$

où Γ est une fonction déterministe ayant pour entrées le vecteur aléatoire de l'échantillonnage par MC $\tilde{\mathbf{X}}$ et le GP conditionné \mathcal{G}_n .

Les contributions de chacune des deux sources d'incertitude sur la variabilité (i.e. variance) de l'estimateur \hat{P}_f peuvent être obtenues en se référant à un outil classique en analyse de sensibilité qui est l'expression de la décomposition de la variance [IOOSS et LEMAÎTRE, 2015] :

$$Var_{\mathcal{G}_n, \tilde{\mathbf{X}}} \left(\hat{P}_f(\tilde{\mathbf{X}}, \mathcal{G}_n) \right) = V_{\tilde{\mathbf{X}}} + V_{\mathcal{G}_n} + V_{\mathcal{G}_n, \tilde{\mathbf{X}}} \quad (2.3)$$

où :

- $V_{\tilde{\mathbf{X}}} = Var_{\tilde{\mathbf{X}}} \left(\mathbb{E}_{\mathcal{G}_n} \left[\hat{P}_f | \tilde{\mathbf{X}} \right] \right)$ est une mesure de l'influence de l'échantillonnage par MC sur la variance de \hat{P}_f ,
- $V_{\mathcal{G}_n} = Var_{\mathcal{G}_n} \left(\mathbb{E}_{\tilde{\mathbf{X}}} \left[\hat{P}_f | \mathcal{G}_n \right] \right)$ donne une mesure de l'influence de l'incertitude du

GP sur la variance de \hat{P}_f ,

- $V_{\mathcal{G}_n, \tilde{\mathbf{X}}} = \text{Var}_{\mathcal{G}_n, \tilde{\mathbf{X}}} \left(\mathbb{E} \left[\hat{P}_f | \mathcal{G}_n, \tilde{\mathbf{X}} \right] \right) - V_{\mathcal{G}_n} - V_{\tilde{\mathbf{X}}}$ est une mesure de la contribution jointe sur la variance de \hat{P}_f des incertitudes due à la fois à l'utilisation de la méthode de MC et à l'approximation par un GP.

2.2.1 Estimation des contributions à la variance de \hat{P}_f

Lors de la phase d'enrichissement, les points d'entraînement du GP étant choisis parmi les points constituant la population de MC, les deux sources d'incertitude ne sont théoriquement pas indépendantes. Cependant, comme les échantillons de la population de MC utilisés pour l'apprentissage du GP ne représentent dans les cas qui nous intéressent qu'une infime partie de la population, les estimateurs développés dans la prochaine section sont basés sur des hypothèses d'indépendance.

2.2.1.1 Estimateur de variance

Supposons que l'on dispose d'un échantillon i.i.d. de taille n_s (Z_1, \dots, Z_{n_s}) d'une variable aléatoire Z de distribution inconnue. La moyenne de Z est approchée par sa moyenne empirique sur les n_s échantillons, notée $\overline{Z_{n_s}}$. La variance empirique d'une variable aléatoire Z , notée $\widehat{Var}(Z)$, est définie par :

$$\widehat{Var}(Z) = \frac{1}{n_s - 1} \sum_{i=1}^{n_s} (Z_i - \overline{Z_{n_s}})^2 \quad (2.4)$$

Néanmoins, une estimation ponctuelle ne prend pas en compte la variabilité due aux variations d'échantillonnage et ne donne pas d'information sur la précision de l'estimation de $\widehat{Var}(Z)$. Afin d'avoir ces précisions, il faut calculer une estimation des intervalles de confiance contenant avec une probabilité définie la valeur correcte de la variance. L'intervalle de confiance asymptotique de la variance de niveau $1 - \alpha$ estimé $\left[\widehat{Var}^{inf}(Z), \widehat{Var}^{sup}(Z) \right]$ est donné à l'aide du théorème centrale limite par :

$$\left[\widehat{Var}(Z) - k \frac{1}{1 - n_s} \sqrt{n_s \widehat{Var} \left((Z_i - \overline{Z_{n_s}})^2 \right)}; \quad \widehat{Var}(Z) + k \frac{1}{1 - n_s} \sqrt{n_s \widehat{Var} \left((Z_i - \overline{Z_{n_s}})^2 \right)} \right] \quad (2.5)$$

où k correspond au quantile d'ordre $1 - \alpha$ de la loi normale centrée réduite.

Cette propriété de l'estimateur de la variance permet de comparer les bornes des intervalles de confiance de $V_{\tilde{\mathbf{X}}}$ et $V_{\mathcal{G}_n}$ et d'ajuster le nombre de simulations à calculer pour obtenir des intervalles de confiance disjoints, permettant ainsi de pouvoir identifier la source d'incertitude ayant le plus d'influence sur la variabilité de \hat{P}_f .

2.2.1.2 Expression de l'estimateur de $V_{\tilde{\mathbf{X}}}$

Rappelons que la variable aléatoire $Y = \mathbb{1}_{\mathcal{G}_n(\mathbf{x}) \leq 0}(\mathbf{x})$ suit une loi de Bernoulli $\mathcal{B}(p(\mathbf{x}))$ de paramètre $p(\mathbf{x})$, la probabilité que \mathbf{x} appartienne au domaine de défaillance \mathcal{D}_f considérant le GP \mathcal{G}_n , qui est défini par :

$$\begin{aligned} p(\mathbf{x}) &= \mathbb{P}[\mathcal{G}_n(\mathbf{x}) \leq 0] \\ &= \Phi\left(-\frac{\mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \end{aligned} \quad (2.6)$$

avec $\Phi(\cdot)$ la fonction de répartition de la loi normale standard.

L'espérance de \hat{P}_f , étant donnée une population de MC de n_{MC} échantillons $\tilde{\mathbf{X}} = (\mathbf{X}_i)_{i=1, \dots, n_{MC}}$, peut s'exprimer comme suit :

$$\begin{aligned} \mathbb{E}_{\mathcal{G}_n}[\hat{P}_f | \tilde{\mathbf{X}}] &= \mathbb{E}_{\mathcal{G}_n} \left[\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mathcal{G}_n(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) | (\mathbf{X}_i)_{i=1, \dots, n_{MC}} \right] \\ &= \mathbb{E}_{\mathcal{G}_n} \left[\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathcal{B}(p(\mathbf{X}_i)) \right] \\ &= \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{E}_{\mathcal{G}_n}[\mathcal{B}(p(\mathbf{X}_i))] \\ &= \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} p(\mathbf{X}_i) \end{aligned} \quad (2.7)$$

En se basant sur l'expression de $\mathbb{E}_{\mathcal{G}_n}[\hat{P}_f | \tilde{\mathbf{X}}]$ donnée par l'Eq. (2.7) et celle de $V_{\tilde{\mathbf{X}}}$ donnée à la suite de l'Eq. (2.3), la variance $V_{\tilde{\mathbf{X}}}$ peut être obtenue en évaluant :

$$\begin{aligned} V_{\tilde{\mathbf{X}}} &= \text{Var}_{\tilde{\mathbf{X}}} \left(\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} p(\mathbf{X}_i) \right) \\ &= \frac{\text{Var}_{\tilde{\mathbf{X}}}(p(\mathbf{X}))}{n_{MC}} \end{aligned} \quad (2.8)$$

On obtient la dernière égalité car $p(\mathbf{X})$ est une variable aléatoire continue entre 0 et 1 dont les $p(\mathbf{X}_i)$ sont des répliques iid.

En pratique, l'Eq. (2.8) est estimée sur la population de MC utilisée pour l'estimation de la probabilité de défaillance. L'estimateur de $V_{\tilde{\mathbf{X}}}$, dénoté $\widehat{V}_{\tilde{\mathbf{X}}}$, sur une réalisation d'une population de MC est ainsi donné par :

$$\widehat{V}_{\tilde{\mathbf{X}}} = \frac{\widehat{\text{Var}}_{\tilde{\mathbf{X}}}(p(\tilde{\mathbf{X}}))}{n_{MC}} = \frac{1}{n_{MC}(n_{MC} - 1)} \sum_{i=1}^{n_{MC}} \left(p(\mathbf{X}_i) - \frac{1}{n_{MC}} \sum_{j=1}^{n_{MC}} p(\mathbf{X}_j) \right)^2 \quad (2.9)$$

et l'estimation de son intervalle de confiance d'ordre $1 - \alpha$ exprimée en utilisant les

opérateurs définis dans la Section 2.2.1.1 est donnée par :

$$\begin{aligned}\widehat{V_{\tilde{\mathbf{X}}}^{inf}} &= \frac{\widehat{Var_{\tilde{\mathbf{X}}}^{inf}}(p(\tilde{\mathbf{x}}))}{n_{MC}} \\ \widehat{V_{\tilde{\mathbf{X}}}^{sup}} &= \frac{\widehat{Var_{\tilde{\mathbf{X}}}^{sup}}(p(\tilde{\mathbf{x}}))}{n_{MC}}\end{aligned}\quad (2.10)$$

2.2.1.3 Expression de l'estimateur de $V_{\mathcal{G}_n}$

Le calcul de l'espérance de \hat{P}_f conditionnée à une réalisation de \mathcal{G}_n peut être interprété comme une simulation de MC classique pour un modèle déterministe. Ainsi elle vérifie l'égalité suivante :

$$\mathbb{E}_{\tilde{\mathbf{X}}} [\hat{P}_f | \mathcal{G}_n] = \mathbb{E}_{\tilde{\mathbf{X}}} \left[\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mathcal{G}_n(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) | \mathcal{G}_n \right] \quad (2.11)$$

$$= P_f(\mathcal{G}_n) \quad (2.12)$$

où $P_f(\mathcal{G}_n)$ est la probabilité de défaillance calculée pour une réalisation de \mathcal{G}_n .

Généralement, cette probabilité de défaillance est estimée à l'aide d'un estimateur de MC :

$$\hat{P}_f^{MC}(\mathcal{G}_n) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mathcal{G}_n(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) \quad (2.13)$$

avec $(\mathbf{X}_i)_{i=1, \dots, n_{MC}}$ une réalisation d'une population de MC.

Ainsi, la variance $V_{\mathcal{G}_n}$ peut être estimée numériquement en simulant différentes trajectoires de \mathcal{G}_n et en calculant l'estimation par MCS de P_f pour chaque trajectoire simulée.

Soient $(G_i)_{1 \leq i \leq n_t}$ n_t réalisations de \mathcal{G}_n , l'estimation de la variance $V_{\mathcal{G}_n}$ correspond à la variance empirique de l'échantillon $\hat{P}_f^{MC}(G_i)_{1 \leq i \leq n_t}$:

$$\widehat{V_{\mathcal{G}_n}} = \widehat{Var_{\mathcal{G}_n}}(\hat{P}_f^{MC}(\mathcal{G}_n)) = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} \left(\hat{P}_f^{MC}(G_i) - \frac{1}{n_t} \sum_{i=1}^{n_t} \hat{P}_f^{MC}(G_i) \right)^2 \quad (2.14)$$

L'estimation des bornes de son intervalle de confiance de niveau $1 - \alpha$, exprimée en utilisant les opérateurs définis dans la Section 2.2.1.1, sont données par :

$$\begin{aligned}\widehat{V_{\mathcal{G}_n}^{inf}} &= \widehat{Var_{\mathcal{G}_n}^{inf}}(\hat{P}_f^{MC}(\mathcal{G}_n)) \\ \widehat{V_{\mathcal{G}_n}^{sup}} &= \widehat{Var_{\mathcal{G}_n}^{sup}}(\hat{P}_f^{MC}(\mathcal{G}_n))\end{aligned}\quad (2.15)$$

En pratique, la génération de réalisations du GP conditionné \mathcal{G}_n est sujet à de nombreux problèmes numériques. Ces problèmes numériques et leurs causes ont été exposés dans la Section 1.1.2.6 ainsi qu'une méthode alternative, présentée dans les travaux de [LE GRATIET et al., 2014; VILLEMONTAIX et al., 2009], basée sur la simulation d'un

GP non conditionné permettant d'éviter ces problèmes numériques.

Remarque Notons que dans le cas où on considérerait uniquement l'incertitude du GP et pas du MC, i.e. $\mathbf{X}_i = \mathbf{x}_i$, la variance de $\hat{P}_f(\mathcal{G}_n)$ par rapport au GP admet une expression analytique. Notons $Y_i = \mathbb{1}_{\mathcal{G}_n(\mathbf{x}_i) \leq 0}(\mathbf{x}_i)$, on a en effet :

$$Var_{\mathcal{G}_n} [\hat{P}_f(\mathcal{G}_n)] = Var_{\mathcal{G}_n} \left[\frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} Y_i \right] \quad (2.16)$$

$$= \frac{1}{n_{MC}^2} \sum_{i=1}^{n_{MC}} \left(Var [Y_i] + \sum_{j \neq i} CoVar [Y_i, Y_j] \right) \quad (2.17)$$

avec :

$$Var [Y_i] = p(\mathbf{x}_i)(1 - p(\mathbf{x}_i)) \quad (2.18)$$

et :

$$CoVar [Y_i, Y_j] = \mathbb{P} [\mathcal{G}_n(\mathbf{x}_i) \leq 0, \mathcal{G}_n(\mathbf{x}_j) \leq 0] - \mathbb{E}_{\mathcal{G}_n} [Y_i] \mathbb{E}_{\mathcal{G}_n} [Y_j] \quad (2.19)$$

$$= \Phi_{2, \mathbf{C}_{n,2}}(0, 0) - p(\mathbf{x}_i)p(\mathbf{x}_j) \quad (2.20)$$

où $\Phi_{2, \mathbf{C}_{n,2}}(\cdot, \cdot)$ est la fonction de répartition de la loi normale de dimension 2 de moyenne $(\mu_n(\mathbf{x}_i), \mu_n(\mathbf{x}_j))$ et de matrice de covariance $\mathbf{C}_{n,2}$ la matrice de covariance conditionnée du GP \mathcal{G}_n calculée aux points \mathbf{x}_i et \mathbf{x}_j .

Ces expressions analytiques sont par exemple utilisées dans [EL HAJ et SOUBRA, 2021; ZHU et DU, 2016] pour proposer une fonction d'apprentissage définie par $c_i = Var [Y_i] + \sum_{j \neq i} CoVar [Y_i, Y_j]$. Or le calcul successif des termes $CoVar [Y_i, Y_j]$ sur une population de MC représente des coûts de calculs importants. L'article [EL HAJ et SOUBRA, 2021] propose de ne calculer les c_i que sur un sous-ensemble de points du MC sélectionnés en se basant sur leurs contributions individuelles. Or la stratégie qui consiste à ne considérer qu'un sous-ensemble de points n'est pas possible pour l'estimation de la variance. De plus, ces expressions ne sont pas applicables dans notre cas puisqu'on tient compte de l'incertitude sur le MC.

2.2.1.4 Expression de l'estimateur de la variance totale

La variance totale de \hat{P}_f peut être estimée comme suit :

$$\widehat{V}_{tot} = \widehat{Var}_{\mathcal{G}_n, \tilde{\mathbf{X}}} \left(\hat{P}_f(\tilde{\mathbf{X}}, \mathcal{G}_n) \right) = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} \left(\hat{P}_f(G_i, \tilde{\mathbf{X}}_i) - \frac{1}{n_t} \sum_{j=1}^{n_t} \hat{P}_f(G_j, \tilde{\mathbf{X}}_j) \right)^2 \quad (2.21)$$

où les $(G_i, \tilde{\mathbf{X}}_i)$, $i = 1, \dots, n_t$ sont n_t réalisations de \mathcal{G}_n et populations de MC $\tilde{\mathbf{X}}$

et $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i)$ est l'estimation de la probabilité de défaillance obtenue pour la $i^{\text{ème}}$ réalisation $(G_i, \tilde{\mathbf{X}}_i)$ de \mathcal{G}_n et de population de MC donnée par :

$$\hat{P}_f(G_i, \tilde{\mathbf{X}}_i) = \frac{1}{n_{MC}} \sum_{j=1}^{n_{MC}} \mathbb{1}_{G_i(\mathbf{X}_{ij}) \leq 0}(\mathbf{X}_{ij}) \quad (2.22)$$

où \mathbf{X}_{ij} est le $j^{\text{ème}}$ échantillon de la population de MC $\tilde{\mathbf{X}}_i$. L'estimateur de la variance totale \widehat{V}_{tot} est la variance empirique de l'échantillon $(\hat{P}_f(G_i, \tilde{\mathbf{X}}_i))_{1 \leq i \leq n_t}$.

En pratique, les n_t estimations des $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i) = \hat{P}_f(G_i(\tilde{\mathbf{X}}_i))$ nécessaires pour avoir une variance suffisamment faible de l'estimateur \widehat{V}_{tot} peut représenter un coût de calcul important. Pour réduire ce coût de calcul, on propose d'utiliser une procédure bootstrap. En effet, la simulation de trajectoires multiples G_i d'un GP sur le même ensemble de points $\tilde{\mathbf{X}}$ est bien plus efficace numériquement que l'obtention des réalisations $G_i(\tilde{\mathbf{X}}_i)$. On peut ainsi s'appuyer sur la méthode bootstrap présentée dans [ARCHER et al., 1997] pour simuler plusieurs populations de MC $\tilde{\mathbf{X}}_i$ à partir de la population $\tilde{\mathbf{X}}$ sur laquelle n_t trajectoires du GP auront été simulées. La méthode suggérée pour estimer les $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i)$ est présentée dans l'Algorithme 3.

Algorithm 3 Procédure d'évaluation des $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i)$

Require: $\tilde{\mathbf{x}}, \mathcal{G}_n(\mathbf{x}), n_t$

- 1: Simulation de $i = 1, \dots, n_t$ trajectoires $G_i(\tilde{\mathbf{x}}) = (G_i(\mathbf{x}_j))_{j=1, \dots, n_{MC}}$ de $\mathcal{G}_n(\mathbf{x})$ aux points de la population de MC $\tilde{\mathbf{x}}$
 - 2: **for** $i = 1, \dots, n_t$ **do**
 - 3: Échantillonnage avec remplacement d'un échantillon \mathbf{z} à partir de $G_i(\tilde{\mathbf{x}})$
 - 4: Estimation de $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i)$ à partir de l'Eq. (2.22) avec $\tilde{\mathbf{X}}_i = \mathbf{z}$
 - 5: **end for**
 - 6: **return** Estimations de $(\hat{P}_f(G_i, \tilde{\mathbf{X}}_i))_{i \in 1, \dots, n_t}$
-

L'estimation de l'intervalle de confiance de niveau $1 - \alpha$ de \widehat{V}_{tot} , exprimée en utilisant les opérateurs définis dans la Section 2.2.1.1, est finalement donnée par :

$$\begin{aligned} \widehat{V}_{tot}^{inf} &= \widehat{Var}_{\mathcal{G}_n, \tilde{\mathbf{X}}}^{inf}(\hat{P}_f(\mathcal{G}_n, \tilde{\mathbf{X}})) \\ \widehat{V}_{tot}^{sup} &= \widehat{Var}_{\mathcal{G}_n, \tilde{\mathbf{X}}}^{sup}(\hat{P}_f(\mathcal{G}_n, \tilde{\mathbf{X}})) \end{aligned} \quad (2.23)$$

Finalement, la probabilité de défaillance est estimée par la moyenne empirique des $(\hat{P}_f(G_i, \tilde{\mathbf{X}}_i))_{1 \leq i \leq n_t}$, c.-à-d. par :

$$\hat{P}_f^t = \frac{1}{n_t} \sum_{i=1}^{n_t} \hat{P}_f(G_i, \tilde{\mathbf{X}}_i) \quad (2.24)$$

Le COV total, noté COV_{tot} , de l'estimateur de MC \hat{P}_f^{MC} de la probabilité de défaillance \hat{P}_f pour une réalisation d'une population de MC est estimé par :

$$\widehat{COV}_{tot} = \frac{\sqrt{\widehat{V}_{tot}}}{\widehat{P}_f^t} \quad (2.25)$$

Son intervalle de confiance de niveau $1 - \alpha$ est alors estimé par :

$$\begin{aligned} \widehat{COV}_{tot}^{inf} &= \frac{\sqrt{\widehat{V}_{tot}^{inf}}}{\widehat{P}_f^t} \\ \widehat{COV}_{tot}^{sup} &= \frac{\sqrt{\widehat{V}_{tot}^{sup}}}{\widehat{P}_f^t} \end{aligned} \quad (2.26)$$

Notons enfin que la contribution jointe $V_{\mathcal{G}_n, \tilde{\mathbf{X}}}$ de l'intégration par **MC** et de l'incertitude de l'approximation par un **GP** peut être estimée en évaluant les trois estimateurs précédents et en utilisant l'expression de la décomposition de la variance Eq. (2.3) :

$$V_{\mathcal{G}_n, \tilde{\mathbf{X}}} = Var_{\mathcal{G}_n, \tilde{\mathbf{X}}} \left(\widehat{P}_f(\tilde{\mathbf{X}}, \mathcal{G}_n) \right) - V_{\tilde{\mathbf{X}}} - V_{\mathcal{G}_n} \quad (2.27)$$

2.2.2 Motivation pour le développement d'une nouvelle approche illustrée sur un cas test

Dans cette section, l'idée est d'expliquer les raisons qui motivent la proposition d'une nouvelle approche permettant un compromis entre l'amélioration du **GP** et l'ajout d'échantillons au **MC** en s'appuyant sur un cas test.

L'application porte sur l'exemple du système en séries à 4 branches, également étudié dans les travaux de [ECHARD et al., 2011; SCHUEREMANS et VAN GEMERT, 2005]. Pour cet exemple, on considère deux variables aléatoires X_1 et X_2 suivant des lois normales standard et une fonction de performance donnée par :

$$y = \min_{x_1, x_2} \left\{ \begin{array}{l} 3 + 0.1(x_1 + x_2)^2 - \frac{(x_1 + x_2)}{\sqrt{(2)}}; \\ 3 + 0.1(x_1 + x_2)^2 + \frac{(x_1 + x_2)}{\sqrt{(2)}}; \\ (x_1 - x_2) + \frac{6}{\sqrt{2}}; \\ (x_2 - x_1) + \frac{6}{\sqrt{2}} \end{array} \right\} \quad (2.28)$$

L'algorithme **AK-MCS+EFF** a été appliqué à cet exemple avec une population de **MC** initiale de taille 10^4 et un **COV** maximum autorisé de 5%. A chaque itération de l'algorithme, les variances imputées au **GP** \mathcal{G}_n et à l'intégration par une méthode de **MC** ont été estimées en se servant respectivement des Eq. (2.14) et Eq. (2.9). L'évolution de la probabilité de défaillance et des variances estimées correspondantes à chaque itération de l'algorithme sont respectivement représentées sur les Figure 2.1 et Figure 2.2.

Sur la Figure 2.2, on peut observer que la variabilité de la probabilité de défaillance est due majoritairement au **GP** sur les 18 premières itérations. Puis, jusqu'à la fin

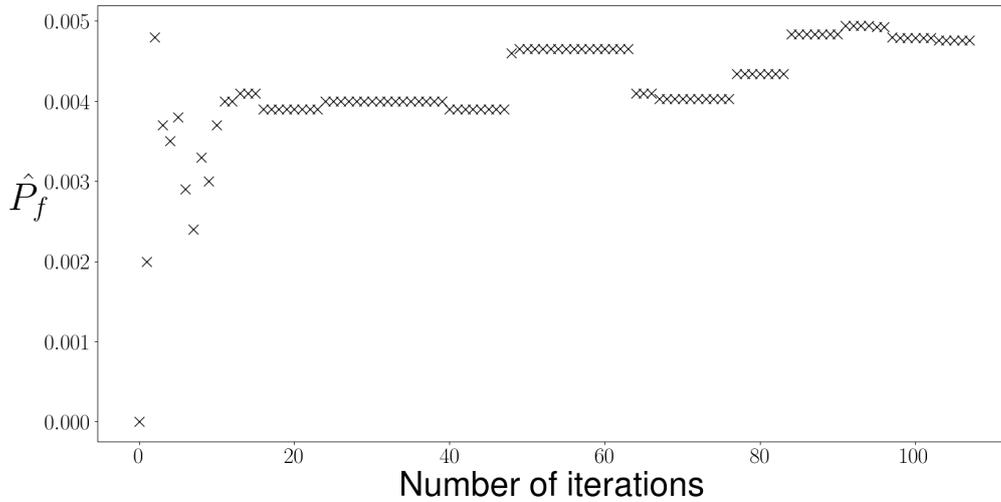


FIGURE 2.1 – Évolution de l'estimation de la probabilité de défaillance en fonction du nombre d'itérations au cours d'une exécution de l'algorithme $AK-MCS+EFF$ ($\hat{P}_f^{MC} = 4.46 \times 10^{-3}$ (1.6%))

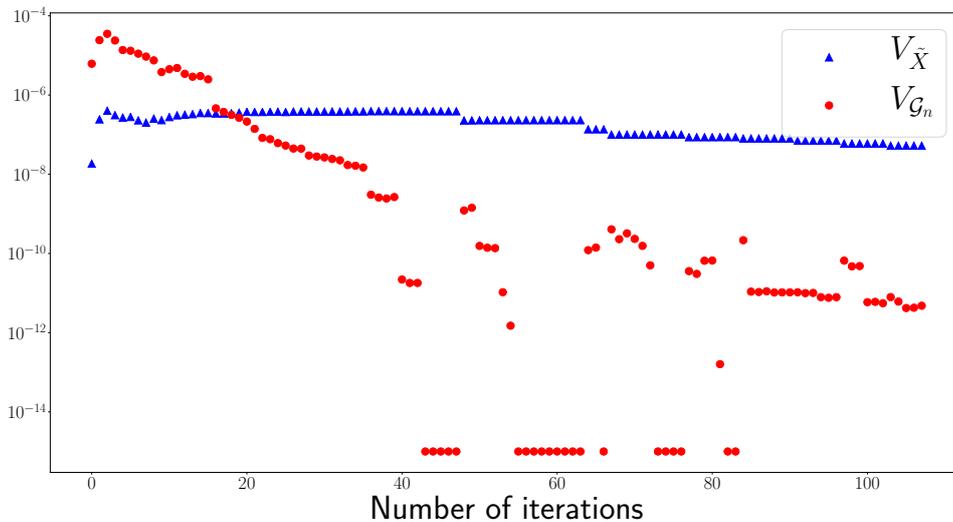


FIGURE 2.2 – Évolution des estimations des variances $V_{\tilde{X}}$ et $V_{\mathcal{G}_n}$ en fonction du nombre d'itérations au cours d'une exécution de l'algorithme $AK-MCS+EFF$

de l'algorithme la source principale de variabilité est l'intégration par MCS . On peut observer sur la Figure 2.1, que l'arrêt de la dispersion des valeurs prises par l'estimation \hat{P}_f est observé approximativement à la 15ème itération de l'algorithme puis que les estimations évoluent autour de la valeur de la bonne valeur de la probabilité de défaillance. Cependant, le GP est encore enrichi après la 18ème itération et à la fin de l'algorithme 99 points ont été ajoutés au total pour apprendre le GP . A la fin de l'algorithme, la part de la variabilité de \hat{P}_f due à l'intégration par MCS est 1.1×10^4

plus grande que celle attribuée au GP.

Cette comparaison entre les variances nous amène à conclure qu'il n'est pas nécessaire d'apprendre de manière aussi précise le GP et qu'il est possible de s'affranchir de quelques simulations qui ne sont pas nécessaires. Il semble donc intéressant d'intégrer ces mesures de variance au sein de la procédure d'apprentissage pour palier à une précision excessive de l'apprentissage du GP. La méthode proposée est détaillée dans la section suivante.

2.3 Approche basée sur les mesures de sensibilité aux incertitudes

2.3.1 Principe général

La nouvelle méthode proposée consiste à utiliser les estimateurs de variance obtenus dans la section précédente à la fois comme critère de décision entre l'enrichissement de la population de MC et l'amélioration du GP et comme critère d'arrêt dans un algorithme d'apprentissage actif.

En effet, les parts de variabilité de \hat{P}_f attribuées à l'intégration par MCS et au GP peuvent être utilisées pour choisir d'améliorer le GP en l'enrichissant ou d'augmenter la taille de la population de MC. Ces contributions à la variance peuvent être quantifiées en s'appuyant sur les estimateurs de variances définis par l'Eq. (2.9) pour la contribution de l'intégration par MC et par l'Eq. (2.14) pour la contribution de l'incertitude de l'approximation par GP.

D'autre part, la variance totale de la variable \hat{P}_f , dont l'estimateur est donnée par l'Eq. (2.21), donne une indication de l'erreur sur l'estimation qui peut être utilisée pour stopper l'algorithme lorsque la précision est suffisante.

2.3.2 Algorithme proposé

La procédure *Apprentissage actif de GP basé sur la variance*, ou *Variance based Active GP learning* (Vb-AGP) proposée est résumée sur la Figure 2.3 et ses différentes étapes sont décrites ci-dessous :

1. Génération d'une population de MC initiale $\tilde{\mathbf{x}}$ de taille n_{MC} .
2. Définition d'un DoE initial de n échantillons en s'appuyant sur une méthode d'Échantillonnage par hypercube latin, ou *Latin Hypercube Sampling* (LHS). La fonction de performance $G(\mathbf{x})$ est évaluée aux n points du DoE.
3. Construction d'un modèle de GP $\mathcal{G}_n(\mathbf{x})$ de la fonction de performance $G(\mathbf{x})$ conditionné aux points du DoE.

4. Estimation de la probabilité de défaillance P_f sur la population de MC $\tilde{\mathbf{x}}$ par :

$$\hat{P}_f^{MC}(\tilde{\mathbf{X}}) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mu_n(\mathbf{x}_i) \leq 0} \quad (2.29)$$

5. Estimation des intervalles de confiance des variances $V_{\tilde{\mathbf{X}}}$ et $V_{\mathcal{G}_n}$ en utilisant respectivement les Eq. (2.9) et Eq. (2.14).

Afin de pouvoir comparer les valeurs des variances, on cherche à obtenir à partir des estimateurs obtenus dans la Section 2.2.1 des intervalles de confiance vérifiant que :

$$\left[\widehat{V}_{\tilde{\mathbf{X}}}^{inf}, \widehat{V}_{\tilde{\mathbf{X}}}^{sup} \right] \cap \left[\widehat{V}_{\mathcal{G}_n}^{inf}, \widehat{V}_{\mathcal{G}_n}^{sup} \right] = \emptyset \quad (2.30)$$

Ainsi, tant que la condition de l'Eq. (2.30) n'est pas vérifiée, des nouvelles réalisations du GP \mathcal{G}_n sont simulées afin d'affiner l'intervalle de confiance de $V_{\mathcal{G}_n}$.

6. Calcul de $\widehat{COV}_{red} = \frac{\sqrt{\widehat{V}_{\mathcal{G}_n}^{sup} + \widehat{V}_{\tilde{\mathbf{X}}}^{sup}}}{\hat{P}_f^{MC}}$.

Si $\widehat{COV}_{red} < COV_{max}$, avec COV_{max} la valeur maximale autorisée par l'utilisateur du COV total, l'algorithme passe à l'étape 7 pour vérifier que le COV total ne dépasse pas la valeur maximale autorisée.

Sinon, l'algorithme passe à l'étape 8 afin de réduire la source d'incertitude majeure.

7. Estimation de l'intervalle de confiance du COV total COV_{tot} à partir de l'Eq. (2.26) : simulations de nouvelles réalisations différentes de population de MC $\tilde{\mathbf{X}}_i$ et du GP G_i tant que $COV_{max} \in \left[\widehat{COV}_{tot}^{inf}, \widehat{COV}_{tot}^{sup} \right]$.

Si $COV_{tot}^{sup} \leq COV_{max}$ l'estimation de \hat{P}_f par \hat{P}_f^t , dont l'expression est rappelée par l'Eq. (2.31), est suffisamment précise et l'algorithme est stoppé.

$$\hat{P}_f^t = \frac{1}{n_t} \sum_{i=1}^{n_t} \hat{P}_f(G_i, \tilde{\mathbf{X}}_i) \quad (2.31)$$

avec n_t le nombre final de simulations réalisées.

Sinon, l'algorithme passe à l'étape suivante.

8. Si $\widehat{V}_{\mathcal{G}_n} < \widehat{V}_{\tilde{\mathbf{X}}}$, de nouveaux échantillons sont ajoutés à la population de MC et l'algorithme retourne à l'étape 4 pour mettre à jour l'estimation de P_f .

Sinon si $\widehat{V}_{\mathcal{G}_n} > \widehat{V}_{\tilde{\mathbf{X}}}$, l'algorithme passe à l'étape 9.

9. La fonction d'apprentissage $EFF(\mathbf{x})$ définie à l'Eq. (1.116) est évaluée sur la population de **MC** pour déterminer le meilleur candidat \mathbf{x}^* à évaluer pour enrichir le modèle de **GP**. La fonction de performance est calculée à l'échantillon \mathbf{x}^* et le **DoE** est enrichi avec le nouveau point \mathbf{x}^* et l'algorithme retourne à l'étape 3.

Suite à cette définition de l'algorithme, faisons maintenant quelques remarques.

L'arrêt de l'algorithme est conditionné à la vérification du critère suivant :

$$\widehat{COV}_{red} = \frac{\sqrt{\widehat{V}_{G_n}^{sup} + \widehat{V}_X^{sup}}}{\widehat{P}_f^{MC}} < COV_{max} \quad (2.32)$$

où COV_{max} est la valeur maximale autorisée par l'utilisateur du **COV** total.

Le critère de l'Eq. (2.32) consiste à vérifier dans un premier temps que la somme des deux contributions ne dépasse pas la valeur maximale admissible en faisant abstraction de leur contribution jointe, qui n'est jamais évaluée au cours de l'apprentissage. Elle permet donc de déterminer, sans coût numérique supplémentaire, si le calcul de la variance totale pour choisir de stopper ou non l'apprentissage est nécessaire. Ainsi, si l'Eq. (2.32) est vérifiée, la variance totale V_{tot} (et par la suite le **COV** total COV_{tot}), comprenant la contribution jointe, peut être estimée pour vérifier qu'elle respecte la valeur maximale de la variance acceptée. Si cette condition est vérifiée, il faut ensuite vérifier que le **COV** total ne dépasse pas la valeur maximale autorisée.

Notons que le critère d'arrêt de l'algorithme **Vb-AGP** dépend d'une valeur limite du **COV** comme pour les méthodes **AK**. Or, le **COV** calculé dans l'algorithme proposé prend en compte l'ensemble des incertitudes, plutôt que seule celle liée au **MC**, et est construit de manière à équilibrer la contribution du **GP** et du **MC** à la variance de l'estimation. Ceci est à prendre en compte lors du choix de la valeur de COV_{max} .

Pour la mise en place de l'algorithme, nous avons fait le choix d'utiliser la fonction $EFF(\mathbf{x})$ pour l'apprentissage du **GP**. En effet, comme expliqué dans la Section 1.2.4.2.1 du chapitre 1, le critère $EFF(\mathbf{x})$ assure un meilleur compromis entre l'exploration et l'exploitation durant l'apprentissage et permet ainsi d'identifier plus rapidement les différentes zones de défaillances dans le cas où la défaillance est multimodale. Ce choix permet d'éviter d'avoir une prédiction du **GP** très précise au voisinage d'une des zones de défaillances alors que d'autres zones n'auraient pas encore été détectées et où l'estimation de la probabilité de défaillance pourrait alors ne différer que très peu pour différentes trajectoires du **GP**, ce qui pourrait mener à une erreur d'estimation par l'algorithme. Plusieurs tests ont en effet été effectués avec la fonction d'apprentissage U et ont conduit à de moins bons résultats.

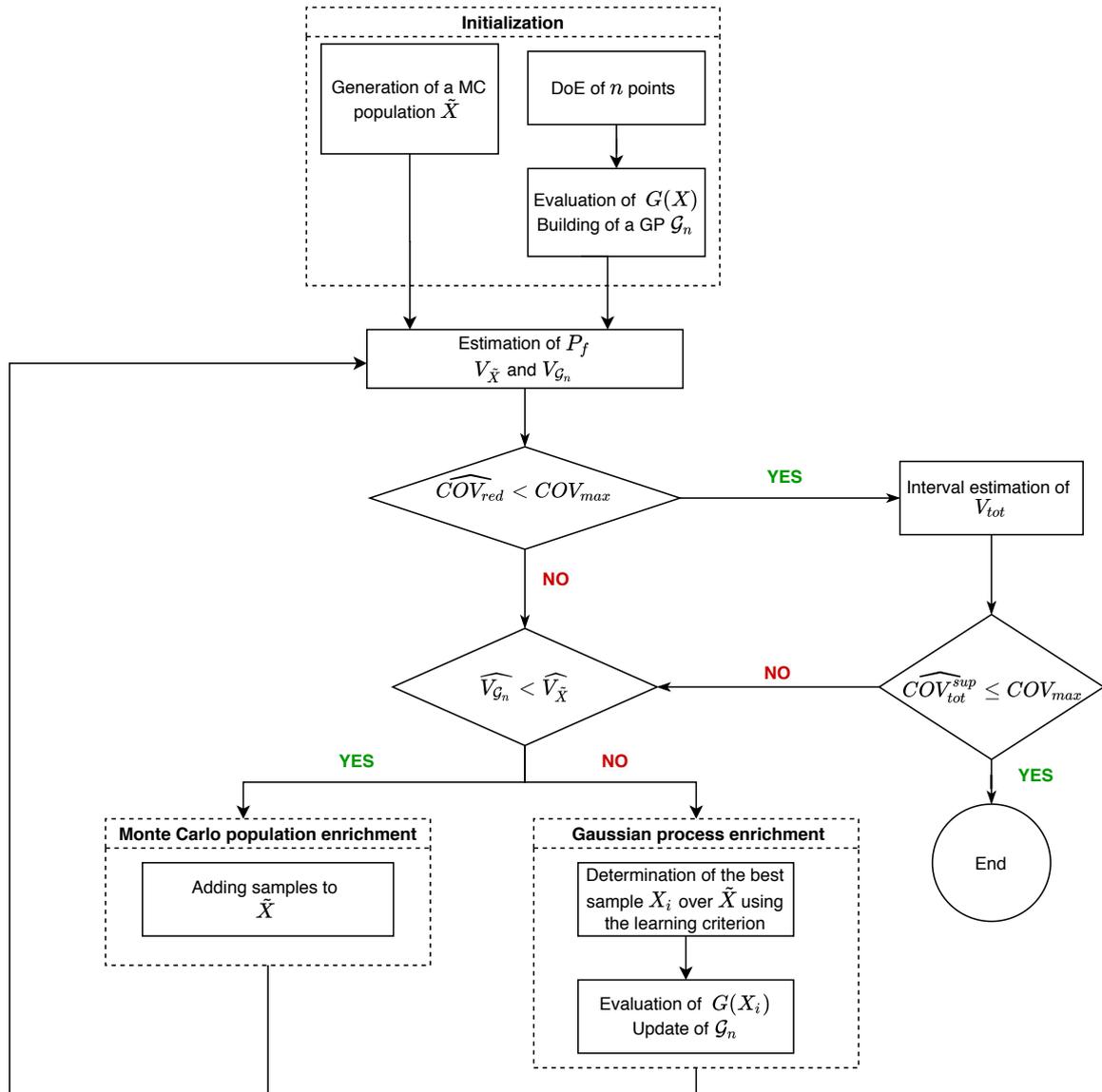


FIGURE 2.3 – Diagramme de l'algorithme d'apprentissage actif Vb-AGP + MCS

2.4 Applications

2.4.1 Méthodologie et mesures de comparaison

Mise en place pratique de l'algorithme

Dans la Section 1.1.2.5 du chapitre 1, les problèmes numériques rencontrés par les GP construits à l'aide de noyaux gaussiens ont été soulignés. En effet, les matrices de covariances construites à partir de ce noyau sont particulièrement mal conditionnées pour des points d'apprentissage fortement corrélés, ce qui est le cas lorsque ces points sont sélectionnés sur la base d'un critère d'apprentissage pour les analyses de fiabilité par apprentissage actif d'un GP. Pour cette raison, le noyau Matérn $5/2$, défini par l'Eq. (1.15), sera utilisé dans ces travaux.

L'estimateur de la variance $V_{\mathcal{G}_n}$ donné par l'Eq. (2.14) est obtenu en simulant des réalisations G_i du GP conditionné \mathcal{G}_n . Or en pratique, la génération de nombreuses réalisations G_i sur des grandes populations de MC pose des problèmes numériques liés au conditionnement et à la mémoire expliqués dans la Section 1.1.2.6 du chapitre 1.

Le recours à une méthode alternative, présentée dans les travaux de [LE GRATIET et al., 2014; VILLEMONTAIX et al., 2009], basée sur la simulation d'un GP non conditionné combinée à l'utilisation d'une représentation de représentation de Karhunen-Loève (KL) de ce GP permet de réduire ces problèmes numériques. La simulation de l'expansion de KL permet également d'être numériquement plus efficace. L'expansion de KL est alors obtenue en utilisant la méthode EOLE décrite dans la Section 1.1.2.7, un cas particulier de la méthode de Nyström, qui est particulièrement efficace pour obtenir des approximations de GP [BETZ et al., 2014]. La méthode EOLE correspond en fait au cas particulier où l'intégration sur laquelle repose la méthode de Nyström est effectuée sur des échantillons distribués uniformément. L'utilisation d'échantillons générés aléatoirement selon une distribution uniforme sur le domaine d'étude permet alors de gérer le nombre n_q de points d'intégration dépendamment de la dimension par rapport à l'utilisation d'une grille uniformément structurée, où le nombre de points est de $n_l^{n_u}$ (n_l le nombre de niveaux de la grille et n_u la dimension du domaine d'étude). Du point de vue de la gestion de la mémoire, cette méthode est donc plus intéressante lorsque le nombre d'entrées n_u augmente.

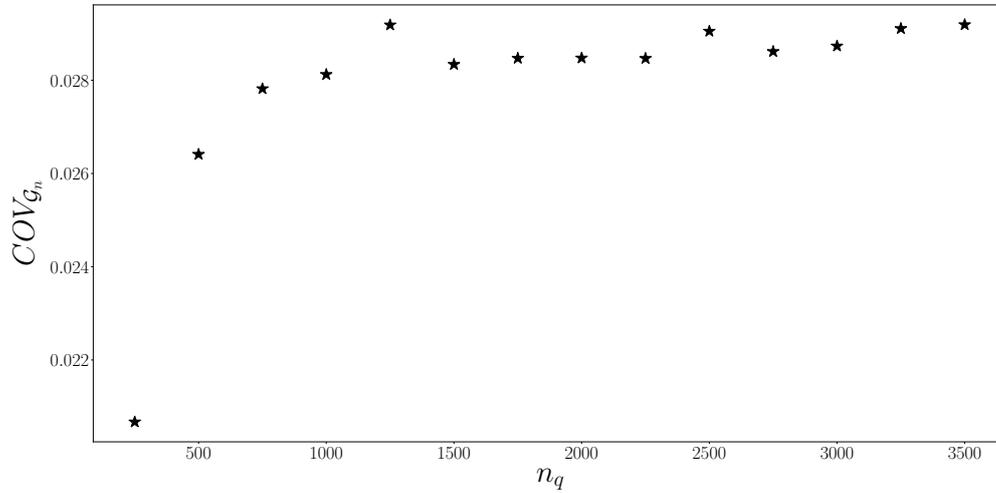
Finalement, pour un nombre de points d'intégration fixé n_q , la représentation de KL est approchée par une expansion $\hat{\mathcal{G}}$ tronquée à $M \leq n_q$ termes dont on rappelle l'expression ci-dessous :

$$\begin{aligned}\hat{\mathcal{G}}(\mathbf{x}, \omega) &= \mu(\mathbf{x}) + \sum_{i=1}^M \sqrt{\hat{\lambda}_i} \hat{\phi}_i(\mathbf{x}) \xi_i(\omega) \\ &= \mu(\mathbf{x}) + \sum_{i=1}^M \frac{1}{\sqrt{\hat{\lambda}_i}} \sum_{j=1}^{n_q} k(\mathbf{x}, \mathbf{x}_j) w_j y_{i,j} \xi_i(\omega)\end{aligned}\tag{2.33}$$

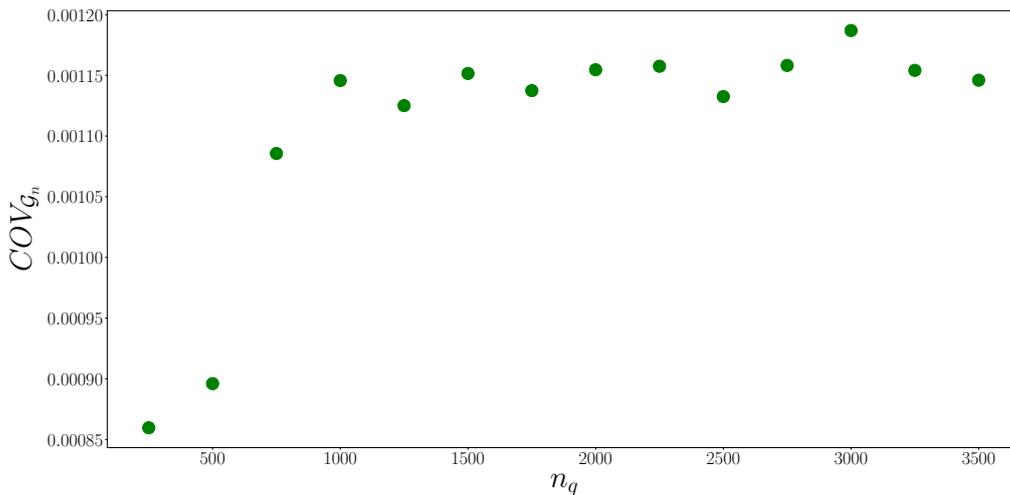
où les $\xi_i(\omega) : \Omega_P \rightarrow \mathbb{R}$ sont des variables aléatoires normales centrées réduites indépendantes, les \mathbf{x}_j sont les n_q points d'intégration. $\lambda_i \geq 0$, $\hat{\phi}_i(\mathbf{x})$, les valeurs propres et les fonctions propres associées au noyau k , w_j les poids d'intégration et y_i un vecteur de taille n_q tel que $y_{i,j} = \hat{\phi}_i(\mathbf{x}_j)$ déterminés suivant la méthode EOLE selon la procédure expliquée dans la Section 1.1.2.7.

Il est important d'avoir une approximation précise de la représentation du GP car l'estimateur de la variance $V_{\mathcal{G}_n}$ est très sensible aux trajectoires simulées. La Figure 2.4 représente l'évolution de l'estimation de la variance $V_{\mathcal{G}_n}$ en fonction du nombre n_q de points d'intégration utilisés pour la méthode EOLE sur l'exemple du cas 4 branches d'une part à l'itération $i = 50$ de l'apprentissage et à la dernière itération de l'algo-

rithme. Pour ces estimations, on a choisi $M = n_q$.



(a) Itération $i = 50$



(b) Dernière itération

FIGURE 2.4 – Estimation de $COV_{G_n} = \frac{\sqrt{V_{G_n}}}{\hat{P}_f}$ pour un nombre de points d'intégration n_q croissant

On observe sur cette figure qu'un nombre insuffisant de points d'intégration, c.-à-d. une précision trop approximative de l'expansion de KL, mène à une mauvaise estimation de la variance V_{G_n} mais qu'à partir d'un certain nombre de points l'estimation de la variance se stabilise.

Il faut alors noter que l'obtention d'une représentation de KL précise nécessite un nombre important de points d'intégrations en grande dimension et peut alors représenter un budget non négligeable en terme de mémoire.

Ainsi, pour un nombre de points d'intégration fixé n_q , la représentation de KL est

approchée par une expansion tronquée à M termes dont l'expression est donnée par l'Eq. (2.33) et telle que M vérifie :

$$1 - \frac{\sum_{i=1}^M \lambda_i}{\sum_{i=1}^{n_q} \lambda_i} < \epsilon_{tronc} \quad (2.34)$$

avec λ_i les valeurs propres et $\phi_i(\mathbf{x})$ les fonctions propres associées rangées par ordre décroissant et ϵ_{tronc} un seuil à fixer. La valeur de ϵ_{tronc} doit être relativement faible pour avoir un nombre de modes suffisamment représentatifs. Pour les applications présentées ensuite, nous avons fixé $\epsilon_{tronc} = 10^{-8}$.

En pratique, on cherche à évaluer la représentation de **KL** du **GP** sur une population de **MC** $\tilde{\mathbf{X}}$ de n_{MC} points. L'Equation (2.33) devient alors le système matriciel suivant :

$$\hat{\mathcal{G}}(\tilde{\mathbf{X}}, \omega) = \boldsymbol{\mu}(\tilde{\mathbf{X}}) + \mathbf{C}(\tilde{\mathbf{X}})\mathbf{W}\mathbf{Y}\boldsymbol{\xi}(\omega) \quad (2.35)$$

Afin de construire de manière numériquement efficace la matrice de covariance $\mathbf{C}(\tilde{\mathbf{X}})$ de taille $M \times n_{MC}$ entre les points du **MC** et les points d'intégration la demande en mémoire est importante. Ainsi, il faut noter que la méthode est limitée d'un point de vue mémoire pour l'évaluation efficace des trajectoires sur une grande population de **MC** lorsque la dimension stochastique augmente et nécessite un nombre de points d'intégration plus important pour la précision.

Mesures de comparaison

La comparaison des performances des différentes méthodes d'apprentissage actif peut être effectuée en se basant sur différents critères ou mesures d'erreur :

- $COV(\hat{P}_f)$: le **COV** des estimations \hat{P}_f obtenues sur un nombre n_{run} de répétitions de l'ensemble de la procédure d'estimation de \hat{P}_f ;
- e_r : l'erreur relative absolue moyenne entre les n_{run} estimations $\hat{P}_{f_i}, i = 1, \dots, n_{run}$ obtenues et une valeur de référence $P_{f_{ref}}$ (obtenue par exemple par **MCS** avec un grand nombre d'échantillons)

$$e_r = \frac{1}{n_{run}} \sum_{i=1}^{n_{run}} \frac{|\hat{P}_{f_i} - P_{f_{ref}}|}{P_{f_{ref}}}; \quad (2.36)$$

- ν_{MC} : un coefficient permettant de comparer l'efficacité numérique de la méthode à une estimation classique par **MCS**, qui est défini comme suit :

$$\nu_{MC} = \frac{N_{sim}^{MC}}{N_{sim}} \quad (2.37)$$

où N_{sim} correspond au nombre d'appels de la méthode d'apprentissage actif à la

fonction de performance et N_{sim}^{MC} est le nombre d'échantillons de MC nécessaires pour obtenir le même coefficient de variation $COV(\hat{P}_f)$ sur l'estimation de la probabilité de défaillance.

L'efficacité ν_{MC} correspond en fait au facteur par lequel le budget de MCS est divisé pour un algorithme donné avec le même niveau de précision.

La méthode Vb-AGP proposée va ensuite être appliquée à deux cas tests pour évaluer ses performances et le gain potentiel atteint par rapport à la méthode classique AK-MCS.

2.4.2 Système en série à 4 branches

La méthode AK-MCS et la méthode Vb-AGP proposée dans ces travaux ont toutes deux été appliquées sur l'exemple du système en série à 4 branches introduits dans la Section 2.2.2.

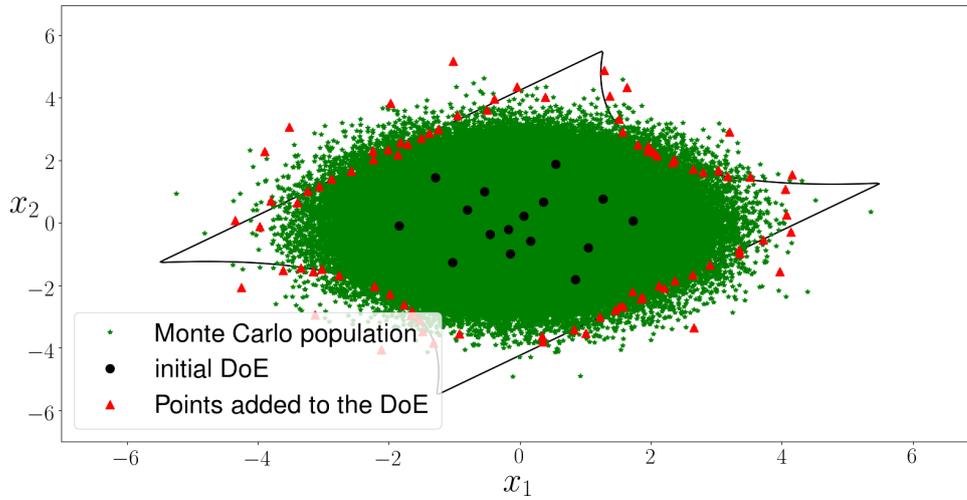
La méthode proposée a été appliquée pour un DoE initial de 16 points, une population de MC initiale de $n_{MC}^{init} = 5 \times 10^4$ échantillons et une valeur du COV_{max} limite de 3%. A chaque enrichissement de la population de MC, n_{MC}^{init} échantillons supplémentaires sont ajoutés à la population.

Les DoE finaux résultant de AK-MCS+EFF et de la méthode Vb-AGP sur les mêmes DoE initiaux et populations de MC sont représentés respectivement sur la Figure 2.5a et la Figure 2.5b. Sur ces Figures, on peut voir que l'algorithme proposé ajoute moins de points d'apprentissage (triangles rouges sur les Figures) pour atteindre le critère d'arrêt d'apprentissage.

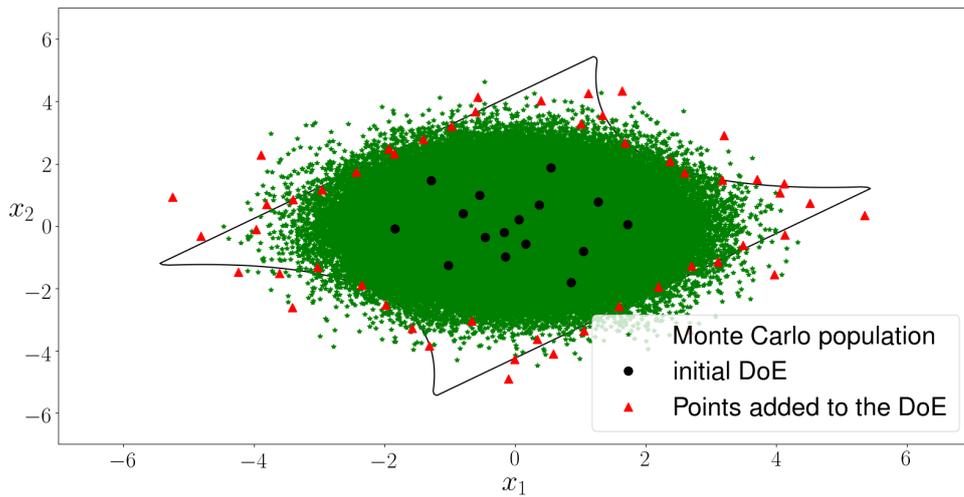
L'évolution des deux variances $V_{\tilde{\mathbf{X}}}$ et V_{G_n} estimées au cours d'une exécution de la méthode proposée est donnée sur la Figure 2.6. On observe sur cette Figure que la variance liée au MC $\tilde{\mathbf{X}}$ subit des chutes importantes. Cela correspond à un ajout d'échantillons dans la population de MC effectué car la valeur de $V_{\tilde{\mathbf{X}}}$ est devenue plus importante que celle de V_{G_n} à l'itération précédent la chute de la variabilité due à $\tilde{\mathbf{X}}$.

Comme pour l'algorithme AK-MCS, l'influence du GP est prédominante au début de l'algorithme et le métamodèle doit être amélioré. Cependant, on peut observer qu'à la fin de l'exécution les valeurs estimées de $V_{\tilde{\mathbf{X}}}$ et V_{G_n} sont plus équilibrées et la valeur de V_{G_n} n'est pas beaucoup plus faible que celle de $V_{\tilde{\mathbf{X}}}$. La valeur estimée du COV total de l'estimation de P_f à la fin de l'algorithme est de 2.9% sur cet exemple et respecte donc la valeur limite de 3%. Ici, la quantité $\frac{\sqrt{V_{\tilde{\mathbf{X}}}}}{P_f}$ correspondante au COV du MC, généralement utilisée comme mesure de la variabilité de l'estimation, vaut 2%.

L'algorithme a ensuite été lancé à 100 reprises pour différents DoE initiaux de 16 points et différentes populations de MC initiales de taille 2×10^4 et un COV total maximal autorisé de 3%.



(a) Résultat de AK-MCS+EFF



(b) Résultat de l'approche Vb-AGP + MCS

FIGURE 2.5 – Comparaison de deux DoE : un résultat d'une exécution de la méthode AK-MCS+EFF et un autre obtenu par la méthode proposée pour les mêmes DoE et population de MC initiaux avec $COV_{max} = 3\%$

L'histogramme des estimations \hat{P}_f de P_f obtenues à partir de l'algorithme Vb-AGP + MCS est donné sur la Figure 2.7. De plus, la probabilité de référence $P_{f_{ref}}$ obtenue à partir des 100 répétitions de MCS avec $n_{MC} = 10^6$ simulations et son intervalle de confiance à $\pm 3\sigma$ (où σ est son écart-type) est représentée sur la figure.

Les résultats de référence obtenus sur une moyenne de 100 répétitions de MCS ($n_{MC} = 10^6$) et les résultats moyens obtenus sur 100 exécutions de chaque méthode sont présentés dans le Tab. 2.1.

Les résultats montrent que la méthode Vb-AGP permet de réduire le nombre de points d'apprentissage nécessaires pour respecter une même précision. La méthode Vb-

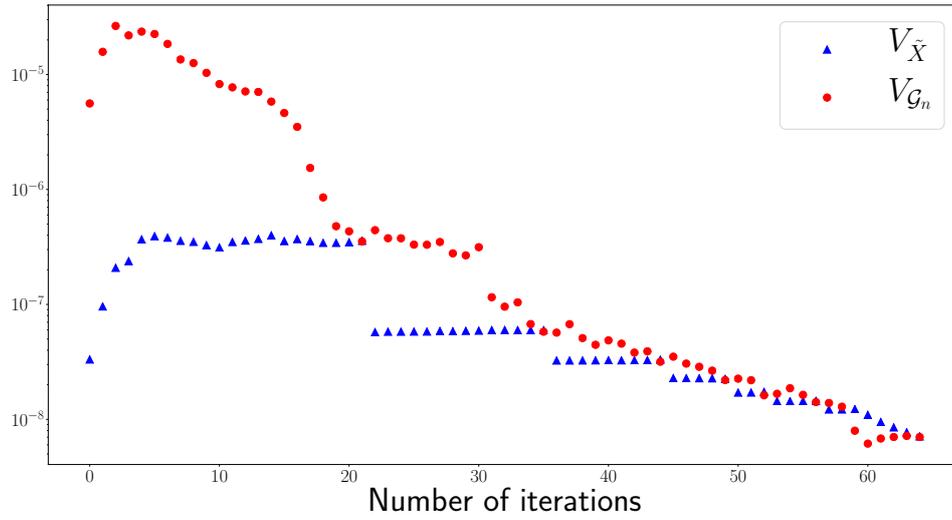


FIGURE 2.6 – Évolution des logarithmes des variances $V_{\tilde{X}}$ et $V_{\tilde{G}_n}$ estimées en fonction du nombre d'itérations au cours de l'algorithme Vb-AGP

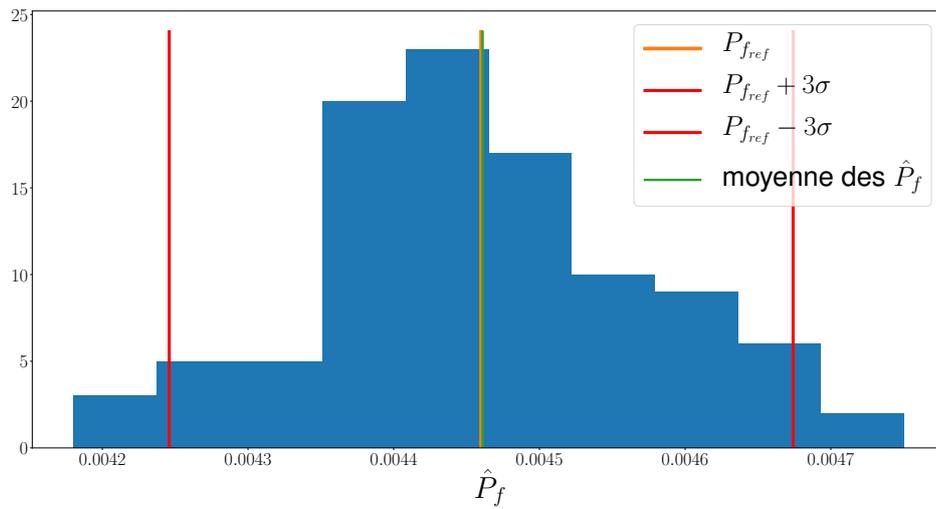


FIGURE 2.7 – Histogramme des estimations \hat{P}_f sur 100 répétitions de Vb-AGP + MCS sur le cas test système en série à 4 branches

Méthode	N_{sim}	$COV(N_{sim})$	\hat{P}_f	$COV(\hat{P}_f)$	e_r	ν_{MC}
MCS	10^6	-	4.46×10^{-3}	1.6%		
AK-MCS + U	128	6.6%	4.48×10^{-3}	3%	2.4%	1976
AK-MCS + EFF	144	6.6%	4.46×10^{-3}	3%	2.5%	1690
Vb-AGP + MCS	68	9.0%	4.46×10^{-3}	2.6%	2.0%	5146

TABEAU 2.1 – Système en série à 4 branches— résultats moyens sur 100 répétitions - paramètres de la méthode adaptative : $n_{MC}^{init} = 5 \times 10^4$, $n_{DoE}^{init} = 16$, $COV_{max} = 0.03$

AGP permet même d'obtenir dans cet exemple un COV plus faible que les autres techniques d'apprentissage actif.

L'indicateur ν_{MC} de la méthode Vb-AGP, indiquant l'efficacité numérique de la méthode montre que celle-ci est 3 fois plus efficace que AK-MCS + EFF et 2.6 fois plus efficace pour que AK-MCS + U.

La méthode a ensuite été appliquée sur d'autres exemples pour valider ces résultats.

2.4.3 Réponse dynamique d'un oscillateur non linéaire

Cet exemple a été utilisé à de nombreuses reprises dans la littérature [ECHARD, 2012; ECHARD et al., 2013; LELIÈVRE et al., 2018] et s'intéresse à la réponse dynamique d'un oscillateur non-linéaire à un degré de liberté et non-amorti illustré sur la Figure 2.8.

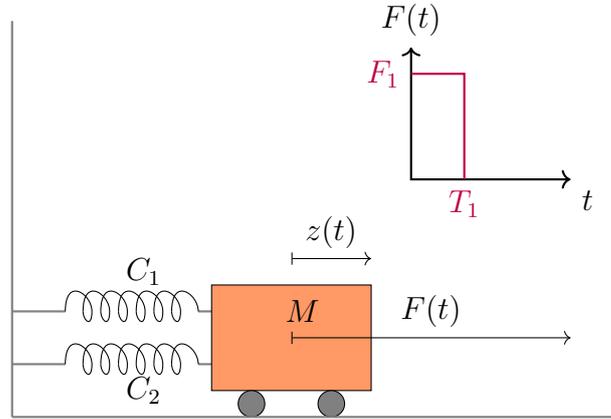


FIGURE 2.8 – Schéma de l'oscillateur non-linéaire à un degré de liberté

L'expression de la fonction de performance étudiée est donnée par :

$$G(C_1, C_2, M, R, T_1, F_1) = 3R - \left| \frac{2F_1}{M\omega_0^2} \sin\left(\frac{\omega_0 T_1}{2}\right) \right| \quad (2.38)$$

avec $\omega_0 = \sqrt{(C_1 + C_2)/M}$.

Ce problème dépend de six variables aléatoires listées dans le Tab. 2.2. Deux cas d'étude sont proposés ici pour deux jeux de paramètres de la distribution de la force F_1 menant à deux valeurs de la probabilité de défaillance d'ordre de grandeur différents.

La méthode proposée Vb-AGP + MCS et la méthode classique AK-MCS ont été appliquées sur le premier cas, pour $F_1 \sim \mathcal{N}(1, 0.04)$, qui correspond à une probabilité de défaillance de $P_{f_{ref}} = 2.86 \times 10^{-2}$ (2%) obtenue pour 100 répétitions de la méthode MCS avec un MC de taille $n_{MC} = 1 \times 10^5$. Les deux méthodes Vb-AGP + MCS et AK-MCS ont chacune été répétées 100 fois pour des DoE initiaux de 12 points et des populations de MC initiale de $n_{MC}^{init} = 10^4$ échantillons en considérant un COV maximum de 3%. De plus, pour chaque enrichissement de la population de MC, $n_{MC}^{init} = 10^4$ échantillons

Variable	Distribution	Moyenne	Écart type
C_1	Gaussienne	1	0.1
C_2	Gaussienne	0.1	0.01
M	Gaussienne	1	0.05
R	Gaussienne	0.5	0.05
T_1	Gaussienne	1	0.2
F_1 – Cas n° 1	Gaussienne	1	0.2
F_1 – Cas n° 2	Gaussienne	0.6	0.1

TABLEAU 2.2 – Oscillateur non-linéaire à un degré de liberté - Distributions des paramètres du problème

supplémentaires sont ajoutés à la population. Les résultats de référence et les résultats moyens sont donnés dans le Tab. 2.3.

Méthode	N_{sim}	$COV(N_{sim})$	\hat{P}_f	$COV(\hat{P}_f)$	e_r	ν_{MC}
MCS	1×10^5	-	2.86×10^{-2}	2%	-	-
AK-MCS + U	59.8	6.4%	2.85×10^{-2}	2.7%	2.2%	792
AK-MCS + EFF	52.5	7.1%	2.87×10^{-2}	2.8%	2.4%	890
Vb-AGP + MCS	22.5	14.4%	2.84×10^{-2}	3.2%	2.6%	1436

TABLEAU 2.3 – Oscillateur non linéaire ($\mu_{F_1} = 1$, $\sigma_{F_1} = 0.2$) — Résultats moyens sur 100 répétitions - Paramètres de l'algorithme : $n_{MC}^{init} = 10^4$, $n_{DoE}^{init} = 12$, $COV_{max} = 0.03$

En terme d'efficacité du coût de calcul par rapport à la précision, la méthode Vb-AGP reste plus avantageuse en comparaison des autres techniques sur cet exemple avec un coefficient ν_{MC} bien plus élevé. En effet, le coefficient ν_{MC} montre que Vb-AGP + MCS est 1.6 fois plus efficace par rapport à AK-MCS + U et 1.8 fois plus efficace par rapport à AK-MCS + EFF .

Dans le Tab. 2.3, l'estimation du COV vaut 3.2% pour la méthode Vb-AGP sur les 100 répétitions effectuées et son intervalle de confiance à 99% est donné par [0.022, 0.041].

De plus, l'histogramme des estimations \hat{P}_f de P_f obtenues à partir de l'algorithme Vb-AGP + MCS ainsi que la probabilité de référence $P_{f_{ref}}$ et son intervalle de confiance à $\pm 3\sigma$ sont représentés sur la Figure 2.9. On observe sur la figure que la majorité des \hat{P}_f estimés se trouvent dans l'intervalle de confiance de la probabilité de référence.

On peut remarquer que le COV de la taille N_{sim} du DoE finale de l'algorithme est plus élevé pour la méthode Vb-AGP que pour les autres méthodes. Une hypothèse expliquant une variabilité plus importante du nombre de simulations est que la qualité du DoE initial (par rapport à la classification) a une influence sur le nombre de points supplémentaires pour atteindre le critère sur l'erreur de l'estimation. Cette idée est renforcée par le fait que le nombre de points ajoutés par l'algorithme, montant au

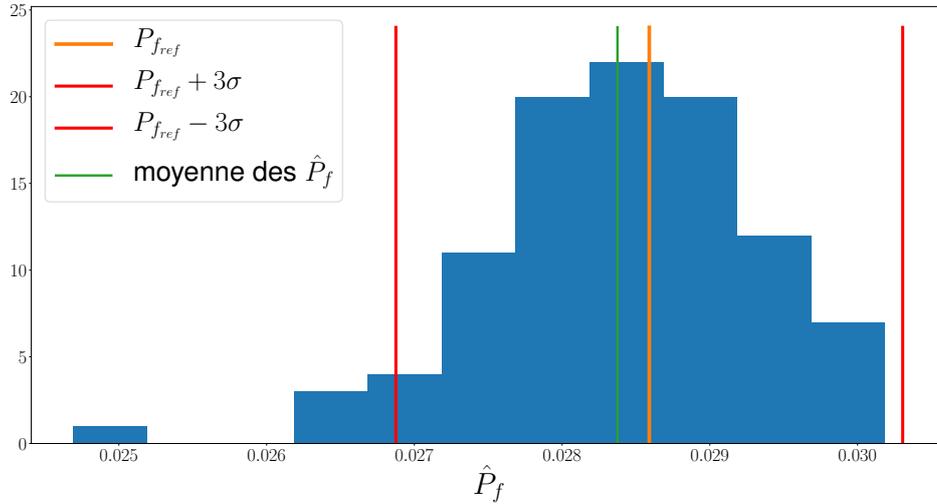


FIGURE 2.9 – Histogramme des estimations \hat{P}_f sur 100 répétitions de Vb-AGP + MCS sur l'exemple de l'oscillateur non linéaire

maximum à 16 points sur les 100 répétitions, n'est pas excessivement élevé par rapport à la taille du DoE initial qui est de 12 points ici.

2.5 Amélioration avec l'Échantillonnage préférentiel, ou *Importance Sampling*

L'utilisation de l'Échantillonnage préférentiel, ou *Importance Sampling* (IS) a pour objectif de réduire fortement la variance d'un estimateur basé sur l'intégration par une méthode de MC. Par conséquent, l'utilisation de l'IS permet de traiter des problèmes d'estimation de probabilités d'événements rares car le nombre d'échantillons nécessaires à l'intégration peut être considérablement réduit.

2.5.1 Estimateur de la probabilité de défaillance dans le cas de l'IS

Comme rappelé dans le chapitre 1, l'estimation de la probabilité de défaillance calculée par IS sur un GP \mathcal{G}_n est donnée par :

$$\hat{P}_f(\tilde{\mathbf{X}}, \mathcal{G}_n) = \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i \mathbb{1}_{\mathcal{G}_n(\mathbf{x}_i) \leq 0}(\mathbf{x}_i) \quad (2.39)$$

avec $w_i = w(\mathbf{x}_i) = \frac{f_{\mathbf{x}}(\mathbf{x}_i)}{f_{aux}(\mathbf{x}_i)}$ les poids des échantillons générés par IS et f_{aux} la densité auxiliaire de l'IS.

Les contributions du GP et de l'intégration par IS à la variance de la quantité d'intérêt peuvent également être estimées. Ces valeurs peuvent être obtenues en réécrivant les indices de variance proposés dans la Section 2.2.1 en les adaptant pour l'IS.

2.5.2 Estimateurs de variance

L'espérance de \hat{P}_f conditionnée à une population de n_{IS} échantillons $\tilde{\mathbf{X}} = (\mathbf{X}_i)_{i=1, \dots, n_{IS}}$ générés par une densité auxiliaire f_{aux} peut être exprimée en écrivant l'expression donnée par l'Eq. (2.7) comme suit :

$$\begin{aligned} \mathbb{E}_{\mathcal{G}_n} [\hat{P}_f | \tilde{\mathbf{X}}] &= \mathbb{E}_{\mathcal{G}_n} \left[\frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i \mathbb{1}_{\mathcal{G}_n(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) | (\mathbf{X}_i)_{i=1, \dots, n_{IS}} \right] \\ &= \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i \mathbb{E}_{\mathcal{G}_n} [\mathcal{B}(p(\mathbf{X}_i))] \\ &= \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i p(\mathbf{X}_i) \end{aligned} \quad (2.40)$$

L'estimateur de la variance $V_{\tilde{\mathbf{X}}}$ est ainsi obtenu en adaptant l'Eq. (2.9) pour l'IS :

$$\begin{aligned} \widehat{V}_{\tilde{\mathbf{X}}} &= \frac{\widehat{Var}_{\tilde{\mathbf{X}}}((\mathbf{w}(\tilde{\mathbf{x}})\mathbf{p}(\tilde{\mathbf{x}}))}{n_{IS}} \\ &= \frac{1}{n_{IS}(n_{IS} - 1)} \sum_{i=1}^{n_{IS}} \left(w(\mathbf{X}_i)p(\mathbf{X}_i) - \frac{1}{n_{IS}} \sum_{j=1}^{n_{IS}} w(\mathbf{X}_j)p(\mathbf{X}_j) \right)^2 \end{aligned} \quad (2.41)$$

et les bornes de son intervalle de confiance d'ordre $1 - \alpha$ peuvent être estimées en utilisant les opérateurs définis dans la Section 2.2.1.1 par :

$$\begin{aligned} \widehat{V}_{\tilde{\mathbf{X}}}^{inf} &= \frac{\widehat{Var}_{\tilde{\mathbf{X}}}^{inf}(\mathbf{w}(\tilde{\mathbf{x}})\mathbf{p}(\tilde{\mathbf{x}}))}{n_{IS}} \\ \widehat{V}_{\tilde{\mathbf{X}}}^{sup} &= \frac{\widehat{Var}_{\tilde{\mathbf{X}}}^{sup}(\mathbf{w}(\tilde{\mathbf{x}})\mathbf{p}(\tilde{\mathbf{x}}))}{n_{IS}} \end{aligned} \quad (2.42)$$

De la même manière que dans la Section 2.2.1 pour MCS, une fois le premier terme $V_{\tilde{\mathbf{X}}}$ estimé, l'idée est de donner une expression du second terme de variance $V_{\mathcal{G}_n}$.

Le calcul de l'espérance de \hat{P}_f conditionnée à une réalisation de \mathcal{G}_n peut être interprété ici comme une estimation classique de probabilité de défaillance par IS pour un modèle déterministe, qui correspond à la réalisation concernée. Cette espérance vérifie

donc l'égalité suivante :

$$\mathbb{E}_{\tilde{\mathbf{X}}} [\hat{P}_f | \mathcal{G}_n] = \mathbb{E}_{\tilde{\mathbf{X}}} \left[\frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i \mathbb{1}_{\mathcal{G}_n(\mathbf{X}_i) \leq 0}(\mathbf{X}_i) | \mathcal{G}_n \right] \quad (2.43)$$

$$= P_f(\mathcal{G}_n) \quad (2.44)$$

La probabilité de défaillance $P_f(\mathcal{G}_n)$ est approchée ici par son estimation par **IS**

$$\hat{P}_f^{IS}(\mathcal{G}_n) = \frac{1}{n_{IS}} \sum_{i=1}^{n_{IS}} w_i \mathbb{1}_{\mathcal{G}_n(\mathbf{x}_i) \leq 0}(\mathbf{x}_i),$$

où $(\mathbf{x}_i)_{i=1, \dots, n_{IS}}$ les points constituant une réalisation de la population d'**IS** $\tilde{\mathbf{X}}$.

Ainsi, de la même façon que pour **MCS**, $V_{\mathcal{G}_n}$ peut être obtenue numériquement en évaluant l'estimateur de l'**IS** de P_f pour différentes trajectoires de \mathcal{G}_n . L'expression de l'estimateur de $V_{\mathcal{G}_n}$ donné par l'Eq. (2.14) pour **MCS** devient :

$$\widehat{V}_{\mathcal{G}_n} = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} \left(\hat{P}_f^{IS}(G_i) - \frac{1}{n_t} \sum_{j=1}^{n_t} \hat{P}_f^{IS}(G_j) \right)^2 \quad (2.45)$$

où n_t est le nombre de réalisations de \mathcal{G}_n .

De plus, les estimations des bornes de son intervalle de confiance d'ordre $1 - \alpha$ sont données par :

$$\begin{aligned} \widehat{V}_{\mathcal{G}_n}^{inf} &= \widehat{Var}_{\mathcal{G}_n}^{inf}(\hat{P}_f^{IS}(\mathcal{G}_n)) \\ \widehat{V}_{\mathcal{G}_n}^{sup} &= \widehat{Var}_{\mathcal{G}_n}^{sup}(\hat{P}_f^{IS}(\mathcal{G}_n)) \end{aligned} \quad (2.46)$$

Finalement, comme pour la méthode **Vb-AGP + MCS**, l'estimation de la probabilité de défaillance \hat{P}_f^t et sa variance totale peuvent être estimées à partir des expressions données respectivement par l'Eq. (2.24) et l'Eq. (2.21).

L'estimation finale de la probabilité de défaillance est alors donnée par :

$$\hat{P}_f^t = \frac{1}{n_t} \sum_{i=1}^{n_t} \hat{P}_f^{IS}(G_i, \tilde{\mathbf{X}}_i) \quad (2.47)$$

où $(G_i, \tilde{\mathbf{X}}_i)_{1 \leq i \leq n_t}$ sont n_t réalisations du **GP** \mathcal{G}_n et de populations d'**IS** $\tilde{\mathbf{X}}$ générées selon la densité auxiliaire f_{aux} et $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i)$ est l'estimation de la probabilité de défaillance obtenue pour la $i^{\text{ème}}$ réalisation $(G_i, \tilde{\mathbf{X}}_i)$ de \mathcal{G}_n et d'une population d'**IS** donnée par :

$$\hat{P}_f^{IS}(G_i, \tilde{\mathbf{X}}_i) = \frac{1}{n_{MC}} \sum_{j=1}^{n_{MC}} w(\mathbf{X}_{ij}) \mathbb{1}_{G_i(\mathbf{X}_{ij}) \leq 0}(\mathbf{X}_{ij}) \quad (2.48)$$

La variance totale de \hat{P}_f peut ainsi être estimée comme suit :

$$\widehat{V}_{tot} = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} \left(\hat{P}_f^{IS}(G_i, \tilde{\mathbf{X}}_i) - \frac{1}{n_t} \sum_{j=1}^{n_t} \hat{P}_f^{IS}(G_j, \tilde{\mathbf{X}}_j) \right)^2 \quad (2.49)$$

L'estimation de l'intervalle de confiance de niveau $1 - \alpha$ de \widehat{V}_{tot} , exprimée en utilisant les opérateurs définis dans la Section 2.2.1.1, est finalement donnée par :

$$\begin{aligned} \widehat{V}_{tot}^{inf} &= \widehat{Var}_{\mathcal{G}_n, \tilde{\mathbf{X}}}^{inf}(\hat{P}_f(\mathcal{G}_n, \tilde{\mathbf{X}})) \\ \widehat{V}_{tot}^{sup} &= \widehat{Var}_{\mathcal{G}_n, \tilde{\mathbf{X}}}^{sup}(\hat{P}_f(\mathcal{G}_n, \tilde{\mathbf{X}})) \end{aligned} \quad (2.50)$$

Puis l'estimation du COV total de l'estimation \hat{P}_f^{IS} de la probabilité de défaillance \hat{P}_f obtenue sur une réalisation d'IS est donnée par :

$$\widehat{COV}_{tot} = \frac{\sqrt{\widehat{V}_{tot}}}{\hat{P}_f^t} \quad (2.51)$$

Son intervalle de confiance de niveau $1 - \alpha$ est alors estimé par :

$$\begin{aligned} \widehat{COV}_{tot}^{inf} &= \frac{\sqrt{\widehat{V}_{tot}^{inf}}}{\hat{P}_f^t} \\ \widehat{COV}_{tot}^{sup} &= \frac{\sqrt{\widehat{V}_{tot}^{sup}}}{\hat{P}_f^t} \end{aligned} \quad (2.52)$$

2.5.3 Méthode étendue à l'échantillonnage préférentiel

Afin de pouvoir traiter des problèmes d'estimation de probabilités faibles, nous proposons d'intégrer l'échantillonnage préférentiel à la méthode Vb-AGP proposée précédemment.

L'idée est de remplacer la population de MC par une population d'IS pour l'estimation de la probabilité de défaillance. Dans ce cas, les variances $V_{\tilde{\mathbf{X}}}$ et $V_{\mathcal{G}_n}$ sont calculées sur les échantillons $\tilde{\mathbf{X}}$ générés par IS et leurs estimations sont obtenues en appliquant les expressions des Eq. (2.41) et Eq. (2.45). Cependant, les échantillons candidats pour l'enrichissement du GP correspondent à l'ensemble des échantillons \mathbf{X}^{aux} générés au cours de l'exécution de l'algorithme d'IS adaptatif effectuée pour la dernière mise à jour de la densité auxiliaire. Ici, l'algorithme d'IS adaptatif utilisé est l'algorithme NAIS, détaillé dans la Section 1.2.3.2.2 du chapitre 1, qui permet notamment de traiter les cas de domaines de défaillances multimodaux.

Au début de l'algorithme, les candidats initiaux pour l'apprentissage sont simplement les échantillons d'une population de MC classique générée selon la distribution $f_{\mathbf{X}}$.

Ensuite les étapes 8 et 9 décrites dans la Section 2.3 pour MCS sont modifiées comme suit :

8. Si $\widehat{V}_{\mathcal{G}_n} < \widehat{V}_{\mathcal{X}}$ alors :

8.1. Si il s'agit du premier passage dans cette boucle ou si \mathcal{G}_n a été mis à jour, alors une nouvelle densité auxiliaire f_{aux} est construite avec le GP \mathcal{G}_n . La population d'IS et les échantillons candidats pour l'apprentissage \mathbf{X}^{aux} sont également remplacés par les plus récemment générés. Puis l'algorithme retourne à l'étape 4.

Sinon, l'algorithme passe à l'étape 8.2.

8.2. Des nouveaux échantillons sont ajoutés à la population d'IS et la méthode retourne à l'étape 4.

Sinon si $\widehat{V}_{\mathcal{G}_n} > \widehat{V}_{\mathcal{X}}$, l'algorithme passe à l'étape 9.

9. La fonction d'apprentissage $EFF(\mathbf{x})$ est évaluée sur l'ensemble de la population de points candidats \mathbf{X}^{aux} pour déterminer le meilleur candidat x^* pour enrichir le métamodèle de GP. La fonction de performance est évaluée au meilleur point x^* et le DoE est enrichi avec cette nouvelle observation. Ensuite la méthode passe à l'étape 3 pour mettre à jour le modèle de GP.

La procédure de l'extension de la méthode Vb-AGP + IS est résumée sur la Figure 2.10.

De plus, pour des problèmes d'estimation de probabilités faibles, la probabilité estimée avec les données initiales pour le DoE et la population de MC a une forte probabilité d'être nulle. L'estimation des variances pour la phase d'enrichissement est alors impossible. Pour résoudre ce problème, l'algorithme NAIS est exécuté une fois en utilisant le GP construit à l'aide du DoE initial. Ceci permet d'une part d'estimer une densité auxiliaire mieux adaptée pour l'estimation de la probabilité de défaillance et également de fournir des candidats plus appropriés pour l'apprentissage du GP.

En effet, dans le cadre de l'estimation de probabilités de défaillance très faibles, pour obtenir un GP initial approprié pour continuer l'apprentissage, un DoE initial adapté devrait être échantillonné dans un certain voisinage de l'état limite. Cependant, le manque d'information concernant le domaine de défaillance à l'initialisation de l'algorithme peut conduire à une estimation de la première densité optimale par NAIS relativement éloignée de la vraie densité optimale. Lorsque l'estimation de la probabilité de défaillance est égale à zéro, un second DoE initial de taille $2n_u$, plus pertinent pour l'apprentissage, est alors généré après que NAIS ait convergé. Les points

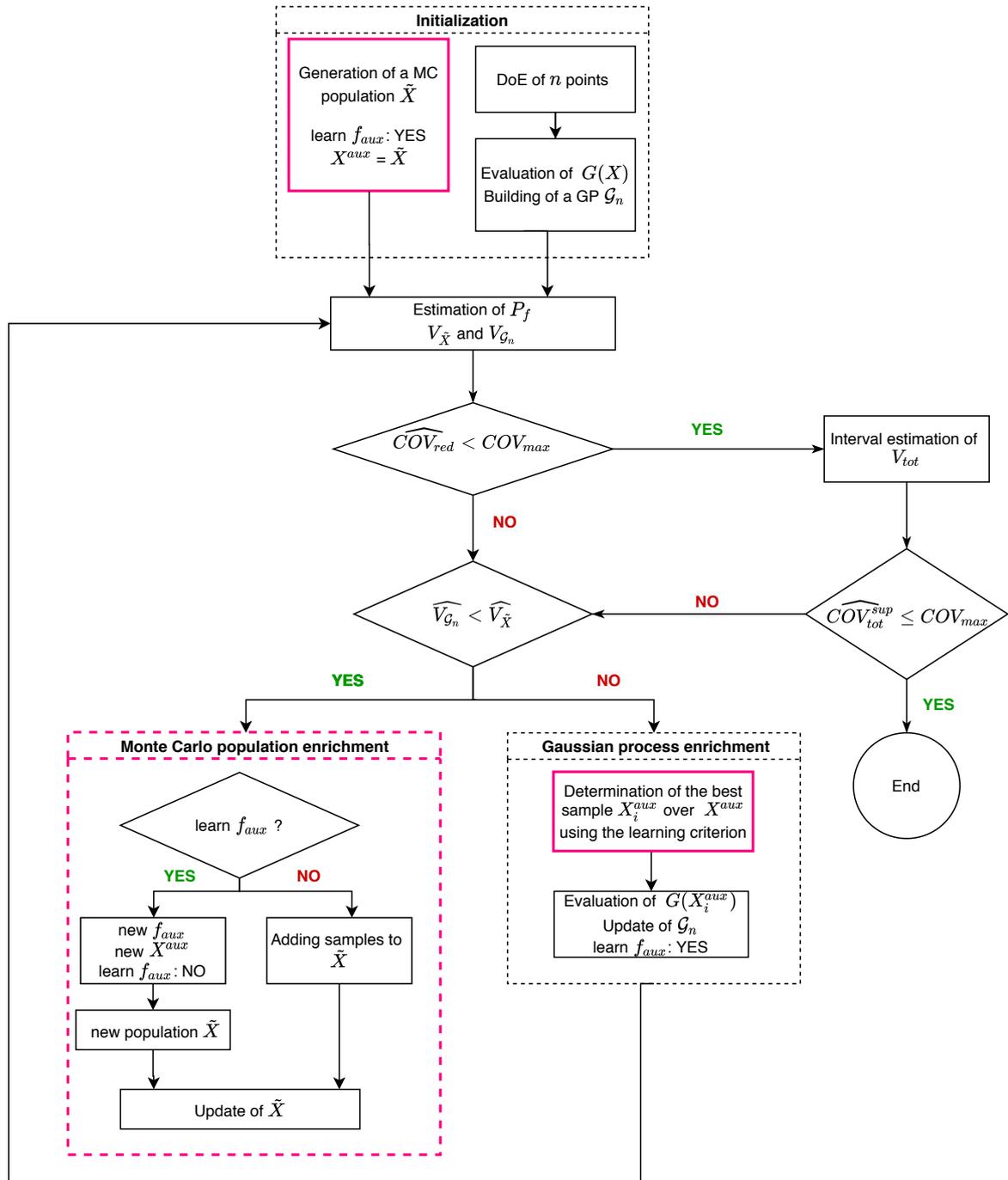


FIGURE 2.10 – Diagramme de l'algorithme d'apprentissage actif Vb-AGP + IS

du DoE sont choisis itérativement parmi l'ensemble des échantillons générés par l'algorithme NAIS à l'aide du critère d'apprentissage EFF, avec une mise à jour du GP après chaque point ajouté au DoE.

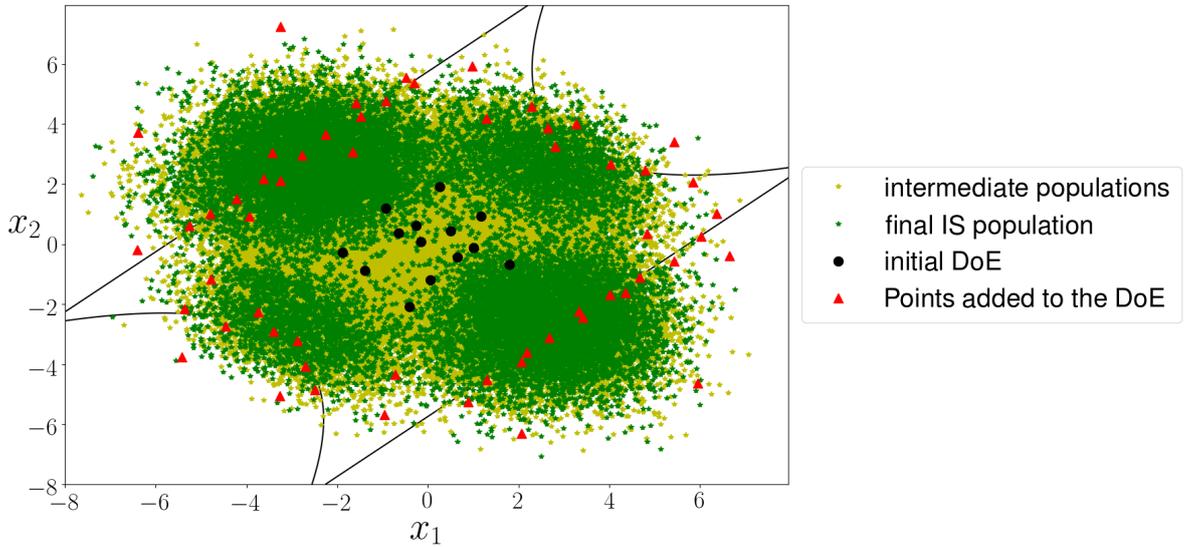


FIGURE 2.11 – Illustration de la méthode **Vb-AGP + IS (NAIS)** sur l'exemple du système série à 4 branches

2.5.4 Applications

2.5.4.1 Système série à 4 branches en probabilité faible

La méthode **Vb-AGP + IS** a été appliquée sur un cas test dérivé de la fonction de performance du système en série à 4 branches défini par l'Eq. (2.28). Ici, la défaillance est définie par $G(\mathbf{x}) \leq 1.5$ et le problème de fiabilité associé correspond à une probabilité de défaillance de 5.29×10^{-5} avec un **COV** de 2.1% estimée par **MCS** (100 exécutions pour $n_{MC} = 5 \times 10^7$).

La méthode basée sur la variance proposée **Vb-AGP + IS** avec la méthode d'**IS** adaptatif **NAIS** appliquée sur ce cas test est illustrée sur la Figure 2.11. La population intermédiaire (points verts clairs), générée au cours de l'algorithme **NAIS** pour la densité auxiliaire finale, utilisée comme échantillons candidats pour l'apprentissage du **GP** et la population d'**IS** (points verts) utilisée pour calculer la probabilité sont représentées sur cette figure.

L'histogramme des estimations \hat{P}_f de P_f obtenues à partir de l'algorithme **Vb-AGP + IS** ainsi que la probabilité de référence $P_{f_{ref}}$ et son intervalle de confiance à $\pm 3\sigma$ sont représentés sur la Figure 2.12.

Les résultats moyens obtenus sur 100 répétitions de l'algorithme sur ce cas test sont donnés dans le Tab. 2.4. Ils montrent que la méthode **Vb-AGP + IS** respecte le **COV** maximal fixé à 3% sur les 100 répétitions de l'algorithme avec un nombre d'appels à la fonction de performance correspondant au nombre de points d'apprentissage nécessaires pour vérifier ce **COV**.

Notons qu'on ne se compare pas à **AK-MCS** ici, car ce type de problème de fiabilité,

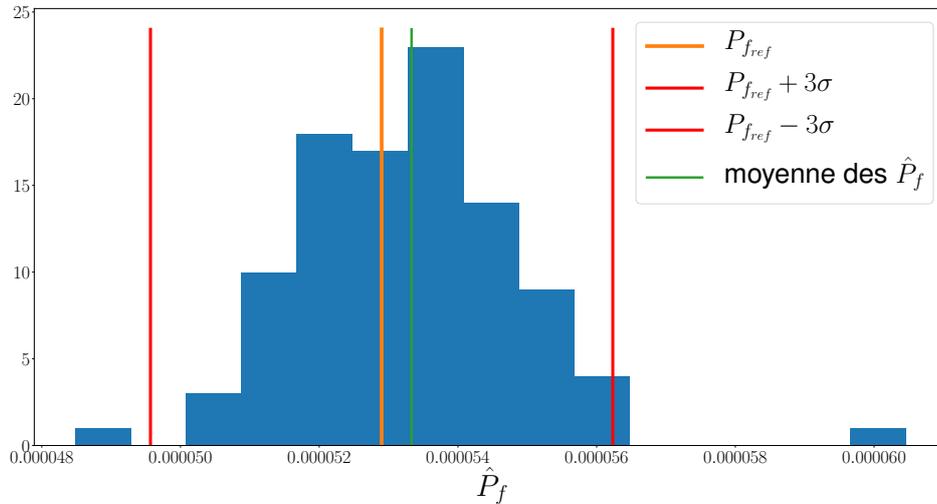


FIGURE 2.12 – Histogramme des estimations \hat{P}_f sur 100 répétitions de **Vb-AGP + IS** sur le cas test système en série à 4 branches - cas probabilité faible

Méthode	N_{sim}	$COV(N_{sim})$	\hat{P}_f	$COV(\hat{P}_f)$	e_r	ν_{MC}
MCS	5×10^7	-	5.29×10^{-5}	2.1%	-	-
Vb-AGP + IS	104	16.2%	5.33×10^{-5}	3.0%	2.4%	201437

TABLEAU 2.4 – Système en série à 4 branches - cas probabilité faible — résultats moyens sur 100 répétitions - paramètres de la méthode adaptative : $n_{MC}^{init} = 1 \times 10^4$, $n_{DoE}^{init} = 12$, $COV_{max} = 0.03$

avec des probabilités de défaillance faibles, est très difficile à traiter par cette méthode. L'application des méthodes adaptatives à base d'**IS** est également délicate, car ces algorithmes sont difficilement applicables directement à des états limites avec des zones de défaillance multiples. On peut tout de même noter que l'indice d'efficacité ν_{MC} de la méthode par rapport à **MCS** est très élevé.

2.5.4.2 Réponse dynamique d'un oscillateur non linéaire

La méthode **Vb-AGP + IS** a ensuite été appliquée au deuxième cas d'étude (cf Section 2.4.3) dérivée de l'exemple basé sur la réponse dynamique d'un oscillateur à un degré de liberté non-linéaire et non-amorti avec $F_1 \sim \mathcal{N}(0.6, 0.01)$, qui correspond à une probabilité de défaillance de 9.08×10^{-6} (2.47%) estimée par **MCS**. La méthode a été exécutée 100 fois pour des **DoE** initiaux de 12 points et un **COV** limite de 3%.

Ce type de problème de fiabilité est très difficile à traiter par la méthode **AK-MCS** mais il a été traité par la méthode **AK-IS** [ECHARD et al., 2013]. L'**IS** est échantillonné via une méthode **FORM+IS** dans cet algorithme. A l'initialisation, une optimisation est donc effectuée pour déterminer le point de conception puis la phase d'enrichissement

est effectuée sur l'IS.

On compare donc ici les résultats obtenus par l'approche **Vb-AGP + IS** avec les résultats obtenus par AK-IS présentés dans le papier [ECHARD et al., 2013]. Ces résultats sont donnés dans le Tab. 2.5.

Méthode	N_{sim}	$COV(N_{sim})$	\hat{P}_f	$COV(\hat{P}_f)$	e_r	ν_{MC}
MCS	1.8×10^9	-	9.08×10^{-6}	2.47%		
FORM+IS	$29 + 10^4$	-	9.13×10^{-6}	2.29%		
AK-IS [ECHARD et al., 2013]	29 + 38	-	9.13×10^{-6}	2.29%		1943812
Vb-AGP + IS	58	20%	9.09×10^{-6}	2.9%	2.3%	2240946

TABLEAU 2.5 – Oscillateur non linéaire ($\mu_{F_1} = 0.6, \sigma_{F_1} = 0.1$) — Résultats moyens sur 50 répétitions - Paramètres de l'algorithme : $n_{MC}^{init} = 10^4, n_{DoE}^{init} = 12, COV_{max} = 0.03$

Sur cet exemple, la méthode **Vb-AGP** divise en moyenne le nombre d'appels à la fonction de performance d'un facteur 1.16 par rapport à AK-IS. Ce gain est moins important que ceux que nous avons pu obtenir sur les exemples précédents avec **AK-MCS**, il reste donc sans doute des améliorations possibles sur le couplage avec **NAIS**. En effet, on peut noter que le **COV** de N_{sim} est relativement élevé pour la méthode **Vb-AGP + IS**. Ceci peut s'expliquer par l'influence du **DoE** initial, qui va avoir un effet important pour cette méthode sur les performances du **NAIS**. De plus, la dimension d'entrées de ce cas d'application est de six paramètres aléatoires et s'approche donc des limites de l'algorithme **NAIS**. En effet, l'efficacité de cette méthode est fortement diminuée lorsque la dimension stochastique augmente comme expliqué dans la Section 1.2.3.2.2.

L'histogramme des estimations \hat{P}_f de P_f obtenues à partir de l'algorithme **Vb-AGP + IS** ainsi que la probabilité de référence $P_{f,ref}$ et son intervalle de confiance à $\pm 3\sigma$ sont représentés sur la Figure 2.13. On observe sur cette Figure que l'ensemble des estimations de la moyenne de \hat{P}_f se trouve dans l'intervalle de confiance de l'estimation obtenue par **MCS**.

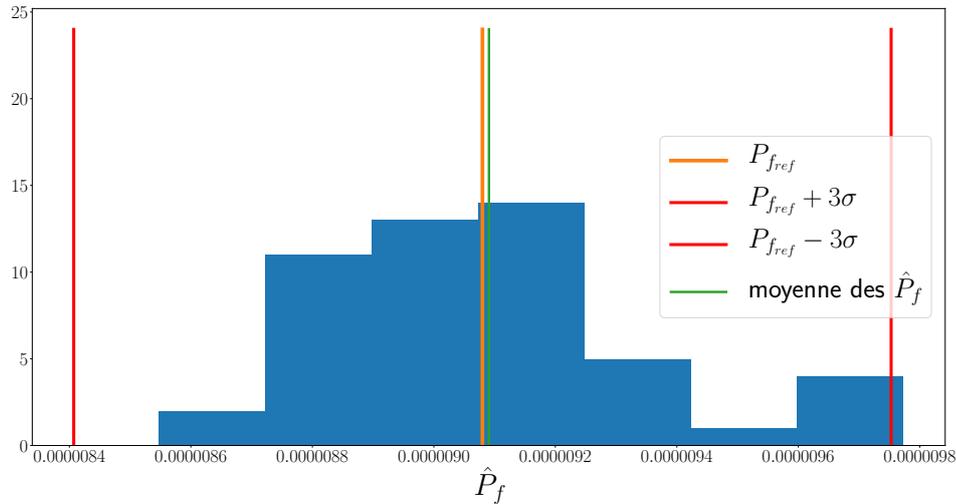


FIGURE 2.13 – Histogramme des estimations \hat{P}_f sur 50 répétitions de Vb-AGP + IS sur l'exemple de l'oscillateur non linéaire - cas probabilité faible

2.6 Conclusion

Dans ce chapitre, nous avons proposé d'analyser la contribution des incertitudes liées à l'approximation par processus gaussien et à l'échantillonnage d'une population de Monte-Carlo à la variabilité de l'estimateur de la probabilité de défaillance. Après avoir montré que cette analyse peut s'écrire comme un problème d'analyse de sensibilité basé sur la décomposition de la variance, des estimateurs des variances dues à l'effet de chaque variable et de la variance totale ont été proposées. Nous avons ensuite montré sur un exemple, en comparant les valeurs obtenues pour les variances, qu'il n'est pas nécessaire d'apprendre de manière aussi précise le processus gaussien que l'impose le critère d'arrêt classique de l'algorithme d'apprentissage actif et qu'il est possible de s'affranchir de quelques simulations qui ne sont pas nécessaires.

L'approche Vb-AGP intégrant cette analyse pour améliorer adaptativement la source d'erreur majeure durant la phase d'apprentissage ainsi qu'un critère d'arrêt basé sur la variance totale de l'estimation de la probabilité de défaillance a ensuite été proposée. Nous avons montré sur deux exemples, que cette approche est prometteuse d'un point de vue de la réduction du nombre de points d'apprentissage, en respectant le COV maximal imposé par l'utilisateur. De plus, elle donne une estimation du COV total, c.-à-d. prenant en compte toutes les incertitudes de la probabilité de défaillance estimée.

Une seconde approche, qui est une extension de la méthode précédente à l'IS, afin d'être adaptée à l'estimation de probabilité d'événements rares, a ensuite été présentée. Dans ces travaux, la méthode d'IS adaptatif NAIS a été utilisée pour pouvoir traiter des domaines de défaillance multimodaux sans connaissances a priori. L'approche a

été illustrée sur un premier exemple puis testée sur un deuxième cas d'application, sur lesquelles elle s'est montrée efficace du point de vue de la précision des estimations et de la réduction du nombre de simulations.

Les approches proposées nous ont ainsi permis de réduire le nombre de points d'apprentissage nécessaires pour estimer la probabilité de défaillance et fournissent également une estimation du COV total sur l'estimation obtenue. Néanmoins, l'application de la méthode Vb-AGP + MCS est limitée pour des problèmes dont la dimension stochastique est peu élevée à cause du coût en terme de mémoire de l'évaluation des trajectoires en utilisant l'expansion de KL sur des populations de MC importantes. L'utilisation de la méthode Vb-AGP + IS permet alors de réduire fortement la taille de la population de MC mais l'algorithme NAIS n'étant efficace que pour des entrées dont la dimension est relativement faible, son utilisation dans l'état n'est pas adaptée non plus pour des problèmes dont la dimension stochastique est importante.

CHAPITRE 3

Apprentissage actif combiné à la modélisation en base réduite

Objectifs du chapitre

- Présenter une méthodologie combinant les analyses de fiabilité par apprentissage actif et la modélisation en base réduite.
- Proposer des critères de couplage efficaces.
- Étudier les performances (gain en temps de calcul et précision) sur des cas d'application concrets.

Sommaire

3.1	Introduction	101
3.2	Approche combinant apprentissage actif et modélisation en base réduite	102
3.2.1	Motivations	102
3.2.2	Concept général de la méthodologie adaptative proposée	103
3.2.3	Estimation de l'erreur de la solution du système projetée sur la base réduite	105
3.2.4	Couplage de AK-MCS et de la modélisation en base réduite	106
3.2.5	Amélioration du critère de couplage par l'utilisation de pré-conditionneurs	109
3.2.6	Etude de l'extension potentielle à la modélisation multi-fidélité	112
3.3	Applications	113
3.3.1	Premier exemple : Analyse de fiabilité sur un problème thermique	113
3.3.2	Second exemple : application à la rupture d'une plaque composite trouée soumise à un chargement complexe	126
3.3.3	Gain en temps de calcul pour des problèmes de grande dimension	134
3.3.4	Application du couplage entre la méthode Vb-AGP et de la modélisation en base réduite	136
3.4	Conclusion	140

3.1 Introduction

Dans le chapitre précédent, nous nous sommes intéressés aux méthodes d'analyse de fiabilité par apprentissage actif. Nous avons proposé d'utiliser la variance de l'estimateur de la probabilité de défaillance en prenant en compte l'incertitude du modèle de processus gaussien, ou *Gaussian process* (GP) pour réduire le nombre de points d'apprentissage ajoutés par les algorithmes. Néanmoins, l'évaluation de chaque point d'apprentissage nécessite de résoudre le modèle numérique. Ainsi, le coût lié à ces évaluations peut encore être élevé, voir prohibitif pour des problèmes complexes, résolus par des codes de calcul coûteux.

Dans ce chapitre, on s'intéresse plus particulièrement au cas de modèles paramétriques de grande taille (c.-à-d. avec un grand nombre de Degrés de liberté (ddl)). Nous nous intéressons ici aux problèmes résolus par la méthode des Éléments Finis (EF) et nous nous restreignons aux problèmes éléments finis. Par exemple, dans le cadre de la mécanique des milieux continus (cf Section 1.1.3.2) ces problèmes s'écrivent sous la forme d'un système linéaire :

$$\mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \quad (3.1)$$

avec $\mathbf{K}(\mathbf{x})$ la matrice de rigidité, $\mathbf{u}(\mathbf{x})$ le vecteur déplacement et $\mathbf{F}(\mathbf{x})$ le vecteur des efforts extérieurs.

Afin de réduire les temps de calcul représentés par les appels successifs au modèle numérique complet, nous proposons dans ce chapitre d'utiliser des solutions en base réduite dès lors que l'estimation obtenue est suffisamment précise. En effet, les techniques de réduction de modèles permettent de diminuer les coûts de calcul liés à la résolution du système donné par l'Eq. (3.1) en le projetant dans un sous-espace de plus petite dimension. Plus précisément la méthode utilisée dans ce travail est la modélisation en base réduite ou réduction de modèles par projection [BENNER et al., 2015; KERSCHEN et al., 2005]. Les modèles par projection, présentés dans le chapitre 1, visent alors à résoudre le système Eq. (3.1) en projection sur une base réduite Ψ de taille n_r construite à partir de solutions du problème complet. Ainsi le problème initial projeté sur cette base peut s'écrire :

$$\Psi^T \mathbf{K}(\mathbf{x}) \Psi \boldsymbol{\alpha}(\mathbf{x}) = \Psi^T \mathbf{F}(\mathbf{x}) \quad (3.2)$$

Ces méthodes permettent d'obtenir une approximation $\mathbf{u}_r(\mathbf{x})$ du vecteur des déplacements $\mathbf{u}(\mathbf{x})$ en l'exprimant dans la base Ψ :

$$\mathbf{u}_r(\mathbf{x}) = \Psi \boldsymbol{\alpha}(\mathbf{x}) \quad (3.3)$$

où les coefficients $\boldsymbol{\alpha}(\mathbf{x})$ du vecteur déplacement $\mathbf{u}(\mathbf{x})$ dans la base Ψ sont obtenus en

résolvant le problème de l'Eq. (3.2). Le gain de calcul permis par la résolution en base réduite est dû au fait que le système linéaire de l'Eq. (3.2) à résoudre est un système de taille n_r bien inférieure à la taille du système matriciel dans l'espace de départ qui est de dimension n_{ddl} .

La suite de ce chapitre est organisée comme suit. Dans une première partie, nous exposerons le concept général de l'approche développée. Les différentes étapes de la nouvelle méthodologie seront ensuite détaillées. Enfin, nous illustrerons l'approche sur deux applications et nous étudierons les gains obtenus en terme de nombre de résolutions du système EF et de temps de calcul.

3.2 Approche combinant apprentissage actif et modélisation en base réduite

Nous proposons ici une méthode d'analyse de fiabilité hybride combinant les approches d'apprentissage actif de modèles de GP et de réduction de modèles. Nous présentons d'abord les motivations qui nous ont poussés à développer cette méthodologie et son concept général. Ensuite le couplage concret de la modélisation en base réduite avec la méthode d'analyse de fiabilité AK-MCS est détaillé avec un premier critère puis avec un second critère amélioré pour répondre aux limitations du premier.

3.2.1 Motivations

Au cours des procédures d'apprentissage des algorithmes d'analyse de fiabilité par apprentissage actif, les points sont en majorité ajoutés dans les mêmes régions de l'espace, au voisinage de l'état limite, et de nombreux points du DoE sont proches les uns des autres. En reprenant l'exemple de l'application de la méthode AK-MCS sur le cas test à 4 branches présenté dans le chapitre 1 et dont le résultat est rappelé sur la Figure 3.1, on observe que la grande majorité des points du DoE final sont concentrés sur les 4 branches de l'état limite.

Considérant un cas d'application EF, un nouveau point d'apprentissage ajouté proche de l'état limite a de fortes chances de se trouver au voisinage d'autres points du DoE ajoutés au cours de la procédure d'apprentissage et déjà évalués en utilisant le modèle numérique complet. Ainsi, l'approximation de la solution calculée en projection sur la base réduite construite à partir des solutions complètes déjà existantes sera relativement précise en ce nouveau point. Une solution réduite pourrait donc être suffisante pour préciser le comportement du système en un certain nombre de nouveaux points du DoE construit au cours de l'algorithme.

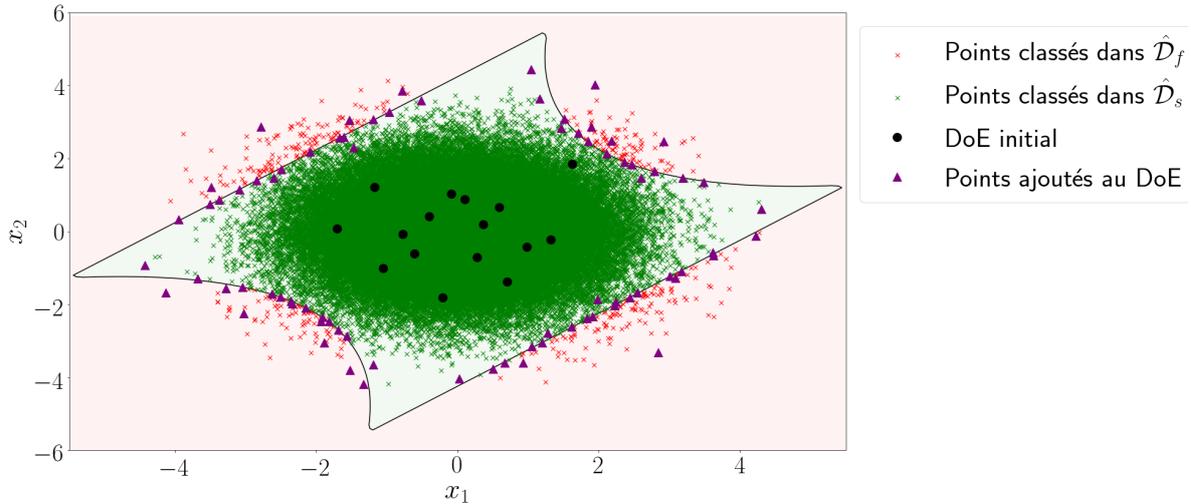


FIGURE 3.1 – Résultat final de AK-MCS à la fin de l'apprentissage

3.2.2 Concept général de la méthodologie adaptative proposée

La procédure proposée débute par la définition d'un DoE initial qui sert à la fois à initialiser le modèle de substitution et la base réduite. Le premier point du DoE est évalué à l'aide du modèle numérique complet et est pris comme premier élément de la base réduite. Ensuite, à chaque point du DoE, la solution du modèle en base réduite et un critère sur la précision de la solution réduite sont calculés. Sur la base de ce critère de précision, le choix est pris à chaque itération d'utiliser la solution en base réduite ou de résoudre le modèle complet. Dans ce dernier cas, la solution obtenue en résolvant le modèle complet est également utilisée pour enrichir la base réduite Ψ . Un métamodèle est ensuite entraîné en utilisant l'ensemble des solutions évaluées, sans faire de différence entre les solutions obtenues en utilisant le modèle réduit et le modèle complet puisque le critère de précision est fixé de manière à avoir des solutions réduites très précises. A chaque itération de la phase d'apprentissage, les points d'apprentissage sont évalués en suivant la même procédure que pour les points du DoE initial. Un schéma récapitulant les étapes de l'algorithme générique, en ne considérant aucune méthode d'apprentissage spécifique, est donné sur la Fig 3.2. Finalement, la probabilité de défaillance est estimée en utilisant la moyenne de la prédiction du métamodèle de GP \mathcal{G}_n sur une population $\tilde{\mathbf{X}} = (\mathbf{X}_i)_{i=1, \dots, n_{MC}}$, avec \mathbf{X}_i i.i.d de même distribution que \mathbf{X} , en suivant l'expression :

$$\hat{P}_f^{AK-MCS+BR} = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mu_n(\mathbf{X}) \leq 0}(\mathbf{X}_i) \quad (3.4)$$

Dans les sections suivantes, la méthode combinant l'algorithme d'apprentissage actif AK-MCS et la modélisation en base réduite avec différents critères est décrite plus en détails. Pour cela, un premier estimateur de la précision de la solution réduite peu

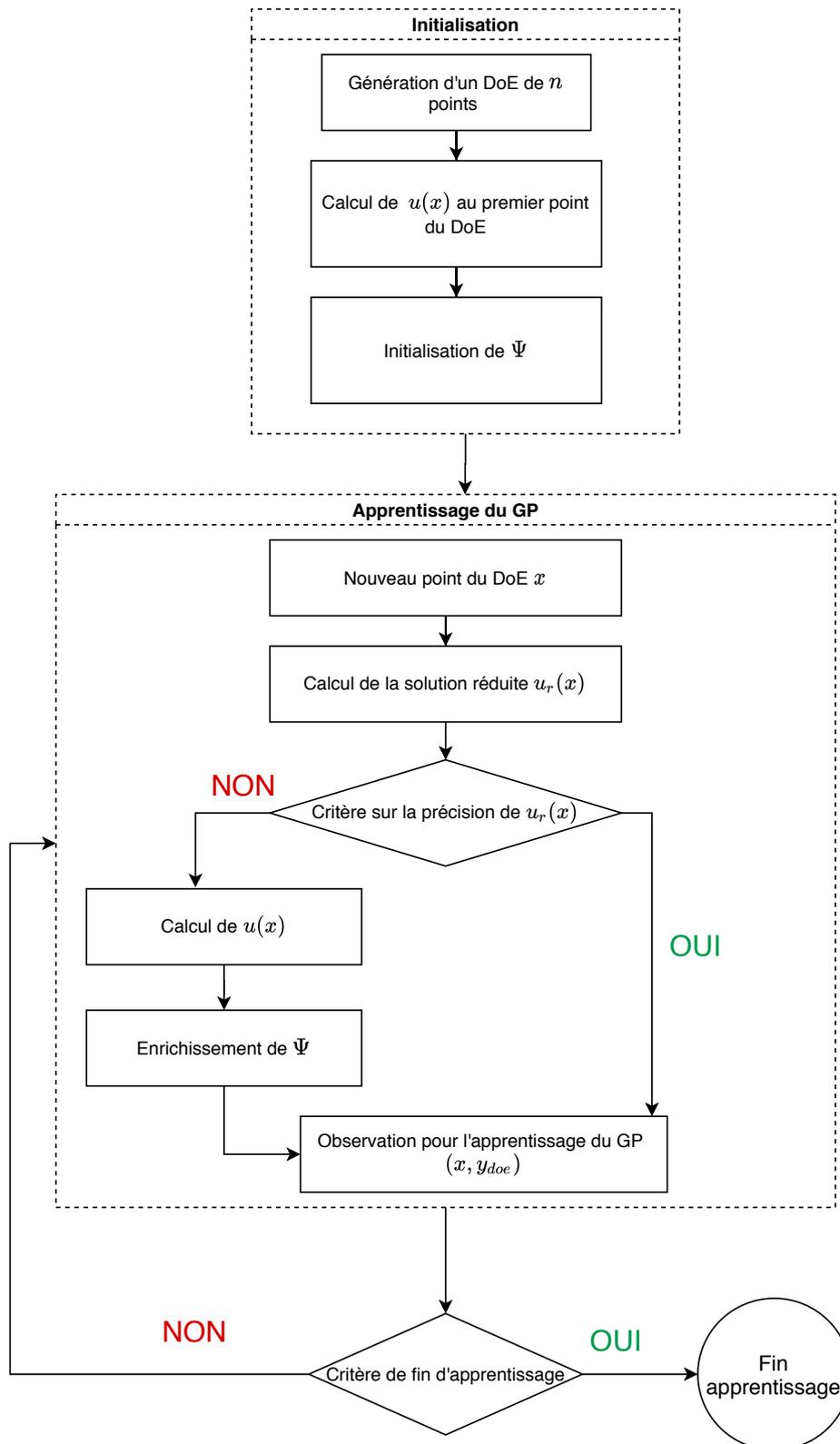


FIGURE 3.2 – Schéma de l’algorithme générique de couplage

coûteux à évaluer sera d’abord donné et la méthode utilisant cet estimateur d’erreur comme critère de couplage sera ensuite détaillée, puis un second estimateur de l’erreur, plus performant, sera proposé. Enfin, nous proposerons d’aborder la méthode comme

une approche multi-fidélité à travers plusieurs propositions qui seront étudiées dans la suite.

3.2.3 Estimation de l'erreur de la solution du système projetée sur la base réduite

Jusqu'ici la base réduite Ψ , sur laquelle le problème était projeté, n'était pas spécifiée puisque de nombreuses méthodes existent pour définir un sous-espace de projection [KERSCHEN et al., 2005; ROZZA et al., 2008; BENNER et al., 2015], parmi lesquelles des méthodes très efficaces existent pour des équations affines par rapport aux paramètres. Cependant, on cherche ici à proposer une méthode applicable à tout problème linéaire, y compris les problèmes pour lesquels la décomposition affine est difficilement accessible et pour lesquels d'autres alternatives plus simples sont préférables. Ainsi, la méthode de construction à la volée [GOGU et PASSIEUX, 2013], décrite dans le chapitre 1, est utilisée pour construire la base réduite. En effet, cette méthode est particulièrement bien adaptée pour le couplage avec l'apprentissage actif puisqu'elle consiste à enrichir de manière adaptative la base réduite au cours de la procédure itérative d'apprentissage actif avec les simulations correspondantes à la résolution du problème EF Eq. (3.1). Un couplage de cette méthode avec une optimisation à base de métamodèles avait ainsi été proposé dans l'article [SOILAHOU DINE et al., 2017].

En pratique, l'erreur exacte entre la solution réduite et la solution du problème EF complet $e(\mathbf{x}) = \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2$ n'est pas calculée puisque l'on cherche justement à éviter l'évaluation de la vraie solution $\mathbf{u}(\mathbf{x})$. De nombreuses techniques existent dans la littérature pour estimer cette erreur sans recourir au calcul de $\mathbf{u}(\mathbf{x})$ [ROZZA et al., 2008; TONN et al., 2011; JANON et al., 2016; GOGU et al., 2016; SMETANA et al., 2019].

Dans ces travaux, on propose dans un premier temps de quantifier la précision des solutions en base réduite en s'appuyant sur un estimateur d'erreur basé sur un résidu tel que proposé dans les travaux de [GOGU et al., 2016] sur la méthode de construction à la volée de la base réduite. Cet estimateur a l'avantage de représenter un faible coût de calcul et d'un point de vue pratique, d'être simple à implémenter. Ainsi la qualité de l'approximation du déplacement par une solution réduite $\mathbf{u}_r(\mathbf{x}) = \Psi\alpha(\mathbf{x})$ peut se quantifier par l'estimateur d'erreur basé sur le résidu suivant :

$$\epsilon_{BR}(\mathbf{x}) = \frac{\|\mathbf{K}(\mathbf{x})\Psi\alpha(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|_2}{\|\mathbf{F}(\mathbf{x})\|_2} \quad (3.5)$$

En pratique, la construction de la base réduite et le calcul du résidu, nécessitent tous les deux d'assembler la matrice de rigidité $\mathbf{K}(\mathbf{x})$. Le coût numérique de cette opération peut être non négligeable mais pour des problèmes de grande dimension, ce coût devient d'autant plus négligeable en comparaison à la résolution du système de

l'Eq. (3.1) que la dimension (nombre de ddl) augmente.

En plus de l'assemblage de la matrice de rigidité $\mathbf{K}(\mathbf{x})$, le calcul du résidu $\epsilon_{BR}(\mathbf{x})$ implique seulement un produit matrice-vecteur et des différences dont le coût est aussi négligeable pour des problèmes de grande dimension. La réduction de modèle est donc une technique puissante pour réduire le coût de calcul et son utilisation au cours de l'algorithme AK-MCS est proposée dans la suite.

3.2.4 Couplage de AK-MCS et de la modélisation en base réduite

Nous proposons ici un algorithme combinant la méthode AK-MCS, décrite dans la Section 1.2.4.2.1 du chapitre 1, et la modélisation en base réduite dans le cadre de la méthode générale présentée dans la section précédente. Cela implique notamment de prendre en compte des solutions de modèles réduits dans les phases d'initialisation et d'apprentissage de AK-MCS.

L'algorithme, noté AK-MCS + BR pour la combinaison de AK-MCS et de la *Base Réduite*, de la méthode proposée est fourni dans l'Algorithme 4 et ces différentes étapes sont décrites ci-dessous.

L'algorithme démarre par l'initialisation de la base réduite. Pour cela, le premier échantillon du DoE initial est calculé et normalisé pour servir de premier élément Ψ_1 à la base réduite. Puis les autres échantillons du DoE, i.e. les échantillons du DoE initial et les points d'enrichissement ajoutés au cours de la phase d'apprentissage, sont évalués suivant la procédure suivante :

1. Calcul de la solution réduite \mathbf{u}_{r_i} au point \mathbf{x}_i par projection sur la base réduite disponible.
2. Calcul du résidu $\epsilon_{BR}(\mathbf{x}_i)$ pour la solution précédente en utilisant l'Eq. (3.5).
3. Évaluation de la précision de la solution réduite \mathbf{u}_{r_i} . Si la valeur de $\epsilon_{BR}(\mathbf{x}_i)$ est inférieure à un seuil limite ϵ défini par l'utilisateur (voir aussi les discussions dans les Section 3.2.5 et Section 3.3 pour le choix de ϵ), la solution réduite est considérée suffisamment précise et est ajoutée au DoE. Sinon, l'algorithme passe à l'étape suivante.
4. Si la solution réduite n'est pas jugée suffisamment précise en se basant sur le critère défini, le modèle numérique complet est résolu en ce point. La solution associée \mathbf{u}_i est ajoutée au DoE et utilisée pour enrichir la base réduite après avoir

été orthonormalisée suivant les Eq. (3.6) et Eq. (3.7).

$$\Psi_i = \mathbf{u}_i - \sum_{k=1}^{i-1} \langle \mathbf{u}_i, \Psi_k \rangle \Psi_k \quad (3.6)$$

$$\Psi_i = \frac{\Psi_i}{\|\Psi_i\|_2} \quad (3.7)$$

avec $\langle \cdot, \cdot \rangle$ le produit scalaire dans L^2 .

La méthodologie adaptative proposée permet de choisir d'utiliser les solutions en base réduite ou celles associées au modèle numérique complet. Ce couplage peut ainsi fortement réduire le temps d'exécution de l'algorithme [AK-MCS](#) puisque le calcul de solutions réduites est bien moins coûteux que la résolution du modèle [EF](#) de départ.

Algorithm 4 Algorithme du couplage AK-MCS + BR

Require: n_{MC}, f_X, G, n

- 1: $\tilde{\mathbf{X}} \leftarrow$ population de taille n_{MC} générée suivant f_X
 - 2: $\mathbf{x}_{doe} \leftarrow$ définition d'un DoE initial de taille n
 - 3: $\mathbf{u} \leftarrow$ solution de $\mathbf{K}(\mathbf{x}_1)\mathbf{u}(\mathbf{x}_1) = \mathbf{F}(\mathbf{x}_1)$
 - 4: $\mathbf{y}_{doe} \leftarrow G(\mathbf{u}, \mathbf{x}_1)$
 - 5: **for** $\mathbf{x} \in \mathbf{x}_{doe} \setminus \{\mathbf{x}_1\}$ **do**
 - 6: $\boldsymbol{\alpha} \leftarrow$ solution de $\boldsymbol{\Psi}^T \mathbf{K}(\mathbf{x}) \boldsymbol{\Psi} \boldsymbol{\alpha}(\mathbf{x}) = \boldsymbol{\Psi}^T \mathbf{F}(\mathbf{x})$
 - 7: $\mathbf{u} \leftarrow \boldsymbol{\Psi} \boldsymbol{\alpha}$
 - 8: $\epsilon_{BR} \leftarrow \frac{\|\mathbf{K} \boldsymbol{\Psi} \boldsymbol{\alpha} - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$
 - 9: **if** $\epsilon_{BR} > \epsilon$ **then**
 - 10: $\mathbf{u} \leftarrow$ solution of $\mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x}) = \mathbf{F}(\mathbf{x})$
 - 11: $\mathbf{V} \leftarrow \mathbf{u} - \boldsymbol{\Psi}(\boldsymbol{\Psi}^T \mathbf{u})$
 - 12: $\boldsymbol{\Psi} \leftarrow [\boldsymbol{\Psi}, \frac{\mathbf{V}}{\|\mathbf{V}\|_2}]$
 - 13: **end if**
 - 14: $\mathbf{y}_{doe} \leftarrow [\mathbf{y}_{doe}, G(\mathbf{u}, \mathbf{x})]$
 - 15: **end for**
 - 16: **while** $COV < COV_{max}$ **do**
 - 17: **while** Critère d'apprentissage non atteint **do**
 - 18: $\mathcal{G}_n \leftarrow$ Entraînement du GP avec $(\mathbf{x}_{doe}, \mathbf{y}_{doe})$
 - 19: $\mu_n, \sigma_n \leftarrow \mathcal{G}_n(\mathbf{x}), \forall \mathbf{x} \in \tilde{\mathbf{X}}$
 - 20: $\hat{P}_f^0 \leftarrow \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \mathbb{1}_{\mu_n(\mathbf{x}) \leq 0}(\mathbf{x}_i)$
 - 21: $COV \leftarrow \sqrt{\frac{1 - \hat{P}_f}{\hat{P}_f n_{MC}}}$
 - 22: $U \leftarrow \frac{|\mu_n|}{\sigma_n}$
 - 23: $\mathbf{x} \leftarrow \arg \min_{\mathbf{x}' \in S} U(\mathbf{x}')$
 - 24: **if** $\min U(\mathbf{x}) < 2$ **then**
 - 25: $\boldsymbol{\alpha}(\mathbf{x}) \leftarrow$ solution de $\boldsymbol{\Psi}^T \mathbf{K}(\mathbf{x}) \boldsymbol{\Psi} \boldsymbol{\alpha}(\mathbf{x}) = \boldsymbol{\Psi}^T \mathbf{F}(\mathbf{x})$
 - 26: $\mathbf{u} \leftarrow \boldsymbol{\Psi} \boldsymbol{\alpha}$
 - 27: $\epsilon_{BR} \leftarrow \frac{\|\mathbf{K} \boldsymbol{\Psi} \boldsymbol{\alpha} - \mathbf{F}\|_2}{\|\mathbf{F}\|_2}$
 - 28: **if** $\epsilon_{BR} > \epsilon$ **then**
 - 29: $\mathbf{u} \leftarrow$ solution de $\mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x}) = \mathbf{F}(\mathbf{x})$
 - 30: $\mathbf{V} \leftarrow \mathbf{u} - \boldsymbol{\Psi}(\boldsymbol{\Psi}^T \mathbf{u})$
 - 31: $\boldsymbol{\Psi} \leftarrow [\boldsymbol{\Psi}, \frac{\mathbf{V}}{\|\mathbf{V}\|_2}]$
 - 32: **end if**
 - 33: $\mathbf{x}_{doe} \leftarrow [\mathbf{x}_{doe}, \mathbf{x}]$
 - 34: $\mathbf{y}_{doe} \leftarrow [\mathbf{y}_{doe}, G(\mathbf{u}, \mathbf{x})]$
 - 35: **end if**
 - 36: Critère d'apprentissage non atteint
 - 37: **end while**
 - 38: **end while**
-

3.2.5 Amélioration du critère de couplage par l'utilisation de préconditionneurs

La corrélation entre la vraie erreur et le résidu défini par l'Eq. (3.5) dépend notamment du conditionnement de la matrice de rigidité. Le résidu peut donner une estimation inexacte de l'erreur de la solution en base réduite dans certains cas où le conditionnement est mauvais. De plus, le choix d'un seuil limite pour l'estimateur d'erreur de l'Eq. (3.5), qui est défini sur le vecteur du chargement (par exemple les forces extérieures ou les flux de chaleur), peut être difficile à estimer car la quantité d'intérêt dont on cherche à estimer l'erreur est en fait le vecteur de la variable d'état (par exemple le vecteur déplacement ou le champ de température).

L'utilisation d'un résidu préconditionné, homogène à la variable d'état \mathbf{u} , peut améliorer la précision de l'estimateur d'erreur basé sur des résidus [ZAHM et NOUY, 2016].

En effet, l'opérateur matriciel $\mathbf{K}(\mathbf{x})$ est un endomorphisme linéaire continu, symétrique, coercif. Il existe donc $\gamma(\mathbf{x}) > 0$ et $\beta(\mathbf{x})$ tels que :

$$\gamma(\mathbf{x}) = \inf_{\mathbf{v} \in V, \mathbf{v} \neq 0} \frac{\langle \mathbf{K}(\mathbf{x})\mathbf{v}, \mathbf{v} \rangle}{\|\mathbf{v}\|_2^2} \quad (\text{coercivité}) \quad (3.8)$$

et

$$\beta(\mathbf{x}) = \sup_{\mathbf{v} \in V, \mathbf{v} \neq 0} \sup_{\mathbf{u} \in V, \mathbf{u} \neq 0} \frac{\langle \mathbf{K}(\mathbf{x})\mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{v}\|_2 \|\mathbf{u}\|_2} \quad (\text{continuité}) \quad (3.9)$$

D'une part, l'application de l'inégalité de Cauchy-Schwartz et l'équation Eq. (3.8) permettent d'écrire :

$$\gamma(\mathbf{x})\|\mathbf{v}\|_2^2 \leq \langle \mathbf{K}(\mathbf{x})\mathbf{v}, \mathbf{v} \rangle \leq \|\mathbf{K}(\mathbf{x})\mathbf{v}\|_2 \|\mathbf{v}\|_2, \quad \forall \mathbf{v} \in V \quad (3.10)$$

Ainsi on peut écrire (en prenant $\mathbf{v} = \mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})$) :

$$\|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2 \leq \frac{1}{\gamma(\mathbf{x})} \|\mathbf{K}(\mathbf{x})\mathbf{u}_r(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|_2 \quad (3.11)$$

D'autre part, la propriété donnée par l'Eq. (3.9) permet d'écrire en considérant les vecteurs $\mathbf{v} \in V$ et $\mathbf{u} = \mathbf{K}(\mathbf{x})\mathbf{v} \in V$ que :

$$\frac{\langle \mathbf{K}(\mathbf{x})\mathbf{v}, \mathbf{K}(\mathbf{x})\mathbf{v} \rangle}{\|\mathbf{v}\|_2 \|\mathbf{K}(\mathbf{x})\mathbf{v}\|_2} = \frac{\|\mathbf{K}(\mathbf{x})\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \beta(\mathbf{x}) \quad (3.12)$$

Finalement, l'erreur $e(\mathbf{x}) = \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2$ entre la solution exacte et la solution

réduite satisfait les inégalités suivantes :

$$\begin{cases} \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2 \leq \frac{1}{\gamma(\mathbf{x})} \|\mathbf{K}(\mathbf{x})\mathbf{u}_r(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|_2 \\ \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2 \geq \frac{1}{\beta(\mathbf{x})} \|\mathbf{K}(\mathbf{x})\mathbf{u}_r(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|_2 \end{cases} \quad (3.13)$$

Ainsi on a :

$$\frac{1}{\beta(\mathbf{x})} \|\mathbf{K}(\mathbf{x})\mathbf{u}_r(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|_2 \leq \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2 \leq \frac{1}{\gamma(\mathbf{x})} \|\mathbf{K}(\mathbf{x})\mathbf{u}_r(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|_2 \quad (3.14)$$

Supposons que l'on dispose d'un bon préconditionneur \mathbf{P} pour $\mathbf{K}(\mathbf{x})$, le conditionnement κ de $\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})$ est alors tel que :

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}} \approx 1 \quad (3.15)$$

où $\lambda_{min} \approx 1$ et $\lambda_{max} \approx 1$ correspondent respectivement à la valeur propre minimale et maximale de $\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})$. De plus, on a dans ce cas :

$$\beta(\mathbf{x}) = \sup_{v \in \mathbb{R}^n, v \neq 0} \frac{\|\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \approx \lambda_{max} \quad (3.16)$$

et de la même façon :

$$\gamma(\mathbf{x}) = \lambda_{min} \quad (3.17)$$

Ainsi, le calcul de $\|\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})\mathbf{u}_r - \mathbf{P}^{-1}\mathbf{F}(\mathbf{x})\|_2$, avec un bon préconditionneur \mathbf{P} , a de fortes chances de fournir une bonne approximation de l'erreur $e(\mathbf{x}) = \|\mathbf{u}(\mathbf{x}) - \mathbf{u}_r(\mathbf{x})\|_2$.

On propose donc d'utiliser un résidu préconditionné donné par la formule suivante pour améliorer les performances du couplage proposé :

$$\epsilon_{BR}^P(\mathbf{x}) = \frac{\|\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})\Psi\alpha(\mathbf{x}) - \mathbf{P}^{-1}\mathbf{F}(\mathbf{x})\|_2}{\|\mathbf{P}^{-1}\mathbf{F}(\mathbf{x})\|_2} \quad (3.18)$$

avec \mathbf{P} un préconditionneur qui pourra prendre différentes formes, proposées ensuite.

Notons que l'évaluation du résidu préconditionné $\epsilon_{BR}^P(\mathbf{x})$ a généralement un coût plus élevé que celle du résidu non préconditionné $\epsilon_{BR}(\mathbf{x})$, car elle nécessite le calcul du terme $\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})\Psi\alpha(\mathbf{x})$. Cependant pour des préconditionneurs indépendants du paramètre \mathbf{x} , la factorisation de \mathbf{P} peut être effectuée une unique fois puis stockée pour toute la durée d'exécution de l'algorithme. Ainsi, le seul coût restant correspond à la résolution du système en fonction du type de factorisation choisie, qui doit être effectuée quel que soit le résidu et qui n'induit donc pas un coût de calcul additionnel significatif.

Il existe de nombreuses méthodes pour construire un préconditionneur, qu'on souhaite ici indépendant de \mathbf{x} . Dans ces travaux, la méthode proposée sera exécutée avec

deux techniques de construction de préconditionneurs possibles :

- le préconditionneur \mathbf{P}_{mean} calculé au point moyen [GHANEM et KRUGER, 1996], dénommé préconditionneur moyen dans la suite : la matrice de rigidité \mathbf{K}^0 est assemblée pour le point moyen du domaine physique et est utilisée comme préconditionneur $\mathbf{P}_{mean} = \mathbf{K}^0$.
- le préconditionneur au plus proche $\mathbf{P}_{nearest}$: chaque fois que le modèle numérique complet est évalué pour enrichir la base réduite, la matrice de rigidité obtenue $\mathbf{K}^i = \mathbf{K}(\mathbf{x}_i)$ est stockée pour servir de préconditionneur.

Notons $\mathcal{X}_{full} = \{\mathbf{x}_i, i = 1, \dots, n_{full}\} \subset \mathcal{X}$ l'ensemble qui contient tous les points auxquels la solution complète a été évaluée. A chaque évaluation d'un point d'apprentissage \mathbf{x}^* , le résidu préconditionné est calculé en utilisant le préconditionneur :

$$\mathbf{P}_{nearest} = \mathbf{K}(\mathbf{x}_i) \quad \text{tel que} \quad \mathbf{x}_i = \min_{\mathbf{x} \in \mathcal{X}_{full}} \|\mathbf{x} - \mathbf{x}^*\|_2 \quad (3.19)$$

Notons que ce préconditionneur n'est pas entièrement indépendant de \mathbf{x} puisque le choix de la matrice utilisée dépend de la distance de \mathbf{x} aux \mathbf{x}_i mais les matrices elles-mêmes sont indépendantes de \mathbf{x} .

Les préconditionneurs présentés ci-dessus sont construits à partir de valeurs de la matrice de rigidité, i.e. $\mathbf{P}^{-1} = \mathbf{K}^{-1}(\mathbf{x}')$ avec \mathbf{x}' un point particulier du domaine des entrées. On peut ainsi supposer qu'en un point \mathbf{x} du domaine la matrice $\mathbf{K}^{-1}(\mathbf{x})$ est relativement proche de \mathbf{P}^{-1} et que le conditionnement du problème avec un tel préconditionneur est ainsi très bon. Ainsi, $\epsilon_{BR}^P(\mathbf{x})$ peut être considéré comme une bonne approximation de l'erreur relative $\mathbf{u}_r(\mathbf{x})$ par rapport à la vraie solution du système EF $\mathbf{u}(\mathbf{x})$. En se basant sur ces résultats, le préconditionnement du résidu facilite le choix d'une valeur seuil ϵ , qui permet d'assurer que l'erreur sur $\mathbf{u}(\mathbf{x})$ ne dépasse pas l'erreur maximale autorisée.

De plus, on peut s'attendre à ce que l'erreur sur la fonction de performance $G(\mathbf{x})$ soit faible si le seuil ϵ est fixé pour avoir une erreur relative faible sur $\mathbf{u}(\mathbf{x})$ (par exemple 10^{-3} , valeur qui sera discutée par la suite). Cependant, le lien entre ces deux erreurs est fortement dépendant de la forme de $G : \mathbf{x} \rightarrow G(\mathbf{u}(\mathbf{x}), \mathbf{x})$. Généralement, on observe ce résultat car G est souvent faiblement ou moyennement non linéaire en $\mathbf{u}(\mathbf{x})$. Si $G(\mathbf{x})$ est fortement non linéaire en $\mathbf{u}(\mathbf{x})$ le choix de ϵ n'est pas trivial et peut nécessiter plusieurs essais pour obtenir une valeur raisonnable.

Ainsi la méthode proposée précédemment pour combiner l'algorithme AK-MCS et la modélisation en base réduite peut être améliorée en utilisant un résidu préconditionné $\epsilon_{BR}^P(\mathbf{x})$. Ce nouvel estimateur de l'erreur permet d'améliorer l'estimateur basé sur un résidu défini par l'Eq. (3.5) utilisé comme critère de décision pour choisir si une solution

réduite est suffisamment précise pour remplacer une simulation du modèle numérique complet.

3.2.6 Etude de l'extension potentielle à la modélisation multi-fidélité

3.2.6.1 Reformulation du problème comme un problème multi-fidélité

L'utilisation des solutions réduites et complètes comme solutions du même modèle pour construire un modèle de substitution se discute. Dans les faits, les données ont différents niveaux de fidélité : les solutions complètes sont des données haute fidélité alors que les solutions réduites sont des données basse fidélité avec n_r (nombre d'éléments de la base réduite) niveaux différents. En effet, à chaque fois que la base réduite est enrichie, le modèle réduit est amélioré et correspond à un nouveau niveau de fidélité. L'utilisation d'un métamodèle comme le co-krigeage, basé sur différents niveaux de fidélité, peut donc être envisagé. Cependant, relativement peu de simulations sont disponibles en pratique pour chaque niveau de fidélité et il est donc impossible de construire un bon modèle de co-krigeage à n_r niveaux de fidélité. Une simplification du problème consiste à considérer deux niveaux de fidélité : un niveau haute fidélité pour les solutions obtenues en résolvant le système complet et un niveau basse fidélité pour les solutions obtenues en résolvant tout modèle réduit du modèle de départ. Ainsi, un modèle de co-krigeage à deux niveaux pourrait être utilisé.

La reformulation comme un problème de co-krigeage en utilisant les solutions du modèle numérique complet comme des données haute fidélité et les solutions du modèle réduit comme données basse fidélité peut être abordée de trois façons différentes :

1. en remplaçant le modèle de GP par un modèle de co-krigeage tout le long de l'algorithme d'apprentissage,
2. en utilisant un GP au cours de la phase d'apprentissage, puis en remplaçant le GP final par un modèle de co-krigeage avec les points des solutions réduites (considérées basse fidélité) et les points des solutions complètes (considérées haute fidélité),
3. en suivant la même stratégie que le point 2, excepté que les données basse fidélité sont recalculées à la fin de l'algorithme en utilisant la base réduite finale pour avoir exactement le même niveau de fidélité pour tous les points issus de solutions réduites.

Ces différentes stratégies seront comparées lors des applications numériques.

Dans la section suivante, une étude numérique des deux stratégies de couplage de la méthode AK-MCS avec la modélisation en base réduite avec ou sans préconditionnement est menée afin de comparer leurs performances sur deux cas d'applications. De plus, une étude des stratégies multi-fidélité basées sur un modèle de co-krigeage sera menée sur une application pour déterminer le potentiel de cette approche du problème. La première application est un problème thermique et la seconde porte sur un problème mécanique.

3.3 Applications

Dans la suite, deux applications numériques de la méthode sont présentées. Pour chaque application, une description du problème physique est donnée et est suivie par une comparaison numérique des différentes stratégies d'estimation de la probabilité de défaillance.

3.3.1 Premier exemple : Analyse de fiabilité sur un problème thermique

3.3.1.1 Description du problème

Dans cette section, l'application considérée est une analyse de fiabilité impliquant le transfert de chaleur à travers la paroi d'une chambre de combustion d'un moteur de fusée refroidi par régénération [GOGU et al., 2016]. De l'hydrogène liquide passant par des canaux de refroidissement à travers la paroi de la chambre de combustion à une température de 40K permet de refroidir le moteur. On considère qu'il y a défaillance lorsque la température maximale de la paroi interne de la chambre de combustion dépasse une valeur critique T_{allow} , ce qui correspond à une rupture de la paroi de la chambre de combustion due aux contraintes induites thermiquement.

La Figure 3.3 fournit un schéma de principe de la chambre de combustion d'un tel moteur de fusée à propulsion liquide hydrogène (LH2) / oxygène (LOX) refroidi par régénération. Comme illustrée sur cette figure, la paroi de la chambre de combustion est formée de deux parties différentes chacune constituée d'un matériau : une partie interne composée d'un alliage de cuivre et une partie externe composée d'un alliage de nickel. Les échanges de chaleur ont lieu par convection entre les parois de la chambre de combustion et les sources de chaleur (gaz de la chambre de combustion), de refroidissement (hydrogène liquide) et aussi avec l'extérieur. Considérant ces conditions aux limites, le transfert thermique dépend des paramètres suivants : la conductivité entre la partie intérieure de la paroi (k_{Cu}), la conductivité de l'enveloppe (k_{Ni}), la température des gaz sur la partie interne de la chambre de combustion (T_{hot}), le coefficient de transfert

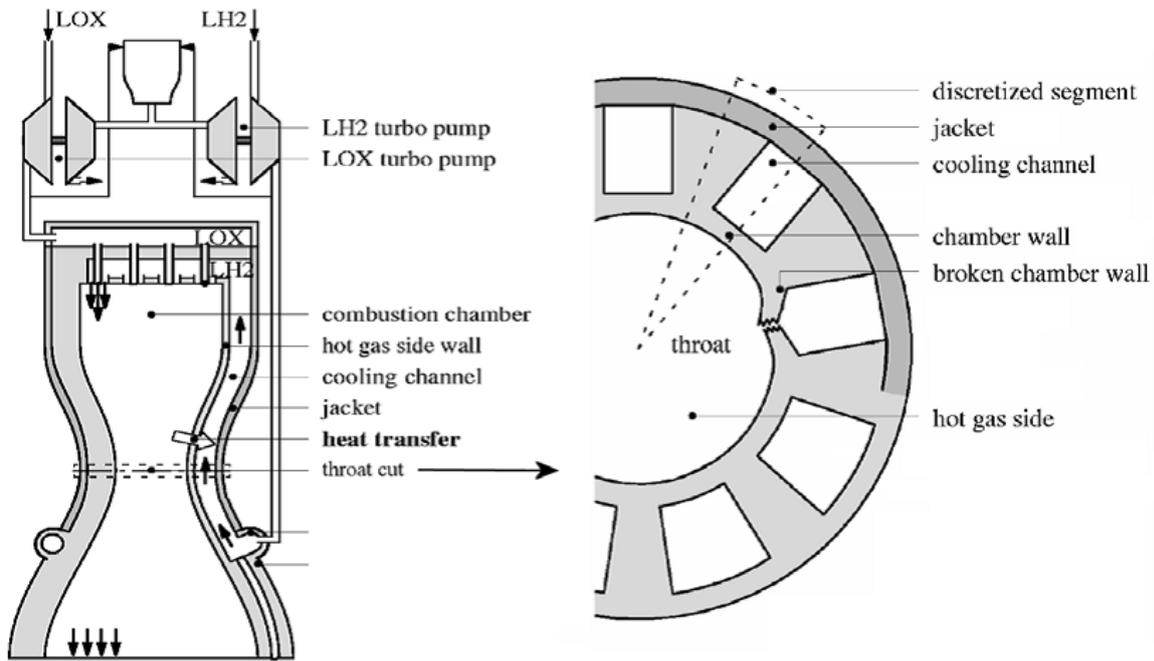


FIGURE 3.3 – Schéma de la chambre de combustion d'une fusée refroidie par régénération.

Entrée	k_{Cu}	k_{Ni}	T_{hot}	h_{hot}	T_{out}	h_{out}	T_{cool}	h_{cool}	T_{allow}
Unité	W/mK	W/mK	K	kW/m^2K	K	kW/m^2K	K	kW/m^2K	K
Distribution	gaussien	gaussien	uniforme	uniforme	uniforme	uniforme	uniforme	uniforme	uniforme
Moyenne	310	75	900	31	293	6	40	250	230
COV	2%	2%							
Moitié de la taille de l'intervalle			10%	10%	5%	5%	5%	10%	7.5%

TABLEAU 3.1 – Lois de probabilité des paramètres du problème thermique.

par convection de la partie interne de la chambre de combustion (h_{hot}), la température de la face externe de la chambre de combustion (T_{out}), le coefficient de transfert par convection de la face externe de la chambre de combustion (h_{out}), la température du liquide de refroidissement (T_{cool}) et le coefficient de transfert par convection du canal de refroidissement (h_{cool}). Ces paramètres ainsi que la température maximale admissible T_{allow} sont supposés incertains et modélisés par des variables aléatoires indépendantes suivant les distributions données dans le Tab. 3.1.

Le champ thermique à l'équilibre stationnaire est obtenu en résolvant une équation de convection-diffusion par une approche EF. Le maillage EF de la paroi de la chambre de combustion et les conditions aux limites sont illustrés sur la Figure 3.4 (par conditions de symétrie seule la partie dénommée "discretized segment" dans la Figure 3.3 est modélisée). Dans ces travaux, un solveur EF maison implémenté en Python est utilisé pour calculer le champ thermique. Un exemple du champ de température calculé au point moyen par rapport aux bornes du Tab. 3.1 est fourni dans la Figure 3.5. Ainsi, la température maximale $T_{max}(\mathbf{x})$ est accessible et nous pouvons en déduire la fonction de performance, définie ici par $G(\mathbf{x}) = T_{allow}(\mathbf{x}) - T_{max}(\mathbf{x})$. La probabilité de défaillance

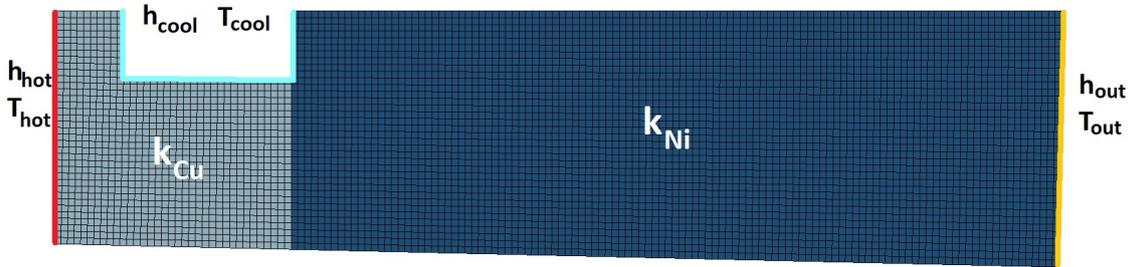


FIGURE 3.4 – Maillage EF de la paroi de la chambre de combustion et conditions aux limites du problème thermique.

peut ensuite être estimée en utilisant une MCS ou en utilisant une approche d'analyse de fiabilité par apprentissage actif telle que la méthode AK-MCS par exemple.

Les résultats de comparaison entre la méthode classique de MC, la méthode AK-MCS et les deux stratégies de couplage AK-MCS + BR (avec et sans préconditionnement) sont présentés dans la suite.

3.3.1.2 Résultats

En premier lieu, la méthode classique MCS nous a permis d'avoir une estimation de référence de la probabilité de défaillance sur une simulation de MC. L'estimation obtenue est alors $\hat{P}_{f_{MC}} = 6.22 \cdot 10^{-3}$ avec un COV de MC estimé $\widehat{COV}_{P_f}_{MC} = 5.65\%$ (pour $n_{MC} = 5 \times 10^4$).

La méthode AK-MCS combinée à la modélisation en base réduite a été implémentée avec les critères basés sur le résidu classique et les résidus préconditionnés avec les deux préconditionneurs (au point moyen \mathbf{P}_{mean} et au plus proche $\mathbf{P}_{nearest}$) proposés dans la Section 3.2.5.

Une étude préalable a d'abord été menée pour s'assurer que l'utilisation du résidu est un critère de couplage valide.

3.3.1.2.1 Étude du critère de couplage

Afin de vérifier si les résidus utilisés comme critères d'enrichissement (cf Eq. (3.5) et Eq. (3.18)) décrivent correctement l'évolution de la vraie erreur (la vraie erreur est définie comme l'erreur entre la solution réduite et la solution calculée par la méthode

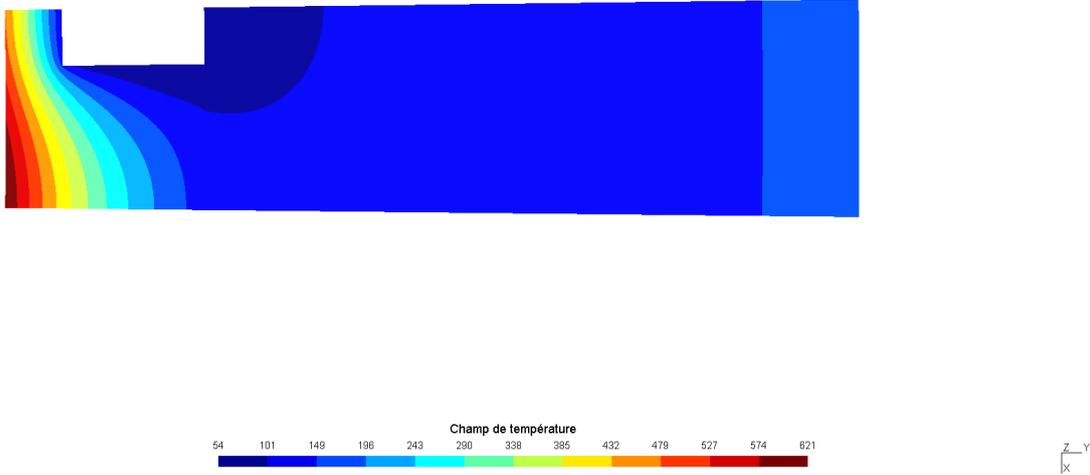


FIGURE 3.5 – Champ de température calculé au point moyen.

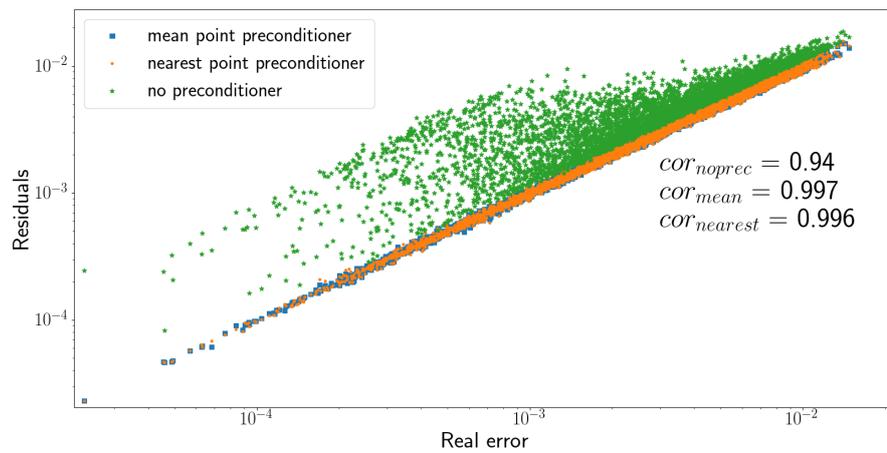
des EF) nous avons dans un premier temps calculé le coefficient de corrélation de Pearson (corrélation linéaire) entre ces résidus et la vraie erreur. Ces estimations sont effectuées pour différentes tailles de base réduite et sont calculées avec les valeurs des résidus et des vraies erreurs sur une population de MC aléatoire.

Sur la Figure 3.6, on peut constater que le résidu préconditionné est quasiment parfaitement corrélé à la vraie erreur, ce qui peut s’expliquer par la formulation du problème. On peut noter que les performances des deux préconditionneurs sont similaires et qu’ils fonctionnent tous deux aussi bien ici. Concernant le résidu non préconditionné, le coefficient de corrélation de Pearson est moins élevé et on observe graphiquement que la corrélation est moins forte, avec notamment l’existence d’un biais, en particulier pour des bases réduites de taille plus importante. A noter cependant qu’étant donnée la formulation du critère d’enrichissement basé sur ce résidu, ce biais n’est pas trop pénalisant, nécessitant uniquement un décalage du seuil.

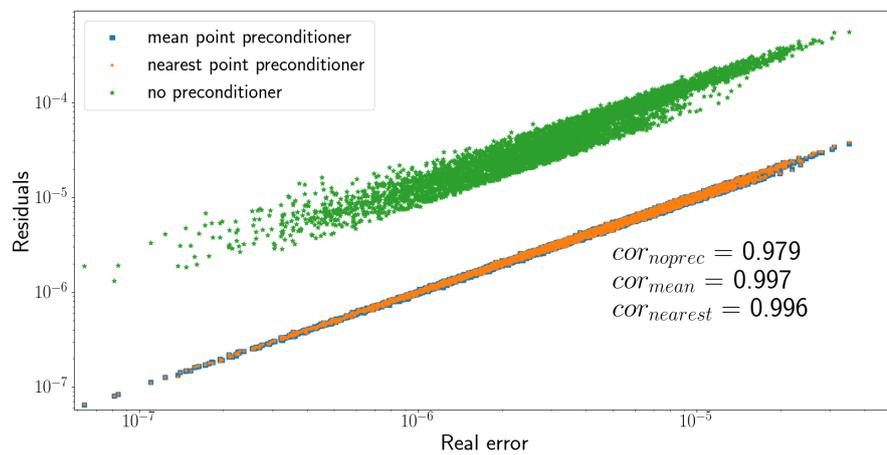
Cette première étude ayant vérifié la validité des résidus préconditionnés (avec les deux préconditionneurs proposés) et non préconditionnés comme estimateurs d’erreur, les performances du couplage pour ces trois résidus vont ensuite être étudiées.

3.3.1.2.2 Étude des performances du couplage

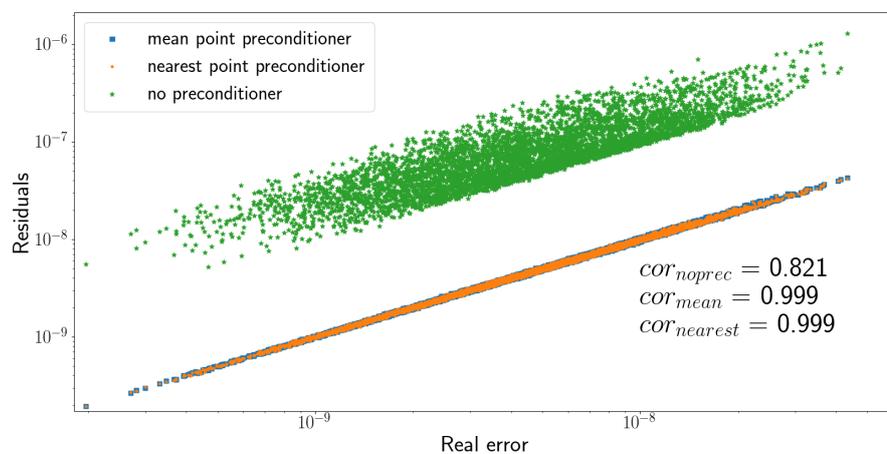
La représentation graphique des trois premiers modes (c.-à-d. les trois premiers vecteurs) d’une base réduite du champ de température construite à la volée durant le couplage basé sur ϵ_{BR} pour $\epsilon = 10^{-3}$ est donnée sur la Figure 3.7. La méthodologie de construction utilisée implique que le premier mode correspond au mode dominant, c’est-à-dire celui qui représente le mieux le comportement général du champ thermique.



Taille de la base réduite = 5



Taille de la base réduite = 10



Taille de la base réduite = 20

FIGURE 3.6 – Corrélation entre les résidus et la vraie erreur pour le problème thermique. Trois tailles de base réduite sont considérées (5, 10 et 20).

Cependant, des vecteurs de base supplémentaires (correspondant à des modes supplémentaires) sont nécessaires pour capturer les variations fines du champ thermique et obtenir des solutions réduites précises. Cela montre bien l'utilité d'un critère de précision, permettant d'ajouter de nouveaux modes à la base réduite lorsque le champ thermique n'est pas suffisamment bien approché avec les éléments disponibles.

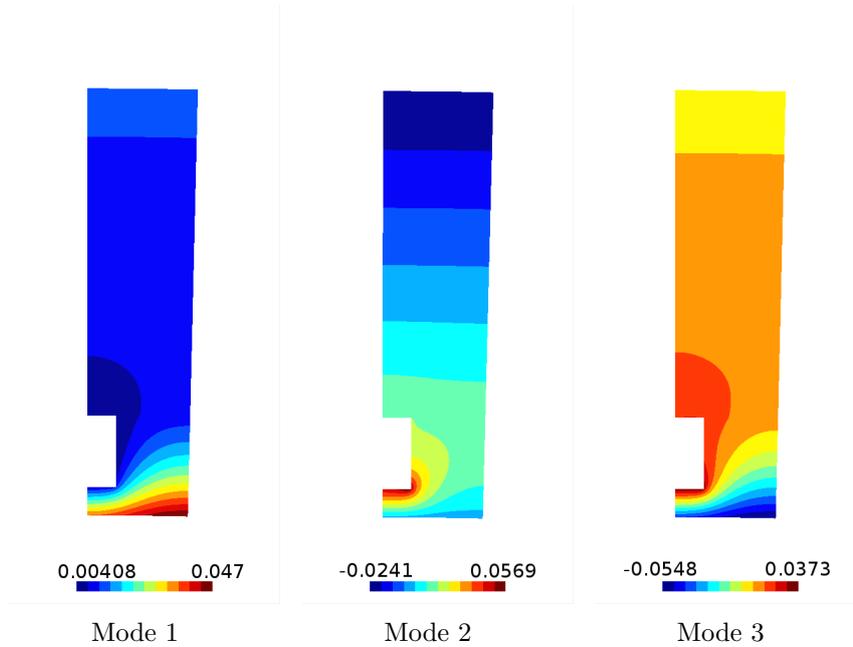


FIGURE 3.7 – Trois premiers modes de température obtenus en utilisant la procédure de construction à la volée au cours de AK-MCS pour le problème thermique.

Il est donc intéressant d'étudier l'influence de la valeur du seuil ϵ utilisé comme critère de couplage sur les performances de la méthode proposée. Dans ce but, la procédure suivante est exécutée :

- Exécution de l'algorithme AK-MCS et sauvegarde du DoE obtenu (i.e. DoE initial et points d'enrichissement)
- Pour des valeurs différentes de ϵ : application du couplage avec la base réduite sur les échantillons du DoE obtenus avec AK-MCS en conservant l'ordre d'ajout des points au cours de l'algorithme AK-MCS

Deux critères différents sont utilisés pour comparer les performances du couplage en fonction de la valeur de ϵ :

- l'accélération atteinte en utilisant l'approche proposée pour la méthode AK-MCS, i.e. le rapport entre le nombre d'évaluations du modèle numérique complet pour la méthode AK-MCS n_{full}^{AK-MCS} et son nombre d'évaluations quand le couplage est utilisé $n_{full}^{AK-MCS+BR}$. On compare uniquement les calculs correspondant aux

appels du modèle numérique complet car le coût d'inversion du système projeté sur la base réduite devient négligeable lorsque la taille (le nombre de `ddl`) du problème augmente.

- la précision de l'estimation de la probabilité de défaillance. Les estimations de $\hat{P}_f^{AK-MCS+BR}$ et \hat{P}_f^{AK-MCS} sur le même DoE initial et sur la même population de MC sont comparées par ce critère.

La précision est obtenue en calculant :

$$1 - \frac{|\hat{P}_f^{AK-MCS} - \hat{P}_f^{AK-MCS+BR}|}{\hat{P}_f^{AK-MCS}}.$$

Ainsi une précision exacte de la solution réduite correspond à une valeur de 1.

Afin de prendre en compte la variabilité stochastique de l'approche AK-MCS dans l'évaluation de l'influence du paramètre ϵ , la procédure décrite précédemment a été répétée 10 fois pour différents DoE initiaux et différentes populations de MC. La moyenne des deux critères de comparaison sur 10 exécutions a été calculée pour différentes valeurs du seuil. L'évolution de ces deux critères est donnée sur la Figure 3.8 pour le résidu non préconditionné ϵ_{BR} et sur les Figure 3.9 et Figure 3.10 pour le résidu préconditionné ϵ_{BR}^P avec respectivement le préconditionneur du point moyen P_{mean} et le préconditionneur du point le plus proche $P_{nearest}$.

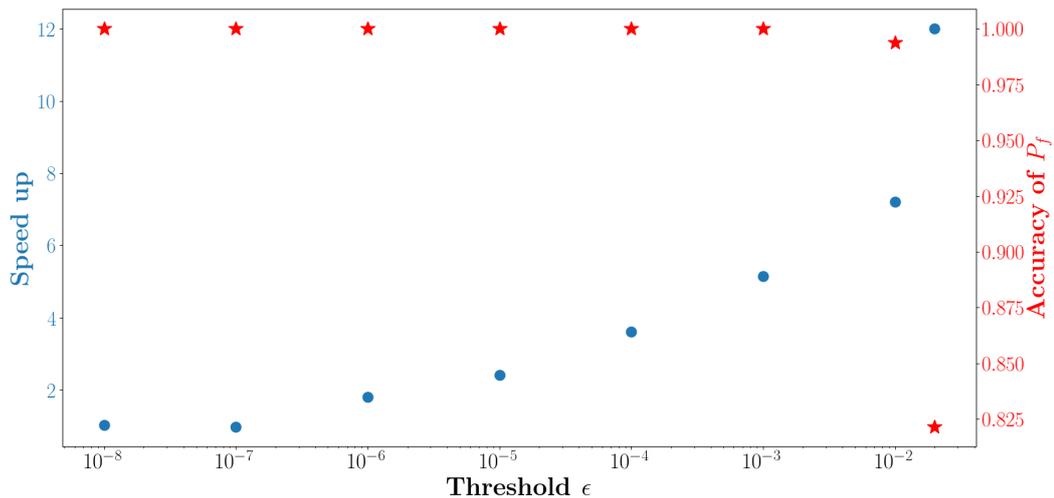


FIGURE 3.8 – Précision (étoiles rouges) de P_f et accélération (points bleus) de la méthode proposée en fonction du seuil ϵ pour le résidu non préconditionné pour différents DoE et populations de MC générés par l'algorithme AK-MCS pour le problème thermique.

A partir de ces résultats (avec ou sans préconditionnement), on peut déduire qu'un bon compromis entre l'accélération obtenue par le couplage et la précision de l'estimation de la probabilité de défaillance est obtenu pour une valeur $\epsilon = 10^{-3}$ sur cette

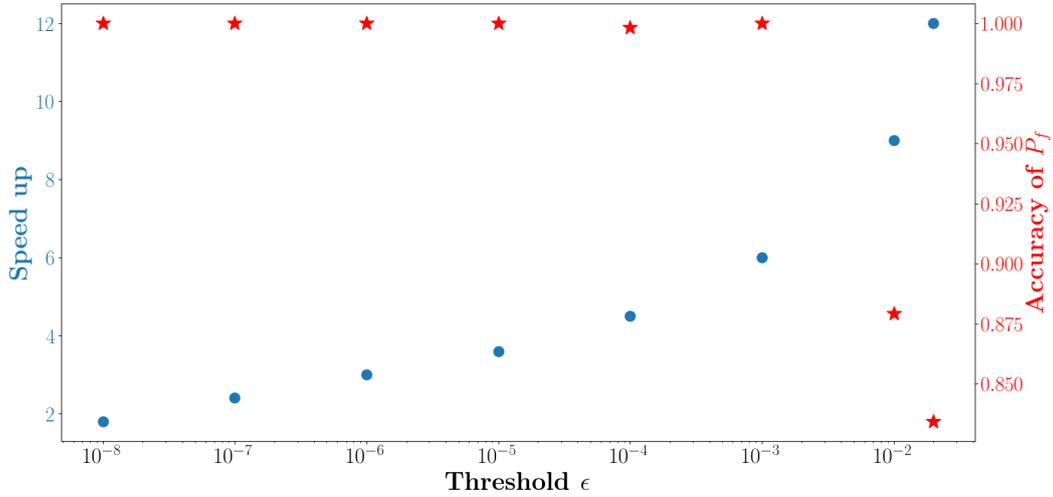


FIGURE 3.9 – Précision (étoiles rouges) de P_f et accélération (points bleus) de la méthode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le preconditionneur du point moyen pour différents DoE et populations de MC générés par l’algorithme AK-MCS pour le problème thermique.

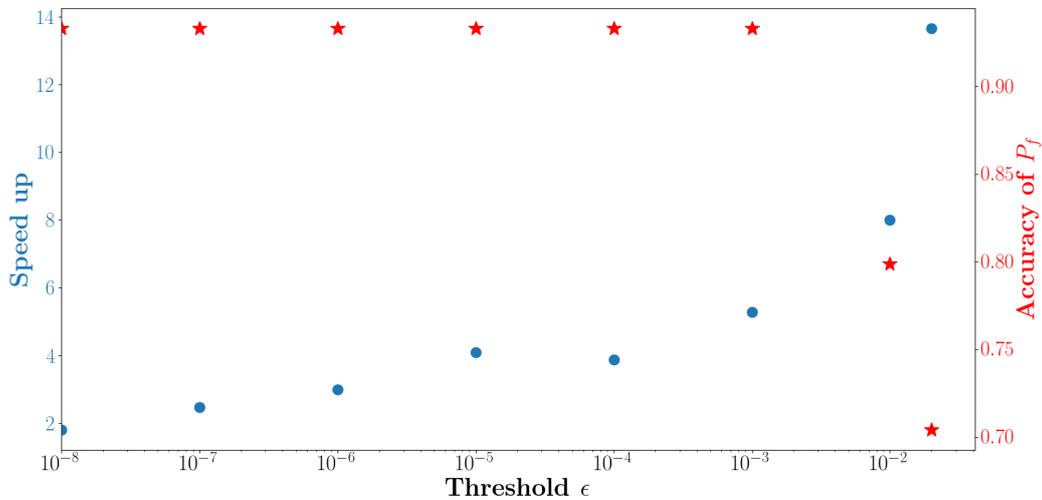


FIGURE 3.10 – Précision (étoiles rouges) de P_f et accélération (points bleus) de la méthode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le preconditionneur du point le plus proche pour différents DoE et populations de MC générés par l’algorithme AK-MCS pour le problème thermique.

application. On peut également remarquer que même si la corrélation entre la vraie erreur et le résidu non préconditionné n’est pas aussi forte qu’avec le preconditionnement (résultats donnés par la Figure 3.6), elle est suffisante sur ce cas pour obtenir des résultats satisfaisants pour l’estimation de la probabilité de défaillance.

Sur la base de ces résultats, une nouvelle étude est menée ensuite pour comparer plus en détails les méthodes AK-MCS et AK-MCS + BR pour $\epsilon = 10^{-3}$. En comparaison

à l'étude précédente, la sélection des points d'enrichissement dans l'approche proposée est potentiellement guidée par des solutions obtenues en utilisant l'approximation en base réduite dans la méthode AK-MCS + BR.

Pour l'étude comparative, chaque version de la méthode AK-MCS + BR proposée (avec et sans préconditionneur) et de AK-MCS pour le même DoE initial et la même population de MC a été exécutée 30 fois. La moyenne $\overline{P_f}$ et l'écart-type $\overline{\sigma_{P_f}}$ de l'échantillon basé sur les résultats obtenus sont donnés dans le Tab. 3.2 pour la version non préconditionnée, le Tab. 3.3 pour le préconditionneur au point moyen et le Tab. 3.4 pour le préconditionneur au plus proche.

	$\overline{P_f}$	$\overline{\sigma_{P_f}}$	n_{sim}	n_{full}
MCS	6.22×10^{-3}	3.5×10^{-4}	5×10^4	5×10^4
AK-MCS	6.52×10^{-3}	1.88×10^{-4}	38.2	38.2
AK-MCS + BR	6.52×10^{-3}	1.90×10^{-4}	38.1	6.7

TABEAU 3.2 – Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le résidu non préconditionné pour le problème thermique.

	$\overline{P_f}$	$\overline{\sigma_{P_f}}$	n_{sim}	n_{full}
MCS	6.22×10^{-3}	3.5×10^{-4}	5×10^4	5×10^4
AK-MCS	6.52×10^{-3}	1.88×10^{-4}	38.2	38.2
AK-MCS + BR - \mathbf{P}_{mean}	6.52×10^{-3}	1.88×10^{-4}	40.7	6

TABEAU 3.3 – Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point moyen pour le problème thermique.

	$\overline{P_f}$	$\overline{\sigma_{P_f}}$	n_{sim}	n_{full}
MCS	6.22×10^{-3}	3.5×10^{-4}	5×10^4	5×10^4
AK-MCS	6.47×10^{-3}	3.03×10^{-4}	38.1	38.1
AK-MCS + BR - $\mathbf{P}_{nearest}$	6.47×10^{-3}	3.02×10^{-4}	39.4	6

TABEAU 3.4 – Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point le plus proche pour le problème thermique.

Premièrement, on remarque que chaque valeur moyenne obtenue avec l'algorithme AK-MCS ou avec la méthode proposée (avec ou sans préconditionneur) est cohérente avec la valeur de référence $\hat{P}_{f_{MC}} = 6.22 \times 10^{-3}$ (avec un COV estimé à 5.65%) obtenue

par MCS. Plus précisément, les moyennes des échantillons de \hat{P}_f obtenues pour chacune des variantes de l'algorithme de couplage sont très proches de la moyenne des échantillons \hat{P}_f^{AK-MCS} résultants de la méthode AK-MCS. De plus, les variances estimées ne prennent pas des valeurs extrêmes lorsque la base réduite est utilisée et leurs valeurs, en particulier pour la stratégie utilisant le préconditionnement, sont même très proches de la référence $\hat{\sigma}_{\hat{P}_f^{AK-MCS}}$.

Pour ces simulations, l'accélération (la moyenne sur les 30 répétitions des ratios $n_{full}^{AK-MCS}/n_{full}^{AK-MCS+BR}$) de l'algorithme en terme de nombre d'évaluations du modèle numérique complet pour l'approche basée sur le résidu préconditionné est de 6.37 pour le préconditionneur au point moyen \mathbf{P}_{mean} et 6.33 pour le préconditionneur au plus proche $\mathbf{P}_{nearest}$. Concernant la méthode avec un résidu non préconditionné, le facteur d'accélération est plus faible et vaut 5.74 ce qui peut s'expliquer par un meilleur choix des points servant à la construction des bases réduites (qui sont par conséquent de plus petite taille) grâce au préconditionnement des résidus. De plus, les évaluations du modèle complet représentent 15% et 15.5% du nombre total de simulations respectivement pour une exécution de l'algorithme de couplage avec les préconditionneurs du point moyen et du point le plus proche. Sans préconditionnement le pourcentage d'évaluations du modèle complet s'élève en moyenne à 17.8%.

Les différentes reformulations du problème, présentées dans la Section 3.2.6.1, comme un problème multi-fidélité modélisé à l'aide d'un modèle de co-krigeage ont ensuite été testées sur cette application.

3.3.1.3 Étude de l'extension avec le co-krigeage

Dans cette section, les trois reformulations du problème avec le co-krigeage (cf Section 3.2.6.1) sont étudiées.

3.3.1.3.1 Prédiction par co-krigeage

Dans cette section, on rappelle brièvement les formulations de la moyenne et de la variance de la prédiction par co-krigeage. Pour davantage de détails sur ce modèle, le lecteur pourra se référer à la Section 1.1.2.8 du chapitre 1.

Supposons que l'on dispose de données haute-fidélité $\{\mathbf{x}_h, y_h = G_h(\mathbf{x}_h)\}$, les points calculés en utilisant le modèle complet ici, et de données basse-fidélité $\{\mathbf{x}_b, y_b = G_b(\mathbf{x}_b)\}$, les points calculés en utilisant un modèle réduit ici. Le modèle de co-krigeage à deux niveaux de fidélité est basé sur deux GP \mathcal{G}_h et \mathcal{G}_b approchant respectivement les réponses y_h et y_b .

La moyenne $\mu_{h,n}(\mathbf{x})$ et la variance $\sigma_{h,n}^2(\mathbf{x})$ de la prédiction de la réponse par le GP $\mathcal{G}_{h,n}$ conditionné à l'ensemble des sorties $\{\mathbf{x}_{doe}, \mathbf{y} = (\mathbf{y}_b \ \mathbf{y}_h)\}$ sont données par les

expressions :

$$\mu_{h,n}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + \mathbf{k}(\mathbf{x})^T \mathbf{V}^{-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\beta}) \quad (3.20)$$

$$\sigma_{h,n}^2(\mathbf{x}) = \rho^2 \sigma_{Z,b}^2 + \sigma_{Z,h}^2 - \mathbf{k}(\mathbf{x})^T \mathbf{V}^{-1} \mathbf{k}(\mathbf{x}) \quad (3.21)$$

où ρ est un facteur d'échelle, $\boldsymbol{\beta}$ et σ_Z sont des hyperparamètres, $\mathbf{f}(\mathbf{x})^T = (\rho \mathbf{f}_b(\mathbf{x})^T \quad \mathbf{f}_h(\mathbf{x})^T)$, $\mathbf{k}(\mathbf{x})^T = \text{CoVar}(\mathcal{G}_{n,h}(\mathbf{x}), \mathbf{y})$, \mathbf{H} est une matrice et \mathbf{V} est la matrice de covariance aux points du DoE.

Dans ces travaux, le modèle de co-krigeage est construit en utilisant la classe MFK de la librairie opensource Python Surrogate Modeling Toolbox [BOUHLEL et al., 2019].

3.3.1.3.2 Première reformulation possible

Dans cette étude, les modèles de co-krigeage sont des modèles à deux niveaux de fidélité en considérant pour les deux niveaux :

- données haute fidélité : solutions du modèle numérique complet
- données basse fidélité : solutions calculées en projection sur la base réduite disponible

Étant donnée la forte variance du modèle de co-krigeage appris sur ce problème et la forme de la fonction d'apprentissage $U(\mathbf{x}) = \frac{|\mu_n(\mathbf{x})|}{\sigma_n(\mathbf{x})}$, l'utilisation du co-krigeage dans l'ensemble de l'algorithme à la place du krigeage (GP) n'a pas permis un apprentissage efficace du modèle.

Les résultats obtenus en se basant sur 20 répétitions de l'algorithme avec le co-krigeage conduisent aux conclusions suivantes :

- l'algorithme d'apprentissage converge en moyenne au bout de 231 itérations (en moyenne 40 itérations sont nécessaires lorsque qu'un GP est utilisé avec un seul niveau de fidélité)
- 4 exécutions (sur les 20) ont été stoppées après avoir atteint le nombre maximal d'itérations autorisés, qui était fixé à 500.

Les causes potentielles de l'échec de cette stratégie sont diverses :

- le co-krigeage fonctionne mieux quand le ratio entre le nombre de points basse fidélité et haute fidélité se trouve dans une certaine gamme de valeurs, comme le montre l'étude de l'article [TOAL, 2015]. Or au début de l'algorithme d'apprentissage, ce ratio est loin des valeurs optimales pour le co-krigeage car on dispose de très peu de solutions basse-fidélité provenant du modèle réduit.

- Le problème considéré est simplifié, il s'agit en réalité d'un problème à n_r niveaux de fidélité et non à seulement deux niveaux de fidélité comme l'hypothèse qui a été faite ici. Lorsque la base réduite est enrichie la corrélation entre la base réduite et la solution complète peut possiblement complètement changer entre un nouveau point et un point se trouvant dans son voisinage qui a été calculé avec une base réduite précédente. Ainsi, le modèle de co-krigeage est amené à apprendre des points très proches mais corrélés de façon très différentes aux points haute fidélité. Ce qui peut perturber le processus d'apprentissage.
- L'algorithme d'apprentissage actif pour les analyses de fiabilité a tendance à concentrer les échantillons d'apprentissage au voisinage proche de l'état limite. La plupart de ces échantillons seront en fait des données basse fidélité évaluées par le modèle réduit. Cet échantillonnage très concentré n'est probablement pas adapté pour un apprentissage raisonnable des hyperparamètres du co-krigeage.

3.3.1.3.3 Deuxième reformulation possible

Les performances du co-krigeage ont été comparées à celles du modèle de GP en suivant la procédure suivante :

- Exécution de l'algorithme de couplage avec un GP à un seul niveau de fidélité (cf Section 3.2.4) et sauvegarde des points haute et basse fidélité.
- Après convergence de l'algorithme d'apprentissage, un modèle de co-krigeage à 2 niveaux est construit à partir des données haute et basse fidélité. Notons toujours qu'en réalité il y a n_r niveaux de fidélité du fait que la base est améliorée au cours de l'algorithme mais le problème est simplifié à seulement 2 niveaux.
- Évaluation des probabilités de défaillance sur une population de MC fixée en utilisant d'une part le modèle de GP et d'autre part le modèle de co-krigeage afin de comparer les estimations de P_f .

La procédure précédemment décrite a été répétée 50 fois pour différents DoE initiaux de taille 14 et différents populations de MC de taille $n_{MC} = 10^6$. Afin de comparer les métamodèles l'erreur relative absolue entre la probabilité de référence \hat{P}_f^{MC} (obtenue par une méthode de MCS) et la probabilité de défaillance estimée $e_r = \frac{|\hat{P}_f^{MC} - \hat{P}_f^{AK-MCS+BR}|}{\hat{P}_f^{MC}}$ a été calculée pour chaque métamodèle. Les résultats obtenus sont donnés dans le Tab. 3.5.

Finalement, les résultats montrent que les probabilités de défaillance estimées avec le co-krigeage sont moins précises en moyenne que celles obtenues avec le modèle de GP.

	GP	co-krigeage
erreur relative absolue maximale	0.15	1.85
erreur relative absolue moyenne	4.18×10^{-2}	2.77×10^{-1}

TABLEAU 3.5 – Résultats de la 2^{ème} stratégie pour le co-krigeage

La mauvaise qualité des prédictions du co-krigeage peut potentiellement s’expliquer par l’erreur importante sur les premières solutions en base réduite (calculées sur des bases réduites disponibles insuffisantes) qui sont très éloignées des vraies solutions du problème et qui perturbent aussi significativement le processus d’apprentissage.

Le problème de la variabilité des corrélations quand la base réduite est enrichie, qui a été discuté pour la première stratégie, peut aussi être une des raisons pour laquelle la qualité des estimations obtenues avec le co-krigeage est médiocre. L’échec de ces deux premières stratégies montre ainsi notamment que le problème est bien un problème à n_r niveaux de fidélité et que considérer toutes les solutions en base réduite comme issues d’un seul niveau de fidélité ne permet pas d’obtenir des résultats acceptables.

3.3.1.3.4 Troisième reformulation possible

Enfin, les performances du co-krigeage ont été comparées à celles de la régression par GP en suivant la procédure suivante :

- Exécution de l’algorithme de couplage avec un GP à un seul niveau de fidélité (cf Section 3.2.4) et sauvegarde des points haute et basse fidélité.
- Après convergence de l’algorithme d’apprentissage : réévaluation des points basse fidélité à partir de la base réduite finale afin d’avoir uniquement deux niveaux de fidélité.
- Construction d’un modèle de co-krigeage à partir des données haute et basse fidélité.
- Évaluation des probabilités de défaillance sur une population de MC fixée en utilisant d’une part le modèle de GP et d’autre part le modèle de co-krigeage afin de comparer les estimations de P_f .

La procédure précédemment décrite a été répétée 50 fois pour différents DoE initiaux de taille 14 et différents populations de MC de taille $n_{MC} = 10^6$. Afin de comparer les métamodèles l’erreur relative absolue $e_r = \frac{|\hat{P}_f^{MC} - \hat{P}_f^{AK-MCS+BR}|}{\hat{P}_f^{MC}}$ entre la probabilité de référence \hat{P}_f^{MC} (obtenue par une méthode de MCS) et la probabilité de défaillance estimée en suivant la procédure a été calculée pour chaque métamodèle. Les résultats obtenus sont donnés dans le Tab. 3.6.

	GP	co-krigeage
erreur relative absolue maximale	0.15	0.15
erreur relative absolue moyenne	4.18×10^{-2}	4.20×10^{-2}

TABEAU 3.6 – Résultats de la 3^{ème} stratégie pour le co-krigeage

Le co-krigeage construit en utilisant cette stratégie permet d’obtenir une estimation précise de la probabilité de défaillance. Cependant, les résultats donnés dans le Tab. 3.6 montre que le co-krigeage n’améliore pas la précision de l’estimation alors que son utilisation augmente le coût de calcul puisque les solution réduites doivent être réévaluées sur la base finale. Ce faible écart peut s’expliquer par le fait que les solutions en base réduite sont in fine très précises, ne justifiant pas ici le recours au co-krigeage.

Cette troisième stratégie a en revanche pour intérêt de fournir un moyen de vérifier l’estimation finale de la probabilité de défaillance obtenue avec l’approche basée sur un GP en contrôlant qu’elle ne diffère pas significativement de celle obtenue avec l’approche utilisant le co-krigeage.

3.3.1.3.5 Bilan

Parmi les trois stratégies d’étude du problème avec le co-krigeage, la seule stratégie donnant des résultats satisfaisants est la troisième stratégie, qui consiste à utiliser un modèle de GP dans la phase d’apprentissage, puis à réévaluer les points basse fidélité à l’aide de la base réduite finale pour construire un modèle de co-krigeage en se basant sur ces données basse fidélité ainsi que les données haute fidélité du modèle complet. Cependant, l’utilisation du co-krigeage n’améliore pas de façon significative la précision de l’estimation pour ce problème alors qu’elle implique une légère augmentation du coût numérique puisque les solutions réduites doivent être réévaluées. Ainsi, on choisit de conserver la stratégie basée sur la modélisation par GP pour la méthode proposée et l’étude du co-krigeage ne sera pas réitérée pour la seconde application.

3.3.2 Second exemple : application à la rupture d’une plaque composite trouée soumise à un chargement complexe

3.3.2.1 Description du problème

Dans cette section, on considère l’application des méthodes proposées à un problème mécanique impliquant une plaque stratifiée en matériaux composites de type carbone-epoxy, potentiellement non-symétrique à cause des incertitudes dans l’empilement des plis du composite. La structure étudiée, illustrée sur la Figure 3.11, est une plaque stratifiée de 6 plis trouée, encadrée sur un côté et soumise à une pression uniformément

répartie sur la partie supérieure et à du cisaillement sur le côté opposé à l'encastrement. L'analyse de fiabilité de cette structure est effectuée en considérant le critère de défaillance de Tsai-Hill, rappelé ci-dessous :

$$\left(\frac{\sigma_L(\boldsymbol{\theta}, \mathbf{h})}{X_{ult}}\right)^2 + \left(\frac{\sigma_T(\boldsymbol{\theta}, \mathbf{h})}{Y_{ult}}\right)^2 + \left(\frac{\tau_{LT}(\boldsymbol{\theta}, \mathbf{h})}{\tau_{ult}}\right)^2 - \frac{\sigma_L(\boldsymbol{\theta}, \mathbf{h})\sigma_T(\boldsymbol{\theta}, \mathbf{h})}{X_{ult}^2} > 1 \quad (3.22)$$

avec X_{ult} , Y_{ult} and τ_{ult} les contraintes ultimes et σ_L , σ_T et τ_{LT} respectivement les contraintes longitudinales, transversales et de cisaillement.

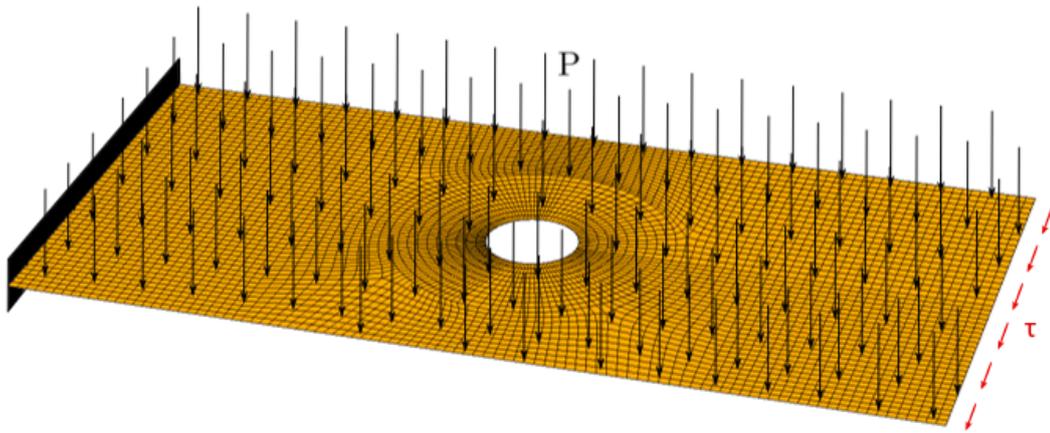


FIGURE 3.11 – Conditions aux limites et chargement imposés à la plaque stratifiée.

Les paramètres $\boldsymbol{\theta}$ et \mathbf{h} représentent respectivement les angles d'orientation des fibres et les épaisseurs de chaque pli. Ces paramètres sont considérés aléatoires du fait de variabilités de fabrication et sont modélisés par des variables aléatoires distribuées selon les lois suivantes :

- les épaisseurs des plis h_i suivent des lois Gamma $\Gamma(\mu, \sigma, \gamma)$ dont les paramètres sont donnés dans le Tab. 3.7,
- les angles d'orientation des plis θ_i suivent des lois uniformes $\mathcal{U}(a, b)$ dont les paramètres sont donnés dans le Tab. 3.8.

μ	$2 \cdot 10^{-4}$
σ	$2 \cdot 10^{-5}$
γ	0

TABLEAU 3.7 – Paramètres de la loi de probabilité gamma utilisée pour modéliser les épaisseurs h des plis.

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
a	-2.5	42.5	-47.5	-47.5	42.5	-2.5
b	2.5	47.5	-42.5	-42.5	47.5	2.5

TABLEAU 3.8 – Paramètres de la loi de probabilité uniforme utilisée pour modéliser les angles d’orientation des fibres.

Notons que le stratifié nominal contient des plis de 0.2 mm d’épaisseur et a une séquence d’empilement $[0, 45, -45]_s$. Cependant, pour le stratifié réel les valeurs des épaisseurs et des orientations peuvent différer à cause des variabilités. Plus particulièrement, le stratifié réel est non-symétrique en général. Un code de calcul **EF** développé en MATLAB est utilisé ici pour calculer le champ de contraintes. Nous avons utilisé des éléments de plaque de type Mindlin à 4 noeuds avec 5 **ddl** par noeud et un facteur de correction en cisaillement calculé en se basant sur [FERREIRA, 2009]. Le maillage **EF** de la plaque stratifiée utilisé ici est illustré sur la Figure 3.12.

Les constantes élastiques d’un pli du stratifié sont données dans le Tab. 3.9. Les contraintes ultimes longitudinales en traction et en compression (resp. X_{ult}^T et X_{ult}^C), les contraintes ultimes transversales en traction et en compression (resp. Y_{ult}^T et Y_{ult}^C) et la contrainte ultime de cisaillement dans le plan (τ_{ult}) du pli sont données dans le Tab. 3.10.

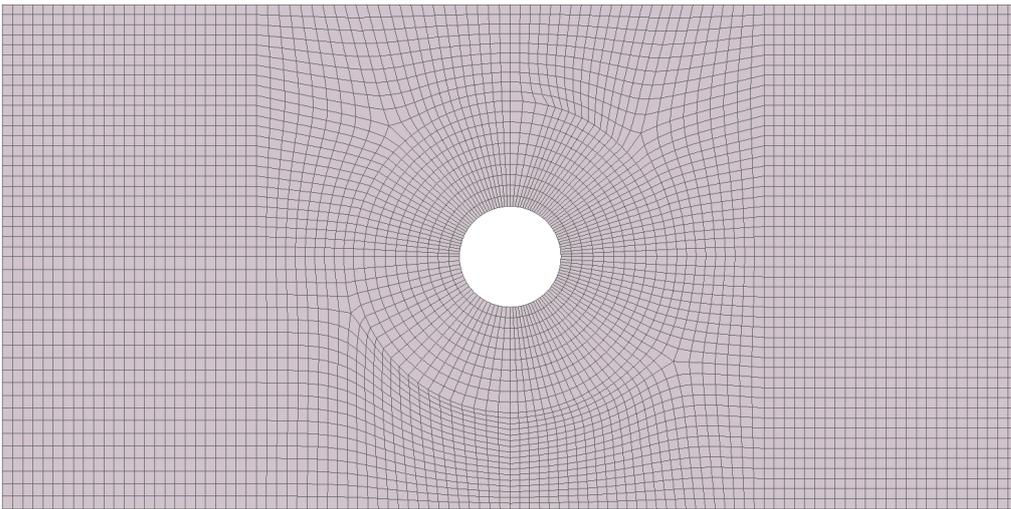


FIGURE 3.12 – Maillage **EF** de la plaque stratifiée trouée.

Dans la section suivante, la probabilité de défaillance sera estimée avec les méthodes **MCS**, **AK-MCS** et les stratégies proposées AK-MCS + BR afin de tester les performances du couplage sur le problème de mécanique présenté ci-dessus.

E_1	E_2	ν_{12}	G_{12}
181 GPa	10.3 GPa	0.28	7.17 GPa

TABLEAU 3.9 – Constantes élastiques d’un pli.

X_{ult}^T	X_{ult}^C	Y_{ult}^T	Y_{ult}^C	τ_{ult}
997 MPa	847 MPa	38 MPa	198 MPa	60 MPa

TABLEAU 3.10 – Contraintes ultimes d’un pli.

Le champ de déplacement du stratifié nominal est illustré dans la Figure 3.13.

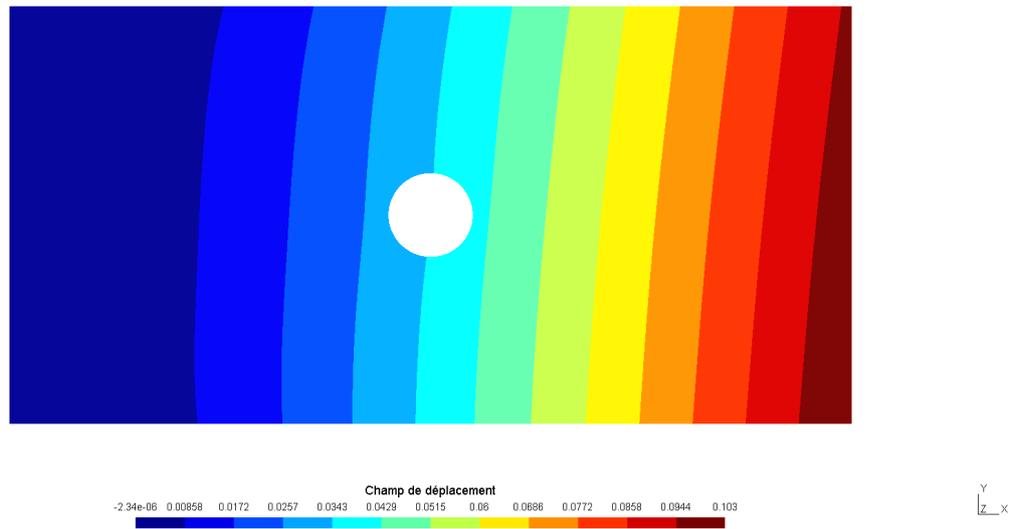


FIGURE 3.13 – Champ de déplacement du stratifié nominal.

3.3.2.2 Résultats

D’abord, une estimation de référence de la probabilité de défaillance à partir de la méthode MCS a été calculée sur ce problème. L’estimation obtenue est $\hat{P}_{f_{MC}} = 7.6 \times 10^{-4}$ avec COV de MC estimé à 9.36% (pour $n_{MC} = 1.5 \times 10^5$).

Une étude similaire à celle effectuée pour le problème thermique de la Section 3.3.1 a donc été menée pour analyser ce problème.

3.3.2.2.1 Étude du critère de couplage

Tout d’abord, la corrélation entre les résidus et la vraie erreur a été estimée.

La Figure 3.14 présente les résultats de cette estimation, obtenus pour différentes tailles de base réduite. Sur cette figure, on observe graphiquement que la corrélation

entre le résidu non préconditionné et la vraie erreur n'est pas significative, ce qui est vérifié par des coefficients de corrélation relativement faibles. De plus, on remarque qu'un grand nombre des résidus relatifs calculés prennent des valeurs supérieures à 1 et ce indépendamment de la taille de la base réduite. Ces valeurs sont certainement liées au mauvais conditionnement du problème, en comparaison au problème thermique précédent. Ce mauvais conditionnement est dû à l'absence de symétrie du stratifié considéré, qui induit un couplage flexion-torsion. Sur la base de ces résultats, il serait difficile de trouver une valeur pertinente pour le seuil ϵ afin d'assurer l'efficacité du couplage.

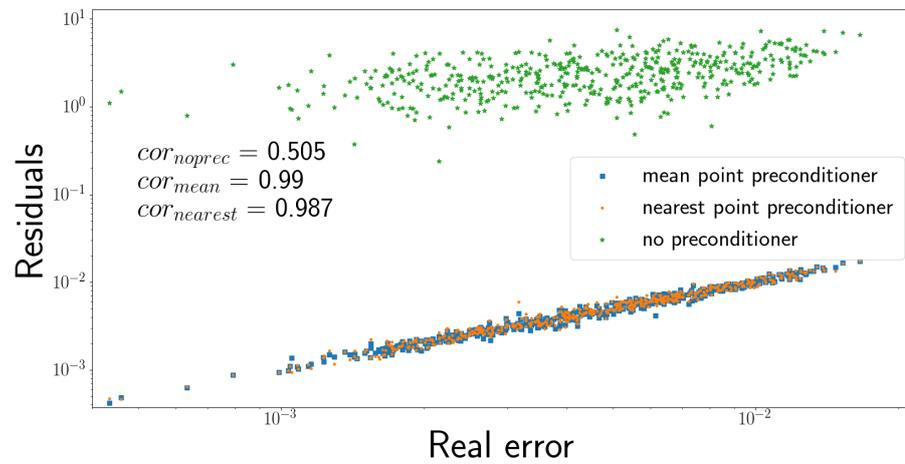
Ensuite, la corrélation avec les résidus préconditionnés a été évaluée et les résultats obtenus sont aussi présentés sur la Figure 3.14. On peut remarquer que les deux préconditionneurs (point moyen et point le plus proche) ont des performances similaires pour ce problème avec des coefficients de corrélation supérieurs à 0.9, ce qui représente une amélioration considérable par rapport aux résultats obtenus avec le résidu non préconditionné.

3.3.2.2 Étude des performances du couplage

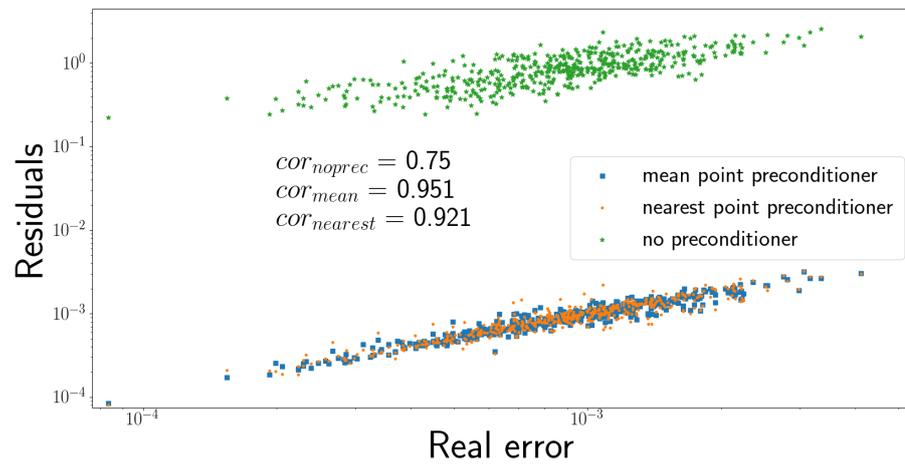
Comme pour le problème thermique, les premiers modes du champ de déplacement construits à la volée au cours d'une exécution de l'algorithme de couplage, basé sur ϵ_{BR}^P pour $\epsilon = 10^{-3}$, peuvent être représentés graphiquement. Les représentations des quatre premiers vecteurs de la base réduite du champ de déplacement dans la direction hors plan sont données sur la Figure 3.15. Les premiers modes ont évidemment beaucoup de similarités avec la forme générale du champ de déplacement, qui est asymétrique à cause du chargement combiné en flexion et cisaillement. Concernant, les modes plus élevés (au sens de l'ordre des éléments de la base) représentés ils décrivent des variations plus complexes des champs de déplacement qui peuvent être liés au couplage flexion-torsion découlant du caractère possiblement non-symétrique du stratifié. Ainsi, le comportement relativement complexe de ce problème amène à penser qu'un nombre plus important de modes est nécessaire pour satisfaire le critère d'erreur $\epsilon = 10^{-3}$.

Les performances des méthodes proposées en fonction du seuil sur le résidu préconditionné ϵ_{BR}^P ont ensuite été étudiées en suivant la même procédure que pour la première application, c.-à-d. en utilisant la stratégie de couplage sur un DoE fixe généré au cours d'une exécution de AK-MCS pour différentes valeurs du seuil. La méthode AK-MCS + BR sans préconditionneur n'a pas été retenue pour la suite de l'étude sur cet exemple étant donné les résultats obtenus lors de l'étude du critère de couplage.

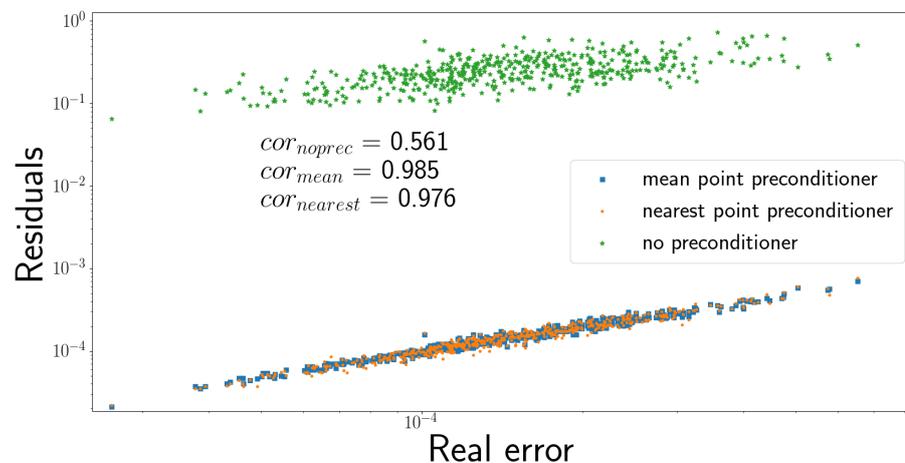
Les valeurs moyennes de l'accélération et de la précision calculées sur 10 exécutions en fonction de ϵ sont données sur la Figure 3.16 pour le préconditionneur au point moyen et sur la Figure 3.17 pour le préconditionneur du point le plus proche.



Taille de la base réduite = 5



Taille de la base réduite = 10



Taille de la base réduite = 20

FIGURE 3.14 – Corrélation entre les résidus et la vraie erreur pour le problème lié à la plaque stratifiée trouée. Trois tailles de base réduite sont considérées (5, 10 et 20).

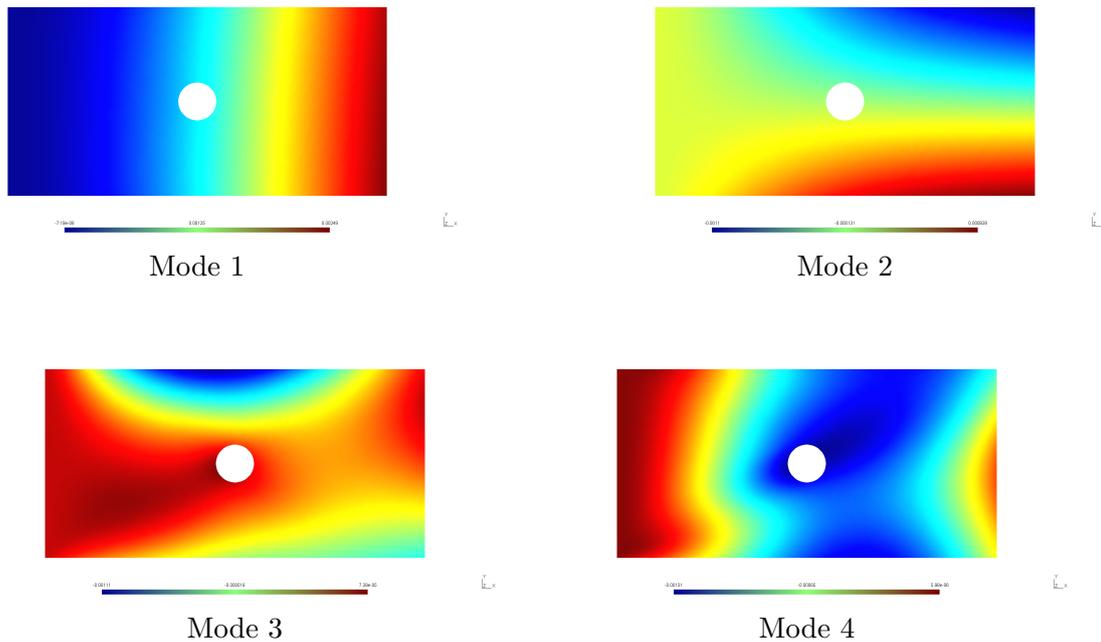


FIGURE 3.15 – Quatre premiers modes de déplacement selon l’axe Z obtenus en utilisant la procédure de construction à la volée au sein de l’algorithme AK-MCS pour le problème de la plaque stratifiée trouée.

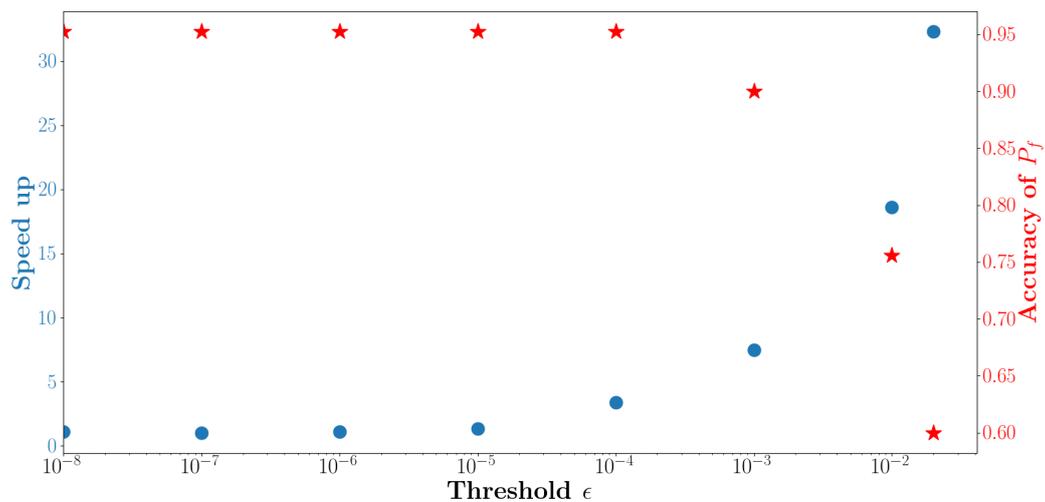


FIGURE 3.16 – Précision (étoiles rouges) de P_f et accélération (points bleus) de la méthode proposée en fonction du seuil ϵ sur le résidu preconditionné pour le preconditionneur au point moyen \mathbf{P}_{mean} pour différents DoE et populations de MC générés par l’algorithme AK-MCS pour le problème de la plaque stratifiée trouée.

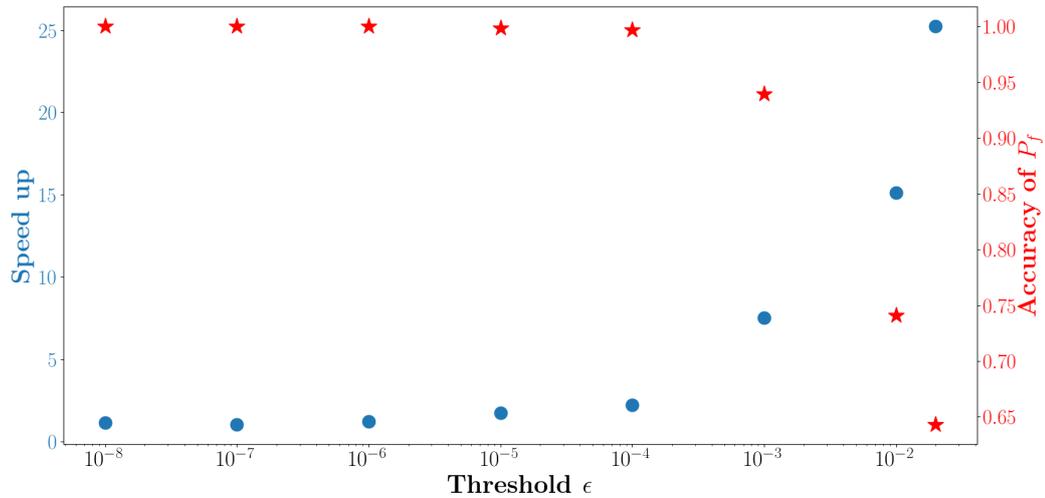


FIGURE 3.17 – Précision (étoiles rouges) de P_f et accélération (points bleus) de la méthode proposée en fonction du seuil ϵ sur le résidu préconditionné pour le préconditionneur calculé au point le plus proche $\mathbf{P}_{nearest}$ pour différents DoE et populations de MC générés par l’algorithme AK-MCS pour le problème de la plaque stratifiée trouée.

En s’appuyant sur la Figure 3.16 et sur la Figure 3.17, le meilleur compromis entre la précision des estimations et une accélération conséquente de la méthode d’analyse de fiabilité semble à nouveau correspondre à un seuil $\epsilon = 10^{-3}$ pour les deux préconditionneurs. Ce résultat s’accorde avec l’hypothèse formulée dans la Section 3.2.5, selon laquelle fixer le seuil ϵ à une valeur jugée acceptable de l’erreur relative sur le champ de déplacement mène à une erreur raisonnable sur l’estimation de la probabilité de défaillance (avec un préconditionneur permettant d’avoir un bon conditionnement de $\mathbf{P}^{-1}\mathbf{K}(\mathbf{x})$). Ainsi, une valeur limite de 10^{-3} pour l’erreur relative sur le champ de déplacement semble être une valeur de départ raisonnable pour un nouveau problème, sur lequel on ne dispose pas de données préalables.

L’algorithme AK-MCS et la méthode proposée AK-MCS + BR ont ensuite été exécutés 40 fois pour différentes populations de MC et différents DoE initiaux avec $\epsilon = 10^{-3}$. La moyenne $\overline{P_f}$ et l’écart type $\overline{\sigma_{P_f}}$ qui en résultent sont donnés dans le Tab. 3.11 pour le préconditionneur au point moyen et dans le Tab. 3.12 pour le préconditionneur au plus proche. La moyenne des estimations de la probabilité de défaillance par la méthode proposée se trouve dans l’intervalle de confiance à 95% obtenus pour les estimations par AK-MCS $[8.20 \times 10^{-4}, 8.80 \times 10^{-4}]$. De plus, les valeurs des écarts types sont du même ordre de grandeur avec ou sans base réduite. En moyenne, le recours à la méthode proposée AK-MCS + BR par rapport à l’utilisation directe de AK-MCS permet une accélération de 6.8 et 6.7 respectivement pour les préconditionneurs du point moyen et du point le plus proche. De plus, les évaluations du modèle complet représentent 16.2% et 14.4% du nombre total de simulations lors d’une exécution de

l'algorithme de couplage avec les préconditionneurs respectivement du point moyen et du point le plus proche.

	\overline{P}_f	$\overline{\sigma}_{P_f}$	n_{sim}	n_{full}
MCS	7.6×10^{-4}	7.1×10^{-5}	1.5×10^5	1.5×10^5
AK-MCS	8.47×10^{-4}	5.58×10^{-5}	99.95	99.95
AK-MCS + BR - \mathbf{P}_{mean}	8.48×10^{-4}	7.14×10^{-5}	98.87	15.97

TABLEAU 3.11 – Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point moyen pour le problème de la plaque stratifiée trouée.

	\overline{P}_f	$\overline{\sigma}_{P_f}$	n_{sim}	n_{full}
MCS	7.6×10^{-4}	7.1×10^{-5}	1.5×10^5	1.5×10^5
AK-MCS	8.50×10^{-4}	6.36×10^{-5}	100.2	100.2
AK-MCS + BR - $\mathbf{P}_{nearest}$	8.23×10^{-4}	6.47×10^{-5}	105.7	14.97

TABLEAU 3.12 – Résultats de MCS et résultats moyens de AK-MCS et AK-MCS + BR avec le préconditionneur du point le plus proche pour le problème de la plaque stratifiée trouée.

3.3.3 Gain en temps de calcul pour des problèmes de grande dimension

Les études numériques précédentes ont été menées pour analyser les performances du couplage avec la modélisation en base réduite du point de vue de la précision des résultats et de la réduction de nombre de résolutions du système complet. Afin de pouvoir comparer les résultats aux vraies valeurs de probabilités de défaillance obtenues par MCS, des modèles avec relativement peu de ddl ont été considérés pour les deux problèmes étudiés.

Cependant, le gain de temps CPU permis par l'utilisation de solutions réduites est particulièrement intéressant pour des problèmes de grande dimension (plusieurs millions de ddl). Dans cette section, des maillages raffinés sont donc utilisés pour comparer la durée en temps CPU du couplage avec le préconditionneur au point moyen et celle de la méthode classique AK-MCS sur les deux problèmes précédents.

3.3.3.1 Premier exemple : Analyse de fiabilité d'un problème thermique

Le problème thermique admettant une décomposition affine simple de la matrice de rigidité $\mathbf{K}(x) = \sum_i \gamma_i(x) \mathbf{K}_i$, l'assemblage des matrices \mathbf{K}_i est effectué une fois pour

toutes. Les temps CPU d’assemblage et de résolution des systèmes par le code de calcul *Code Aster* [ELECTRICITÉ DE FRANCE, 1989–2017] sont fournis, afin d’avoir une valeur représentative du temps CPU nécessaire à l’assemblage de la matrice par un code de calcul EF industriel (et non pas par notre code Python). Le temps CPU donné dans le Tab. 3.13 correspond à un problème à 1,321,537 ddl.

	AK-MCS + BR	AK-MCS
Temps d’assemblage de \mathbf{K}_i (une seule fois)	3.20	3.20
Nombre d’évaluations de $G(\mathbf{x})$	40	38
Nombre de solutions réduites $\mathbf{u}_r(\mathbf{x})$ évaluées	40	-
Nombre de solutions complètes $\mathbf{u}(\mathbf{x})$ évaluées	6	38
Temps de projection $\Phi^T \mathbf{K}_i \Phi$ avec la base réduite initiale	0.6	-
Temps de projection $\Phi^T \mathbf{K}_i \Phi$ avec la base réduite finale	1.8	-
Temps moyen d’évaluation d’une solution complète $\mathbf{u}(\mathbf{x})$	43.6	43.6
Temps moyen d’évaluation d’une solution réduite $\mathbf{u}_r(\mathbf{x})$	0.01	-
Temps moyen d’évaluation du résidu préconditionné $\epsilon_{BR}^P(\mathbf{x})$	1.3	-
Temps CPU total	324	1657

TABLEAU 3.13 – Temps CPU (en secondes) des opérations pour l’analyse de fiabilité du problème de thermique en grande dimension.

Ainsi, l’accélération en terme de temps CPU est environ égale à 5 pour ce problème.

3.3.3.2 Second exemple : application à la rupture d’une plaque composite trouée soumise à un chargement complexe

Concernant cette application, la matrice de rigidité $\mathbf{K}(\mathbf{x})$ doit être assemblée à chaque point \mathbf{x} évalué durant la procédure d’apprentissage car la décomposition affine n’est pas immédiate. En revanche, le vecteur des forces $\mathbf{F}(\mathbf{x}) = \mathbf{F}$ ne dépend pas de \mathbf{x} dans ce cas. Ainsi la valeur $\mathbf{P}^{-1}\mathbf{F}$ peut être évaluée une fois pour toutes au début de l’algorithme. Le temps CPU donné dans le Tab. 3.14 correspond à une étude pour un modèle à 1,029,700 ddl.

Finalement, l’accélération en terme de temps CPU est de 5.4 pour ce problème.

Jusqu’ici le couplage entre la modélisation en base réduite et les approches d’apprentissage actif de GP a été effectué en se basant sur la méthode AK-MCS. Dans la section suivante, le couplage entre l’approche Vb-AGP présentée dans le chapitre 2 et la réduction de modèles est étudiée sur un cas d’application.

	AK-MCS + BR	AK-MCS
Temps de calcul de $\mathbf{P}^{-1}\mathbf{F}$ (une seule fois)	4.84	-
Temps moyen de l'assemble de $\mathbf{K}(\mathbf{x})$	86	86
Nombre d'évaluations de $G(\mathbf{x})$	101	101
Nombre de solutions réduites $\mathbf{u}_r(\mathbf{x})$ évaluées	101	-
Nombre de solutions complètes $\mathbf{u}(\mathbf{x})$ évaluées	15	101
Temps moyen d'évaluation d'une solution complète $\mathbf{u}(\mathbf{x})$	2041	2041
Temps moyen d'évaluation d'une solution réduite $\mathbf{u}_r(\mathbf{x})$	1.28	-
Temps moyen d'évaluation du résidu préconditionné $\epsilon_{BR}^P(\mathbf{x})$	2.92	-
Temps CPU total	39734	214838

TABLEAU 3.14 – Temps CPU (en secondes) des opérations pour l'analyse de fiabilité du problème mécanique en grande dimension.

3.3.4 Application du couplage entre la méthode Vb-AGP et de la modélisation en base réduite

Dans cette section, on s'intéresse au couplage entre les approches Vb-AGP (version MCS ou IS) et la modélisation en base réduite implémentée selon le concept proposé dans la Section 3.2.2 avec le critère de couplage basé sur le résidu préconditionné (voir Section 3.2.5). Les performances obtenues en couplant ces deux approches sont étudiées sur le cas test thermique, décrit dans la Section 3.3.1. Néanmoins, comme nous l'avons expliqué dans le chapitre 2, l'efficacité de l'algorithme Vb-AGP est fortement réduite lorsque la dimension d'entrée du problème augmente. Ainsi, afin de pouvoir appliquer la méthode efficacement sur ce cas d'application, une analyse de sensibilité fiabiliste [LEMAÎTRE, 2014; RAGUET et MARREL, 2018] est d'abord effectuée afin de pouvoir réduire le nombre d'entrées du problème.

Les indices de Sobol S_i du 1^{er} ordre [PERRIN et DEFAUX, 2019; GAMBOA et al., 2016] et les indices totaux S_{T_i} [JANON, 2012] sur la variable aléatoire $\mathbb{1}_{G(\mathbf{X}) \leq 0}$ sont utilisés pour mesurer la sensibilité de l'estimateur de la probabilité de défaillance aux paramètres d'entrée :

$$S_i = \frac{Var_{\mathbf{X}^i} \left(\mathbb{E}_{\mathbf{X}^{-i}} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} | \mathbf{X}^i \right] \right)}{Var_{\mathbf{X}} \left(\mathbb{1}_{G(\mathbf{X}) \leq 0} \right)} \quad (3.23)$$

et :

$$S_{T_i} = 1 - \frac{Var_{\mathbf{X}^{-i}} \left(\mathbb{E}_{\mathbf{X}^i} \left[\mathbb{1}_{G(\mathbf{X}) \leq 0} | \mathbf{X}^{-i} \right] \right)}{Var \left(\mathbb{1}_{G(\mathbf{X}) \leq 0} \right)} \quad (3.24)$$

où \mathbf{X}^{-i} correspond au vecteur aléatoire des paramètres \mathbf{X} sans le i -ème paramètre \mathbf{X}^i .

Les indices de Sobol du 1^{er} ordre et les indices totaux calculés par *pick freeze* sur le métamodèle de la fonction de performance sont donnés dans le Tab. 3.15.

	k_{Cu}	k_{Ni}	T_{hot}	h_{hot}	T_{out}	h_{out}	T_{cool}	h_{cool}	T_{allow}
Sobol 1 ^{er} ordre	1.37×10^{-5}	1.2×10^{-3}	3.76×10^{-2}	8.14×10^{-3}	7.53×10^{-5}	8.75×10^{-5}	9.9×10^{-5}	1.03×10^{-4}	3.4×10^{-2}
Sobol total	5.9×10^{-3}	0.24	0.89	0.64	7.7×10^{-3}	3.7×10^{-3}	2.3×10^{-2}	4.9×10^{-3}	0.87

TABLEAU 3.15 – Résultats de l’analyse de sensibilité fiabiliste.

Au vu des résultats du Tab. 3.15, on considérera pour l’étude du couplage entre l’approche **Vb-AGP** et la modélisation en base réduite que les paramètres k_{Cu} et h_{out} sont fixés à leur valeur nominale. Le problème considéré comporte donc 7 variables d’entrées incertaines.

Dans un premier temps, la précision de l’estimation de la probabilité de défaillance P_f par la méthode **Vb-AGP + IS** avec une valeur de COV_{max} fixée à 3% couplée à la modélisation en base réduite avec le résidu préconditionné \mathbf{P}_{mean} a été étudiée pour différentes valeurs du seuil ϵ . La méthode **Vb-AGP + MCS** n’a pas été utilisée car pour la dimension stochastique et le **COV** fixé considérant l’ordre de grandeur de la probabilité à estimer (c.-à-d. la taille de la population de **MC** nécessaire), la génération de trajectoires suffisamment précises (plus précisément la construction de la matrice $\mathbf{C}(\tilde{\mathbf{X}})$, comme discuté dans la Section 2.4.1) était trop gourmande en mémoire.

La Figure 3.18 représente les intervalles de confiance à 98% de la précision de l’estimation par la méthode **Vb-AGP + IS + BR** pour des **DoE** initiaux et des échantillonnages différents par rapport à la moyenne obtenue sur 70 répétitions de la méthode **AK-MCS** en fonction de différentes valeurs de ϵ . Notons que chaque intervalle est évalué à partir des différents échantillons $\hat{P}_f(G_i, \tilde{\mathbf{X}}_i)$ (définis par l’Eq. (2.22)), où les $(G_i, \tilde{\mathbf{X}}_i)$ sont des réalisations de \mathcal{G}_n et de populations de **MC** $\tilde{\mathbf{X}}$, calculées pour l’estimation finale de la probabilité de défaillance par une exécution de la méthode par **Vb-AGP + IS + BR**.

On peut observer sur cette figure que pour une valeur de ϵ inférieure à 5×10^{-3} , l’erreur introduite par l’utilisation de solutions réduites a un impact faible par rapport à l’incertitude due au **GP** et au **MC** pour un critère d’arrêt $COV_{max} = 0.03$.

La méthode **Vb-AGP + IS + BR** avec le résidu préconditionné \mathbf{P}_{mean} a été appliquée à 70 reprises sur le problème thermique à 7 variables pour un seuil $\epsilon = 10^{-3}$, un **DoE** initial de 14 points et une valeur du **COV** limite de 3% en faisant varier les **DoE** initiaux et populations de **MC** initiales. De même les méthodes **AK-MCS + U**, **AK-MCS + U + BR** et **AK-MCS + EFF + BR** ont été répétées 70 fois pour des **DoE** initiaux de 14 points et une valeur du **COV** de **MC** COV_{MC} limite de 3%.

Pour ce cas test, l’efficacité numérique est mesurée par deux indices ν_{MC} (voir définition Section 2.4.1) en terme de nombre total de simulations $\nu_{MC}^{n_{sim}}$ et en terme

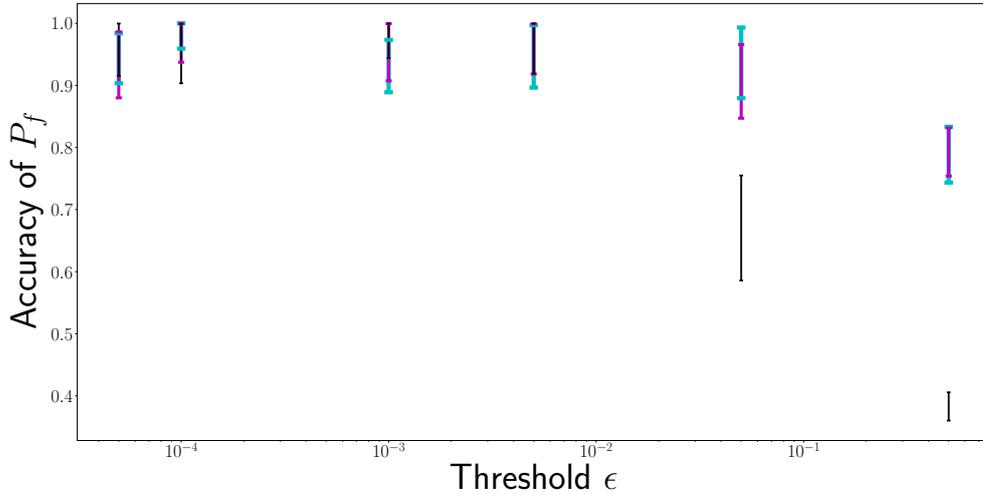


FIGURE 3.18 – Intervalles de confiance obtenus par l’approche **Vb-AGP + IS + BR** ($COV_{max} = 0.03$) de la précision de P_f en fonction du seuil ϵ sur le résidu preconditionné pour le preconditionneur calculé au point moyen \mathbf{P}_{mean} pour le problème thermique.

d’appels au modèle numérique complet $\nu_{MC}^{n_{full}}$ dont les expressions sont donnés par :

$$\nu_{MC}^{n_{sim}} = \frac{N_{sim}^{MC}}{n_{sim}} \quad (3.25)$$

et :

$$\nu_{MC}^{n_{full}} = \frac{N_{sim}^{MC}}{n_{full}} \quad (3.26)$$

où n_{sim} correspond au nombre d’appels de la méthode d’apprentissage actif à la fonction de performance, n_{full} est le nombre de points évalués par le modèle numérique complet et N_{sim}^{MC} est le nombre d’échantillons de **MC** nécessaires pour obtenir le même coefficient de variation $COV(\hat{P}_f)$ sur l’estimation de la probabilité de défaillance.

Les résultats moyens obtenus sur ces 70 répétitions de ces méthodes sont donnés dans le Tab. 3.16.

	\hat{P}_f	$COV(\hat{P}_f)$	n_{sim}	n_{full}	$\nu_{MC}^{n_{sim}}$	$\nu_{MC}^{n_{full}}$
AK-MCS	6.48×10^{-3}	2.85%	32.1 (9%)	32.1	5880	5880
AK-MCS + EFF + BR	6.54×10^{-3}	2.96%	34.9 (13%)	5.99	4980	29072
AK-MCS + U + BR	6.50×10^{-3}	2.98%	32.3 (12%)	6.03	5336	28580
Vb-AGP + IS + BR - \mathbf{P}_{mean}	6.40×10^{-3}	3.46%	15.3 (8%)	6	8480	21563

TABEAU 3.16 – Résultats moyens de **AK-MCS** et **Vb-AGP + IS + BR** (\mathbf{P}_{mean}) avec $\epsilon = 10^{-3}$ pour le problème thermique - paramètres de la méthode adaptative : $n_{MC}^{init} = 10^4$, $n_{DoE}^{init} = 14$, $COV_{max} = 0.03$

Les résultats présentés dans le Tab. 3.16 montrent qu'en terme d'efficacité numérique la méthode permet d'obtenir des indices d'efficacité ν_{MC} plus élevés en terme de nombre de simulations et d'appels au modèle numérique complet. D'une part, la méthode **Vb-AGP + IS** permet d'obtenir un coefficient $\nu_{MC}^{n_{sim}}$ 1.4 fois plus important en terme de nombre total de simulations que **AK-MCS**, 1.7 fois plus important que **AK-MCS + EFF + BR** et 1.6 fois plus important que **AK-MCS + U + BR**. D'autre part, l'utilisation de la réduction de modèles permet d'avoir un indice $\nu_{MC}^{n_{full}}$ 3.7 fois plus élevé en terme d'appels au modèle numérique complet en comparaison à **AK-MCS**. Cependant, les indices $\nu_{MC}^{n_{full}}$ des méthodes **AK-MCS + EFF + BR** et **AK-MCS + U + BR** sont plus élevés que celui de la méthode **Vb-AGP + IS + BR** pour laquelle $COV(\hat{P}_f)$ est plus élevé. Or ces **COV** sont des estimations obtenues sur 70 exécutions, cette constatation est donc à nuancer puisque l'erreur commise sur ces estimations ne permet pas de les comparer rigoureusement.

Il est donc intéressant d'observer la distributions des 70 estimations \hat{P}_f obtenues par la méthode **Vb-AGP + IS + BR**. L'histogramme des estimations \hat{P}_f de P_f obtenues à partir de l'algorithme **Vb-AGP + IS + BR** (\mathbf{P}_{mean}) est donné sur la Figure 3.19. De plus, la probabilité de référence $P_{f_{ref}}$ obtenue à partir des 70 répétitions de **AK-MCS** avec $COV_{MC} = 0.03$ et son intervalle de confiance à $\pm 3\sigma$ (où σ est son écart-type) est représentée sur la figure. On peut observer sur cet histogramme que les estimations centrées autour de la moyenne des \hat{P}_f , dont la valeur est proche de la probabilité de référence obtenue par **AK-MCS**, et que la majorité des estimations obtenues sont dans l'intervalle de confiance à $\pm 3\sigma$ de la valeur de référence.

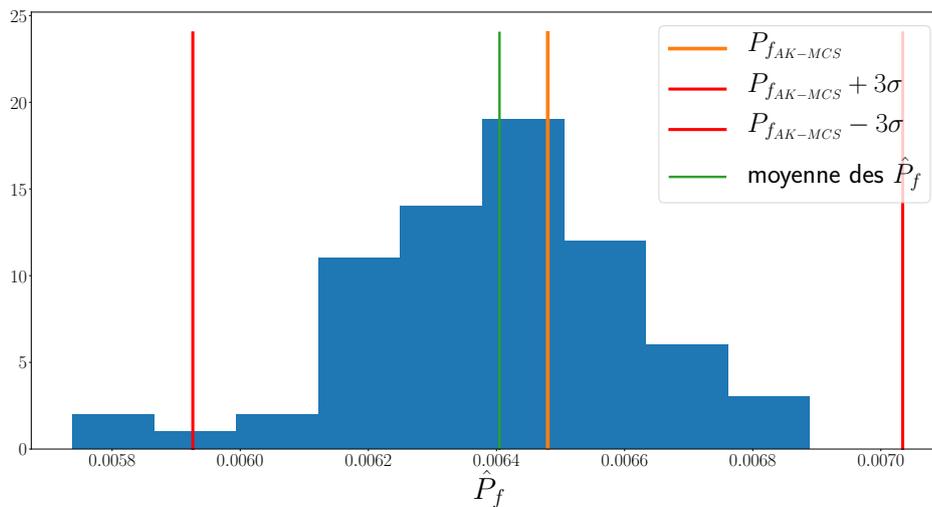


FIGURE 3.19 – Histogramme des estimations \hat{P}_f sur 70 répétitions de **Vb-AGP + IS + BR** (\mathbf{P}_{mean}) avec $\epsilon = 10^{-3}$ et $COV_{max} = 0.03$ sur le cas test thermique

3.4 Conclusion

Dans ce chapitre, des stratégies basées sur l'utilisation de solutions réduites pour l'amélioration de l'efficacité numérique des approches d'apprentissage actif pour les analyses de fiabilité ont été présentées. L'hybridation des approches d'apprentissage actif et de réduction de modèles est possible en considérant une construction adaptative des bases réduites.

La méthode proposée débute par l'initialisation d'une base réduite au cours de la phase initiale d'un algorithme d'apprentissage actif, i.e. lors de l'évaluation des points du DoE initial. Puis, pour chaque point ajouté au cours de l'apprentissage, l'algorithme décide d'utiliser une solution réduite ou de résoudre le système complet en se basant sur un critère sur la précision de la solution réduite obtenue. Les critères d'apprentissage de l'échantillonnage adaptatif étant construits pour apprendre précisément l'état limite, une large partie des points d'apprentissage se trouve alors concentrée au voisinage de l'état limite. Ainsi, la modélisation en base réduite est très efficace puisque les nouveaux points à évaluer ont de fortes chances de se trouver à proximité de points où le modèle complet a été résolu auparavant.

Deux critères de couplage basés sur une mesure de l'erreur entre la solution réduite et la vraie solution du problème par des résidus ont été proposés. Les performances d'un premier critère s'appuyant sur un résidu relatif classique puis de l'amélioration de ce critère en utilisant un préconditionnement ont été étudiées. L'application de la méthode avec le premier critère de couplage sur deux cas d'application a montré les limitations de l'utilisation d'un résidu classique pour des problèmes pour lesquels le conditionnement n'est pas suffisamment bon. Le second critère de couplage, basé sur un résidu préconditionné, a également été testé. Dans ces travaux, nous avons proposé d'utiliser deux types de préconditionneurs construits à partir de valeurs de la matrice de rigidité : le premier correspond à sa valeur au point moyen et le second correspond à sa valeur au point le plus proche (parmi les points évalués au préalable). Finalement, l'utilisation du second critère a permis d'obtenir de bons résultats pour les deux préconditionneurs proposés, dus à une bonne estimation de l'erreur, même lorsque le conditionnement de la matrice de rigidité est mauvais.

L'utilisation du résidu classique ne permet pas en général de choisir efficacement une valeur du seuil puisqu'elle ne permet pas de représenter correctement le comportement de l'erreur sur la solution réduite. L'utilisation du préconditionnement permet alors de fixer le seuil en fonction de l'erreur tolérée sur l'approximation de la solution réduite, puisque le résidu préconditionné est alors homogène au déplacement $\mathbf{u}(\mathbf{x})$. En considérant alors que la fonction de performance $G(\mathbf{u}(\mathbf{x}), \mathbf{x})$ est relativement linéaire en $\mathbf{u}(\mathbf{x})$, nous avons observé sur les applications considérées que l'on peut considérer un seuil ϵ de l'ordre de 10^{-3} . Néanmoins, dans le cas où la fonction de performance est

fortement non linéaire en $\mathbf{u}(\mathbf{x})$, une étude du seuil peut être nécessaire en débutant par la valeur recommandée de $\epsilon = 10^{-3}$.

L'étude des performances de l'approche proposée sur deux applications a montré que celle-ci avait un potentiel important de réduction des coûts de calcul, tout en donnant une estimation précise pour un seuil bien choisi. L'accélération dépend naturellement du problème étudié mais sur les applications considérées précédemment, l'accélération théorique (en terme de nombre d'évaluations du modèle complet) a atteint un facteur 7.1 et l'accélération pratique (en terme de temps CPU) a atteint un facteur 5.4. Plus le nombre de `ddl` du modèle `EF` augmente, plus l'accélération pratique de l'algorithme est susceptible d'approcher l'accélération théorique de l'algorithme.

Finalement, la méthode a été couplée à l'approche `Vb-AGP` proposée dans le [chapitre 2](#) et appliquée sur un cas thermique. Le couplage des deux méthodes a montré un fort potentiel de réduction du nombre d'appels au modèle numérique complet d'une part et d'autre part du nombre d'évaluations de solutions réduites.

Une piste d'amélioration du critère de couplage consisterait à s'intéresser à la phase d'exploration de l'apprentissage. En effet, l'exploration des régions moins connues du domaine pour clarifier les performances du système mène à ajouter des points parfois éloignés de l'état limite et donc des autres points d'apprentissage. Considérant le critère de couplage actuel, de tels points seront probablement évalués avec le modèle numérique complet alors qu'une précision approximative de la solution en ces points pourrait être suffisante pour clarifier le comportement du système dans la région concernée du domaine. L'intégration de ces considérations dans le critère de couplage pourrait être une perspective pour réduire encore davantage le nombre de résolutions du système complet.

Conclusion générale et perspectives

Conclusion

Dans cette thèse, nous nous sommes intéressés au développement de nouvelles méthodes permettant de réduire le coût numérique des méthodes d'analyse de fiabilité par apprentissage actif de GP. Dans un premier temps, nous nous sommes intéressés à l'erreur introduite par l'incertitude du métamodèle de GP sur l'estimation de la probabilité de défaillance par une méthode d'apprentissage actif de GP. Puis, nous avons étudié le couplage entre les méthodes d'analyse de fiabilité par apprentissage actif de GP et les approches de réduction de modèles.

Dans le chapitre 2, nous avons proposé d'analyser la sensibilité de l'estimateur de la probabilité de défaillance aux incertitudes liées à l'approximation par processus gaussien et à l'échantillonnage d'une population de MC. Des estimateurs permettant d'estimer numériquement les variances dues à l'effet de chaque source d'incertitudes et de la variance totale ont été proposés. En comparant les valeurs de ces variances sur un exemple, nous avons montré que la précision de l'apprentissage du GP imposée par un critère d'arrêt classique est trop conservative. L'approche proposée Vb-AGP intègre cette analyse pour améliorer adaptativement la source d'erreur majeure durant la phase d'apprentissage ainsi qu'un critère d'arrêt basé sur la variance totale de l'estimation de la probabilité de défaillance. Les résultats obtenus ont montré que les estimateurs de la variance permettent d'obtenir une estimation de la variance totale cohérente avec la valeur de la variance empirique obtenue sur plusieurs répétitions de l'algorithme. Ils ont de plus montré que cette approche est prometteuse d'un point de vue de la réduction du nombre de points d'apprentissage. L'approche a ensuite été étendue via l'IS à l'estimation de probabilités d'évènements rares, en utilisant plus précisément la

méthode **NAIS**. Les résultats ont montré que l'approche permet de respecter la précision fixée sur l'estimation et a un potentiel de réduction du nombre de simulations.

Dans le **chapitre 3**, des stratégies basées sur l'utilisation de solutions réduites pour l'amélioration de l'efficacité numérique des approches d'apprentissage actif pour les analyses de fiabilités ont été présentées. L'hybridation des approches d'apprentissage actif et de réduction de modèles par projection est possible en considérant une construction adaptative des bases réduites. La méthode proposée débute par l'initialisation d'une base réduite au cours de la phase initiale d'un algorithme d'apprentissage actif. Puis pour chaque nouveau point d'apprentissage, l'algorithme décide d'utiliser une solution réduite ou de résoudre le système complet en se basant sur un critère sur la précision de la solution réduite obtenue. Dans ce cadre, la modélisation en base réduite est très efficace puisque les nouveaux points d'apprentissage se trouvent majoritairement au voisinage de l'état limite et donc à proximité de points où le modèle complet a été résolu auparavant. Deux critères de couplage basés sur une mesure de l'erreur entre la solution réduite et la vraie solution du problème par des résidus ont été proposés : un premier critère s'appuyant sur un résidu relatif classique et un critère amélioré en utilisant un préconditionnement. Les résultats ont montré les limitations de l'utilisation du résidu classique et l'utilité de recourir à un résidu préconditionné, d'une part car il facilite le choix de la valeur seuil et d'autre part car il donne une meilleure estimation du comportement de l'erreur. L'étude des performances de l'approche proposée sur deux applications a montré que celle-ci avait un potentiel important de réduction des coûts de calcul, tout en donnant une estimation précise pour un seuil bien choisi. Finalement, le couplage de l'approche d'apprentissage actif de **GP** proposée dans le **chapitre 2** et de la modélisation en base réduite a été étudié sur un cas d'application.

Dans la suite, les perspectives induites par les différentes proposées dans ces travaux de thèse sont présentées.

Perspectives

Perspectives pour l'approche **Vb-AGP**

Pour la mise en place de la méthode **Vb-AGP + IS** présentée dans ces travaux de thèse, nous avons fait le choix d'utiliser la méthode d'**IS** adaptatif **NAIS** [**MORIO, 2012**] pour sa capacité à traiter des domaines de défaillance multimodaux sans connaissance a priori. Or, nous avons vu que cette méthode souffre du fléau de la dimension, c.-à-d. qu'elle perd beaucoup en efficacité en grande dimension, du fait de l'utilisation de l'estimation par noyau des densités auxiliaires. En ce sens, la méthode pourrait être améliorée en s'intéressant à d'autres méthodes de réduction de variance pour l'échantillonnage. Dans le cas où des informations sur l'état limite sont disponibles, par expérience par exemple,

une autre méthode d'IS telle que la CE [RUBINSTEIN et KROESE, 2004] ou une méthode multi-FORM peut être utilisée puisque les estimateurs de variances proposés dans ces travaux sont valables pour toute méthode d'IS.

Une autre perspective serait de proposer des estimateurs de variance dans le cas où la méthode d'échantillonnage utilisée serait la méthode des subset simulations (SS) [AU et BECK, 2001a]. L'utilisation des SS à la place de l'IS permettrait de calculer des probabilités de défaillance faibles en grande dimension. Néanmoins, l'expression donnée par cette méthode de l'estimation de la probabilité de défaillance comme un produit de probabilités intermédiaires ne permet pas une déduction directe des estimateurs de variance à partir de ceux proposés pour les MCS ou l'IS.

Une perspective d'utilisation des approches Vb-AGP pour le Dimensionnement optimal sous contrainte de fiabilité, ou *Reliability Based Design Optimisation* (RBDO), plus précisément pour les méthodes de RBDO assistés par métamodélisation [MOUSTAPHA et SUDRET, 2019], est également possible. En effet, puisque ces méthodes permettent de fournir une estimation précise du COV total de la probabilité de défaillance, il serait possible de régler judicieusement l'arrêt de l'algorithme par rapport à la valeur limite de la contrainte de fiabilité à partir de l'estimation de la probabilité et la valeur du COV total. En effet, lorsque l'intervalle de confiance de l'estimation de la probabilité de défaillance, calculé à partir du COV total, permet de montrer que sa valeur est bien inférieure ou supérieure à la valeur seuil il n'est pas nécessaire de continuer l'apprentissage pour obtenir un COV très bas.

Perspectives pour le couplage entre les méthodes d'analyse de fiabilité par apprentissage actif et la modélisation en base réduite

Le critère de couplage proposé dans ces travaux pour combiner les méthodes d'analyse de fiabilité par apprentissage actif de GP et la modélisation en base réduite prend uniquement en compte l'erreur sur la solution et non son intérêt par rapport à l'amélioration de la connaissance de l'état limite. Une piste d'amélioration du critère de couplage consisterait à s'intéresser à la phase d'exploration de l'apprentissage pour prendre en compte l'amélioration potentielle de l'apprentissage de l'état limite. En effet, l'exploration des régions moins connues du domaine pour clarifier les performances du système mène à ajouter des points parfois éloignés de l'état limite et donc des autres points d'apprentissage. En se basant sur le critère actuel, de tels points seront probablement évalués avec le modèle numérique complet alors qu'une précision approximative de la solution en ces points pourrait être suffisante pour clarifier le comportement du système dans la région concernée du domaine. L'intégration de ces considérations dans le critère de couplage pourrait être une perspective pour réduire encore davantage le

nombre de résolutions du système complet.

Dans la méthodologie proposée l'erreur sur la solution réduite est uniquement utilisée pour le critère de couplage et n'est pas propagée, c.-à-d. qu'elle n'est pas prise en compte dans l'estimation de la probabilité de défaillance. Notons, qu'une première étude pour prendre en compte les différentes fidélité des solutions a été effectuée dans ce manuscrit en utilisant un modèle de co-Krigeage. Différentes méthodologies de construction ont été considérées et les résultats n'ont pas montré d'amélioration notable de l'estimation de la probabilité de défaillance. Une perspective d'amélioration pour propager l'erreur due à l'utilisation de solutions réduites consisterait à construire un modèle de GP bruité de la fonction de performance, en considérant un bruit correspondant à l'erreur sur la fonction de performance aux points d'observations évalués à partir du modèle en base réduite. La mise en place de cette méthodologie nécessiterait un estimateur précis de l'erreur pouvant être évalué avec un coût de calcul raisonnable. En effet, le critère d'erreur actuel se base sur un estimateur de l'erreur de la solution du système linéaire car cet estimateur offre une bonne estimation de cette erreur avec des coûts de calcul relativement faibles et une mise en place facile.

Finalement, la méthode peut être utilisée pour réduire le coût de calcul des RBDO. Elle est pertinente pour les méthodes de type *nested RBDO* [MOUSTAPHA et SUDRET, 2019; MOUSTAPHA et al., 2016], c.-à-d. lorsqu'une analyse de fiabilité est effectuée à chaque itération d'une optimisation. Dans ce cas, des travaux ont proposé d'utiliser un seul GP pour toute l'optimisation qui sera construit dans un espace augmenté [AU, 2005] afin de réutiliser les informations des itérations précédentes à chaque nouvelle analyse de fiabilité par apprentissage actif. Le même principe pourrait alors être utilisé pour la construction de la base réduite, c.-à-d. qu'une seule base réduite pourrait être construite et enrichie dans l'espace augmenté au cours de l'optimisation.

Références

- ARCHER, G., A. SALTELLI et I. SOBOL. 1997, «Sensitivity measures, anova-like techniques and the use of bootstrap», *Journal of Statistical Computation and Simulation*, vol. 58, n° 2, p. 99–120. 73
- AU, S. K. 2005, «Reliability-based design sensitivity by efficient simulation», *Computers & Structures*, vol. 83, n° 14, doi: 10.1016/j.compstruc.2004.11.015, p. 1048–1061, ISSN 0045-7949. URL <http://www.sciencedirect.com/science/article/pii/S0045794905000398>. 146
- AU, S.-K. et J. L. BECK. 2001a, «Estimation of small failure probabilities in high dimensions by subset simulation», *Probabilistic Engineering Mechanics*, vol. 16, n° 4, doi: 10.1016/S0266-8920(01)00019-4, p. 263–277, ISSN 0266-8920. URL <http://www.sciencedirect.com/science/article/pii/S0266892001000194>. 41, 145
- AU, S. K. et J. L. BECK. 2001b, «Subset Simulation and its application to seismic risk based on dynamic analysis», *Journal of Engineering Mechanics*, vol. 129, n° 8, p. 1–17. 49
- BALESDENT, M., J. MORIO et J. MARZAT. 2013, «Kriging-based adaptive Importance Sampling algorithms for rare event estimation», *Structural Safety*, vol. 44, n° Supplement C, doi: 10.1016/j.strusafe.2013.04.001, p. 1–10, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473013000350>. 58
- BASUDHAR, A. et S. MISSOUM. 2013, «Reliability assessment using probabilistic support vector machines», *International Journal of Reliability and Safety*, vol. 7, n° 2, doi: 10.1504/IJRS.2013.056378, p. 156–173, ISSN 1479-389X. URL <https://www.inderscienceonline.com/doi/abs/10.1504/IJRS.2013.056378>. 50

- BECT, J., D. GINSBOURGER, L. LI, V. PICHENY et E. VAZQUEZ. 2012, «Sequential design of computer experiments for the estimation of a probability of failure», *Statistics and Computing*, vol. 22, n° 3, doi: 10.1007/s11222-011-9241-4, p. 773–793, ISSN 0960-3174, 1573-1375. URL <https://link.springer.com/article/10.1007/s11222-011-9241-4>. 50, 52, 58
- BENNER, P., S. GUGERCIN et K. WILLCOX. 2015, «A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems», *SIAM Review*, vol. 57, n° 4, doi: 10.1137/130932715, p. 483–531, ISSN 0036-1445. URL <http://epubs.siam.org/doi/abs/10.1137/130932715>. 101, 105
- BETZ, W., I. PAPAIOANNOU et D. STRAUB. 2014, «Numerical methods for the discretization of random fields by means of the Karhunen–Loève expansion», *Computer Methods in Applied Mechanics and Engineering*, vol. 271, doi: 10.1016/j.cma.2013.12.010, p. 109–129, ISSN 0045-7825. URL <http://www.sciencedirect.com/science/article/pii/S0045782513003502>. 20, 80
- BICHON, B. J., M. S. ELDRED, L. P. SWILER, S. MAHADEVAN et J. M. MCFARLAND. 2008, «Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions», *AIAA Journal*, vol. 46, n° 10, doi: 10.2514/1.34321, p. 2459–2468, ISSN 0001-1452, 1533-385X. URL <http://arc.aiaa.org/doi/10.2514/1.34321>. 3, 50, 52, 53, 54
- BICHON, B. J., J. M. MCFARLAND et S. MAHADEVAN. 2011, «Efficient surrogate models for reliability analysis of systems with multiple failure modes», *Reliability Engineering & System Safety*, vol. 96, n° 10, doi: 10.1016/j.res.2011.05.008, p. 1386–1395, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832011001062>. 50, 58
- BOUHLEL, M. A., J. T. HWANG, N. BARTOLI, R. LAFAGE, J. MORLIER et J. R. R. A. MARTINS. 2019, «A python surrogate modeling framework with derivatives», *Advances in Engineering Software*, doi: <https://doi.org/10.1016/j.advengsoft.2019.03.005>, p. 102 662, ISSN 0965-9978. 123
- BOURINET, J. M., F. DEHEEGER et M. LEMAIRE. 2011, «Assessing small failure probabilities by combined subset simulation and Support Vector Machines», *Structural Safety*, vol. 33, n° 6, doi: 10.1016/j.strusafe.2011.06.001, p. 343–353, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473011000555>. 11, 50
- BUHMANN, M. D. 2000, «Radial basis functions», *Radial basis functions*, p. 38. 11

- CADINI, F. et A. GIOLETTA. 2016, «A Bayesian Monte Carlo-based algorithm for the estimation of small failure probabilities of systems affected by uncertainties», *Reliability Engineering & System Safety*, vol. 153, doi: 10.1016/j.ress.2016.04.003, p. 15–27, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832016300175>. 58
- CADINI, F., F. SANTOS et E. ZIO. 2014, «An improved adaptive kriging-based importance technique for sampling multiple failure regions of low probability», *Reliability Engineering & System Safety*, vol. 131, doi: 10.1016/j.ress.2014.06.023, p. 109–117, ISSN 09518320. URL <http://linkinghub.elsevier.com/retrieve/pii/S0951832014001537>. 58
- CHAUDHURI, A. et R. T. HAFTKA. 2015, *Effectiveness Indicators for Stopping Criteria based on Minimum Required Improvement*, doi: 10.2514/6.2015-0901. URL <https://arc.aiaa.org/doi/abs/10.2514/6.2015-0901>. 59
- CHEVALIER, C., J. BECT, D. GINSBOURGER, E. VAZQUEZ, V. PICHENY et Y. RICHET. 2014, «Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set», *Technometrics*, vol. 56, n° 4, doi: 10.1080/00401706.2013.860918, p. 455–465. URL <https://hal.archives-ouvertes.fr/hal-00641108>. 52
- CHINESTA, F., A. AMMAR, A. LEYGUE et R. KEUNINGS. 2011, «An overview of the proper generalized decomposition with applications in computational rheology», *Journal of Non-Newtonian Fluid Mechanics*, vol. 166, n° 11, doi: 10.1016/j.jnnfm.2010.12.012, p. 578–592, ISSN 0377-0257. URL <http://www.sciencedirect.com/science/article/pii/S0377025711000061>. 28
- CRAIG, R. R. et M. C. C. BAMPTON. 1968, «Coupling of Substructures for Dynamic Analyses», *AIAA Journal*, vol. 6, n° 7, doi: 10.2514/3.4741, p. 1313–1319. URL <https://hal.archives-ouvertes.fr/hal-01537654>. 27
- DE BOER, P.-T., D. P. KROESE, S. MANNOR et R. Y. RUBINSTEIN. 2005, «A tutorial on the cross-entropy method», *Annals of operations research*, vol. 134, n° 1, p. 19–67. 43, 44
- DER KIUREGHIAN, A. et T. DAKESSIAN. 1998, «Multiple design points in first and second-order reliability», *Structural Safety*, vol. 20, n° 1, p. 37–49. 38
- DITLEVSEN, O., P. BJERAGER, R. OLESEN et A. HASOFER. 1988, «Directional simulation in gaussian processes», *Probabilistic Engineering Mechanics*, vol. 3, n° 4, doi: [https://doi.org/10.1016/0266-8920\(88\)90013-6](https://doi.org/10.1016/0266-8920(88)90013-6), p. 207 – 217, ISSN 0266-8920. URL

- <http://www.sciencedirect.com/science/article/pii/0266892088900136>. 3, 49
- DITLEVSEN, O. et H. O. MADSEN. 1996, *Structural reliability methods*, Wiley, Chichester ; New York, ISBN 978-0-471-96086-7. 35, 36
- DUBOURG, V., B. SUDRET et J.-M. BOURINET. 2011, «Reliability-based design optimization using kriging surrogates and subset simulation», *Structural and Multidisciplinary Optimization*, vol. 44, n° 5, doi: 10.1007/s00158-011-0653-8, p. 673–690, ISSN 1615-1488. URL <https://doi.org/10.1007/s00158-011-0653-8>. 60
- DUBOURG, V., B. SUDRET et F. DEHEEGER. 2013, «Metamodel-based importance sampling for structural reliability analysis», *Probabilistic Engineering Mechanics*, vol. 33, doi: 10.1016/j.probengmech.2013.02.002, p. 47–57, ISSN 0266-8920. URL <http://www.sciencedirect.com/science/article/pii/S0266892013000222>. 50, 58
- ECHARD, B. 2012, *Kriging-based reliability assessment of structures submitted to fatigue*, PhD Thesis, Université Blaise Pascal. 86
- ECHARD, B., N. GAYTON et M. LEMAIRE. 2011, «AK-MCS : An active learning reliability method combining Kriging and Monte Carlo Simulation», *Structural Safety*, vol. 33, n° 2, doi: 10.1016/j.strusafe.2011.01.002, p. 145–154, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473011000038>. 3, 52, 53, 54, 62, 74
- ECHARD, B., N. GAYTON, M. LEMAIRE et N. RELUN. 2013, «A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models», *Reliability Engineering & System Safety*, vol. 111, n° Supplement C, doi: 10.1016/j.ress.2012.10.008, p. 232–240, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832012002086>. 50, 58, 86, 95, 96
- EL HAJ, A.-K. et A.-H. SOUBRA. 2021, «Improved active learning probabilistic approach for the computation of failure probability», *Structural Safety*, vol. 88, doi: 10.1016/j.strusafe.2020.102011, p. 102011, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473020300904>. 72
- ELECTRICITÉ DE FRANCE. 1989–2017, «Finite element *code_aster*, analysis of structures and thermomechanics for studies and research», . 135
- FAURIAT, W. et N. GAYTON. 2014, «AK-SYS : An adaptation of the AK-MCS method for system reliability», *Reliability Engineering & System Safety*, vol. 123,

- doi: 10.1016/j.ress.2013.10.010, p. 137–144, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832013002949>. 50, 58
- FERREIRA, A. J. M. 2009, *MATLAB Codes for Finite Element Analysis*, Solid Mechanics and Its Applications, Springer Netherlands, ISBN 978-1-4020-9199-5. URL <http://www.springer.com/gp/book/9781402091995>. 128
- FORRESTER, A. I. J., A. SBESTER et A. J. KEANE. 2008, *Engineering Design via Surrogate Modelling*, John Wiley & Sons, Ltd, Chichester, UK, ISBN 978-0-470-77080-1 978-0-470-06068-1, doi: 10.1002/9780470770801. URL <http://doi.wiley.com/10.1002/9780470770801>. 12, 23
- FRISWELL, M. I., S. D. GARVEY et J. E. T. PENNY. 1995, «Model reduction using dynamic and iterated IRS techniques», *Journal of Sound and Vibration*, vol. 186, n° 2, doi: 10.1006/jsvi.1995.0451, p. 311–323, ISSN 0022-460X. URL <http://www.sciencedirect.com/science/article/pii/S0022460X85704518>. 27
- FUGLSTAD, G.-A., D. SIMPSON, F. LINDGREN et H. RUE. 2019, «Constructing priors that penalize the complexity of gaussian random fields», *Journal of the American Statistical Association*, vol. 114, n° 525, p. 445–452. 13
- GAMBOA, F., A. JANON, T. KLEIN, A. LAGNOUX-RENAUDIE et C. PRIEUR. 2016, «Statistical inference for Sobol pick freeze Monte Carlo method», *Statistics*, vol. 50, n° 4, doi: 10.1080/02331888.2015.1105803, p. 881–902. URL <https://hal.inria.fr/hal-00804668>. 136
- GASPAR, B., A. TEIXEIRA et C. G. SOARES. 2015, «A study on a stopping criterion for active refinement algorithms in kriging surrogate models», *Safety and reliability of complex engineered systems*, p. 1219–1227. 62
- GASPAR, B., A. P. TEIXEIRA et C. GUEDES SOARES. 2017, «Adaptive surrogate model with active refinement combining Kriging and a trust region method», *Reliability Engineering & System Safety*, vol. 165, doi: 10.1016/j.ress.2017.03.035, p. 277–291, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832016301892>. 58
- GHANEM, R. G. et R. M. KRUGER. 1996, «Numerical solution of spectral stochastic finite element systems», *Computer Methods in Applied Mechanics and Engineering*, vol. 129, n° 3, doi: 10.1016/0045-7825(95)00909-4, p. 289–303, ISSN 0045-7825. URL <http://www.sciencedirect.com/science/article/pii/0045782595009094>. 111
- GHANEM, R. G. et P. D. SPANOS. 1991, «Stochastic Finite Element Method: Response Statistics», dans *Stochastic Finite Elements : A Spectral Approach*,

- Springer, New York, NY, ISBN 978-1-4612-7795-8 978-1-4612-3094-6, p. 101–119, doi: 10.1007/978-1-4612-3094-6_4. URL https://link.springer.com/chapter/10.1007/978-1-4612-3094-6_4. 11
- GOGU, C., A. CHAUDHURI et C. BES. 2016, «How Adaptively Constructed Reduced Order Models Can Benefit Sampling-Based Methods for Reliability Analyses», *International Journal of Reliability, Quality and Safety Engineering*, vol. 23, n° 05, p. 1650019. 4, 105, 113
- GOGU, C. et J.-C. PASSIEUX. 2013, «Efficient surrogate construction by combining response surface methodology and reduced order modeling», *Structural and Multidisciplinary Optimization*, vol. 47, n° 6, doi: 10.1007/s00158-012-0859-4, p. 821–837, ISSN 1615-147X, 1615-1488. URL <https://link.springer.com/article/10.1007/s00158-012-0859-4>. 31, 105
- GUYAN, R. J. 1965, «Reduction of stiffness and mass matrices», *AIAA Journal*, vol. 3, n° 2, doi: 10.2514/3.2874, p. 380–380, ISSN 0001-1452. URL <https://doi.org/10.2514/3.2874>. 27
- HASOFER, A. M. 1974, «An exact and invariant first order reliability format», *J. Eng. Mech. Div., Proc. ASCE*, vol. 100, n° 1, p. 111–121. 36
- HASTIE, T., R. TIBSHIRANI et J. FRIEDMAN. 2009, *The elements of statistical learning: data mining, inference, and prediction*, Springer Science & Business Media. 11
- HAYKIN, S., éd.. 1994, *Neural networks, A comprehensive Foundation*, Macmillan, New York, ISBN 978-0-02-352761-6. 11
- HOPFIELD, J. J. 1982, «Neural networks and physical systems with emergent collective computational abilities», *Proceedings of the National Academy of Sciences*, vol. 79, n° 8, doi: 10.1073/pnas.79.8.2554, p. 2554–2558, ISSN 0027-8424, 1091-6490. URL <http://www.pnas.org/content/79/8/2554>. 11
- HUANG, X., J. CHEN et H. ZHU. 2016, «Assessing small failure probabilities by AK-SS : An active learning method combining Kriging and Subset Simulation», *Structural Safety*, vol. 59, n° Supplement C, doi: 10.1016/j.strusafe.2015.12.003, p. 86–95, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473016000035>. 58
- HURTADO JORGE E. et ALVAREZ DIEGO A. 2003, «Classification Approach for Reliability Analysis with Stochastic Finite-Element Modeling», *Journal of Structural Engineering*, vol. 129, n° 8, doi: 10.1061/(ASCE)0733-9445(2003)129:

- 8(1141), p. 1141–1149. URL [https://ascelibrary.org/doi/abs/10.1061/\(ASCE\)0733-9445\(2003\)129:8\(1141\)](https://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9445(2003)129:8(1141)). 11
- IOOSS, B. et P. LEMAÎTRE. 2015, «A Review on Global Sensitivity Analysis Methods», dans *Uncertainty Management in Simulation-Optimization of Complex Systems : Algorithms and Applications*, édité par G. Dellino et C. Meloni, Operations Research/Computer Science Interfaces Series, Springer US, Boston, MA, ISBN 978-1-4899-7547-8, p. 101–122, doi: 10.1007/978-1-4899-7547-8_5. URL https://doi.org/10.1007/978-1-4899-7547-8_5. 68
- JANON, A. 2012, *Analyse de sensibilité et réduction de dimension. Application à l'océanographie*, PhD Thesis. 136
- JANON, A., M. NODET, C. PRIEUR et C. PRIEUR. 2016, «Goal-oriented error estimation for fast approximations of nonlinear problems», Research Report, GIPSA-lab. URL <https://hal.archives-ouvertes.fr/hal-01290887>. 105
- JONES, D. R. 2001, «A Taxonomy of Global Optimization Methods Based on Response Surfaces», *Journal of Global Optimization*, vol. 21, n° 4, doi: 10.1023/A:1012771025575, p. 345–383, ISSN 1573-2916. URL <https://doi.org/10.1023/A:1012771025575>. 13
- JONES, D. R., M. SCHONLAU et W. J. WELCH. 1998, «Efficient global optimization of expensive black-box functions», *Journal of Global Optimization*, vol. 13, p. 455 – 492. 11
- KARHUNEN, K. 1947, «Über lineare methoden in der wahrscheinlichkeitsrechnung», *Über Lineare Methoden in der Wahrscheinlichkeitsrechnung*, p. 3–79. 19
- KENNEDY, M. C. et A. O'HAGAN. 2000, «Predicting the output from a complex computer code when fast approximations are available», *Biometrika*, vol. 87, n° 1, doi: 10.1093/biomet/87.1.1, p. 1–13, ISSN 0006-3444. URL <https://academic.oup.com/biomet/article/87/1/1/221217>. 23
- KERSCHEN, G., J.-C. GOLINVAL, A. F. VAKAKIS et L. A. BERGMAN. 2005, «The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview», *Nonlinear Dynamics*, vol. 41, n° 1-3, doi: 10.1007/s11071-005-2803-2, p. 147–169, ISSN 0924-090X, 1573-269X. URL <https://link.springer.com/article/10.1007/s11071-005-2803-2>. 4, 101, 105
- KOSTINSKI, A. B. et A. C. KOIVUNEN. 2000, «On the condition number of Gaussian sample-covariance matrices», *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, n° 1, doi: 10.1109/36.823928, p. 329–332, ISSN 1558-0644. 15

- KRIGE, D. G. 1951, «A statistical approach to some basic mine valuation problems on the Witwatersrand», *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, n° 6, p. 119–139, ISSN 0038-223X. URL https://journals.co.za/content/saimm/52/6/AJA0038223X_4792. 12
- KULLBACK, S. et R. A. LEIBLER. 1951, «On information and sufficiency», *The Annals of Mathematical Statistics*, vol. 22, n° 1, p. 79–86, ISSN 00034851. URL <http://www.jstor.org/stable/2236703>. 44
- KURTZ, N. et J. SONG. 2013, «Cross-entropy-based adaptive importance sampling using gaussian mixture», *Structural Safety*, vol. 42, doi: <https://doi.org/10.1016/j.strusafe.2013.01.006>, p. 35 – 44, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473013000131>. 46
- LADEVÈZE, P. 1985, «New algorithms: mechanical framework and development», *Compte rendu de l'académie des Sciences*, vol. 300, n° 2, p. 41–44. 28
- LADEVEZE, P. 2012, *Nonlinear Computational Structural Mechanics: New Approaches and Non-Incremental Methods of Calculation*, Springer Science & Business Media, ISBN 978-1-4612-1432-8. 28
- LE GRATIET, L. 2013, *Multi-fidelity Gaussian process regression for computer experiments*, PhD Thesis. 23
- LE GRATIET, L., C. CANNAMELA et B. IOOSS. 2014, «A Bayesian Approach for Global Sensitivity Analysis of (Multifidelity) Computer Codes», *SIAM/ASA Journal on Uncertainty Quantification*, vol. 2, n° 1, doi: 10.1137/130926869, p. 336–363. URL <https://epubs.siam.org/doi/abs/10.1137/130926869>. 17, 71, 80
- LELIÈVRE, N., P. BEAUREPAIRE, C. MATTRAND et N. GAYTON. 2018, «AK-MCSi : A Kriging-based method to deal with small failure probabilities and time-consuming models», *Structural Safety*, vol. 73, doi: 10.1016/j.strusafe.2018.01.002, p. 1–11, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473017301558>. 58, 60, 86
- LEMAIRE, M., A. CHATEAUNEUF et J.-C. MITTEAU. 2009, *Structural reliability*, Wiley Online Library. 2, 35, 36
- LEMAÎTRE, P. 2014, *Analyse de sensibilité en fiabilité des structures*, thèse de doctorat. 136
- LEON, A., S. MUELLER, P. DE LUCA, R. SAID, J.-L. DUVAL et F. CHINESTA. 2019, «Non-intrusive proper generalized decomposition involving space and parameters:

- application to the mechanical modeling of 3d woven fabrics», *Advanced Modeling and Simulation in Engineering Sciences*, vol. 6, n° 1, p. 13. 28
- LI, C.-C. et A. DER KIUREGHIAN. 1993, «Optimal discretization of random fields», *Journal of engineering mechanics*, vol. 119, n° 6, p. 1136–1154. 21
- LI, L., J. BECT et E. VAZQUEZ. 2012, «Bayesian Subset Simulation : a kriging-based subset simulation algorithm for the estimation of small probabilities of failure», dans *11th International Probabilistic Assessment and Management Conference (PSAM11) and The Annual European Safety and Reliability Conference (ES-REL 2012)*, Helsinki, Finland, p. CD-ROM Proceedings (10 pages). URL <https://hal-supelec.archives-ouvertes.fr/hal-00715316>. 50, 58
- LOEVE, M. 1948, «Fonctions aléatoires du second ordre, supplement to p», *Lévy, Processus Stochastiques et Mouvement Brownien, Gauthier-Villars, Paris*. 19
- LUMLEY, J. L. 1967, «The structure of inhomogeneous turbulent flows», *Atmospheric turbulence and wave propagation*, p. 166–178. URL <https://ci.nii.ac.jp/naid/10029969783/en/>. 28
- LV, Z., Z. LU et P. WANG. 2015, «A new learning function for Kriging and its applications to solve reliability problems in engineering», *Computers & Mathematics with Applications*, vol. 70, n° 5, doi: 10.1016/j.camwa.2015.07.004, p. 1182–1197, ISSN 0898-1221. URL <http://www.sciencedirect.com/science/article/pii/S0898122115003399>. 58
- MACKAY, M. D. 1992, «Latin hypercube sampling as a tool in uncertainty analysis of computer models», dans *WSC '92 Proceedings of the 24th conference on Winter simulation*, ACM, New York, p. 557–564. 55
- MATHERON, G. 1963, «Principles of geostatistics», *Economic Geology*, vol. 58, n° 8, doi: 10.2113/gsecongeo.58.8.1246, p. 1246–1266, ISSN 0361-0128. URL <https://pubs.geoscienceworld.org/segweb/economicgeology/article-abstract/58/8/1246/17275/principles-of-geostatistics>. 12
- MCKAY, M. D., R. J. BECKMAN et W. CONOVER. 1979, «A Comparison of Three Methods for selecting values of input variables in the analysis of output from a computer code», *Technometrics*, vol. 21, p. 239 – 245. 55
- MELCHERS, R. E. 1989, «Importance sampling in structural systems», *Structural Safety*, vol. 6, n° 1, doi: 10.1016/0167-4730(89)90003-9, p. 3–10, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/0167473089900039>. 3, 41, 43

- MENZ, M., C. GOGU, S. DUBREUIL, N. BARTOLI et J. MORIO. 2020, «Adaptive coupling of reduced basis modeling and Kriging based active learning methods for reliability analyses», *Reliability Engineering & System Safety*, vol. 196, doi: 10.1016/j.res.2019.106771, p. 106 771, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S095183201930184X>. 7
- MORIO, J. 2012, «Extreme quantile estimation with nonparametric adaptive importance sampling», *Simulation Modelling Practice and Theory*, vol. 27, doi: 10.1016/j.simpat.2012.05.008, p. 76–89. 43, 46, 144
- MORIO, J., M. BALESSENT, D. JACQUEMART et C. VERGÉ. 2014, «A survey of rare event simulation methods for static input–output models», *Simulation Modelling Practice and Theory*, vol. 49, p. 287–304. 39
- MOUSTAPHA, M. et B. SUDRET. 2019, «Surrogate-assisted reliability-based design optimization : a survey and a new general framework», *arXiv :1901.03311 [stat]*. URL <http://arxiv.org/abs/1901.03311>, arXiv : 1901.03311. 59, 145, 146
- MOUSTAPHA, M., B. SUDRET, J.-M. BOURINET et B. GUILLAUME. 2016, «Quantile-based optimization under uncertainties using adaptive Kriging surrogate models», *arXiv :1605.08978 [stat]*. URL <http://arxiv.org/abs/1605.08978>, arXiv : 1605.08978. 146
- MURPHY, K. P. 2012, *Machine learning: a probabilistic perspective*, MIT press. 11
- MYERS, R. H., D. C. MONTGOMERY, G. G. VINING et T. J. ROBINSON. 2012, *Generalized Linear Models: with Applications in Engineering and the Sciences*, John Wiley & Sons, ISBN 978-0-470-55697-9. Google-Books-ID : LBiT207QLZgC. 11
- NATAF, A. 1962, «Distribution des distributions dont les marges sont données (in French)», *Comptes rendus de l'Académie des Sciences*, vol. 225, p. 42–43. 35
- PAPAIOANNOU, I., C. PAPADIMITRIOU et D. STRAUB. 2016, «Sequential importance sampling for structural reliability analysis», *Structural safety*, vol. 62, p. 66–75. 43
- PERRIN, G. 2016, «Active learning surrogate models for the conception of systems with multiple failure modes», *Reliability Engineering & System Safety*, vol. 149, doi: 10.1016/j.res.2015.12.017, p. 130–136, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832015003695>. 58
- PERRIN, G. et G. DEFAUX. 2019, «Efficient Evaluation of Reliability-Oriented Sensitivity Indices», *Journal of Scientific Computing*, vol. 79, n° 3, doi: 10.1007/s10915-018-00898-8, p. 1433–1455, ISSN 1573-7691. URL <https://doi.org/10.1007/s10915-018-00898-8>. 136

- PICHENY, V. 2009, *Improving accuracy and compensating for uncertainty in surrogate modeling*, PhD Thesis, University of Florida. 52
- PICHENY, V., D. GINSBOURGER, O. ROUSTANT, R. T. HAFTKA et N.-H. KIM. 2010, «Adaptive Designs of Experiments for Accurate Approximation of a Target Region», *Journal of Mechanical Design*, vol. 132, n° 7, doi: 10.1115/1.4001873, p. 071 008–071 008–9, ISSN 1050-0472. URL <http://dx.doi.org/10.1115/1.4001873>. 52
- QIAN, P. Z. et C. J. WU. 2008, «Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments», *Technometrics*, vol. 50, n° 2, p. 192–204. 23
- RAGUET, H. et A. MARREL. 2018, «Target and conditional sensitivity analysis with emphasis on dependence measures», *arXiv preprint arXiv :1801.10047*. 136
- RANJAN, P., D. BINGHAM et G. MICHAILIDIS. 2008, «Sequential Experiment Design for Contour Estimation From Complex Computer Codes», *Technometrics*, vol. 50, n° 4, doi: 10.1198/004017008000000541, p. 527–541, ISSN 0040-1706, 1537-2723. URL <http://www.tandfonline.com/doi/abs/10.1198/004017008000000541>. 52, 53
- RASMUSSEN, C. E. et C. K. I. WILLIAMS. 2006, *Gaussian processes for machine learning*, Adaptive computation and machine learning, MIT Press, Cambridge, Mass, ISBN 978-0-262-18253-9. OCLC : ocm61285753. 11, 12, 13, 17
- RIXEN, D. J. 2004, «A dual Craig–Bampton method for dynamic substructuring», *Journal of Computational and Applied Mathematics*, vol. 168, n° 1, doi: 10.1016/j.cam.2003.12.014, p. 383–391, ISSN 0377-0427. URL <http://www.sciencedirect.com/science/article/pii/S0377042703010045>. 27
- ROCHES, G. V. D., J. P. BIANCHI, E. BALMÈS, R. LEMAIRE et P. THIERRY. 2010, «Using component modes in a system design process», p. 75. URL <https://hal.archives-ouvertes.fr/hal-00457484>. 27
- ROSENBLATT, M. 1952, «Remarks a multivariate transformation», *The Annals of Mathematical Statistics*, vol. 23, n° 3, p. 470–472. 35
- ROZZA, G., D. B. P. HUYNH et A. T. PATERA. 2008, «Reduced Basis Approximation and a Posteriori Error Estimation for Affinely Parametrized Elliptic Coercive Partial Differential Equations: Application to Transport and Continuum Mechanics», *Archives of Computational Methods in Engineering*, vol. 15, n° 3, doi: 10.1007/s11831-008-9019-9, p. 229–275, ISSN 1134-3060, 1886-1784. URL <http://link.springer.com/10.1007/s11831-008-9019-9>. 105

- RUBINSTEIN, R. 1999, «The cross-entropy method for combinatorial and continuous optimization», *Methodology and Computing in Applied Probability*, vol. 2, p. 127–129. 43, 44
- RUBINSTEIN, R. et D. KROESE. 2004, *The Cross-Entropy Method : A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning (Information Science and Statistics)*, Springer. 44, 45, 145
- RUBINSTEIN, R. Y. 1997, «Optimization of computer simulation models with rare events», *European Journal of Operations Research*, vol. 99, p. 89–112. 44
- SACKS, J., W. J. WELCH, T. J. MITCHELL et H. P. WYNN. 1989, «Design and analysis of computer experiments», *Statistical Science*, vol. 4, n° 4, p. 409–423, ISSN 08834237. URL <http://www.jstor.org/stable/2245858>. 12
- SCHMIDHUBER, J. 2015, «Deep learning in neural networks: An overview», *Neural Networks*, vol. 61, doi: 10.1016/j.neunet.2014.09.003, p. 85–117, ISSN 0893-6080. URL <http://www.sciencedirect.com/science/article/pii/S0893608014002135>. 11
- SCHÖBI R., SUDRET B. et MARELLI S. 2017, «Rare Event Estimation Using Polynomial-Chaos Kriging», *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A : Civil Engineering*, vol. 3, n° 2, doi: 10.1061/AJRUA6.0000870, p. D4016 002. URL <https://ascelibrary.org/doi/abs/10.1061/AJRUA6.0000870>. 50, 60, 67
- SCHUEREMANS, L. et D. VAN GEMERT. 2005, «Benefit of splines and neural networks in simulation based structural reliability analysis», *Structural Safety*, vol. 27, n° 3, doi: 10.1016/j.strusafe.2004.11.001, p. 246–261, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473004000529>. 74
- SCHUËLLER, G. et R. STIX. 1987, «A critical appraisal of methods to determine failure probabilities», *Structural Safety*, vol. 4, n° 4, doi: [https://doi.org/10.1016/0167-4730\(87\)90004-X](https://doi.org/10.1016/0167-4730(87)90004-X), p. 293 – 309, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/016747308790004X>. 43
- SMETANA, K., O. ZAHM et A. PATERA. 2019, «Randomized Residual-Based Error Estimators for Parametrized Equations», *SIAM Journal on Scientific Computing*, vol. 41, n° 2, doi: 10.1137/18M120364X, p. A900–A926, ISSN 1064-8275. URL <https://epubs.siam.org/doi/abs/10.1137/18M120364X>. 105
- SMOLA, A. J. et B. SCHÖLKOPF. 2004, «A tutorial on support vector regression», *Statistics and Computing*, vol. 14, n° 3, doi: 10.1023/B:STCO.0000035301.49549.

- 88, p. 199–222, ISSN 0960-3174, 1573-1375. URL <https://link.springer.com/article/10.1023/B:STCO.0000035301.49549.88>. 11
- SOILAHOUINE, M., C. GOGU et C. BES. 2017, «Accelerated Adaptive Surrogate-Based Optimization Through Reduced-Order Modeling», *AIAA Journal*, vol. 55, n° 5, doi: 10.2514/1.J055252, p. 1681–1694, ISSN 0001-1452. URL <https://doi.org/10.2514/1.J055252>. 105
- SUDRET, B. 2007, «Uncertainty propagation and sensitivity analysis in mechanical models—contributions to structural reliability and stochastic spectral methods», *Habilitations dirigées des recherches, Université Blaise Pascal, Clermont-Ferrand, France*, vol. 147. 33
- TOAL, D. J. 2015, «Some considerations regarding the use of multi-fidelity kriging in the construction of surrogate models», *Structural and Multidisciplinary Optimization*, vol. 51, n° 6, p. 1223–1245. 123
- TONN, T., K. URBAN et S. VOLKWEIN. 2011, «Comparison of the reduced-basis and POD *a posteriori* error estimators for an elliptic linear-quadratic optimal control problem», *Mathematical and Computer Modelling of Dynamical Systems*, vol. 17, n° 4, doi: 10.1080/13873954.2011.547678, p. 355–369, ISSN 1387-3954, 1744-5051. URL <http://www.tandfonline.com/doi/abs/10.1080/13873954.2011.547678>. 105
- VAPNIK, V. N. 1998, *Statistical learning theory*, Adaptive and learning systems for signal processing, communications, and control, Wiley, New York, ISBN 978-0-471-03003-4. 11
- VAZQUEZ, E. et J. BECT. 2009, «A sequential Bayesian algorithm to estimate a probability of failure», *IFAC Proceedings Volumes*, vol. 42, n° 10, doi: 10.3182/20090706-3-FR-2004.00090, p. 546–550, ISSN 1474-6670. URL <http://www.sciencedirect.com/science/article/pii/S1474667016387043>. 50, 58
- VEROY, K., C. PRUD'HOMME, D. ROVAS et A. PATERA. 2003, «A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations», dans *16th AIAA Computational Fluid Dynamics Conference*, p. 3847. 30
- VILLEMONTAIX, J., E. VAZQUEZ et E. WALTER. 2009, «An informational approach to the global optimization of expensive-to-evaluate functions», *Journal of Global Optimization*, vol. 44, n° 4, doi: 10.1007/s10898-008-9354-2, p. 509–534. URL <https://hal-supelec.archives-ouvertes.fr/hal-00354262>. 17, 71, 80

- WANG, Z. et A. SHAFIEEZADEH. 2019a, «ESC: an efficient error-based stopping criterion for kriging-based reliability analysis methods», *Structural and Multidisciplinary Optimization*, vol. 59, n° 5, doi: 10.1007/s00158-018-2150-9, p. 1621–1637, ISSN 1615-1488. URL <https://doi.org/10.1007/s00158-018-2150-9>. 61, 67
- WANG, Z. et A. SHAFIEEZADEH. 2019b, «REAK: Reliability analysis through Error rate-based Adaptive Kriging», *Reliability Engineering & System Safety*, vol. 182, doi: 10.1016/j.ress.2018.10.004, p. 33–45, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832017312504>. 58
- WANG, Z. et J. SONG. 2016, «Cross-entropy-based adaptive importance sampling using von mises-fisher mixture for high dimensional reliability analysis», *Structural Safety*, vol. 59, doi: <https://doi.org/10.1016/j.strusafe.2015.11.002>, p. 42 – 52, ISSN 0167-4730. URL <http://www.sciencedirect.com/science/article/pii/S0167473015000909>. 46
- ZAHM, O., M. BILLAUD-FRIESS et A. NOUY. 2016, «Projection based model order reduction methods for the estimation of vector-valued variables of interest», *arXiv :1603.00336 [math]*. URL <http://arxiv.org/abs/1603.00336>, arXiv : 1603.00336. 30
- ZAHM, O. et A. NOUY. 2016, «Interpolation of inverse operators for preconditioning parameter-dependent equations», *SIAM Journal on Scientific Computing*, vol. 38, n° 2, p. A1044–A1074. 30, 109
- ZHANG, P. 1996, «Nonparametric importance sampling», *Journal of the American Statistical Association*, vol. 91, n° 435, p. 1245–1253. 43, 46
- ZHANG, X., L. WANG et J. D. SØRENSEN. 2019, «Reif: A novel active-learning function toward adaptive Kriging surrogate models for structural reliability analysis», *Reliability Engineering & System Safety*, vol. 185, doi: 10.1016/j.ress.2019.01.014, p. 440–454, ISSN 0951-8320. URL <http://www.sciencedirect.com/science/article/pii/S0951832018305969>. 58
- ZHU, Z. et X. DU. 2016, «Reliability Analysis With Monte Carlo Simulation and Dependent Kriging Predictions», *Journal of Mechanical Design*, vol. 138, n° 12, doi: 10.1115/1.4034219, ISSN 1050-0472. URL <https://asmedigitalcollection.asme.org/mechanicaldesign/article/138/12/121403/376322/Reliability-Analysis-With-Monte-Carlo-Simulation>. 72
- ZIMMERMANN, R. 2015, «On the condition number anomaly of Gaussian correlation matrices», *Linear Algebra and its Applications*, vol. 466, doi: 10.1016/j.laa.2014.10.

038, p. 512–526, ISSN 0024-3795. URL <http://www.sciencedirect.com/science/article/pii/S0024379514007174>. 15

ZOU, X., M. CONTI, P. DÍEZ et F. AURICCHIO. 2018, «A nonintrusive proper generalized decomposition scheme with application in biomechanics», *International Journal for Numerical Methods in Engineering*, vol. 113, n° 2, p. 230–251. 28