HAL open science

Estimation of failure probability and its local sensitivity in a high-dimensional standard elliptical space

Marie Chiron

To cite this version:

Marie Chiron. Estimation of failure probability and its local sensitivity in a high-dimensional standard elliptical space. Physics [physics]. Institut Supérieur de l'Aéronautique et de l'Espace - ISAE, 2023. English. ⟨NNT : 2023ESAE0039⟩. ⟨tel-04659998⟩

HAL Id: tel-04659998

https://hal.science/tel-04659998v1

Submitted on 23 Jul 2024

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.

Autorisation HAL

HAL Authorization

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut Supérieur de l'Aéronautique et de l'Espace

Présentée et soutenue par :
Marie CHIRON
le mardi 4 juillet 2023

Titre :
Estimation de probabilité de défaillance et de sa sensibilité locale dans un espace elliptique standard de grande dimension

Estimation of failure probability and its local sensitivity in a high-dimensional standard elliptical space

---

École doctorale et discipline ou spécialité :
ED MITT : Mathématiques appliquées

Unité de recherche :
Équipe d'accueil ISAE-ONERA MOIS

Directeur(s) de Thèse :

M. Jérôme MORIO (directeur de thèse)
M. Michel SALAÜN (co-directeur de thèse)

Jury :
M. Josselin GARNIER Professeur École Polytechnique - Président
M. Sylvain DUBREUIL Chargé de recherche ONERA - Co-encadrant
Mme Céline HELBERT Maître de conférences École Centrale Lyon - Rapporteure
Mme Cécile MATTRAND Maître de conférences Université Clermont Auvergne - Examinatrice
M. Jérôme MORIO Directeur de recherche ONERA - Directeur de thèse
M. Guillaume PERRIN Chercheur Université Gustave Eiffel - Examinateur
M. Carsten PROPPE Professeur Karlsruhe Institute of Technology - Rapporteur
M. Michel SALAÜN Professeur ISAE-SUPAERO - Co-directeur de thèse

Remerciements

Beaucoup de personnes m'ont apporté leur soutien tout au long de ces 3 années et demi de thèse ponctuées par de multiples confinements et quelques passages aux urgences. Aussi je prends ici le temps de les remercier chaleureusement.

Je voudrais tout d'abord remercier les rapporteurs qui ont pris le soin de lire ce manuscrit et de m'autoriser à défendre ma thèse, Céline Helbert et Carsten Proppe. Merci pour vos rapports détaillés et vos remarques, qui étaient très instructives et bienveillantes. Je remercie également Josselin Garnier, qui a présidé mon jury de thèse ainsi que Cécile Mattrand et Guillaume Perrin, les deux autres membres de mon jury de soutenance. Toutes les suggestions et questions posées pendant la soutenance m'ont permis de prendre du recul sur le travail effectué et j'ai pris grand plaisir à échanger avec vous tous.

Je tiens ensuite à remercier mon encadrement de thèse en France: Sylvain Dubreuil, Jérôme Morio et Michel Salaün. Ils le savent, cette thèse aura été assez éprouvante pour moi et j'ai toujours pu compter sur leur soutien. Michel tu as été de très bon conseil et bien que nos échanges furent assez succincts, j'ai apprécié toutes tes remarques et tes suggestions. Sylvain, je te remercie pour toutes ces heures passées sur les codes, pour ta disponibilité au labo et ta bonne humeur. Jérôme, merci pour ton soutien du début jusqu'à la fin. Tu as toujours eu les bons mots au bon moment pour me redonner du courage.

Cette thèse a été effectuée entre l'ONERA centre de Toulouse et l'ISAE Supaéro. Je tiens donc à exprimer ma gratitude envers ces deux institutions scientifiques, qui m'ont permis d'effectuer cette thèse dans de très bonnes conditions. En particulier, je souhaite remercier tous les permanents, doctorants et stagiaires que j'ai pu rencontrer au DTIS de Toulouse. Vous avez contribué à une super ambiance au laboratoire, et j'espère vous revoir bientôt !

J'ai effectué au milieu de ma thèse une mobilité doctorale à l'Université McGill à Montréal sous la supervision de Christian Genest. Aussi je tiens à remercier chaleureusement Christian, qui m'a aidé à prendre un nouveau départ, en pleine crise COVID. Sans ce passage de 5 mois à Montréal en ta compagnie, je ne serais sans doute pas allée au bout de cette thèse, merci pour tout. Cette mobilité a été financée par l'ISAE Supaéro et l'institut EURMINT, aussi je les remercie pour cette merveilleuse opportunité. A Montréal j'ai eu la chance d'être accueillie par la Famille Fon ainsi que la Famille Caire. Je tiens donc à vous remercier particulièrement Nathalie, Ted et Noah pour le temps passé avec vous. Je remercie aussi Ginette, qui a toujours été là pour moi.

Enfin je tiens à remercier ma famille et mes amis. C'est dans les moments les plus difficiles que l'on apprécie le soutien de ses proches et je sais que la réussite de cette thèse n'aurait pas été possible sans eux. Aussi je tiens à remercier particulièrement Alice, Aurélie, Raphaël, Diana, Youssr, Sereine, Maud, Claire, Benjamin et Abir. Je remercie aussi Marie-Laure et Marc, ma famille en France et au Mexique, spécialement mes grand parents qui sont venus à ma soutenance. Et bien sûr, mes parents et mon frère pour leur soutien inconditionnel.

Contents

1 Introduction **11**

2 Mathematical modeling of a system **19**
2.1 Input-output black-box model of the system . . . 19
2.2 Probabilistic modeling of the random inputs . . . 20
2.2.1 Elements of probability theory . . . 20
2.2.2 Standard elliptical laws in high-dimensional space . . . 27
2.3 Performing a reliability analysis of the system . . . 30
2.3.1 Estimation of the failure probability of the system . . . 31
2.3.2 Probability of failure sensitivity analysis . . . 31

3 Failure probability estimation in a high-dimensional space **39**
3.1 Introduction . . . 40
3.2 Design-point-based techniques: FORM and SORM . . . 40
3.2.1 FORM/SORM in the standard elliptical space . . . 40
3.2.2 Multiple design points and multiple failure regions . . . 42
3.2.3 Design points in high-dimensional standard normal space . . . 45
3.3 Simulation methods . . . 45
3.3.1 Crude Monte Carlo . . . 46
3.3.2 Subset Simulation . . . 46
3.3.3 Importance Sampling . . . 50
3.3.4 Directional Sampling . . . 62
3.4 Simulation techniques based on the design point . . . 64
3.4.1 Radial-Based Importance Sampling . . . 65
3.4.2 Line Sampling . . . 65
3.4.3 Importance Sampling using design points . . . 68
3.5 Conclusion . . . 70

4 Local reliability-based sensitivity analysis **75**
4.1 Introduction . . . 76
4.1.1 Brief presentation of the reliability-based sensitivity analysis . . . 76
4.1.2 Local RSA with respect to deterministic inputs . . . 77
4.2 Local RSA in the original input space . . . 79
4.2.1 Score function method for sensitivity analysis . . . 79
4.2.2 Statistical properties of the score function estimate with crude MC . . . 79
4.2.3 Score function method with advanced sampling techniques . . . 80
4.3 Local RSA in the standard elliptical space . . . 82
4.3.1 Sensitivity through the design point of the system . . . 82
4.3.2 Sensitivity with direct simulation method . . . 84
4.3.3 Sensitivity through an approximation of the failure domain indicator function . . . 86
4.3.4 Sensitivity with finite difference schemes . . . 90
4.4 Conclusion . . . 91

5 New failure probability estimation method with MIS **101**
5.1 Introduction . . . 102
5.2 Adaptive method to estimate the failure probability . . . 103
5.2.1 Presentation of the proposed method . . . 103
5.2.2 Identification problem . . . 104
5.2.3 Sampling problem . . . 106
5.2.4 Adaptation step . . . 111
5.2.5 Search for missed failure regions . . . 112
5.2.6 Estimation of the global failure probability . . . 114
5.3 Numerical investigations with high-dimensional input space . . . 117
5.3.1 Series system of four linear limit state functions . . . 118
5.3.2 Duffing oscillator . . . 124
5.3.3 Portfolio loss . . . 128
5.3.4 Airfoil in inviscid transonic flow . . . 132
5.4 Conclusion . . . 138
5.4.1 Necessity of standard elliptical inputs . . . 138
5.4.2 Influence of IS auxiliary density in high-dimensional spaces . . . 139
5.4.3 Proposed method without the gradient of the limit state function . . . 139

6 Sensitivity estimation through polynomial regression **147**
6.1 Introduction . . . 148
6.2 New sensitivity expression through Taylor series expansion . . . 148
6.2.1 Baseline of the mathematical framework of the approach . . . 149
6.2.2 Illustration with a one-dimensional toy example . . . 151
6.2.3 Combining sampling methods and polynomial regression to derive the failure probability sensitivity . . . 152
6.3 Heteroscedastic polynomial regression . . . 155
6.3.1 Linear least squares method in our specific context . . . 155
6.3.2 Settings of the regression parameters . . . 157
6.3.3 Proposed approach with importance sampling methods . . . 162
6.4 Numerical investigation . . . 165
6.4.1 Cantilever beam . . . 166
6.4.2 Roof truss . . . 168
6.4.3 Duffing oscillator . . . 171
6.5 Conclusion . . . 173
6.5.1 Discussion on the heteroscedastic polynomial regression . . . 173
6.5.2 High-dimensional elliptical space with multiple failure regions . . . 174
6.5.3 Generalization of the proposed approach . . . 174

7 Conclusion **177**

A MCMC algorithms **181**
A.1 Original Metropolis algorithm . . . 181
A.2 Modified Metropolis algorithm . . . 181
A.3 MCMC algorithm developed by Katafygiotis et al. . . . 181
A.4 Statistical properties of the Metropolis algorithms . . . 182

B MSE for the IS estimate in the score function method **183**

C Discussion on the optimizations in the important ring **185**

D Calculus for the new sensitivity estimate **187**
D.1 Expression of the derivative in the Taylor series expansion . . . 187
D.2 Expression of the failure probability derivative functions . . . 187
D.3 Expression of the other Taylor coefficients of the toy example . . . 188

E Polynomial regressions in the OLS framework **191**
E.1 Introduction of the Ordinary Least Squares (OLS) method 191
E.2 Toy example with various MC in the OLS framework 191
E.3 Toy example with NAIS in the OLS framework 192

Bibliography **202**

Introduction

Dans de nombreux domaines scientifiques, le comportement d'un système est souvent simulé grâce à un modèle numérique. Ce modèle numérique dépend le plus souvent d'un système d'équations aux dérivées partielles, dont la résolution peut être numériquement coûteuse. Le modèle numérique est vu dans ce manuscrit comme une boîte noire, dont peu d'information est disponible. Les entrées de cette boîte noire représentent les variables influençant le comportement du système. La sortie de cette boîte noire est la réponse observée du système. Les entrées de la boîte noire sont alors regroupées en deux ensembles : les entrées aléatoires et les entrées déterministes. Il est supposé dans ce manuscrit que les entrées aléatoires sont définies avec une loi de probabilité et représentent les incertitudes qui affectent le système. Les entrées déterministes regroupent les paramètres de conception du système ainsi que les paramètres de distribution des entrées aléatoires. La sortie de la boîte noire est ainsi également aléatoire. Ce modèle numérique permet d'étudier la performance du système ainsi que sa fiabilité. Une fonction de performance est en effet définie afin d'apprécier le comportement de la sortie du modèle numérique. Selon la valeur de cette fonction de performance, le système est soit dans un état sûr soit dans un état défaillant, la défaillance étant un évènement rare.

La défaillance du système entrainant souvent des conséquences économiques ou environnementales, la probabilité de défaillance du système est calculée dans un but d'analyse de fiabilité,. Cette probabilité de défaillance est exprimée grâce à une intégrale définie sur le domaine de défaillance des entrées aléatoires. Cette intégrale peut être estimée de différentes manières, soit par des méthodes de simulation basée sur la technique de Monte-Carlo [Morio et al., 2014], soit par des méthodes d'approximation comme le FORM/SORM [Madsen et al., 2006] par exemple. Selon la dimension du système ainsi que la forme du domaine de défaillance, les différentes méthodes sont plus ou moins efficaces. La probabilité de défaillance obtenue dépend alors fortement des entrées déterministes du système. Le calcul de la dérivée de cette probabilité par rapport à ces paramètres est essentielle en optimisation fiabiliste (dans le cas des paramètres de conception), ainsi que pour comprendre l'influence du modèle probabiliste choisi sur la fiabilité du système (dans le cas des paramètres de distribution). Selon la nature de ces paramètres, plusieurs méthodes d'estimation de la dérivée sont possibles [Papaioannou et al., 2019], qui ne requièrent pas un budget de simulation très élevé. Néanmoins, des approximations sont souvent nécessaires ce qui amène à des estimateurs de dérivée biaisés.

Nous supposons dans ce manuscrit que la dimension du système est élevée et que le domaine de défaillance possède plusieurs régions de défaillance. De plus, les entrées aléatoires sont distribuées selon une loi elliptique standard. Dans un tel contexte, le calcul de la probabilité de défaillance est particulièrement difficile. Récemment, des méthodes présentées dans [Wang and Song, 2016] et [Papaioannou et al., 2019] proposent des algorithmes robustes à la grande dimension des entrées elliptiques standards. Néanmoins, l'identification du nombre de régions de défaillance est alors particulièrement complexe dans un contexte de boîte noire. **Le premier objectif du manuscrit consiste alors à développer une nouvelle méthode pour estimer la probabilité de défaillance avec de telles hypothèses**. Les régions de défaillance sont identifiées au fur et à mesure, et la probabilité est estimée avec un échantillonnage préférentiel multiple. Dans l'espace standard elliptique, les dérivées de la probabilité sont nécessairement définies par une intégrale surfacique. L'approche Faible [Torii, 2020] consiste alors à calculer une approximation de la dérivée avec une intégrale de domaine plus facile à estimer. Cela introduit un biais théorique qu'il est difficile d'examiner. **Le second objectif du manuscrit consiste donc à introduire une nouvelle approche pour estimer la dérivée de la probabilité, présentée comme une amélioration de l'approche Faible**. L'estimateur de la dérivée est obtenu grâce à une régression polynomiale hétéroscédastique. Le biais de cet estimateur

est alors contrôlable grâce aux différents paramètres de la régression.

Le manuscrit se divise en 7 chapitres (dont celui-ci d'introduction). Le Chapitre 2 rappelle les bases mathématiques utiles au modèle probabiliste choisi dans la thèse: les lois elliptiques standards en grande dimension. Les Chapitres 3 et 4 présentent l'état de l'art en estimation de probabilité et en estimation de dérivée de probabilité dans le contexte spécifique du manuscrit. Les Chapitres 5 et 6 décrivent en détail les deux algorithmes développés durant la thèse. Enfin le Chapitre 7 conclut et propose quelques perspectives.

Chapter 1

Introduction

Context

In many scientific or economical fields, the behavior of a complex system is typically simulated with a numerical model. Several hypothesis are made in order to build a numerically calculable model as close as possible to the system. A fixed set of parameters is thus defined and presumably gathers the most influential variables of the system. The numerical model is also built under several assumptions, simplifying the environmental context for instance. The numerical model of the system can rely on an analytical formula or on a system of equations. In general, a system of partial differential equations is employed to model complex systems. The resolution of these equations is then numerically expensive as it can require the use of finite element or finite volume methods for example. This model of the system makes it possible to study its performance and its reliability by simulating its behavior multiple times with different settings. As an illustration in an aerodynamic engineering context, one can be interested in the behavior of an airfoil in inviscid transonic flow. The environment of the airfoil is then defined with specific flight conditions (atmospheric conditions, Mach number, angle of attack, etc...). The quantity of interest, for example the drag coefficient, is obtained from the numerical resolution of the Euler equation (generally by a finite volume method, cf [Economon et al., 2016]). This quantity of interest depends on many parameters, such as the shape of the airfoil which could be parameterized by hundreds of variables. Even if the sensitivity of the quantity of interest with respect to these parameters is nowadays accessible in modern computational codes (through adjoint differentiation for example), it should be noted that the accurate numerical resolution of these equations leads to a high numerical cost. The statistical analysis (e.g. estimation of a failure probability) of such systems is thus nearly impossible with a limited computational budget. This general problematic motivates the research development presented in this thesis by focusing on the estimation of failure probability and its local sensitivity in high-dimensional spaces.

In a general context, expressing the numerical model by means of a deterministic black-box function $\mathcal{M}$, the behavior of the system is thus represented by the output Y of the black-box function, while its inputs represent all the identified variables that have an influence on the behavior of the system. As the system is subject to uncertainty and randomness, such as natural hazards, measurement errors or numerical approximations for instance, the inputs of this black-box function can be separated in two categories: random inputs $\mathbf{X}$ and deterministic inputs $\mathbf{s}$. It is assumed here that a probabilistic model allows to define the random inputs $\mathbf{X}$ by their joint probability distribution, while the deterministic inputs $\mathbf{s}$ gather fixed design parameters of the system or distribution parameters defining the random vector $\mathbf{X}$ for example. The output of the black-box function is consequently random, and a performance function is usually introduced to define the state of the system, according to the value of Y: it can either be in a safe state or in a failure state. This performance function commonly depends on a threshold, which defines the maximum or minimum acceptable value of the output Y. Taking the airfoil system introduced above, it is thus supposed that the shape of the airfoil relies on fixed design parameters, as well as random variables $\mathbf{X}$ which represent small deformations of the airfoil due to manufacturing uncertainties. These deformations are assumed to be independent and normally distributed. The failure of the system corresponds then to the drag coefficient of the airfoil exceeding a certain maximum threshold value. A large drag induces a larger fuel consumption which leads to

environmental and economical consequences.

Evaluating the system's probability of failure is therefore essential in order to assess the reliability analysis of the system. As the failure of the system is usually not a consequence of a normal behavior but rather an extreme response of the black-box model, this probability is small (typically inferior to 10^{-3}). The probability of failure is expressed as an integral defined over the failure domain of the random inputs $\mathbf{X}$ of the system. This integral can be estimated by various methods such as simulation methods based on the Monte Carlo method [Owen, 2013], or approximation methods such as FORM/SORM [Madsen et al., 2006]. This probability estimation can be quite cumbersome, depending on the shape of the failure domain of the random inputs, especially in high-dimensional space. Furthermore, the failure probability estimation has to require as few evaluations of the black-box function as possible, since this function is potentially numerically expensive, as for the airfoil system for instance.

The failure probability significantly depends on the deterministic parameters $\mathbf{s}$ of the system. The estimation of the sensitivities of the failure probability with respect to the design parameters of the system is then crucial in reliability-based design optimization (RBDO) [Moustapha and Sudret, 2019]. For example, the derivatives of the failure probability with respect to the design parameters of the airfoil presented above give enlightening insights on the optimal shape of the airfoil minimizing the drag coefficient. On the other hand, the derivatives of the failure probability with respect to the distribution parameters of the random deformations $\mathbf{X}$ allow to appreciate better the influence of the selected probabilistic model on the reliability of the system. Both sensitivities make it possible to understand the critical failure mechanisms of the airfoil. Depending on the nature of the deterministic inputs $\mathbf{s}$, several methods have been developed to estimate the failure probability sensitivities. For distribution parameters, simulation methods based on the score function method [Rubinstein, 1986] allow to obtained an unbiased sensitivity estimate in the original space of the random inputs. For design parameters, different approaches are available and most of them result in an approximation of the derivatives, as the direct differentiation leads to surface integrals. The evaluation of these sensitivities is performed at the same time or after the failure probability estimation. It must not lead to a significant additional amount of costly black-box function evaluations.

This manuscript focuses on systems with a large number of random inputs (typically superior to 50), whose failure domain can possibly encompass several failure regions. We further assume the random inputs $\mathbf{X}$ to follow a standard elliptical distribution [Fang et al., 2002], which may be the result of an isoprobabilistic transformation of the original inputs. The family of elliptical distributions contains the normal distribution and the Student distribution among others, and is very popular in many scientific and economical fields to model the random inputs.

Problem statement and objectives

In such a context, the number of methods available to estimate the failure probability considerably narrows. Indeed, many numerical techniques are efficient when the dimension of $\mathbf{X}$ is small (less than 50) but become inapplicable or irrelevant when the dimension increases. The *curse of dimensionality* is a well-know phenomenon which affects numerous approaches, such as the importance sampling (IS) method, through the degeneracy of the IS weights [Tabandeh et al., 2022]. Furthermore, as there may be multiple failure regions in the failure domain, classical approach to estimate the failure probability must be adapted. In the FORM/SORM approach with multiple design points, the standard elliptical space is thus searched to find all the failure regions [Der Kiureghian and Dakessian, 1998]. In the parametric IS framework, the auxiliary density is then constructed as a mixture of parametric densities, to take into account the multiple failure regions [Kurtz and Song, 2013]. However, few of these methods are efficient for both high-dimensional inputs and multiple failure regions, especially if the black-fox function is strongly nonlinear.

In addition, amongst the elliptical family, the behavior of the standard normal distribution singularly evolves as the dimension increases. Indeed, using the stochastic representation of the standard normal inputs, a particularly narrow important ring is defined with two hyperspheres and it contains most of the probability mass of the random inputs. This specificity of the high-dimensional standard normal law has to be taken into account in the estimation of the failure probability. Recently, the authors in [Wang and Song, 2016] presented a new parametric importance sampling approach with a mixture of von Mises–Fisher densities as auxiliary density for the directional component of the standard normal inputs, taking into account the important ring. Their approach has been enhanced with

the parametric importance sampling mixture presented in [Papaioannou et al., 2019], which assigns an auxiliary density to both the radial component of the standard normal inputs as well as their directional component. Nevertheless, the number of densities in the mixture, corresponding to the number of failure regions of the failure domain, has to be set in advanced and is particularly difficult to derive in a black-box context.

The first objective of this thesis is thus to **present a new method to estimate the failure probability of such a complex, high-dimensional system, efficient in the standard normal space, as well as for other standard elliptical distributions**. Without any hypothesis on the number of failure regions, the proposed method searches the input space for failure regions located in the important ring, with an approach inspired by the FORM/SORM framework. For each identified failure region, an auxiliary importance sampling density is constructed, using the stochastic representation of the inputs. The failure probability is finally estimated in a multiple importance sampling context, with a mixture of the auxiliary densities previously constructed.

As the estimation of the derivatives of the failure probability with respect to the deterministic variables of the system is performed during or after the failure probability estimation, the same framework is usually employed. Therefore, the sensitivity analysis is performed in the standard elliptical space as well. In this particular space, no matter the nature of the deterministic inputs, the derivatives of the failure probability lead to a surface integral difficult to compute. This surface integral can be approximated with a domain integral in the Weak approach framework [Torii, 2020], where the failure domain indicator function is replaced by a smoother function in the probability integral. However, a theoretical bias is then introduced in the sensitivity estimation. The resulting bias of the sensitivity estimate is then difficult to control. Other methods do not introduce bias but they require a heavier simulation budget, as the evaluation of the gradient of the black-box function is then necessary.

Consequently, the second objective of this thesis is to **introduce a new approach to estimate the failure probability derivatives with respect to deterministic inputs, in which the bias can be controlled and the simulation budget is kept low**. The sensitivity estimate is obtained as a byproduct of a heteroscedastic polynomial regression. The polynomial comes from the Taylor series expansion of the approximated sensitivity domain integral obtained with the Weak approach. In order to perform this heteroscedastic polynomial regression, the sample used for the failure probability estimation is reused and replicated to obtain other independent samples with bootstrap [Horowitz, 2001]. The bias of the sensitivity estimate is controlled with the degree of the polynomial used for the regression. The simulation budget is the same as in the Weak approach framework.

Outline of the thesis

This manuscript is composed of seven chapters including this one. The first three chapters present the state of the art in failure probability estimation and sensitivity estimation in the specific context mentioned above; the next two chapters describe in detail two algorithms developed during the thesis to achieve the two main objectives.

Chapter 2 describes the mathematical framework used to define the system and its random inputs. Some elements of the probability theory are introduced, with a special focus on the properties of the standard elliptical laws in high-dimensional space. The particular context of the thesis is further detailed as well as the two main objectives.

Chapter 3 details the state of the art in failure probability estimation. This presentation focuses on methods robust to high-dimensional spaces as well as multiple failure regions. The approximation methods based on the design points of the system are first described. Then the simulation methods based on the Monte Carlo method are presented, with an emphasis on the various cross-entropy based adaptive importance sampling methods. Finally, the approaches combining the information based on the design points location and simulation methods are presented. For all of the technique mentioned above, the way each method deals with multiple failure regions in a high-dimensional standard elliptical space is underlined.

Chapter 4 presents an overview of the state of the art in local reliability-based sensitivity analysis, with a focus on the estimation of the derivatives of the failure probability with respect to deterministic

inputs. Depending on the nature of the deterministic inputs, the local reliability-based sensitivity analysis can either be performed in the original space of the inputs or in the standard elliptical space. For each method, the statistical properties of the estimates are derived, as well as the amount of additional evaluations of the black-box function and its derivatives required, compared to the simulation budget necessary for the failure probability estimation only.

Chapter 5 describes in detail the first algorithm developed during the thesis, whose objective is to estimate the failure probability in a high-dimensional elliptical space, with a failure domain which possibly encompasses several failure regions. The outline of the method is first presented, with the introduction of the three major steps of the algorithm: identification problem, sampling problem and adaptation step. These three steps are gradually repeated until all the failure regions are (presumably) found. The proposed algorithm is tested with four numerical applications in which the inputs follow a standard normal distribution or a multivariate Student distribution. The first three examples are inspired by the literature while the last one is a realistic aerodynamic engineering application. For the first two numerical applications, the proposed algorithm is compared with the method presented in [Papaioannou et al., 2019] as well as a Subset sampling algorithm.

Chapter 6 presents in detail the second algorithm developed during the thesis, whose objective is to estimate the failure probability sensitivity with respect to deterministic inputs, with a controlled bias. The mathematical framework of the proposed approach is first described and starts with the approximation of the failure domain indicator function with a cumulative distribution function, as in the Weak approach. A change of random variable in the image measure, a differentiation and a Taylor series expansion are derived and make it possible to express the failure probability sensitivity as the constant coefficient of a polynomial. Employing simulation methods to build the polynomial regression database, the heteroscedastic polynomial regression framework of the linear least squares methods is required and presented. The proposed algorithm is tested with three numerical applications, all of them in the standard normal space. The first two examples are inspired by the literature and allow to test the proposed method in small dimensions. The last example makes it possible to assess the performance of the proposed method in high-dimensional space. For each example, the Weak approach estimates are used as benchmarks.

A short conclusion summarizes the main contributions of the thesis and presents some outlooks for future studies in Chapter 7.

Modélisation mathématique d'un système

Dans de nombreux domaines scientifiques, un système complexe est souvent modélisé grâce à une fonction $\mathcal{M}$, censée simuler le comportement du système. La sortie de cette fonction est la réponse observée tandis que les entrées représentent les différents paramètres influençant le comportement du système, comme des variables environnementales ou physiques. Une telle fonction, couramment appelée modèle entrée-sortie, est définie comme suit

$$\begin{aligned} \mathcal{M} : \quad \mathcal{Z} \subset \mathbb{R}^m &\longrightarrow \quad \mathcal{Y} \subset \mathbb{R} \\ \boldsymbol{\zeta} &\longmapsto \quad y \end{aligned}$$

où $\boldsymbol{\zeta} = (\zeta_1, ..., \zeta_m)^\top$ est le vecteur de taille m des variables d'entrée et y est la sortie scalaire. Cette fonction $\mathcal{M}$ est supposée déterministe dans la suite du manuscrit : il n'y a pas de stochasticité intrinsèque. De plus, l'évaluation de cette fonction est numériquement coûteuse, elle peut par exemple nécessiter la résolution d'un système d'équations aux dérivées partielles. Le comportement de $\mathcal{M}$ est assimilé à celui d'une une boîte noire et très peu d'information concernant $\mathcal{M}$ est disponible. Le gradient de $\mathcal{M}$ par rapport aux entrées du système $\boldsymbol{\zeta}$ est néanmoins supposé calculable dans la suite.

Plusieurs sources d'incertitudes et d'aléas peuvent agir sur le système. En conséquence, les entrées de la boîte noire $\mathcal{M}$ sont divisés en deux groupes : les variables déterministes rassemblées dans le vecteur $\mathbf{s} \in \mathbb{R}^p$ d'une part et les variables aléatoires rassemblées dans le vecteur aléatoire $\mathbf{X} \in \mathbb{R}^d$ d'autre part. Ainsi $m = p + d$. Néanmoins, la *dimension du système* mentionnée ensuite fait seulement référence à la valeur de la dimension d des variables aléatoires du système.

Selon la quantité d'information disponible concernant le vecteur aléatoire $\mathbf{X}$, différents cadres mathématiques existent pour modéliser les variables aléatoires du système [Ferson and Ginzburg, 1996, Qiu et al., 2008]. Dans ce manuscrit, le vecteur $\mathbf{X}$ est supposé absolument continu, de densité de probabilité $f_{\mathbf{X}}$ connue. La famille des lois elliptiques [Fang et al., 2018] jouant un rôle essentiel dans la suite de ce manuscrit, leurs propriétés en grande dimension ($d > 50$) sont brièvement présentées ici.

Supposons que le vecteur $\mathbf{X}$ suit une loi elliptique standard, la représentation stochastique du vecteur $\mathbf{X}$ s'écrit alors [Genest and Nešlehová, 2012]

$$\mathbf{X} = R\mathbf{T},$$

où $\mathbf{T}$ est un vecteur aléatoire uniformément distribué sur la sphère unité $\mathcal{T}^d = \{(t_1, \ldots, t_d) \in \mathbb{R}^d : t_1^2 + \cdots + t_d^2 = 1\}$ et R est une variable aléatoire positive et indépendante de $\mathbf{T}$. Le vecteur $\mathbf{T}$ est le composant directionnel de $\mathbf{X}$, alors que R est son composant radial, égal à sa norme Euclidienne $\|\mathbf{X}\| = R$. La densité de R permet de distinguer les lois standards elliptiques entre elles.

La loi standard normale multivariée fait partie de la famille des lois elliptiques; c'est la seule loi de cette famille qui mène à un vecteur aléatoire indépendant [Arnold and Lynch, 1982]. Le composant radial d'un vecteur normal standard suit une loi du chi, dont les propriétés en grande dimension ont été étudiées dans de nombreux travaux [Katafygiotis and Zuev, 2008, Canal, 2005]. Ainsi, lorsque la dimension du système est élevée, R peut en fait être approché par une loi normale de moyenne $\sqrt{d}$ et de variance $1/2$: $R \approx \mathcal{N}(\sqrt{d}, 1/2)$, pour $d \to +\infty$. De ce fait, la majeure partie de sa masse probabiliste se situe autour de $\sqrt{d}$. Un anneau d'importance est alors défini grâce à deux hypersphères de part et d'autre de l'hypersphère de rayon $\sqrt{d}$. Cet anneau d'importance englobe donc la région la plus essentielle de la densité de probabilité de la loi normale standard dans $\mathbb{R}^d$. Pour les autres lois

elliptiques standards, un anneau d'importance peut aussi être défini, il n'est néanmoins pas aussi étroit que celui de la loi normale.

Les lois standards elliptiques sont particulièrement courantes pour modéliser les variables aléatoires. Il existe de fait de nombreuses transformations isoprobabilistes permettant de retrouver les lois standards elliptiques [Bourinet, 2018, Sarazin, 2021]. Nous supposerons donc dans la suite du manuscrit que le vecteur aléatoire $\mathbf{X}$ suit une loi elliptique standard.

Afin de réaliser l'analyse de fiabilité du système, une fonction de performance g définie sur l'espace de sortie de la boîte noire $\mathcal{M}$ permet de déterminer l'état du système

- Le système est dans un état sûr pour les entrées $\boldsymbol{\zeta} = [\mathbf{s}, \mathbf{X}] \in \mathbb{R}^p \times \mathbb{R}^d$ vérifiant $g(\mathcal{M}(\boldsymbol{\zeta})) > 0$, le domaine de sûreté est alors défini par $D_s = \{\boldsymbol{\zeta} \in \mathbb{R}^m \,|g(\mathcal{M}(\boldsymbol{\zeta})) > 0\}$.
- Le système est dans un état défaillant pour les entrées $\boldsymbol{\zeta} = [\mathbf{s}, \mathbf{X}] \in \mathbb{R}^p \times \mathbb{R}^d$ vérifiant $g(\mathcal{M}(\boldsymbol{\zeta})) \leq 0$, le domaine de défaillance est défini par $D_f = \{\boldsymbol{\zeta} \in \mathbb{R}^m \,|g(\mathcal{M}(\boldsymbol{\zeta})) \leq 0\}$.

Cette fonction de performance est couramment définie à l'aide d'une valeur de réponse seuil y_{th} $g(\mathcal{M}(\boldsymbol{\zeta})) = y_{th} - \mathcal{M}(\boldsymbol{\zeta})$. La défaillance du système représente alors le dépassement de cette valeur seuil, qui est la valeur maximale acceptable pour la sûreté du système. L'hypersurface dans $\mathbb{R}^m$ vérifiant $\{\boldsymbol{\zeta} \,|\, g(\mathcal{M}(\boldsymbol{\zeta})) = 0\}$ est appelé surface d'état limite et permet de séparer en deux l'espace d'entrée entre le domaine de sûreté et le domaine de défaillance. Par simplification, la fonction $g \circ M : \mathbb{R}^m \to \mathbb{R}$ est par la suite appelée *fonction d'état limite* et est simplement notée g. Le gradient de cette fonction est supposé disponible. Puisque la fonction d'état limite est numériquement coûteuse, le nombre d'appel à cette fonction nécessaire à l'analyse de fiabilité du système doit rester le plus bas possible. Le nombre total d'appel à cette fonction et ses dérivées est appelé budget de simulation.

Dans ce manuscrit, nous allons nous intéresser tout d'abord au calcul de la probabilité de défaillance du système. En utilisant la représentation stochastique du vecteur $\mathbf{X}$, cette quantité s'écrit

$$P_f(\mathbf{s}) = \int_{D_f(\mathbf{s})} f_R(r) f_{\mathbf{T}}(\mathbf{t}) \,\mathrm{d}r\mathrm{d}\mathbf{t} = \int_{\mathbb{R}^+} \int_{\mathcal{T}^d} \mathbb{I}_{D_f(\mathbf{s})}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{t}) \,\mathrm{d}r\mathrm{d}\mathbf{t} = \mathbb{E}_{f_R, f_{\mathbf{T}}}\left[\mathbb{I}_{D_f(\mathbf{s})}(R\mathbf{T})\right],$$

où $D_f(\mathbf{s}) = \{r\mathbf{t} \in \mathbb{R}^d \,|\, g(\mathbf{s}, r\mathbf{t}) \leq 0\}$ est le domaine de défaillance pour un vecteur $\mathbf{s}$ fixé, et $\mathbb{I}_{D_f(\mathbf{s})}$ est la fonction indicatrice égale à 1 lorsque $g(\mathbf{s}, r\mathbf{t}) \leq 0$ et 0 sinon. L'opérateur $\mathbb{E}_{f_R, f_{\mathbf{T}}}$ dénote l'espérance par rapport à R et $\mathbf{T}$ simultanément. La fiabilité du système Re est définie par $Re = 1 - P_f$. Puisque la défaillance du système est présumée être un évènement rare, la probabilité de défaillance P_f est faible, i.e. inférieure à 10^{-3}.

L'estimation de cette probabilité de défaillance est particulièrement ardue. Dans ce manuscrit, deux hypothèses expliquent cette difficulté. Tout d'abord, la dimension d du système est supposée élevée, $d > 50$. L'intégrale $P_f(\mathbf{s})$ est alors définie sur un grand espace aléatoire. Ensuite, le système étant complexe, le domaine de défaillance $D_f(\mathbf{s})$ est supposé posséder plusieurs régions de défaillance. Plusieurs configurations différentes du vecteur $\mathbf{X}$ peuvent ainsi mener à la défaillance. Ces différentes régions de défaillance doivent toutes être prises en compte dans l'estimation de la probabilité de défaillance. En conclusion, l'estimation de cette intégrale est donc particulièrement difficile et des méthodes numériques adaptées doivent être considérées.

La probabilité de défaillance dépend également du vecteur de variables déterministes $\mathbf{s}$. Ainsi, dans ce manuscrit nous nous intéressons aussi à la dérivée de P_f par rapport à $\mathbf{s}$. En supposant que le gradient $\nabla_{r\mathbf{t}} g(\mathbf{s}, r\mathbf{t}) \neq \mathbf{0}$ pour tous $r\mathbf{t}$ et $\mathbf{s}$ sur la surface d'état limite $\{g(\mathbf{s}, r\mathbf{t}) = 0\}$, la dérivée de $P_f(\mathbf{s})$ par rapport à s_ℓ pour $\ell = 1, \ldots, p$ est égale à l'intégrale surfacique suivante [Breitung, 2006]

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = -\int_{g(\mathbf{s}, r\mathbf{t})=0} \frac{1}{\parallel \nabla_{r\mathbf{t}} g(\mathbf{s}, r\mathbf{t}) \parallel} \frac{\partial g(\mathbf{s}, r\mathbf{t})}{\partial s_\ell} f_R(r) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}s(r\mathbf{t}),$$

où $\mathrm{d}s(r\mathbf{t})$ dénote l'intégration surfacique sur la surface d'état limite $\{g(\mathbf{s}, r\mathbf{t}) = 0\}$. L'estimation d'une telle intégrale permets d'évaluer l'influence de $\mathbf{s}$ sur $P_f(\mathbf{s})$ et fait partie de l'analyse de sensibilité locale du système [Torii and Novotny, 2021].

L'estimation de cette intégrale surfacique peut être problématique, surtout pour une fonction d'état limite g englobant plusieurs régions de défaillance dans un espace d'entrée de grande dimension. Le principal défi de cette estimation est de recourir le moins possible à des nouvelles évaluations de

la fonction d'état limite. En effet, il s'agit de réutiliser les évaluations de la fonction d'état limite nécessaires à l'estimation de la probabilité de défaillance afin d'augmenter le budget de simulation de manière négligeable.

Chapter 2

Mathematical modeling of a system

Contents

2.1 Input-output black-box model of the system 19
2.2 Probabilistic modeling of the random inputs 20
2.2.1 Elements of probability theory 20
2.2.1.1 Moments of a random variable 21
2.2.1.2 Dependence structure of a random vector 21
2.2.1.3 Examples of probability distribution 24
2.2.2 Standard elliptical laws in high-dimensional space 27
2.2.2.1 The important ring of elliptical laws 27
2.2.2.2 Probabilistic transformation to elliptical laws 28
2.3 Performing a reliability analysis of the system 30
2.3.1 Estimation of the failure probability of the system 31
2.3.2 Probability of failure sensitivity analysis 31

2.1 Input-output black-box model of the system

In many scientific fields, a complex system is often modeled with a function $\mathcal{M}(\cdot)$ expected to simulate the behavior of the system. The output of this function is the observed response while the inputs represent the different parameters that have an influence over the behavior of the system, such as environmental or physical variables. This function is typically named input-output model and is mathematically defined as such :

$$\begin{aligned} \mathcal{M} : \quad \mathcal{Z} \subset \mathbb{R}^m &\longrightarrow \mathcal{Y} \subset \mathbb{R} \\ \boldsymbol{\zeta} &\longmapsto y \end{aligned}$$

where $\boldsymbol{\zeta} = (\zeta_1, ..., \zeta_m)^\top$ is a m-length vector of input variables and y is a scalar output. The function $\mathcal{M}$ can be given by an analytical formula or a numerically expensive computational code. For instance, it can require the resolution of a partial differential equations system by a numerical method such as finite volume or finite element. It is assumed here that $\mathcal{M}$ is deterministic. Consequently, there is no intrinsic stochasticity in the function $\mathcal{M}$: for the same set of inputs, the output will remain the same.

In this manuscript, the function $\mathcal{M}$ is seen as a black-box: no information about it is given. However, it is assumed that its gradient with respect to the parameter $\boldsymbol{\zeta}$ can be computed, which is nowadays often possible; see [Kenway et al., 2019, Yu et al., 2018] for recent examples in the aerospace field. Thus, for a particular set of inputs $\boldsymbol{\zeta}$, the response y and the gradient $[\partial\mathcal{M}/\partial\zeta_i]_{i=1,..,m}$ are supposed to be available. As the black-box function is numerically expensive (and can require several hours or days to run), the number of time $\mathcal{M}$ is evaluated has to be as low as possible.

The inputs of the system are divided into two groups : deterministic inputs and random inputs. The deterministic inputs represent variables which have fixed values, such as design parameters for instance. On the other hand, the random inputs represent variables which have uncertain measures or

whose value cannot be set due to a lack of knowledge or simply due to their nature. Amongst the m inputs variable, the deterministic variables are denoted $\mathbf{s} \in \mathbb{R}^p$ while the random variable are denoted $\mathbf{X} \in \mathbb{R}^d$. The total number of inputs m is thus the sum of the cardinal of those two groups : $m = p + d$. However, in the following chapters, when referring to the *dimension of the system*, only the number of random variables d is taken into account. It is assumed in this manuscript that the dimension of the system is high (d is above 50, say), in order to properly account for all sources of uncertainty. The random variables are typically denoted with uppercase letters, while deterministic variables are denoted with lowercase letters. The variables of dimension superior to one are written in bold.

2.2 Probabilistic modeling of the random inputs

When modeling the random inputs of the system, several frameworks exist depending on the amount of input information available and on the nature of the inputs [Ferson and Ginzburg, 1996, Qiu et al., 2008]. If the joint probability distribution of the random vector $\mathbf{X}$ can be easily obtained, the probabilistic approach [Apostolakis, 1990, Paté-Cornell, 1996] is employed to model the inputs of the system. On the contrary, if the probability distribution of some of the random variables cannot be obtained precisely, due to a lack of data for fitting their probabilistic models for example, non-probabilistic approaches (also called imprecise probabilities [Augustin et al., 2014]), such as possibility theory, fuzzy sets, interval theory (and others), are employed to model $\mathbf{X}$. In this manuscript, it is assumed that enough information is available to construct a probabilistic model for the random vector $\mathbf{X}$, thus non-probabilistic approaches are not presented.

2.2.1 Elements of probability theory

In order to define the probabilistic model, a probabilistic space is considered with the following triplet: $(\Omega, \mathcal{A}, \mathbb{P})$. Where Ω is a sample space, an arbitrary non-empty set, $\mathcal{A}$ is a σ-algebra, a set of subsets of Ω called events, and $\mathbb{P} : \mathcal{A} \to [0, 1]$ is a probability measure such that the measure of the entire sample space is equal to 1, $\mathbb{P}(\Omega) = 1$. A continuous random vector $\mathbf{X}$ of dimension d is then a (measurable) function of $(\Omega, \mathcal{A}, \mathbb{P})$ in the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, where $\mathcal{B}(\mathbb{R}^d)$ stands for the Borel set of $\mathbb{R}^d$.

Let $\mathbf{X}$ be a random vector of dimension d. The distribution of $\mathbf{X}$ is uniquely defined in the following manner

- by its law $\mathbb{P}_{\mathbf{X}}$
$$\forall\, B \in \mathcal{B}(\mathbb{R}^d), \qquad \mathbb{P}_{\mathbf{X}}(B) = \mathbb{P}(\mathbf{X} \in B),$$
it is thus assumed here that the law $\mathbb{P}_X$ is absolutely continuous with respect to Lebesgue's measure.

- by its joint cumulative distribution function (cdf) $F_{\mathbf{X}}$
$$\forall\, \mathbf{x} \in \mathbb{R}^d, \qquad F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}}(x_1, \ldots, x_d) = \mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d),$$
or its survival function $H_{\mathbf{X}}$
$$\forall\, \mathbf{x} \in \mathbb{R}^d, \qquad H_{\mathbf{X}}(\mathbf{x}) = H_{\mathbf{X}}(x_1, \ldots, x_d) = \mathbb{P}(X_1 > x_1, \ldots, X_d > x_d),$$

- or (in the absolutely continuous case) by its joint probability density function (pdf) $f_{\mathbf{X}}$
$$\forall\, \mathbf{x} \in \mathbb{R}^d, \qquad f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}}(x_1, \ldots, x_d) = \frac{\partial^d F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \cdots \partial x_d},$$
it is assumed here that the random vector $\mathbf{X}$ is absolutely continuous in the following, thus its pdf $f_{\mathbf{X}}$ is available.

A few interesting attributes of the distribution of $\mathbf{X}$ are presented in the following sections.

2.2.1.1 Moments of a random variable

The moments of a random variable X, when they exist, are characteristics that give additional information about the law of X. The moment of order α is defined, thanks to the transport theorem, by the following quantity

$$\forall \alpha \in \mathbb{N} \quad \mathbb{E}[X^\alpha] = \int_{\mathbb{R}} x^\alpha f_X(x) \mathrm{d}x,$$

where f_X is the pdf of the random variable X. The quantity $\mathbb{E}\left[(X - \mathbb{E}[X])^\alpha\right]$ is called centered moment of order α. In particular, the moment of order 1 written μ_X is called mean value of X and the centered moment of order 2 written $\mathbb{V}\mathrm{ar}(X)$ is called variance of X. The variance of X is always positive. The standard deviation is defined by the square root of the variance and written $\sigma_X = \sqrt{\mathbb{V}\mathrm{ar}(X)}$. Finally, the coefficient of variation (CV) is equal to the ratio between the standard deviation of X and its mean

$$\mathrm{CV}(X) = \frac{\sigma_X}{\mu_X}.$$

If $X \geq 0$, a definition of $\mathbb{E}[X^\alpha]$ is possible when α is not a natural number and is called fractional moments.

When dealing with a random vector $\mathbf{X}$ of dimension d, the mean value $\boldsymbol{\mu}_\mathbf{X}$ and the covariance matrix written $\boldsymbol{\Sigma}_\mathbf{X}$ are given by

$$\begin{aligned} \boldsymbol{\mu}_\mathbf{X} &= (\mathbb{E}[X_1], \ldots, \mathbb{E}[X_d]) \quad \in \quad \mathbb{R}^d, \\ \boldsymbol{\Sigma}_\mathbf{X} &= \begin{pmatrix} \mathrm{Cov}(X_1, X_1) & \cdots & \mathrm{Cov}(X_1, X_d) \\ \vdots & & \vdots \\ \mathrm{Cov}(X_d, X_1) & \cdots & \mathrm{Cov}(X_d, X_d) \end{pmatrix} \quad \in \quad M^d(\mathbb{R}), \end{aligned}$$

where $\mathrm{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]$ is the covariance between the margins X_i and X_j and $M^d(\mathbb{R})$ designates the set of square matrix of size $d \times d$. The covariance between the same margin X_i is thus equal to its variance $\mathrm{Cov}(X_i, X_i) = \mathbb{E}[(X_i - \mathbb{E}[X_i])^2] = \mathbb{V}\mathrm{ar}(X_i)$. Any covariance matrix is symmetric and positive semi-definite. In probabilistic modeling, the covariance matrix $\boldsymbol{\Sigma}_\mathbf{X}$ is often associated with the linear correlation matrix $R_\mathbf{X}$. The coefficient of this matrix are equal to Pearson's correlation (written ρ) and defined by

$$[R_\mathbf{X}]_{i,j} = \rho_{i,j} = \frac{\mathrm{Cov}(X_i, X_j)}{\sigma_{X_i} \sigma_{X_j}}.$$

It is assumed in this manuscript from now on that $\mathbf{X}$ is a second order random vector, thus $\mathbf{X} \in L^2(\Omega, \mathbb{R}^d)$ which ensures that the mean value $\boldsymbol{\mu}_\mathbf{X}$ and covariance matrix $\boldsymbol{\Sigma}_\mathbf{X}$ are well-defined and finite.

2.2.1.2 Dependence structure of a random vector

An essential information about the distribution of the random vector $\mathbf{X}$ is the dependence relationship between the univariate margins X_i. If the univariate margins $X_1, \ldots, X_d$ are independent, the distribution of the random vector $\mathbf{X}$ is given by the product of the univariate marginal probability functions

$$\begin{aligned} \text{CDF}: \quad & \forall \mathbf{x} \in \mathbb{R}^d \quad F_\mathbf{X}(\mathbf{x}) = F_\mathbf{X}(x_1, \ldots, x_d) = \prod_{i=1}^d F_{X_i}(x_i), \\ \text{PDF}: \quad & \forall \mathbf{x} \in \mathbb{R}^d \quad f_\mathbf{X}(\mathbf{x}) = f_\mathbf{X}(x_1, \ldots, x_d) = \prod_{i=1}^d f_{X_i}(x_i). \end{aligned}$$

However, when there is a dependence between the univariate margins, the marginal laws are not enough to describe the joint distribution. Copulas [Sklar, 1959] are then used to define the dependence structure of the random vector $\mathbf{X}$. A copula is a multivariate cdf with uniform univariate margins on the interval $[0, 1]$. Sklar [Sklar, 1959] showed that given d random variables $X_1, \ldots, X_d$ with univariate cdf $F_{X_1}, \ldots, F_{X_d}$ respectively, there always exists a copula $C_\mathbf{X}$ such that

$$\forall \mathbf{x} \in \mathbb{R}^d \quad F_\mathbf{X}(x_1, \ldots, x_d) = \mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d) = C_\mathbf{X}(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)). \tag{2.1}$$

When $F_{X_1}, \dots, F_{X_d}$ are continuous, then $C_{\mathbf{X}}$ is unique and can be retrieved from the joint distribution of $\mathbf{X}$ as

$$\begin{aligned}\forall\, \mathbf{u} \in [0,1]^d \quad C_{\mathbf{X}}(u_1, \dots, u_d) &= \mathbb{P}(X_1 \leq F_{X_1}^{-1}(u_1), \dots, X_d \leq F_{X_d}^{-1}(u_d)) \\ &= F_{\mathbf{X}}(F_{X_1}^{-1}(u_1), \dots, F_{X_d}^{-1}(u_d)),\end{aligned}$$

where $F_{X_i}^{-1}$ denotes the generalized inverse of F_{X_i}. Taking the derivative of Eq. (2.1), Sklar's representation of copulas can be derived for the density functions as

$$\forall\, \mathbf{x} \in \mathbb{R}^d \quad f_{\mathbf{X}}(x_1, \dots, x_d) = c_{\mathbf{X}}(F_{X_1}(x_1), \dots, F_{X_d}(x_d)) \times \prod_{i=1}^{d} f_{X_i}(x_i),$$

where $c_{\mathbf{X}}$ is the copula density of $\mathbf{X}$ defined by

$$\forall\, \mathbf{u} \in [0,1]^d \quad c_{\mathbf{X}}(\mathbf{u}) = \frac{\partial^d C_{\mathbf{X}}(\mathbf{u})}{\partial u_1 \cdots \partial u_d}.$$

Therefore, the independent case is just a particular case where the copula $C_{\mathbf{X}}$ and the copula density are defined by

$$\forall\, \mathbf{u} \in [0,1]^d \quad C_{\mathbf{X}}(u_1, \dots, u_d) = \prod_{i=1}^{d} u_i,$$

thus $c_{\mathbf{X}}(\mathbf{u}) = 1$. The theorem of Sklar written above is essential in the copula theory since it stipulates that the distribution of a random vector $\mathbf{X}$ is entirely characterized with its univariate margins f_{X_i} and its copula $C_{\mathbf{X}}$. In the next paragraphs, we present the meta-elliptical copulas and the Archimedean copulas, which are the most famous copula families. However, there exist many other copula families like the extreme-value copula family [Galambos, 1978] or the Bernstein copula family [Sancetta and Satchell, 2004] for example.

Meta-elliptical copulas

The family of meta-elliptical copulas was first presented in [Fang et al., 2002] and comes from the family of elliptical distributions [Fang et al., 2018]. A random vector $\mathbf{X}$ of dimension d is said to have an elliptical distribution $\mathcal{E}(\boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{D}_{\mathbf{X}}, h)$ with mean vector $\boldsymbol{\mu}_{\mathbf{X}}$, dispersion matrix $\boldsymbol{D}_{\mathbf{X}}$ and density generator $h : \mathbb{R}^+ \to \mathbb{R}^+$ if it can be expressed in the form

$$\mathbf{X} = RA\mathbf{T} + \boldsymbol{\mu}_{\mathbf{X}}, \tag{2.2}$$

where $AA^\top$ is the Cholesky decomposition of $\boldsymbol{D}_{\mathbf{X}}$, $\mathbf{T}$ is a random vector uniformly distributed on the unit sphere $\mathcal{T}^d = \{(t_1, \dots, t_d) \in \mathbb{R}^d : t_1^2 + \dots + t_d^2 = 1\}$ and R is a positive random variable independent of $\mathbf{T}$ with density

$$\forall\, r \in \mathbb{R}^+ \quad f_R(r) = \frac{2\pi^{d/2}}{\Gamma(d/2)} r^{d-1} h\left(r^2\right),$$

where Γ is the gamma function. The random vector $\mathbf{T}$ is named directional component while the random variable R is named radial component and Eq. (2.2) is called stochastic representation of $\mathbf{X}$. The multivariate pdf of $\mathbf{X}$ is then given by

$$\forall\, \mathbf{x} \in \mathbb{R}^d \quad f_{\mathbf{X}} = \frac{1}{\sqrt{\det \boldsymbol{D}_{\mathbf{X}}}} h\left((\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^\top \boldsymbol{D}_{\mathbf{X}}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})\right). \tag{2.3}$$

The underlying copula of $\mathbf{X}$ is said to be meta-elliptical, but is not available in closed form [Genest and Nešlehová, 2012]. The covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}}$ of the random vector $\mathbf{X}$ can be retrieved thanks to the stochastic representation of $\mathbf{X}$

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{X}} = \mathrm{Cov}(\mathbf{X}) &= \mathbb{E}\left[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^\top\right] = \mathbb{E}\left[RA\mathbf{T}(RA\mathbf{T})^\top\right] \\ &= \mathbb{E}\left[R^2 A\mathbf{T}\mathbf{T}^\top A^\top\right] = \mathbb{E}\left[R^2\right] A\mathbb{E}\left[\mathbf{T}\mathbf{T}^\top\right] A^\top,\end{aligned}$$

since R and $\mathbf{T}$ are independent. Then, as $\mathbb{E}\left[\mathbf{T}\mathbf{T}^\top\right] = \frac{1}{d}\mathbf{I}_d$ [Lapuyade-Lahorgue, 2016], where $\mathbf{I}_d$ is the identity matrix of $M^d(\mathbb{R})$, the following relationship holds

$$\mathbf{\Sigma}_\mathbf{X} = \frac{\mathbb{E}\left[R^2\right]}{d}\boldsymbol{D}_\mathbf{X}. \tag{2.4}$$

Thus the covariance matrix is proportional to the dispersion matrix and the constant depends on the distribution of R^2 [Lebrun and Dutfoy, 2009]. The name *elliptical* comes from the isodensities of the elliptical random vector being ellipsoids. If $\boldsymbol{\mu}_\mathbf{X} = \mathbf{0}$ and $\boldsymbol{D}_\mathbf{X} = \mathbf{I}_d$, then the random vector $\mathbf{X}$ is said to follow a standard elliptical distribution, or a *spherical* distribution, because its isodensities are spherical, centered on $\mathbf{0}$. The stochastic representation of a standard elliptical vector $\mathbf{X}$ is then

$$\mathbf{X} = R\mathbf{T}, \tag{2.5}$$

with the same notation as Eq. (2.2), and $\|\mathbf{X}\| = R$, with $\|\cdot\|$ the Euclidean norm. Its probability density function $f_\mathbf{X}$ is then symmetric and a function of $r = \|\mathbf{x}\|$ written

$$\forall\, \mathbf{x} \in \mathbb{R}^d \quad f_\mathbf{X} = h\left(\mathbf{x}^\top\mathbf{x}\right) = h\left(r^2\mathbf{t}^\top\mathbf{t}\right) = h\left(r^2\right) = h\left(\|\mathbf{x}\|^2\right). \tag{2.6}$$

The meta-elliptical family encompasses various well-known distributions, such as the Normal distribution, the Student distribution and the Cauchy distribution for example. For each of these three elliptical laws [Genest et al., 2007], the respective distribution of R^2 and the density generator h are displayed in Table 2.1, where $\chi^2(d)$ is the chi-squared distribution with d degrees of freedom and $\mathcal{F}$ is the Fisher–Snedecor distribution. The definition of these distributions will be given in detail in the following section. The meta-elliptical copulas enables to have tail dependence amongst the univariate

Table 2.1: Distributions of R^2 and density generators of three elliptical laws: the normal distribution, the Student distribution with ν degree of freedom and the Cauchy distribution.

Copula	Distribution of R^2	Generator h
Normal	$R^2 \sim \chi^2(d)$	$h : t \to (2\pi)^{-d/2}\exp\left(-\frac{t}{2}\right)$
Student	$R^2/d \sim \mathcal{F}(d,\nu)$	$h : t \to \frac{(\pi\nu)^{-d/2}\Gamma\left(\frac{d+\nu}{2}\right)}{\Gamma(\nu/2)}(1+\frac{t}{\nu})^{-(d+\nu)/2}$
Cauchy	$R^2/d \sim \mathcal{F}(d,1)$	$h : t \to \frac{(\pi)^{-d/2}\Gamma\left(\frac{d+1}{2}\right)}{\Gamma(1/2)}(1+t)^{-(d+1)/2}$

margins which can be very practical to model certain phenomenon [Joe, 1997].

Archimedean copulas

The family of Archimedean copulas first appeared in the context of probabilistic metric spaces [Schweizer and Sklar, 1983]. It has been proven in [McNeil and Nešlehová, 2009] that the Archimedean copulas are strongly linked with the ℓ_1-norm symmetric (or simplex) distribution. A random vector $\mathbf{X}$ of dimension d is said to have an ℓ_1-norm symmetric distribution if it can be expressed in the form

$$\mathbf{X} = R\mathbf{S}, \tag{2.7}$$

where R is a positive random variable of cdf F_R and $\mathbf{S}$ is a random vector, independent of R, and uniformly distributed on the unit simplex $\mathcal{S}^d = \{(s_1,\ldots,s_d) \in [0,1]^d : s_1 + \cdots + s_d = 1\}$. Eq. (2.7) is the stochastic representation of the random vector $\mathbf{X}$. The distribution of $\mathbf{X}$ is given by the following survival function [McNeil and Nešlehová, 2009]

$$\forall\, \mathbf{x} \in \mathbb{R}^d \quad H_\mathbf{X}(\mathbf{x}) = \mathcal{W}_d F_R\left(\|\max(\mathbf{x},\mathbf{0})\|_1\right) + F_R(0)\mathbb{I}_{\mathbf{x}<0}(\mathbf{x}),$$

where $\mathcal{W}_d$ is the Williamson d-transform [Williamson, 1956] and $\mathbb{I}_{\mathbf{x}<0}$ is the indicator function equal to 1 when $\mathbf{x} < 0$ and 0 otherwise. The random vector $\mathbf{X}$ has then an underlying Archimedean copula, which can be expressed [Ling, 1965] in the simple algebraic form

$$\forall\, \mathbf{u} \in [0,1]^d \quad C_\mathbf{X}(u_1,\ldots,u_d) = \psi(\psi^{-1}(u_1) + \ldots + \psi^{-1}(u_d)), \tag{2.8}$$

where $\psi : \mathbb{R}^+ \to [0,1]$ is called generator function of $C_{\mathbf{X}}$. This generator function is related to the cdf F_R of the radial part R of $\mathbf{X}$ as the following equalities holds [McNeil and Nešlehová, 2009]

$$\psi = \mathcal{W}_d F_R \qquad F_R = \mathcal{W}_d^{-1}\psi$$

where $\mathcal{W}_d^{-1}$ is the inverse Williamson d-transform. The Archimedean family encompasses various well-known copulas such as the Gumbel copula, the Frank copula and the Clayton copula for example. The interested reader may find their respective generator function in [Sarazin, 2021]. In Table 2.2 is displayed the example of the independence copula introduced above, which belongs to the Archimedean family, as an illustration of the links between $C_{\mathbf{X}}, \psi, H_{\mathbf{X}}$ and F_R. The resulting survival function $H_{\mathbf{X}}$ is the one of the independent standard exponential vector, a well known ℓ_1-norm symmetric distribution.

Table 2.2: Cumulative distribution function of R, generator function and survival function associated to the independence copula, an example of the links between ψ, F_R and $H_{\mathbf{X}}$ for the Archimedean copulas.

Copula	$\forall\, \mathbf{u} \in [0,1]^d \;\; C_{\mathbf{X}}(u_1,\dots,u_d) = \prod_{i=1}^d u_i$
Generator ψ	$\forall\, x \in \mathbb{R}^+ \;\; \psi(x) = \exp(-x)$
CDF of R	$\forall\, x \in \mathbb{R}^+ \;\; F_R(x) = 1 - \sum_{n=0}^{d-1} \frac{x^n}{n!}\exp(-x)$ (Erlang distribution)
Survival function	$\forall\, \mathbf{x} \in \mathbb{R}^d \;\; H_{\mathbf{X}}(\mathbf{x}) = \exp\left(-\sum_{i=1}^d x_i\right)$

The symmetry of Eq. (2.8) implies that all pairs of variables share the same dependence structure, which can be a major drawback when modeling random inputs. Thus [Joe, 1996] introduced the hierarchical Archimedean copulas, which are nested classes of Archimedean copulas that can involve up to $p-1$ generators $\psi_1,\dots,\psi_{p-1}$ and thus introduce new dependence. They are not derived here but the interested reader can find more details in [Grimaldi and Serinaldi, 2006].

2.2.1.3 Examples of probability distribution

Some distributions already mentioned are further detailed in this section. As previously noted, it is assumed that the random variables and random vectors are continuous. This list of probability distribution is not exhaustive but it allows a better understanding of the rest of the manuscript.

1. Univariate distributions

Uniform distribution $\mathcal{U}([a,b])$
Let a and b be two real numbers such that $a \neq b$. A random variable X follows a uniform distribution on the interval $[a,b]$, written $X \sim \mathcal{U}([a,b])$, if its density function f_X is given by

$$\forall\, x \in [a,b] \qquad f_X(x) = \frac{1}{b-a},$$

which is constant on the interval $[a,b]$. The mean and the variance of X are given by $\mu_X = \frac{1}{2}(a+b)$ and $\mathbb{V}\mathrm{ar}(X) = \frac{1}{12}(b-a)^2$.

Exponential distribution $E(\lambda)$
Let $\lambda \in \mathbb{R}^+\backslash\{0\}$ be a strictly positive scalar. A random variable X follows an exponential distribution of rate λ if its density function f_X is given by

$$\forall\, x \in \mathbb{R}^+ \qquad f_X(x) = \lambda\exp(-\lambda x).$$

The mean and variance of X are given by $\mu_X = \frac{1}{\lambda}$ and $\mathbb{V}\mathrm{ar}(X) = \frac{1}{\lambda^2}$.

Chi distribution $\chi(\nu)$
Let $\nu \in \mathbb{N}^*$ be an integer. The χ distribution with ν degrees of freedom is the distribution of the

square root of a sum of squares of ν independent standard normal random variables

$$X = \sqrt{\sum_{i=1}^{\nu} Z_i^2},$$

where $Z_1, \ldots, Z_\nu \sim \mathcal{N}(0,1)$. The distribution is written $X \sim \chi(\nu)$, and its density function f_X is given by

$$\forall\, x \in \mathbb{R}^+ \qquad f_X(x) = \frac{1}{2^{(\nu/2)-1}\Gamma(\nu/2)} x^{\nu-1} \exp\left(-\frac{x^2}{2}\right).$$

The mean and the variance of X are given by $\mu_X = \sqrt{2}\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)}$ and $\mathbb{V}\mathrm{ar}(X) = \nu - \mu_X^2$.

Chi-squared distribution $\chi^2(\nu)$
Let $\nu \in \mathbb{N}^*$ be an integer. The χ^2 distribution with ν degrees of freedom is simply the distribution of the square of a χ-distributed random variable with ν degrees of freedom

$$X = \sum_{i=1}^{\nu} Z_i^2,$$

where $Z_1, \ldots, Z_\nu \sim \mathcal{N}(0,1)$. The distribution is written $X \sim \chi^2(\nu)$, and its density function f_X is given by

$$\forall\, x \in \mathbb{R}^+ \qquad f_X(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} \exp\left(-\frac{x}{2}\right).$$

The mean and the variance of X are given by $\mu_X = \nu$ and $\mathbb{V}\mathrm{ar}(X) = 2\nu$. If $\nu = 1$ then the definition domain is restricted to $\mathbb{R}^+\backslash\{0\}$.

Fisher–Snedecor distribution $\mathcal{F}(\kappa, \nu)$
Let $\kappa, \nu \in \mathbb{N}^*$ be two integers. The Fisher–Snedecor distribution, or F-distribution, with κ and ν degrees of freedom is the distribution of the ratio of two independent χ^2-distributed variables S_1 and S_2 with respective degrees of freedom κ and ν

$$X = \frac{S_1/\kappa}{S_2/\nu},$$

where $S_1 \sim \chi^2(\kappa)$ and $S_2 \sim \chi^2(\nu)$. The distribution is written $X \sim \mathcal{F}(\kappa, \nu)$, and its density function f_X is given by

$$\forall\, x \in \mathbb{R}^+ \qquad f_X(x) = \frac{1}{\mathrm{B}\left(\frac{\kappa}{2}, \frac{\nu}{2}\right)} \left(\frac{\kappa}{\nu}\right)^{\kappa/2} x^{\kappa/2-1} \left(1 + \frac{\kappa}{\nu} x\right)^{-(\kappa+\nu)/2},$$

where B is the Beta function. The mean and the variance of X are not necessarily well-defined. If $\nu > 2$ then the mean is finite and $\mu_X = \frac{\nu}{\nu-2}$. If $\nu > 4$ then the variance is finite and $\mathbb{V}\mathrm{ar}(X) = \frac{2\nu^2(\kappa+\nu-2)}{d(\nu-2)^2(\nu-4)}$. Otherwise, μ_X and $\mathbb{V}\mathrm{ar}(X)$ are undefined. If $\kappa = 1$ then the definition domain is restricted to $\mathbb{R}^+\backslash\{0\}$.

2. Multivariate distributions

Uniform distribution on the unit sphere $\mathcal{U}(\mathcal{T}^d)$
The unit sphere of dimension $d-1$ is the set $\mathcal{T}^d = \{(t_1, \ldots, t_d) \in \mathbb{R}^d : t_1^2 + \cdots + t_d^2 = 1\}$ defined in the d-dimensional space. Let $\mathbf{X}$ be a random vector following the uniform distribution on $\mathcal{T}^d$, written $\mathbf{X} \sim \mathcal{U}(\mathcal{T}^d)$. The uniform distribution on the unit sphere is the most natural standard elliptical distribution. Indeed, the stochastic decomposition of $\mathbf{X}$ is simply $\mathbf{X} = \mathbf{T}$ since R is a constant equal to 1. The density function of $\mathbf{X}$ is given by

$$\forall\, \mathbf{x} \in \mathcal{T}^d \qquad f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{A_d} \qquad \text{where} \qquad A_d = \frac{d\,\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)},$$

and is thus constant with A_d the surface area of $\mathcal{T}^d$. The mean and the covariance matrix of $\mathbf{X}$ are given by $\boldsymbol{\mu}_{\mathbf{X}} = \mathbf{0}$, $\boldsymbol{\Sigma}_{\mathbf{X}} = \frac{1}{d}\mathbf{I}_d$, where $\mathbf{I}_d$ is the identity matrix. It should be noted that although the

correlations between each margin are equal to zero, the random vector is, of course, not independent.

Von Mises–Fisher distribution on the unit sphere $\mathcal{T}^d$
The von Mises–Fisher distribution depends on two parameters, namely the concentration parameter $\kappa \in \mathbb{R}^+$ and the mean direction parameter $\boldsymbol{\nu} \in \mathcal{T}^d$. Let $\mathbf{X}$ be a random vector following a von Mises–Fisher distribution, its density function $f_{\mathbf{X}}$ is given by

$$\forall \mathbf{x} \in \mathcal{T}^d \qquad f_{\mathbf{X}}(\mathbf{x}) = c_d(\kappa)\exp\left(\kappa\, \boldsymbol{\nu}^\top \mathbf{x}\right) \qquad \text{with} \qquad c_d(\kappa) = \frac{\kappa^{d/2-1}}{(2\pi)^{d/2} I_{d/2-1}(\kappa)},$$

where $I_{d/2-1}$ denotes the modified Bessel function of the first kind of order $d/2-1$. If $\kappa = 0$, the distribution boils down to the uniform distribution on $\mathcal{T}^d$. As κ grows, the distribution is more and more concentrated around the mean direction $\boldsymbol{\nu}$, as illustrated in Figure 2.1. The mean value of the von Mises–Fisher distribution does not belong to the unit sphere and is equal to $\boldsymbol{\mu}_{\mathbf{X}} = \frac{I_{d/2}(\kappa)}{I_{d/2-1}(\kappa)}\boldsymbol{\nu}$.

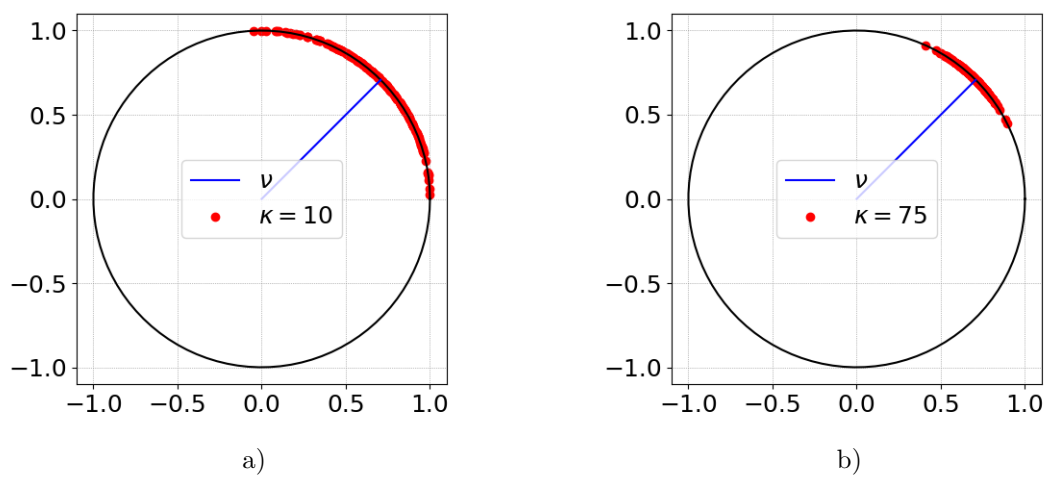


Figure 2.1: The von Mises–Fisher distribution on the unit circle $\mathcal{T}^2$ with $\boldsymbol{\nu} = [0.71, 0.71]$ and a) $\kappa = 10$ b) $\kappa = 75$. The blue line represents the mean direction $\boldsymbol{\nu}$. The red dots represent a sample generated with the chosen κ.

Multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\Sigma}_{\boldsymbol{X}})$
The multivariate normal distribution is an elliptical distribution as previously mentioned. Let $\mathbf{X}$ be a random vector of mean $\mu_{\mathbf{X}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}}$ following a multivariate normal distribution, written $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{\Sigma}_{\mathbf{X}})$. Then its stochastic representation is $\mathbf{X} = RA\mathbf{T} + \boldsymbol{\mu}_{\mathbf{X}}$ where $R^2 \sim \chi^2(d)$ (thus $R \sim \chi(d)$), $\mathbf{T} \sim \mathcal{U}(\mathcal{T}^d)$ and $AA^\top = \boldsymbol{D}_{\mathbf{X}}$. The covariance matrix is equal to the dispersion matrix $\boldsymbol{\Sigma}_{\mathbf{X}} = \boldsymbol{D}_{\mathbf{X}}$ as $\mathbb{E}(R^2) = d$ in the equation Eq. (2.4). The density function $f_{\mathbf{X}}$ is thus given by

$$\forall \mathbf{x} \in \mathbb{R}^d \qquad f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \boldsymbol{\Sigma}_{\mathbf{X}}}} \exp\left(-\frac{(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{X}})^\top \boldsymbol{\Sigma}_{\mathbf{X}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{X}})}{2}\right), \tag{2.9}$$

which is the special case of Eq. (2.3) with $\boldsymbol{D}_{\mathbf{X}} = \boldsymbol{\Sigma}_{\mathbf{X}}$ and the generator function $h : t \to (2\pi)^{-d/2}\exp\left(-\frac{t}{2}\right)$. The univariate version of this distribution is straightforward, as the covariance matrix is reduced to the variance of X equal to σ_X^2 and d is set to 1 in Eq. (2.9). When $\boldsymbol{\mu}_{\mathbf{X}} = \mathbf{0}$ and $\boldsymbol{\Sigma}_{\mathbf{X}} = \mathbf{I}_d$, the identity matrix of dimension d, the distribution is called standard normal distribution. In the rest of the manuscript, the univariate standard normal cdf will be written Ψ and its pdf will be written ϕ.

Multivariate Student distribution $t_\nu(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{D}_{\boldsymbol{X}})$
The multivariate Student distribution is also an elliptical distribution as previously mentioned. The Student distribution of mean $\boldsymbol{\mu}_{\mathbf{X}}$ and dispersion matrix $\boldsymbol{D}_{\mathbf{X}}$, written $\mathbf{X} \sim t_\nu(\boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{D}_{\mathbf{X}})$, is the distribution of the ratio of a centered normal vector $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{D}_{\mathbf{X}})$ and a χ^2-distributed variable U with

degree of freedom $\nu > 0$

$$\mathbf{X} = \frac{\mathbf{Z}}{\sqrt{U/\nu}} + \boldsymbol{\mu}_{\mathbf{X}},$$

where $\mathbf{Z}$ and U are independent. Its stochastic representation is written $\mathbf{X} = RA\mathbf{T} + \boldsymbol{\mu}_{\mathbf{X}}$ where $R^2/d \sim \mathcal{F}(d,\nu)$, $\mathbf{T} \sim \mathcal{U}(\mathcal{T}^d)$ and $AA^\top = \boldsymbol{D}_{\mathbf{X}}$. When $\nu > 2$, the covariance matrix is equal to $\boldsymbol{\Sigma}_{\mathbf{X}} = \frac{\nu}{\nu-2}\boldsymbol{D}_{\mathbf{X}}$ since $\mathbb{E}(R^2) = d \times \frac{\nu}{\nu-2}$ in the equation Eq. (2.4). The density function $f_{\mathbf{X}}$ is thus given by

$$\forall\, \mathbf{x} \in \mathbb{R}^d \qquad f_{\mathbf{X}}(\mathbf{x}) = \frac{(\pi\nu)^{-d/2}\Gamma\left(\frac{d+\nu}{2}\right)}{\Gamma(\nu/2)\sqrt{\det \boldsymbol{D}_{\mathbf{X}}}}\left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{X}})^\top \boldsymbol{D}_{\mathbf{X}}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{X}})}{\nu}\right)^{-(d+\nu)/2}, \tag{2.10}$$

which is the special case of Eq. (2.3) with generator function $h : t \to \frac{(\pi\nu)^{-d/2}\Gamma\left(\frac{d+\nu}{2}\right)}{\Gamma(\nu/2)}(1+\frac{t}{\nu})^{-(d+\nu)/2}$. The univariate version of this distribution is straightforward, as the dispersion matrix is reduced to 1, the mean value is equal to 0 and d is set to 1 in Eq. (2.10).

2.2.2 Standard elliptical laws in high-dimensional space

In this manuscript, standard elliptical distributions play a major role. Consequently, these distributions are now further described. Assume $\mathbf{X}$ follows a standard elliptical law. This distribution family has been introduced in the previous section: the stochastic representation of the random vector $\mathbf{X}$ is thus given by Eq (2.5). It should be noted that although $\boldsymbol{D}_{\mathbf{X}} = \mathbf{I}_d$, only the multivariate standard normal distribution has independent components, all the other standard elliptical distributions include dependence [Arnold and Lynch, 1982]. In this section, the behavior of those laws when the dimension d of the system is high is depicted.

2.2.2.1 The important ring of elliptical laws

The multivariate normal distribution

In the literature, many reliability studies are performed in the high-dimensional standard normal space [Wang and Song, 2016, Papaioannou et al., 2019, Wang and Song, 2018]. This space has specific properties that are derived here. Let $\mathbf{X}$ be a standard normal vector of dimension d. Then, as previously mentioned, $\mathbf{X} = R\mathbf{T}$, where $R^2 \sim \chi^2(d)$ and $\mathbf{T} \sim \mathcal{U}(\mathcal{T}^d)$ and R and $\mathbf{T}$ are independent. This stochastic representation is sometimes called polar representation. Since $R^2 \sim \chi^2(d)$ means that $R \sim \chi(d)$, the properties of the chi distribution apply to the radial component of $\mathbf{X}$. When d is sufficiently large, the pdf of the χ distribution is highly concentrated around $r = \sqrt{d}$, which is illustrated in the Figure 2.2.

It has been proven in [Katafygiotis and Zuev, 2008, Canal, 2005], that if the dimension d is sufficiently large, R can be approximated with a univariate normal distribution of mean $\sqrt{n}$ and variance 1/2: as $d \to +\infty$, $R \approx \mathcal{N}(\sqrt{d}, 1/2)$. Therefore, almost all the probability mass of R is concentrated around $\sqrt{d}$. As the distribution of $\mathbf{T}$ is the uniform distribution on $\mathcal{T}^d$, the *important ring* is then the name given to the region of spherical symmetry which encompasses most of the probability mass of $\mathbf{X}$ in the standard normal space. The important ring is defined with two hyperspheres of radii LB for lower bound and UB for upper bound and [LB, UB] is named Important Interval (II). The important interval can be defined in several ways. For example, if one wishes to have the important ring centered around the hypersphere of radius $\sqrt{d}$, the important interval can be defined with the minimum value ϵ such as $\mathbb{P}(\sqrt{d} - \epsilon \leq R \leq \sqrt{d} + \epsilon) \geq 1 - 10^{-n}$ with n a set integer. Then $\mathrm{II}_1 = [\sqrt{d} - \epsilon, \sqrt{d} + \epsilon]$. Another possibility is to use the quantile of the chi distribution and to define the important interval as $\mathrm{II}_2 = [q_{10^{-n}/2}(\chi(d)), q_{1-10^{-n}/2}(\chi(d))]$. Those two important intervals are displayed in Table 2.3, for several values of d and n. The higher the dimension of the system, the more similar those two important intervals are. Thus when performing the reliability analysis, one must take into account this specificity of the standard normal space.

Other standard elliptical distributions

For the other standard elliptical laws, the same properties of the radial component in the high-dimensional space do not necessarily apply. As previously noted in Table 2.1, both the Student distribution and the Cauchy distribution share the Fisher–Snedecor law as the distribution of their radial component. The density of $R \sim \sqrt{d \times \mathcal{F}(d,\nu)}$ is illustrated for different values of dimension d and degrees of freedom ν in Figure 2.3. The pdfs displayed in Figure 2.3 do not have the same

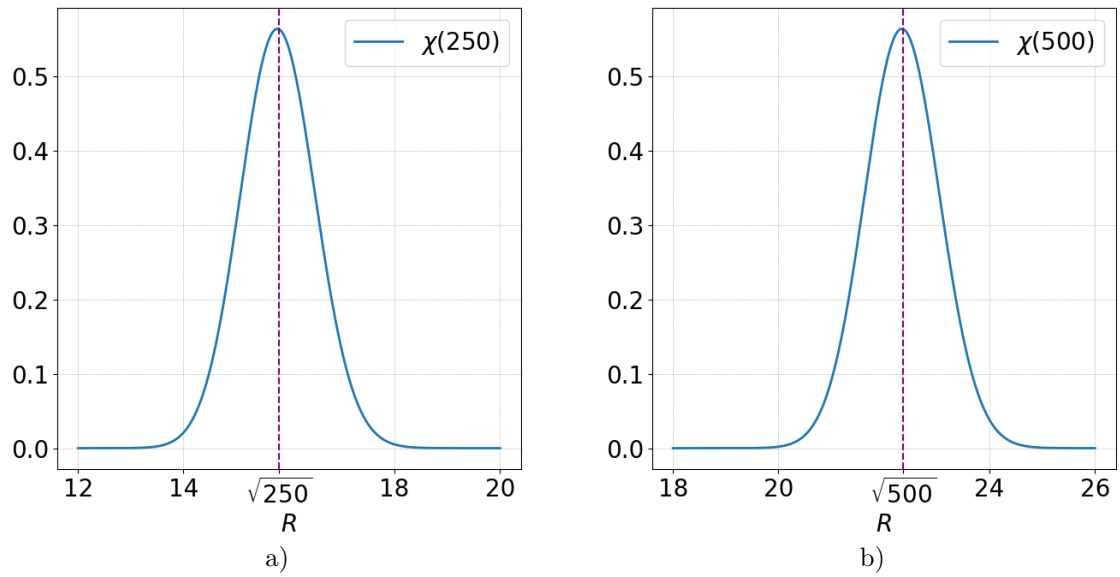


Figure 2.2: The density function of the radial component R when a) $d = 250$ b) $d = 500$. When $d = 250$, $\mathbb{E}[R] = 15.7956$, while $\sqrt{250} = 15.8114$, and $\mathbb{V}\text{ar}(R) = 0.4995$

Table 2.3: Example of important interval definitions for the standard normal vector, according to the dimension d and the integer n such as $\mathbb{P}(R \in \text{II}) \geq 1 - 10^{-n}$.

	$n = 4$	$n = 6$	$n = 8$
$d = 250$	$\text{II}_1 = [13.06, 18.56]$	$\text{II}_1 = [12.33, 19.29]$	$\text{II}_1 = [11.70, 19.92]$
	$\text{II}_2 = [13.12, 18.62]$	$\text{II}_2 = [12.47, 19.37]$	$\text{II}_2 = [11.93, 20.00]$
$d = 500$	$\text{II}_1 = [19.61, 25.11]$	$\text{II}_1 = [18.89, 25.83]$	$\text{II}_1 = [18.28, 26.44]$
	$\text{II}_2 = [19.65, 25.15]$	$\text{II}_2 = [18.98, 25.89]$	$\text{II}_2 = [18.42, 26.51]$
$d = 1000$	$\text{II}_1 = [28.87, 34.37]$	$\text{II}_1 = [28.16, 35.09]$	$\text{II}_1 = [27.55, 35.69]$
	$\text{II}_2 = [28.90, 34.40]$	$\text{II}_2 = [28.22, 35.13]$	$\text{II}_2 = [27.65, 35.75]$

behavior than the χ distribution as the probability mass of R is not concentrated around any particular value. The F-distribution appears to have a right heavy tail in some settings, which means that the right tail is heavier than the one of the exponential distribution. Therefore, an important ring can still be defined by quantiles, but the resulting important interval will be a lot wider than for the standard normal case. For instance, the lower bound can be taken as $\sqrt{d \times q_{10^{-n}/2}(\mathcal{F}(d,\nu))}$ and the upper bound as $\sqrt{d \times q_{1-10^{-n}/2}(\mathcal{F}(d,\nu))}$, with n an integer. In dimension $d = 250$ with $\nu = 1$, the interval $\text{II} = [3.83, 2.52 \times 10^5]$ satisfies $\mathbb{P}(R \in \text{II}) \geq 1 - 10^{-4}$ while if $\nu = 4$, this interval becomes $\text{II} = [6.18, 2.24 \times 10^2]$ for the same probability.

2.2.2.2 Probabilistic transformation to elliptical laws

Assuming the random inputs can be modeled with standard elliptical distributions is not a strong assumption as several probabilistic transformations allow to change a random input space into a standard elliptical space. These transformations have been developed in different contexts, to serve different purposes [Chabridon, 2018]. First of all, the random input variables can be of various nature, as they may represent different physical quantities of specific unit and magnitude. Thus transforming the inputs into a centered and scaled random vector can allow a better understanding of the input

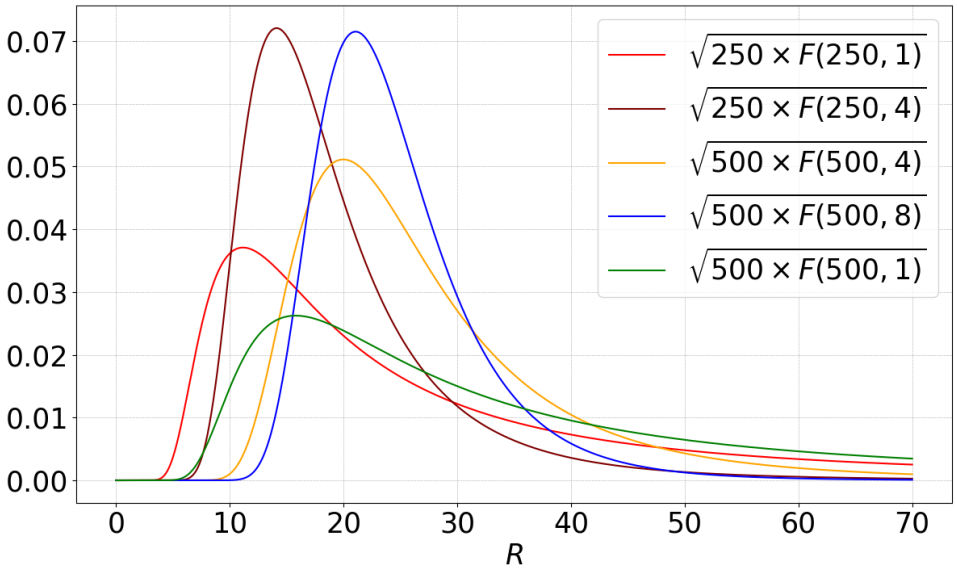


Figure 2.3: The density function of the radial component $R \sim \sqrt{d \times \mathcal{F}(d,\nu)}$ of a Student distribution for different values of d and ν and of a Cauchy distribution for different values of d.

space. Another advantage is that, when transformed in the standard normal space, the random vector is then independent. The independence makes some theoretical and numerical computation easier as explicit formulas can then be derived. The transformation is defined thanks to a diffeomorphism $T : \mathbb{R}^d \to \mathbb{R}^d$, also referred to as isoprobabilistic transformation, such that

$$\mathbf{X} = T(\mathbf{Z}) \quad \text{and} \quad \mathbf{Z} = T^{-1}(\mathbf{X}),$$

where $\mathbf{Z}$ is the original random input vector while $\mathbf{X} \sim \mathcal{E}(\mathbf{0}, \mathbf{I}_d, h)$. Different transformations are accessible depending on the nature of the random vector $\mathbf{Z}$ and the amount of information available about its probabilistic model

- If the random vector $\mathbf{Z}$ is already independent, and its univariate marginal cdfs are known, then the probability integral transform [Sarazin, 2021, Bourinet, 2018] is a basic transformation from $\mathbf{Z}$ to $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$.

- If the random vector $\mathbf{Z}$ has an elliptical copula with known univariate marginal cdf and known Pearson linear correlations between the margins, then the generalized Nataf transformation [Lebrun and Dutfoy, 2009] allows to transform $\mathbf{Z}$ to its standard equivalent $\mathbf{X} \sim \mathcal{E}(\mathbf{0}, \mathbf{I}_d, h)$, with h the density generator associated to the elliptical copula of $\mathbf{Z}$.

- If the random vector $\mathbf{Z}$ is known through its joint distribution (cdf, or pdf), then the Rosenblatt transformation [Rosenblatt, 1952] allows to transform $\mathbf{Z}$ to $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ without any assumption on the dependent structure of $\mathbf{Z}$.

Those transformations are not detailed here but the interested reader may find more information about them in [Sarazin, 2021, Bourinet, 2018].

However, it should be noted that if the dimension d is high, those transformations can be numerically more demanding, especially if the inputs have a strong dependence. Indeed as mentioned in [Martinez et al., 2018], in the Nataf transformation for example, $d(d-1)/2$ Pearson correlation coefficients must be transformed and thus the computational burden increases with the square of d. An example of solution to address this computational burden is therefore presented in [Lin et al., 2020]. As the Rosenblatt transformation requires the knowledge of the conditional cdfs of every univariate margins, if d is high then the computational effort increases as well [Bourinet, 2018].

2.3 Performing a reliability analysis of the system

Several quantities of interest can be studied in order to perform the reliability analysis of the system. Indeed, since the uncertainties of the inputs $\mathbf{X}$ are spreading through the deterministic black-box model to the output, one can be interested in constructing the cdf/pdf of the random output variable Y, or in estimating its moment $\mathbb{E}[Y]$ and $\mathbb{V}\mathrm{ar}(Y)$. One can also be interested in the performance of the system, which is evaluated with a deterministic scalar function $g : \mathbb{R} \to \mathbb{R}$, called the performance function. Traditionally, the performance function g is built such as the behavior of the system is split into two states:

- The system is safe for values of $\boldsymbol{\zeta} = [\mathbf{s}, \mathbf{X}] \in \mathbb{R}^p \times \mathbb{R}^d$ such as $g(\mathcal{M}(\boldsymbol{\zeta})) > 0$, and the safety domain is defined as $D_s = \{\boldsymbol{\zeta} \in \mathbb{R}^m \,|g(\mathcal{M}(\boldsymbol{\zeta})) > 0\}$.
- The system is failing for values of $\boldsymbol{\zeta} = [\mathbf{s}, \mathbf{X}] \in \mathbb{R}^p \times \mathbb{R}^d$ such as $g(\mathcal{M}(\boldsymbol{\zeta})) \leq 0$, and the failure domain is defined as $D_f = \{\boldsymbol{\zeta} \in \mathbb{R}^m \,|g(\mathcal{M}(\boldsymbol{\zeta})) \leq 0\}$.

For example, if one is interested in the exceedance of a given threshold level y_{th} of the system, the performance function takes the form $g(\mathcal{M}(\boldsymbol{\zeta})) = y_{th} - \mathcal{M}(\boldsymbol{\zeta})$. The hypersurface in $\mathbb{R}^m$ such that $\{\boldsymbol{\zeta} \,|\, g(\mathcal{M}(\boldsymbol{\zeta})) = 0\}$ is called the limit state surface and splits the input space into the safety domain and the failure domain. An illustration of the map of the input space with regard to the state of the system is displayed in Figure 2.4, for a system of dimension $d = 2$.

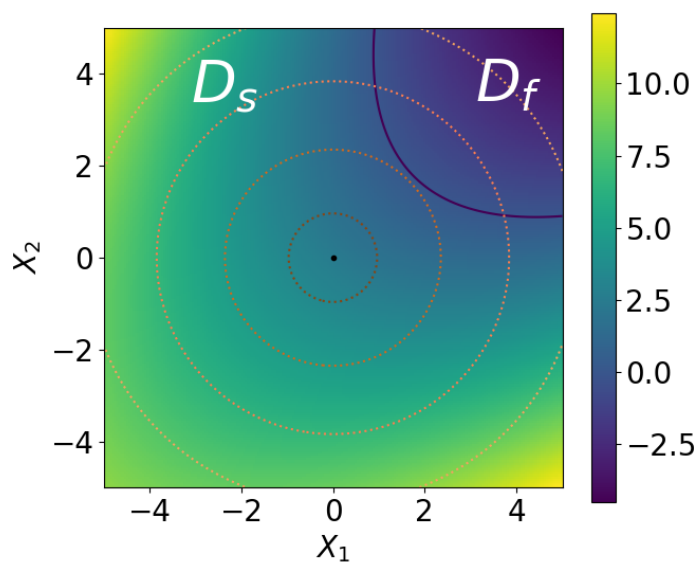


Figure 2.4: Illustration on a two-dimensional example of the map of the input space, split in the failure domain D_f and the safety domain D_s by the limit state surface in blue. The dotted lines represent isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g.

For simplicity, the function $g \circ \mathcal{M} : \mathbb{R}^m \to \mathbb{R}$ will be written g and referred to as *limit state function* (lsf) for the rest of the manuscript. Since the gradient of the black-box model $\mathcal{M}$ is available, so is the gradient of the limit state function g. As previously noted, the number of times $\mathcal{M}$ is evaluated has to remain as low as possible, therefore so is the number of evaluations of the lsf g. This number of evaluations is called the simulation budget. The lsf g can only be evaluated pointwise, thus the limit state surface cannot be explicitly defined.

In this manuscript, the quantities of interest studied to perform the reliability analysis of the system are the failure probability of the system, denoted P_f, which is defined as the probability that the system is in the failure state, and its derivatives with respect to decision parameters, being the deterministic inputs of the system $\mathbf{s}$.

2.3.1 Estimation of the failure probability of the system

In our specific context, the inputs are divided into two groups, namely the deterministic one with vector $\mathbf{s}$ and the random one with vector $\mathbf{X}$. In addition, it is assumed that the random input $\mathbf{X}$ follows a standard elliptical distribution, thus $\mathbf{X} = R\mathbf{T}$, as previously noted in Section 2.2.1.2. Since the radial component R and the directional component $\mathbf{T}$ are independent, the joint pdf of the random variable $\mathbf{X}$ is written $f_{\mathbf{X}} = f_R f_{\mathbf{T}}$. The probability of failure of the system is then defined by the following d-fold integral

$$P_f(\mathbf{s}) = \int_{D_f(\mathbf{s})} f_R(r) f_{\mathbf{T}}(\mathbf{t})\,\mathrm{d}r\mathrm{d}\mathbf{t} = \int_{\mathbb{R}^+}\int_{\mathcal{T}^d} \mathbb{I}_{D_f(\mathbf{s})}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{t})\,\mathrm{d}r\mathrm{d}\mathbf{t} = \mathbb{E}_{f_R,f_{\mathbf{T}}}\left[\mathbb{I}_{D_f(\mathbf{s})}(R\mathbf{T})\right],$$

where $D_f(\mathbf{s}) = \{r\mathbf{t} \in \mathbb{R}^d \,|\, g(\mathbf{s}, r\mathbf{t}) \leq 0\}$ is the failure domain for a specific vector $\mathbf{s}$, and $\mathbb{I}_{D_f(\mathbf{s})}$ is the indicator function equal to 1 when $g(\mathbf{s}, r\mathbf{t}) \leq 0$ and 0 otherwise. The operator $\mathbb{E}_{f_R,f_{\mathbf{T}}}$ stands for the mathematical expectation with respect to R and $\mathbf{T}$ simultaneously. Since the failure of the system is supposed to be a rare event, it is assumed that the failure probability P_f is small, i.e. inferior to 10^{-3}.

It is recalled here that we suppose the dimension of the system d high, above 50. Moreover, it is assumed in the rest of the manuscript that the failure domain $D_f(\mathbf{s})$ encompasses possibly several failure modes [Ditlevsen and Madsen, 1996]. This is particularly true with high-dimensional systems, where the failure can come from different extreme configurations of the random vector $\mathbf{X}$. Thus there may be several regions in the input space, disjoint or not, that belong to the failure domain. Therefore the two main challenges of the estimation of the probability concern the *high dimension* d and the *multiple failure regions*. An illustration in dimension $d = 2$ is given in Figure 2.5, where two failure regions exist amongst the failure domain D_f.

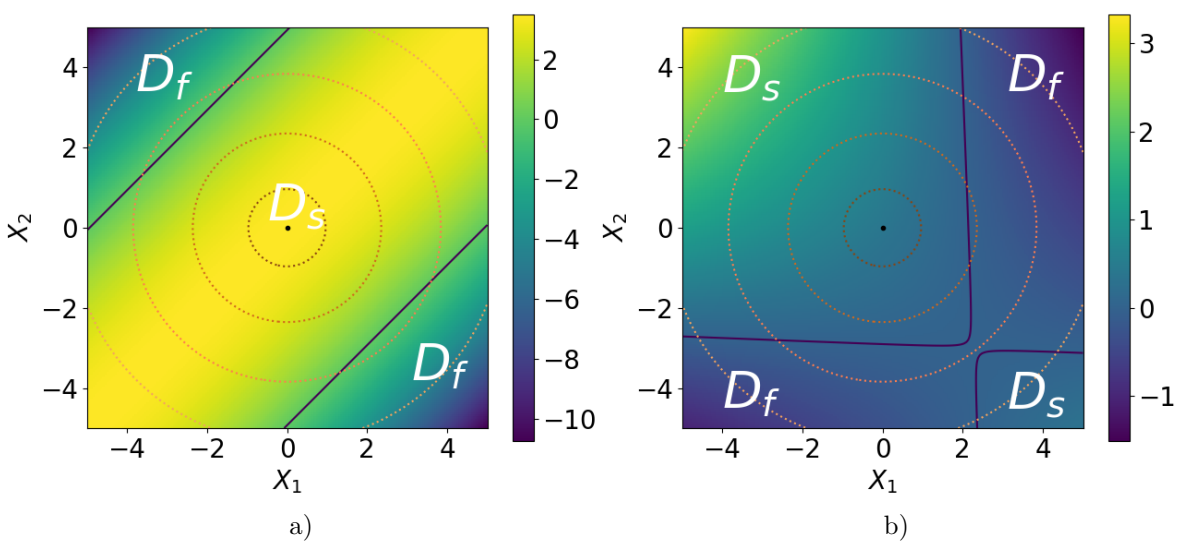


Figure 2.5: Illustration of the location of the two failure regions on an example in dimension 2 a) where the failure regions are disjoint b) where the failure regions are joint. The blue line is the limit state surface. The dotted lines represent isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g.

2.3.2 Probability of failure sensitivity analysis

The probability of failure depends upon the specific value of $\mathbf{s}$. Thus, the other quantity of interest studied in this manuscript is the derivatives of P_f according to $\mathbf{s}$. Assuming that the gradient $\nabla_{r\mathbf{t}} g(\mathbf{s}, r\mathbf{t}) \neq \mathbf{0}$ for all $r\mathbf{t}$ and $\mathbf{s}$ on the limit state surface $\{g(\mathbf{s}, r\mathbf{t}) = 0\}$, the derivative of $P_f(\mathbf{s})$ according to s_ℓ with $\ell = 1, \ldots, p$ is equal to the surface integral [Breitung, 2006]

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = -\int_{g(\mathbf{s},r\mathbf{t})=0} \frac{1}{\parallel \nabla_{r\mathbf{t}} g(\mathbf{s}, r\mathbf{t}) \parallel} \frac{\partial g(\mathbf{s}, r\mathbf{t})}{\partial s_\ell} f_R(r) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}s(r\mathbf{t}),$$

where $\mathrm{d}s(r\mathbf{t})$ stands for surface integration over the limit state surface $\{g(\mathbf{s}, r\mathbf{t}) = 0\}$. The estimation of such an integral allows to evaluate the influence of $\mathbf{s}$ upon $P_f(\mathbf{s})$ and is part of the local sensitivity analysis of the system [Torii and Novotny, 2021]. In the literature, the vector $\mathbf{s}$ is often referred to as decision parameters [Lacaze et al., 2015]. If the vector $\mathbf{s}$ influences the design of the system then it is also called design parameters [Torii, 2020, Papaioannou et al., 2018].

The estimation of such a surface integral is not straightforward, especially for a complex limit state function g encompassing several failure regions in a high-dimensional input space. The main challenge of the estimation of the derivatives is to increase as little as possible the *simulation budget* needed for the estimation of the probability of failure.

Estimation de probabilité de défaillance en grande dimension dans un espace standard elliptique

Contexte d'étude

Dans ce manuscrit, afin d'analyser la fiabilité du système, nous nous intéressons tout d'abord à l'estimation de sa probabilité de défaillance. Les variables d'entrée du système sont divisées en deux groupes : les variables déterministes $\mathbf{s} \in \mathbb{R}^p$ et les variables aléatoires $\mathbf{X} \in \mathbb{R}^d$, avec $d \geq 50$. La probabilité de défaillance dépend alors des variables déterministes et est notée $P_f(\mathbf{s})$. Néanmoins, pour alléger les notations de ce chapitre, cette dépendance sera supposée implicite dans la suite, pour la probabilité de défaillance P_f ainsi que la fonction d'état limite g. En conséquence, la probabilité de défaillance est définie à l'aide de l'intégrale suivante

$$P_f = \mathbb{P}(g(\mathbf{X}) \leq 0) = \int_{D_f} f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} \mathbb{I}_{D_f}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f}(\mathbf{X})\right],$$

où $D_f = \{\mathbf{x} \in \mathbb{R}^d \,|\, g(\mathbf{x}) \leq 0\}$ est le domaine de défaillance avec $g : \mathbb{R}^d \to \mathbb{R}$ la fonction d'état limite restreinte aux variables aléatoires $\mathbf{X}$. Dans le contexte spécifique du manuscrit, puisque le vecteur $\mathbf{X}$ suit une distribution standard elliptique, l'intégrale est réécrite avec la représentation stochastique

$$P_f = \mathbb{P}(g(R\mathbf{T}) \leq 0) = \int_{D_f} f_R(r) f_{\mathbf{T}}(\mathbf{t})\,\mathrm{d}r\,d\mathbf{t} = \int_{\mathbb{R}^+}\int_{\mathcal{T}^d} \mathbb{I}_{D_f}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{t})\,\mathrm{d}r\,\mathrm{d}\mathbf{t} = \mathbb{E}_{f_R, f_{\mathbf{T}}}\left[\mathbb{I}_{D_f}(R\mathbf{T})\right].$$

Comme précédemment mentionné, la défaillance est un évènement rare, ainsi $P_f \leq 10^{-3}$. De plus, plusieurs configurations des entrées entraînent la défaillance du système, ainsi D_f peut englober plusieurs régions de défaillance.

Le but de ce chapitre est alors de présenter différentes techniques développées dans la littérature afin d'estimer la probabilité de défaillance d'un tel système. L'enjeu est de limiter le nombre d'évaluations de la fonction d'état limite g, celle-ci étant calculée par un solveur boîte noire numériquement coûteux.

Présentation de plusieurs méthodes d'estimation de la probabilité de défaillance

Méthodes d'approximation utilisant les points de conception

Une première manière d'estimer la probabilité de défaillance est d'utiliser une méthode d'approximation de la fonction d'état limite g. Nous présentons ici les techniques FORM (First-Order Reliability Method) et SORM (Second-Order Reliability Method). Ces méthodes, développées dans l'espace standard elliptique [Lebrun and Dutfoy, 2009], utilisent la localisation du *point de conception* du système afin de construire un substitut de la fonction d'état limite.

Le point de conception du système, noté P^* est le point de norme euclidienne minimale se trouvant dans le domaine de défaillance du système. C'est le point défaillant maximisant la densité de la

loi standard elliptique, puisqu'il est le plus proche de l'origine. Ce point de conception s'obtient en résolvant le problème d'optimisation quadratique sous contrainte suivant

$$P^* = \underset{\mathbf{x}\in D_f}{\operatorname{argmin}}\ \mathbf{x}^\top \mathbf{x} = \underset{\mathbf{x}\in D_f}{\operatorname{argmin}}\ \|\mathbf{x}\|^2$$

avec $\|\cdot\|$ la norme euclidienne. Les algorithmes permettant d'effectuer cette optimisation sont nombreux [Lemaire, 2013] et nécessitent en général le gradient $\nabla_{\mathbf{x}} g$ de la fonction d'état limite que nous supposons disponible dans ce manuscrit. L'approximation FORM consiste alors à remplacer la fonction d'état limite g par son développement de Taylor à l'ordre 1 au voisinage de P^*, tandis que l'approximation SORM utilise son développement de Taylor d'ordre 2. Des approximations de la probabilité de défaillance sont alors calculables avec un budget de simulation très faible, ce qui constitue le point fort de ces méthodes.

Lorsque le domaine de défaillance possède plusieurs régions de défaillance, une méthode décrite dans [Der Kiureghian and Dakessian, 1998] permet de trouver les différents points de conception de chaque région de défaillance, en adaptant l'optimisation après chaque région trouvée. Le probabilité de défaillance finale s'obtient alors généralement en termes de borne, même si une formule explicite est possible dans le cas gaussien avec FORM [Ditlevsen and Madsen, 1996].

Néanmoins, si les approximations de la fonction d'état limite peuvent être plutôt fidèles à la réalité au voisinage des points de conception, ce n'est pas forcément le cas dans des régions éloignées des points de conception. Il est souligné dans [Katafygiotis and Zuev, 2008] que les points de conception des systèmes modélisés par des variables aléatoires gaussiennes ne sont généralement pas localisés dans l'anneau d'importance de ces dernières. Ainsi, puisque la probabilité de défaillance est principalement calculée avec la masse probabiliste se situant dans cet anneau d'importance, l'approximation FORM/SORM peut être erronée dans cette région.

Méthodes de simulation

Les méthodes de simulations sont basées sur la technique de Monte-Carlo (MC) [Robert and Casella, 1999, Sobol, 2018]. En remarquant que $P_f = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f}(\mathbf{X})\right]$, la méthode de MC consiste tout d'abord à générer un échantillon indépendant et identiquement distribué $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$ de taille N grâce à la densité $f_{\mathbf{X}}$. La probabilité de défaillance est ensuite estimée par la moyenne de la fonction indicatrice du domaine de défaillance évaluée sur l'échantillon

$$\hat{P}_f^{\mathrm{MC}} = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}_{D_f}\left(\mathbf{X}^{(j)}\right).$$

Cette méthode est non biaisée et converge peu importe la dimension d du système ou le nombre de régions de défaillance de g. De plus elle est adéquate pour n'importe quelle densité $f_{\mathbf{X}}$, et convient donc à toutes les lois elliptiques standards. Cependant, sa vitesse de convergence est lente : la variance de l'estimateur est proportionnelle à $1/N$. Cette méthode n'est donc pas adaptée à l'estimation de probabilité d'évènements rares. Plusieurs méthodes ont alors été proposées afin d'accélérer cette vitesse de convergence.

La méthode de "Subset Simulation" [Au and Beck, 2001] repose sur l'idée d'écrire la probabilité de défaillance de valeur très faible en tant que produit de probabilités de défaillance intermédiaires plus facile à estimer car de valeurs plus élevées. Des domaines de défaillance intermédiaires F_i sont introduits à l'aide d'une séquence décroissante de seuils $s_0 > s_1 > \cdots > s_m = 0$ de telle sorte que $F_i = \{\mathbf{x} \in \mathbb{R}^d \,|\, g(\mathbf{x}) \leq s_i\}$. Les probabilités conditionnelles $\mathbb{P}(F_i|F_{i-1})$ sont ensuite calculées, pour retrouver $P_f = \mathbb{P}(F_m)$ grâce au produit de ces probabilités conditionnelles pour $i = 1, \dots, m$. En pratique, ces probabilités sont estimées par méthode de MC grâce à des échantillons de lois conditionnelles obtenus par chaînes de Markov [Metropolis et al., 1953]. Plusieurs versions des chaînes de Markov ont été développées afin de rendre l'algorithme de subset simulation particulièrement performant en grande dimension, surtout pour les lois normales standards [Katafygiotis and Zuev, 2008]. Lorsque le domaine de défaillance possède plusieurs régions distinctes, un budget de simulation plus conséquent peut être nécessaire pour atteindre chacune de ces zones grâce aux chaînes de Markov.

Les méthodes d'échantillonnage préférentiel consistent à introduire une densité h dite *auxiliaire* dans le calcul de la probabilité de défaillance. Cette densité h, dont le support doit inclure celui de $\mathbb{I}_{D_f} f_{\mathbf{X}}$ permet de définir l'intégrale comme

$$P_f = \int_{\mathbb{R}^d} \mathbb{I}_{D_f}(\mathbf{x}) \frac{f_{\mathbf{X}}(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \mathbb{E}_h \left[\mathbb{I}_{D_f}(\mathbf{X}) \frac{f_{\mathbf{X}}(\mathbf{X})}{h(\mathbf{X})} \right].$$

L'estimateur d'échantillonnage préférentiel de P_f est ensuite donné par la formule suivante

$$\hat{P}_f^{\mathrm{IS}} = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{D_f}\left(\mathbf{X}^{(j)}\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}^{(j)}\right)}{h\left(\mathbf{X}^{(j)}\right)},$$

où les variables aléatoires $\left(\mathbf{X}^{(j)}\right)$ sont indépendantes et identiquement distribuées selon h. Avec un choix pertinent de densité auxiliaire, la variance de cet estimateur peut être grandement réduite comparée à celle de l'estimateur de MC. La densité théorique optimale h_{opt}, conduisant à un estimateur de variance nulle, est donnée par [Bucklew, 2004]

$$\forall \mathbf{x} \in \mathbb{R}^d \qquad h_{\mathrm{opt}}(\mathbf{x}) = \frac{\mathbb{I}_{D_f}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{P_f}.$$

Cependant, elle n'est pas utilisable en pratique puisqu'elle dépend de la quantité d'intérêt P_f. Le but de l'échantillonnage préférentiel est alors de construire une densité auxiliaire aussi proche de h_{opt} que possible afin de réduire au maximum la variance de l'estimateur.

Plusieurs algorithmes ont été développés afin de construire de manière adaptative une densité auxiliaire optimale [Tabandeh et al., 2022]. L'algorithme adaptatif non paramétrique typique [Zhang, 1996] (NAIS) utilise l'estimation par noyau afin de construire des densités à partir d'échantillons. Pour chaque observation de l'échantillon, son appartenance ou non au domaine de défaillance est pris en compte dans la construction de la densité de noyau grâce à un poids dépendant de la fonction indicatrice du domaine de défaillance. Des seuils intermédiaires sont introduits afin de se rapprocher progressivement du domaine de défaillance, comme pour le subset sampling. Cet algorithme non paramétrique est particulièrement efficace pour les domaines de défaillance possédant plusieurs régions de défaillance. Néanmoins son efficacité n'est plus vérifiée lorsque la dimension du système est grande.

De nombreux algorithmes adaptatifs paramétriques ont été développés dans la littérature. Une famille de densité est alors sélectionnée et le but est d'optimiser graduellement les paramètres de cette famille afin de se rapprocher le plus possible de h_{opt}. Pour mesurer la distance entre la densité paramétrique construite et h_{opt}, la divergence de Kullback–Leibler [Kullback and Leibler, 1951] est souvent utilisée. L'algorithme est alors appelé d'entropie croisée (Cross-Entropy Adaptive Importance Sampling CE-AIS). Afin d'atteindre le domaine de défaillance, des seuils intermédiaires s_i sont une fois de plus introduits. En notant $h(\cdot\,;\boldsymbol{\Theta})$ la densité paramétrique sélectionnée avec $\boldsymbol{\Theta}$ ses paramètres, la densité auxiliaire $h(\cdot\,;\boldsymbol{\Theta}_{i+1})$ construite à l'étape i résulte de l'optimisation suivante

$$\boldsymbol{\Theta}_{i+1} \approx \underset{\boldsymbol{\Theta} \in \mathbb{R}^n}{\operatorname{argmax}} \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{Y<s_i}\left(Y_i^{(j)}\right) \ln\left(h\left(\mathbf{X}_i^{(j)};\boldsymbol{\Theta}\right)\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_i^{(j)}\right)}{h\left(\mathbf{X}_i^{(j)},\boldsymbol{\Theta}_i\right)},$$

où les variables aléatoires $\left(\mathbf{X}_i^{(j)}\right)$ sont indépendantes et identiquement distribuées selon la densité construite à l'étape précédente $h(\cdot\,;\boldsymbol{\Theta}_i)$ et $Y_i^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)$. Cet algorithme possède plusieurs versions différentes, selon les besoins spécifiques. Dans notre contexte de domaine de défaillance possédant plusieurs régions distinctes, un mélange de densités paramétriques est nécessaire dans la construction la densité auxiliaire h. Aussi, le nombre total de paramètres de ce mélange doit rester faible par rapport à la dimension du système, puisque nous nous intéressons particulièrement aux grandes dimensions. Afin de rentabiliser au mieux les évaluations de la fonction d'état limite g, des adaptations de l'algorithme CE-AIS ont été proposées, notamment avec l'échantillonnage préférentiel multiple (Cross-Entropy Adaptive Multiple Importance Sampling CE-AMIS) [Cornuet et al., 2012] et l'entropie croisée améliorée (improved Cross-Entropy iCE) [Papaioannou et al., 2019].

Les algorithmes adaptatifs paramétriques présentés dans [Wang and Song, 2016] et [Papaioannou et al., 2019] sont alors particulièrement appropriés au contexte du manuscrit, puisque la densité auxiliaire proposée est constituée d'un mélange de densités, prenant en compte le comportement des lois normales standards en dimension élevée. En effet, la représentation stochastique des entrées est prise en considération et chaque densité du mélange est construite comme un produit d'une densité sur la variable radiale R et une densité sur la variable directionnelle $\mathbf{T}$. Cette décomposition favorise une meilleure appréhension de l'anneau d'importance des variables standard normales. De plus, le mélange de densités permet de considérer toutes les régions de défaillance. Néanmoins, le nombre de densité dans le mélange est toujours difficile à paramétrer lorsqu'aucune information n'est disponible sur la fonction d'état limite g, ce qui est le cas dans notre contexte.

La représentation stochastique des variables standards elliptiques est aussi le point de départ d'une autre méthode de simulation : l'échantillonnage directionnel [Bjerager, 1988]. En effet, la probabilité de défaillance est calculée grâce à une formule de MC, où les observations sont tirées non pas selon $f_{\mathbf{X}}$ mais selon une distribution définie sur la sphère unité $\mathcal{T}^d$. Pour chaque observations $\mathbf{T}^{(j)}$, il s'agit ensuite d'identifier avec une recherche linéaire une *racine* r_j, dans le cas où la fonction d'état limite a une forme dite d'*étoile*. La racine r_j est alors définie comme la distance entre la surface d'état limite et l'origine dans la direction $\mathbf{T}^{(j)}$ correspondante. L'échantillonnage directionnel permet d'envisager l'estimation de la probabilité de défaillance sous un autre angle, puisqu'il s'agit alors de calculer des racines dans des directions données. Cependant, cette méthode nécessite quelques améliorations en grande dimension.

Méthodes de simulation basées sur les points de conception

En utilisant les coordonnées des points de conception du système, certaines méthodes de simulation permettent d'échantillonner directement dans les zones d'intérêt, afin d'estimer la probabilité de défaillance de manière plus efficace. Supposons que les différents points de conception du système ont été trouvés grâce à une procédure semblable à celle mentionnée dans les méthodes d'approximation.

Une méthode d'échantillonnage préférentiel appelée "radial-based" [Harbitz, 1986] consiste alors à exclure du domaine d'échantillonnage l'hypersphère centrée de rayon β où β est calculé comme le minimum des normes euclidiennes des points de conception du système. Afin d'exclure cette hypersphère, la représentation stochastique des variables aléatoires elliptiques standards est une fois de plus utilisée. Cette méthode a donné lieu à plusieurs versions pour prendre en compte la multitude des régions de défaillance. Néanmoins, des adaptations sont encore nécessaires pour rendre ces méthodes convenables à la grande dimension.

L'échantillonnage dit de ligne "line sampling" [Hohenbichler and Rackwitz, 1988] est une méthode spécifiquement développée pour les espaces standards gaussiens. En utilisant la direction du point de conception, l'espace d'échantillonnage est orienté pour rendre cette direction particulièrement importante. Les observations sont tirées selon une loi normale standard de dimension $d-1$ définie sur l'hyperplan $v_d = 0$, où v_d est une direction parallèle à celle du point de conception. Pour chaque observation, il faut alors trouver la racine (supposée unique), correspondant à la distance entre l'origine et la surface d'état limite dans la direction donnée par l'observation. Dans notre cas, les points de conception du système sont multiples et une version adaptée à ce contexte a été développée dans [Lu et al., 2008]. Chaque point de l'espace est alors assigné au point de conception le plus proche (grâce à un critère de distance euclidienne). Les régions de défaillance étant alors complètement disjointes, la probabilité de défaillance est calculée comme la somme des probabilités de défaillance locales associées à chaque point de conception. Cependant, en grande dimension, les points de conception n'appartiennent généralement pas à l'anneau d'importance des variables standards normales [Katafygiotis and Zuev, 2008]. Ainsi, les directions des points de conception, qui déterminent grandement l'efficacité de cette méthode, peuvent ne pas être pertinentes vis-à-vis de l'anneau d'importance. Ceci est d'autant plus probable si la fonction d'état limite est non linéaire.

Enfin, plusieurs méthodes d'échantillonnage préférentiel, dont la densité auxiliaire est basée sur les coordonnées des points de conception du système, sont décrites dans la littérature [Schuëller and Stix, 1987]. Il s'agit alors d'utiliser les points de conception comme centres des densités présentes dans le mélange utilisé en tant que densité auxiliaire (dans le cas de régions de défaillance multiples). Chaque densité dans le mélange est prise de même nature que la densité des entrées de départ $f_{\mathbf{X}}$. Dans le cas de l'espace standard normal [Au and Beck, 2003], il est montré que la méthode reste efficace

en grande dimension tant que les variances des densités du mélange sont les mêmes que celles des variables d'entrées $\mathbf{X}$. Néanmoins, cette performance n'est vérifiée que pour les fonctions d'état limite à tendance linéaire. Pour les fonctions d'état limite non linéaire, cette méthode est alors infructueuse. Cela est dû au même argument que pour la méthode d'échantillonnage de ligne; la localisation des points de conception peut être insignifiante dans l'estimation de la probabilité de défaillance si les régions de défaillance dans l'anneau d'importance en sont éloignées.

Conclusion

Plusieurs méthodes d'estimation de probabilité de défaillance ont été abordées dans ce chapitre. La performance de chacune de ces méthodes par rapport au contexte du manuscrit a été soulignée. En effet, nous supposons que les entrées aléatoires du système sont modélisées par des variables elliptiques standards de grande dimension. Nous supposons de même que le domaine de défaillance du système possède plusieurs régions de défaillance. Néanmoins, la fonction d'état limite est une boîte noire, donc le nombre de régions de défaillance est inconnu.

Les méthodes d'approximation basées sur les points de conception du système permettent d'estimer la probabilité de défaillance avec un budget de simulation très faible. Lorsqu'il y a plusieurs points de conception, une technique a été développée afin de les trouver au fur et à mesure, en adaptant la recherche par optimisation. Néanmoins, il est difficile de mesurer l'erreur des probabilités de défaillance obtenues, en particulier lorsqu'il y a plusieurs points de conception. De plus, les approximations de la fonction d'état limite g sont valides aux voisinages des points de conception seulement. Ainsi, ces approximations peuvent être erronées dans des régions éloignées des points de conception, comme dans l'anneau d'importance des lois normales standards par exemple. L'estimateur de probabilité obtenu peut alors être incorrect.

Les méthodes de simulations requièrent un budget de simulation plus élevé. Cependant, il est alors possible d'évaluer l'erreur d'estimation de la probabilité grâce à la variance de l'estimateur. Plusieurs algorithmes d'échantillonnage préférentiel paramétrique adaptatif présentés sont particulièrement pertinents dans le contexte de ce manuscrit. La représentation stochastiques des entrées est prise en compte pour proposer une densité auxiliaire constituée d'un mélange de densités assemblées comme le produit d'une densité sur la variable radiale R et d'une densité sur le vecteur directionnel $\mathbf{T}$. Cependant, le nombre de densités dans le mélange est difficile à paramétrer, puisque la fonction d'état limite est une boîte noire.

En combinant les méthodes de simulation avec l'information apportée par la localisation des points de conception du système, les avantages des deux processus peuvent être exploités. Néanmoins, l'efficacité des méthodes résultantes n'est pas garantie pour un système de grande dimension possédant plusieurs zones de défaillance, en particulier si la fonction d'état limite est non linéaire.

Afin de répondre à ce besoin, un nouvel algorithme d'échantillonnage préférentiel multiple adaptatif paramétrique (CE-AMIS) a été développé au cours de la thèse et est présenté Chapitre 5. Les régions de défaillance sont cherchées avec une méthode proche de celle présentée dans les méthodes d'approximation. Le cadre du CE-AMIS permet de rentabiliser les évaluations de la fonction d'état limite g qui est numériquement coûteuse. Plusieurs méthodes de simulations basées sur les points de conception du système, comme l'échantillonnage préférentiel radial-based, ont inspiré la construction de la densité auxiliaire proposée. Enfin, cette algorithme est aussi adapté à toutes les distributions elliptiques standards.

Chapter 3

Failure probability estimation in a high-dimensional space

Contents

3.1 Introduction **40**
3.2 Design-point-based techniques: FORM and SORM **40**
3.2.1 FORM/SORM in the standard elliptical space 40
3.2.1.1 The FORM approximation 41
3.2.1.2 The SORM approximation 42
3.2.1.3 Remarks on the FORM/SORM approximations 42
3.2.2 Multiple design points and multiple failure regions 42
3.2.2.1 Search for the design points 43
3.2.2.2 Estimation of the failure probability 44
3.2.3 Design points in high-dimensional standard normal space 45
3.3 Simulation methods **45**
3.3.1 Crude Monte Carlo 46
3.3.2 Subset Simulation 46
3.3.2.1 SS in the standard elliptical space 46
3.3.2.2 MCMC algorithms for high-dimensional standard normal inputs . 48
3.3.2.3 Statistical properties of SS 49
3.3.2.4 SS algorithm for multiple failure regions 49
3.3.3 Importance Sampling 50
3.3.3.1 Non parametric Adaptive Importance Sampling 51
3.3.3.2 Cross-Entropy based Adaptive Importance Sampling 53
3.3.3.3 Cross-Entropy based Adaptive Multiple Importance Sampling . . 58
3.3.3.4 Improved Cross Entropy-based Adaptive Importance sampling . . 60
3.3.4 Directional Sampling 62
3.3.4.1 DS in the standard elliptical space 62
3.3.4.2 DS/DIS for star shaped limit state surfaces 63
3.3.4.3 Variants of the DIS method 63
3.4 Simulation techniques based on the design point **64**
3.4.1 Radial-Based Importance Sampling 65
3.4.1.1 RBIS in the standard elliptical space 65
3.4.1.2 RBIS without the design point 65
3.4.2 Line Sampling 65
3.4.2.1 LS in the standard normal space 65
3.4.2.2 LS for multiple design points 67
3.4.2.3 LS for high-dimensional standard normal inputs 68
3.4.2.4 LS without the design point 68

3.4.3 Importance Sampling using design points 68
3.4.3.1 IS using design points in the standard elliptical space 68
3.4.3.2 IS using design points in the high-dimensional standard normal space 69
3.4.3.3 IS using random failing points in the high-dimensional standard normal space 69
3.5 Conclusion 70

3.1 Introduction

In this thesis, the safety measure studied for the reliability analysis of the system is its probability of failure. The inputs of the system are divided into two groups: deterministic inputs $\mathbf{s} \in \mathbb{R}^p$ and random inputs $\mathbf{X} \in \mathbb{R}^d$, with $d \geq 50$. The probability of failure depends on the value of the deterministic vector and is written $P_f(\mathbf{s})$. However in this chapter, we will drop the expression of the dependence on $\mathbf{s}$ for the sake of clarity, for both P_f and the limit state function g. Thus the probability of failure is defined as the d-fold integral

$$P_f = \mathbb{P}(g(\mathbf{X}) \leq 0) = \int_{D_f} f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} \mathbb{I}_{D_f}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f}(\mathbf{X})\right], \tag{3.1}$$

where $D_f = \{\mathbf{x} \in \mathbb{R}^d \,|\, g(\mathbf{x}) \leq 0\}$ is the failure domain with $g : \mathbb{R}^d \to \mathbb{R}$ the limit state function restricted to the random inputs $\mathbf{X}$. And in our specific context, since $\mathbf{X}$ are assumed to be standard elliptical inputs, using the stochastic representation of the inputs, the integral becomes

$$P_f = \mathbb{P}(g(R\mathbf{T}) \leq 0) = \int_{D_f} f_R(r) f_{\mathbf{T}}(\mathbf{t})\,\mathrm{d}r\,d\mathbf{t} = \int_{\mathbb{R}^+}\int_{\mathcal{T}^d} \mathbb{I}_{D_f}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{t})\,\mathrm{d}r\,\mathrm{d}\mathbf{t} = \mathbb{E}_{f_R, f_{\mathbf{T}}}\left[\mathbb{I}_{D_f}(R\mathbf{T})\right]. \tag{3.2}$$

As noted in the previous chapter, it is assumed that the failure is a rare event, thus $P_f \leq 10^{-3}$. Furthermore, several configurations of the inputs can lead to the failure of the system, thus D_f possibly encompasses several failure regions.

The aim of this chapter is to present some of the techniques that have been developed in order to estimate accurately the integral Eq. (3.1) with as few limit state function evaluations as possible. The presentation is not exhaustive and tries to focus on methods that are specifically tailored for multiple failure regions in the high-dimensional standard elliptical space. We will first present the approximation methods based on the design points, where the limit state function is replaced by a numerically inexpensive approximation, in Section 3.2. Then we will present several simulation methods, also called sampling methods, which are all based on the crude Monte Carlo method, in Section 3.3. Finally methods combining the design points approach with sampling techniques are presented in Section 3.4.

3.2 Design-point-based techniques: FORM and SORM

In this section we present approximation methods based on the design points. First a general introduction of the methods is derived in the standard elliptical space. Then, the specific case of failure domains encompassing multiple failure regions is studied. Finally, the performance of those methods in a high-dimensional context is briefly reviewed.

3.2.1 FORM/SORM in the standard elliptical space

The First-Order Reliability Method (FORM) and Second-Order Reliability Method (SORM) are two approximation methods part of the design-point-based techniques. These two methods were initially specifically designed for the standard normal space [Madsen et al., 2006]; however, Lebrun and Dutfoy in [Lebrun and Dutfoy, 2009] extended both methods for the standard elliptical space. These generalized FORM/SORM approximations are presented here. It is recalled that the pdf of a standard elliptical random vector $\mathbf{X}$ is written

$$\forall \mathbf{x} \in \mathbb{R}^d \quad f_{\mathbf{X}} = h\left(\mathbf{x}^\top \mathbf{x}\right) = h\left(r^2 \mathbf{t}^\top \mathbf{t}\right) = h\left(r^2\right) = h\left(\|\mathbf{x}\|^2\right) \quad \text{with} \quad r = \|\mathbf{x}\|, \tag{3.3}$$

with $\|\cdot\|$ the Euclidean norm and where $h : \mathbb{R}^+ \to \mathbb{R}^+$ is the density generator of $\mathbf{X}$, a rapidly decreasing function (see Chapter 2); it should be noted that $f_{\mathbf{X}}$ has a spherical symmetry.

The first step of the FORM/SORM is to find the point P^* in the failure domain D_f that is closest to the origin, cf Figure 3.1. Indeed, as h is decreasing, this point maximizes the density function $f_{\mathbf{X}}$ of Eq. (3.3) and is known as the design point or the most probable failure point. Finding it requires performing the following constrained quadratic optimization problem

$$P^* = \underset{\mathbf{x}\in D_f}{\operatorname{argmin}}\ \mathbf{x}^\top\mathbf{x} = \underset{\mathbf{x}\in D_f}{\operatorname{argmin}}\ \|\mathbf{x}\|^2. \tag{3.4}$$

In a realistic context, the solution is located on the limit state surface. To perform this optimization, several algorithms were specifically tailored and [Lemaire, 2013] presents an overview of them. The similarity between these algorithms is that they all require evaluations of the gradient of the limit state function. Since it is assumed in this manuscript that the gradient $\nabla_{\mathbf{x}} g$ of the lsf is available analytically or at a feasible computational cost (adjoint approaches for example), the optimization Eq. (3.4) is then supposed to be achievable. However, it should be noted that if the gradient must be approximated by finite differentiation, then the optimization becomes quite cumbersome and numerically expensive, cf [Bourinet, 2018], especially if the dimension d is high. The minimum distance between P^* and the origin is then written $\beta = \|P^*\|$, see Figure 3.1 (and it is named the Hasofer-Lind reliability index [Hasofer and Lind, 1974] in the standard normal case).

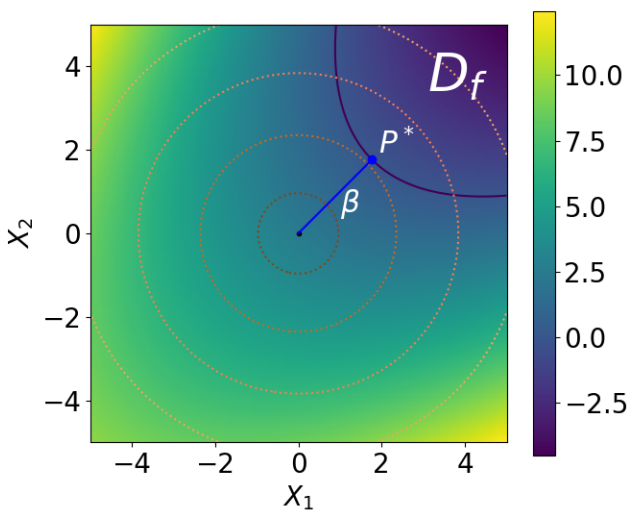


Figure 3.1: Illustration on a two-dimensional example of the design point of the failure domain D_f. The dark blue line is the limit state surface. The dotted lines represent the isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g.

Then, assuming the design point is unique, a simplified version of the failure domain is obtained by an approximation of the limit state surface at the vicinity of P^*, namely its Taylor series expansion.

3.2.1.1 The FORM approximation

The FORM approximation results from a linear approximation of the limit state surface at P^* (first-order Taylor series expansion). The approximated limit state function is then written

$$\forall\, \mathbf{x} \in \mathbb{R}^d \qquad g^{\text{FORM}}(\mathbf{x}) = g(P^*) + \nabla_{\mathbf{x}} g(P^*)^\top(\mathbf{x} - P^*) = \nabla_{\mathbf{x}} g(P^*)^\top(\mathbf{x} - P^*)$$

because $g(P^*) = 0$ by definition of P^*. Since the standard elliptical distributions are invariant by orthogonal transformation [Lebrun and Dutfoy, 2009], one can assume that the design point P^* has the following components $P^* = (\beta, 0, \ldots, 0)^\top$. In the FORM approximation, the failure domain is thus approximated with the hyperplane tangent to D_f at P^*, perpendicular to P^*, thus $\bar{D}_f = \{\mathbf{x} \in$

$\mathbb{R}^d \,|\, \mathbf{x}^\top P^* \geq \beta^2\}$, and the probability of failure is written

$$\hat{P}_f^{\text{FORM}} = \mathbb{P}(\mathbf{X} \in \bar{D}_f) = \mathbb{P}(\mathbf{X}^\top P^* \geq \beta^2) = \mathbb{P}(X_1 \geq \beta) \tag{3.5}$$

$$= H_{X_1}(\beta) = 1 - F_{X_1}(\beta) = F_{X_1}(-\beta), \tag{3.6}$$

where F_{X_1} is the cdf of the marginal X_1.

3.2.1.2 The SORM approximation

The SORM approximation results from a quadratic approximation of the limit state surface at P^* (second order Taylor series expansion). The approximated limit state function is then written

$$\forall \mathbf{x} \in \mathbb{R}^d \quad g^{\text{SORM}}(\mathbf{x}) = \nabla_{\mathbf{x}} g(P^*)^\top (\mathbf{x} - P^*) + \frac{1}{2}(\mathbf{x} - P^*)^\top \nabla^2_{\mathbf{x},\mathbf{x}} g(P^*)(\mathbf{x} - P^*),$$

with $\nabla^2_{\mathbf{x},\mathbf{x}}$ being the Hessian operator. Contrary to FORM, the probability of failure has unfortunately no simple analytical formulation with the SORM approximation. There exist then several methods amongst the SORM framework (curvature-fitting SORM and point-fitting SORM, see [Bourinet, 2018]) to estimate the failure probability integral over the approximated failure domain. It was proven in [Lebrun and Dutfoy, 2009] that Breitung's formula [Breitung et al., 1984] in the curvature-fitting SORM context could be generalized to spherical distributions with the following expression

$$P_f \overset{\beta \to \infty}{=} F_{X_1}(-\beta) \prod_{i=1}^{d-1} \frac{1}{\sqrt{1 + \beta \kappa_i}} + o(1),$$

where the κ_i for $i = 1, \ldots, d-1$ are the principal curvatures of the limit state surface, that have to be estimated. Therefore

$$\hat{P}_f^{\text{SORM}} = F_{X_1}(-\beta) \prod_{i=1}^{d-1} \frac{1}{\sqrt{1 + \beta \kappa_i}}, \tag{3.7}$$

but other formulas exist for the standard normal case [Tvedt, 1989, Hohenbichler and Rackwitz, 1988]. It should be noted that the SORM approximation requires the evaluation of the Hessian matrix at P^*. In high-dimensional input spaces, the computation of such matrix can become numerically cumbersome and is a major drawback. Point-fitting SORM is more adequate when an accurate computation of the Hessian matrix cannot be achieved [Bourinet, 2018]. An illustration of the FORM and SORM approximations of the limit state surface is displayed in Figure 3.2.

3.2.1.3 Remarks on the FORM/SORM approximations

The FORM approximation can be very efficient and practical when the gradient of the limit state function is available, as it can evaluate very small probability of failures with negligible simulation budget (inferior to a hundred lsf evaluations). However, as illustrated by Figure 3.2, the limit state function has to be linear, or else the estimation can be heavily biased [Valdebenito et al., 2010], especially when the dimension of the system is high. If the limit state function is not linear, then SORM approximation can provide a better estimate with a low simulation budget, assuming the Hessian matrix is available.

Furthermore, it should be noted that the error of the probability estimates of Eq. (3.6) and Eq. (3.7) is difficult if not impossible to estimate. It is indeed arduous to validate the hypothesis made on the behavior of limit state surface at the vicinity of the design points when the limit state function is considered a black-box function. This lack of error measurement is the major weakness of the FORM/SORM approximations.

3.2.2 Multiple design points and multiple failure regions

As previously mentioned, the FORM/SORM approximations rely on the hypothesis of the uniqueness of the design point P^*. However, in our specific context, we suppose that the failure domain encompasses several failure regions. Thus there could be several design points, or at least several local design points. Assuming P^* is unique could lead to severe errors and underestimation of the probability of failure P_f. In order to solve this issue, Der Kiureghian and Dakessian proposed a method in [Der Kiureghian and Dakessian, 1998] whose basic idea is to repeat the FORM/SORM approximation for all the identified design points. The main difficulty lies in finding all of the design points.

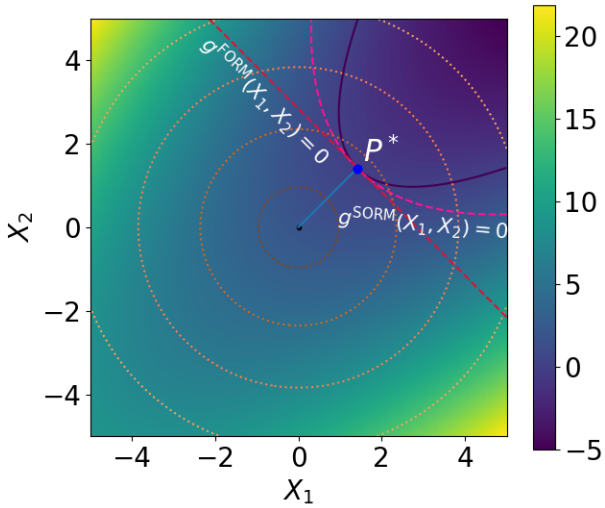


Figure 3.2: Illustration on a two-dimensional example of the FORM and SORM approximations of the limit state surface. The dark blue line is the limit state surface. The dotted lines represent the isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g.

3.2.2.1 Search for the design points

Their scheme to search for the design points, denoted P_i^* for $i = 1, \ldots, K$, relies on the modification of the limit state function each time a design point is found. Indeed, a *bulge* built around a design point P_i^* acts like a barrier and allows the optimizer to explore the rest of the input space at the next optimization Eq. (3.4), as the neighborhood of P_i^* is then in the safety domain. Assuming that K design points have already been found, the expression of the modified limit state function is the following

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \tilde{g}(\mathbf{x}) = g(\mathbf{x}) + \sum_{i=1}^{K} \text{Bulge}_i(\mathbf{x}),$$

where

$$\text{Bulge}_i(\mathbf{x}) = \begin{cases} s_i \left((\alpha_i)^2 - \|\mathbf{x} - P_i^*\|^2\right)^2 & \text{if } \|\mathbf{x} - P_i^*\| \leq \alpha_i, \\ 0 & \text{elsewhere,} \end{cases}$$

The scalar α_i is the radius of the bulge equal to $\alpha_i = \gamma \times \beta_i$ and s_i is the scale of the bulge, which controls the height of the bulge, defined by

$$s_i = \frac{\delta \beta_i \|\nabla g(P_i^*)\|}{\left[(\gamma \beta_i)^2 - (\delta \beta_i)^2\right]^2},$$

with γ and δ two user-defined parameters such that $0 < \delta < \gamma$, and $\beta_i = \|P_i^*\|$. The bulge is built to have a strong outward curvature, so that the limit state surface around P_i^* is sufficiently moved away from the origin. It is continuous and differentiable. With every design point identified, the limit state function is updated and the optimizer cannot converge back to it. In order to further help the optimizer explore the rest of the input space, it is recommended in [Der Kiureghian and Dakessian, 1998] to use as starting point of the optimization the point $\mathbf{x}_0$ located at the opposite direction of the previously found design points

$$\mathbf{x}_0 = -\epsilon \left(\sum_{i=1}^{K} P_i^* \right),$$

where ϵ is a small positive number between 0.2 and 0.5.

The procedure is repeated until a *spurious* design point is found. There are two criteria to determine if a design point is spurious or not. First, if it is found at a distance smaller or equal to the radius

α_i from any design point P_i^* for $i = 1, \ldots, K$, then it is spurious. Second, defining ϕ_i as the angle between P_i^* and a line from the origin to the foot of the Bulge$_i$, it was shown in [Der Kiureghian and Dakessian, 1998] that

$$\phi_i \leq \cos^{-1}(1 - 0.5\gamma^2),$$

with the equality holding when the limit state surface is spherical. Therefore any design point found within these cones centered on the directions of P_i^* for $i = 1, \ldots, K$ is spurious as well. The convergence of the optimizer to a spurious design point is an indication that no more genuine design points exist and thus the algorithm stops.

In order to illustrate the method and the bulges, the Figure 3.3 represents the different steps of the algorithm for the search for all design points, with $\gamma = 1.1$, $\delta = 0.75$ and $\epsilon = 0.5$.

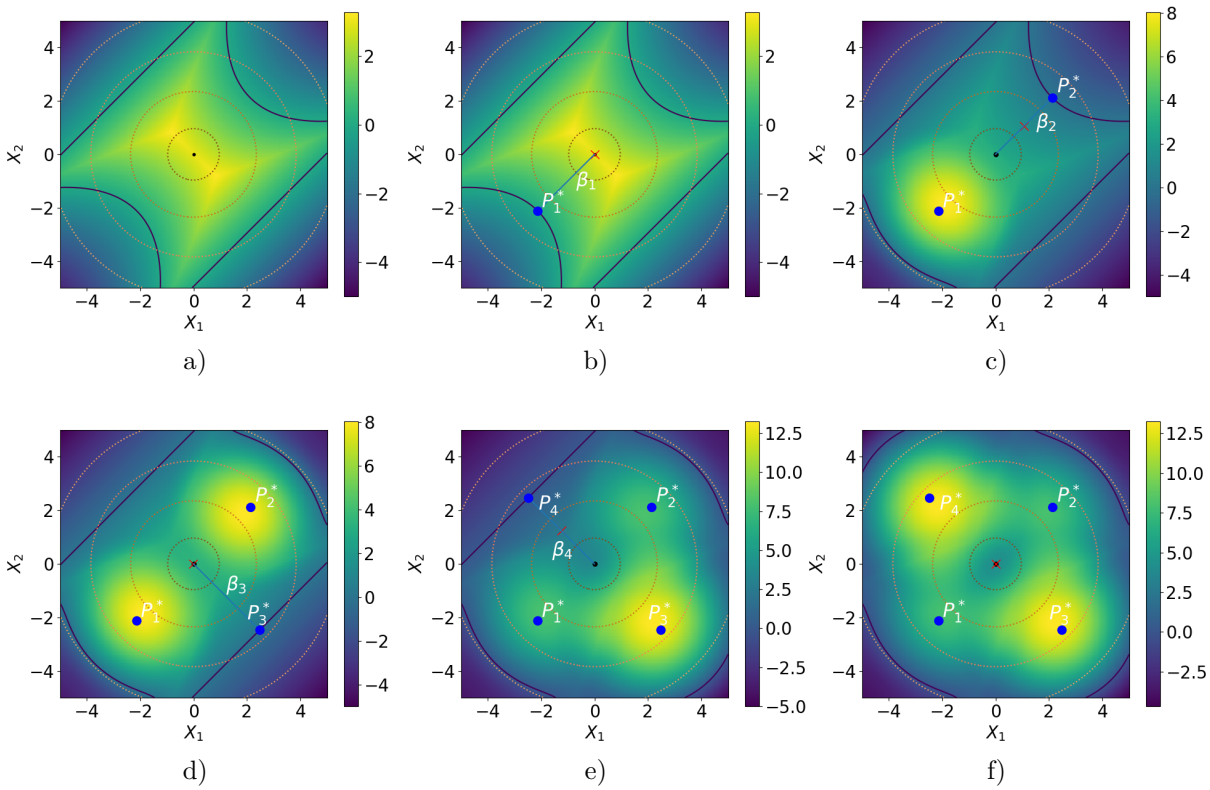


Figure 3.3: Illustration of the search for the design points on an example in dimension 2 with four failure regions a) the original limit state function, b) the first design point found P_1^*, c) the second design point found P_2^*, d) the third design point found P_3^*, e) the fourth design point found P_4^*, f) the final modified limit state function; as the optimization finds a spurious design point, the procedure stops. The dark blue line is the limit state surface. The color map represents the value of the limit state function $\tilde{g}$. The red cross represents the starting point $\mathbf{x}_0$ of the optimization. The dotted lines represent the isovalues of the joint pdf of the inputs.

3.2.2.2 Estimation of the failure probability

When all design points have been identified, the different failure regions approximated by FORM/SORM can be overlapping. Therefore, one must consider the intersections of the resulting approximated failure regions [Schueller et al., 2004]. These expressions are usually hard to derive [Schuëller and Stix, 1987]. Consequently, the final failure probability estimate is usually obtained in terms of bounds [Kounias, 1968]. Nevertheless, in the case of a series-system, a formula can be derived for standard normal inputs [Hohenbichler, 1984] in the FORM context. Suppose the system is composed of K limit state functions such that

$$g(\mathbf{x}) = \min_{k \in [1, \ldots, K]} g_k(\mathbf{x}),$$

the failure event corresponds then to the union of the failure events of the lsfs g_k

$$P_f = \mathbb{P}(g(\mathbf{x}) \leq 0) = \mathbb{P}\left(\cup_{k=1}^{K} g_k(\mathbf{x}) \leq 0\right).$$

This probability can also be written $P_f = 1 - \mathbb{P}\left(\cap_{k=1}^{K} g_k(\mathbf{x}) > 0\right)$. In the FORM context, each g_k is approximated with a linear function l_k of the inputs $\mathbf{X}$ in the vicinity of P_k^*. In the standard normal framework, since all the inputs X_i are independent, the resulting random variable $l_k(\mathbf{X})$ is a linear combination of independent standard normal variables, therefore $l_k(\mathbf{X})$ is also normally distributed. Noticing that

$$1 - \mathbb{P}\left(\cap_{k=1}^{K} g_k(\mathbf{x}) > 0\right) = 1 - \mathbb{P}(l_1(\mathbf{X}) > 0, \ldots, l_K(\mathbf{X}) > 0) = 1 - \mathbb{P}(-l_1(\mathbf{X}) < 0, \ldots, -l_K(\mathbf{X}) < 0)$$

and $\mathbb{P}(-l_1(\mathbf{X}) < 0, \ldots, -l_K(\mathbf{X}) < 0)$ is the joint cumulative distribution function of the vector $(-l_1(\mathbf{X}), \ldots, -l_K(\mathbf{X}))$ evaluated at $(0, \ldots, 0)$, the failure probability can be approximated with

$$\hat{P}_f^{\text{FORM}} = 1 - \Psi_{K,\mathbf{C}}(\boldsymbol{\beta}),$$

where Ψ is the cdf of a centered normal random vector of dimension K with covariance matrix $\mathbf{C}$ and $\boldsymbol{\beta}$ is the vector of the norms of all the design points $(\beta_k = \|P_k^*\|)_{k=1,\ldots,K}$, see [Ditlevsen and Madsen, 1996] for more details. The covariance matrix $\mathbf{C}$ is the covariance of the random vector $(l_1(\mathbf{X}), \ldots, l_K(\mathbf{X}))$. This approximation could not be derived for the other standard elliptical spaces, as the inputs X_i are not independent, therefore the distribution of the vector of linear combinations $(l_k(\mathbf{X}))_{k=1,\ldots,K}$ is more difficult to identify.

In a broader context, computing the failure probability of a system with multiple failure regions with FORM/SORM approximations is thus challenging. It should be noted that the estimation error of the final failure probability estimate encompasses all the estimation errors of the FORM/SORM approximations for each design point. Consequently, the resulting error could be very large for lsf with many failure regions.

3.2.3 Design points in high-dimensional standard normal space

In our specific context, the dimension of the system d is particularly high. In standard elliptical spaces, the probability mass of the inputs is concentrated in the important ring (cf Chapter 2). In the case of the standard normal distribution, the important ring is particularly narrow and located close the hypersphere of radius $\sqrt{d}$. In [Katafygiotis and Zuev, 2008], Katafygiotis and Zuev pointed out that the design points typically do not belong to this important ring. Therefore, although the design points are the failing points with maximum likelihood, the probability masses associated with their neighborhood are negligible. It is the failure domain located amongst the important ring that contributes most to the failure probability integral of Eq. (3.1).

Since the FORM and SORM techniques construct approximations of the limit state surface around the design points, these approximations may be quite accurate in the vicinity of those points but erroneous in regions far from the design points in the important ring. Thus, these methods may lead to biased estimations of P_f. An example of such inaccuracy in large dimensions is presented in [Valdebenito et al., 2010].

3.3 Simulation methods

In this section, simulation methods are described, starting with the Monte Carlo method. The subset sampling method is next presented as it performs well in the particular context of this manuscript. A review of several adaptive importance sampling algorithm is presented afterwards, since the main contributions of this manuscript are strongly linked to some them. Finally, the directional sampling method is detailed, as this method takes into account the stochastic representation of the standard elliptical inputs. For each method presented, its efficiency regarding failure domains encompassing multiple failure regions is underlined, as well as its performance in high-dimensional standard elliptical spaces.

3.3.1 Crude Monte Carlo

The crude Monte Carlo (MC) method is the original sampling method and the most basic one to estimate a mathematical expectation [Robert and Casella, 1999, Sobol, 2018]. Using the fact that $P_f = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f}(\mathbf{X})\right]$, the MC method consists in generating a sample of size N of independent and identically distributed (iid) observations $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$ from their joint pdf $f_{\mathbf{X}}$ to estimate the failure probability by the sample mean of the failure domain indicator function

$$\hat{P}_f^{\mathrm{MC}} = \frac{1}{N}\sum_{j=1}^{N}\mathbb{I}_{D_f}\left(\mathbf{X}^{(j)}\right). \tag{3.8}$$

The convergence of this estimator is a consequence of the law of large numbers. The Monte Carlo method leads to an unbiased estimator of the failure probability $\mathbb{E}\left[\hat{P}_f^{MC}\right] = P_f$, which converges regardless of the complexity of the limit state function g and the dimension of the system d. Furthermore, the family distribution of the inputs does not matter, as long as one can generate sample from $f_{\mathbf{X}}$. Thus, in our specific context of high-dimensional standard elliptical space with multiple failure regions, the MC method is relevant.

The error of the estimate Eq. (3.8) can be measured with its variance and its Coefficient of Variation (CV). The theoretical variance of the estimate is obtained by noticing that $\mathbb{I}_{D_f}(\mathbf{X})$ follows a Bernoulli distribution of parameter P_f

$$\mathbb{V}\mathrm{ar}\left(\hat{P}_f^{\mathrm{MC}}\right) = \frac{1}{N}\mathbb{V}\mathrm{ar}\left(\mathbb{I}_{D_f}(\mathbf{X})\right) = \frac{P_f(1-P_f)}{N}.$$

The CV is then defined as the ratio between the standard error and the expected value of the estimate $\hat{P}_f^{\mathrm{MC}}$

$$\mathrm{CV}\left(\hat{P}_f^{\mathrm{MC}}\right) = \frac{\sqrt{\mathbb{V}\mathrm{ar}\left(\hat{P}_f^{\mathrm{MC}}\right)}}{\mathbb{E}\left[\hat{P}_f^{\mathrm{MC}}\right]} = \sqrt{\frac{1-P_f}{NP_f}}. \tag{3.9}$$

The ratio Eq. (3.9) shows that the convergence speed of the MC estimate is quite slow, as it is proportional to $1/\sqrt{N}$. As P_f tends to 0, for a set N, the coefficient of variation is unbounded. For a failure probability of value 10^{-p}, one should have a simulation budget of order 10^{p+2} so that the CV reaches 10%. This heavy simulation budget cannot be afforded if the limit state function g is computationally expensive and thus crude MC method cannot be directly applied to the rare event probability estimation of this manuscript. However, the crude MC method will be considered as the reference method in the rest of the manuscript, with its estimate of the failure probability $\hat{P}_f^{\mathrm{MC}}$ being considered the reference value. The next presented simulation method, called subset sampling, is also used for comparison purposes later in the manuscript.

3.3.2 Subset Simulation

3.3.2.1 SS in the standard elliptical space

The Subset Simulation (SS) method, or subset sampling method was developed for reliability assessment purpose by Au and Beck in [Au and Beck, 2001], although it had been introduced before in other scientific fields. Here we present the algorithm as developed in [Au and Beck, 2001]. The basic idea of SS is to gradually move towards the failure regions of the failure domain D_f in the original input space through the introduction of a decreasing sequence of subsets, or intermediate failure events F_i, with $i = 0, \dots, m$ such that $F_0 \supset F_1 \supset \cdots \supset F_m = D_f$ and $F_k = \bigcap_{i=0}^{k} F_i$ for $k = 0, \dots, m$. The failure probability can then be expressed as

$$P_f = \mathbb{P}(F_m) = \mathbb{P}(F_0)\prod_{i=0}^{m-1}\mathbb{P}(F_{i+1}|F_i),$$

where $\mathbb{P}(F_0) = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{F_0}(\mathbf{X})\right]$ and $\mathbb{P}(F_{i+1}|F_i)$ is the conditional probability of failure F_{i+1} given that the failure F_i has already occurred, thus $\mathbb{P}(F_{i+1}|F_i) = \mathbb{E}_{f_{\mathbf{X}|F_i}}\left[\mathbb{I}_{F_{i+1}}(\mathbf{X})\right]$, where the conditional density is

Algorithm 1 Subset simulation

Require: The joint pdf $f_{\mathbf{X}}$, the probability level p_0 and the sample size N.

Initialization

Generate an iid sample $\left(\mathbf{X}_0^{(j)}\right)_{j=1,\dots,N}$ from the joint pdf $f_{\mathbf{X}}$. Compute the response of the system $\left(Y_0^{(j)} = g\left(\mathbf{X}_0^{(j)}\right)\right)_{j=1,\dots,N}$. Select the first intermediate threshold s_0 as the p_0-quantile of the response values so that the estimation of $\mathbb{P}(F_0) = \mathbb{P}(g(\mathbf{X} \leq s_0)) = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{F_0}(\mathbf{X})\right]$ is equal to p_0 with crude MC. Set the index i equal to 0.

while $s_i > 0$ **do**

Markov Chain Monte Carlo for the intermediate failure events

Amongst the N-sample $\left(\mathbf{X}_i^{(j)}\right)_{j=1,\dots,N}$, p_0 percent of the observations have a corresponding response value $Y_i^{(j)}$ smaller than s_i and hence lie in F_i. These observations are distributed as $f_{\mathbf{X}|F_i}$ and provide seeds for simulating additional observations with MCMC. A N-sample $\left(\mathbf{X}_{i+1}^{(j)}\right)_{j=1,\dots,N}$ lying in F_i is then obtained. From this sample, the new responses $\left(Y_{i+1}^{(j)} = g\left(\mathbf{X}_{i+1}^{(j)}\right)\right)_{j=1,\dots,N}$ are computed. The new intermediate threshold s_{i+1} is selected as the p_0-quantile of the response values so that the estimation of $\mathbb{P}(F_{i+1}|F_i) = \mathbb{P}(g(\mathbf{X} \leq s_{i+1}))$ is equal to p_0 with crude MC. The index i is incremented: $i = i + 1$.

end while

Estimation of the failure probability

Set m equal to i. The probability of failure is estimated with the following formula

$$\hat{P}_f^{\mathrm{SS}} = p_0^m \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{Y \leq 0}\left(Y_m^{(j)}\right)$$

return $\hat{P}_f^{\mathrm{SS}}$.

written

$$\forall \mathbf{x} \in \mathbb{R}^d \qquad f_{\mathbf{X}|F_i}(\mathbf{x}) = \frac{\mathbb{I}_{F_i}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{\mathbb{P}(F_i)}.$$

Thus the small probability P_f is computed as a product of $(m+1)$ larger conditional probabilities, which are easier to estimate than P_f with crude Monte Carlo. The intermediate failure events are typically defined with intermediate threshold s_i, such that $F_i = \{\mathbf{x} \in \mathbb{R}^d \,|\, g(\mathbf{x}) \leq s_i\}$. Therefore $s_0 > s_1 > \cdots > s_m = 0$.

The thresholds are adaptively selected so that all of the probabilities $\mathbb{P}(F_0)$, $\mathbb{P}(F_1|F_0)\dots$, and $\mathbb{P}(F_{m-1}|F_{m-2})$ are equal to the same constant with a MC estimation, which is often taken as $p_0 = 0.1$. Targeting equal probability levels for the conditional probabilities was proven in [Guyader, 2011] to be the optimal choice for a minimal asymptotic variance of the estimate. The value of p_0 is also not randomly selected as it was proven in [Zuev et al., 2012] that an appropriate range for p_0 is between 0.1 and 0.3, to minimize the variance of the estimate for a given simulation budget. This probability p_0 sets the number m of intermediate failure domain F_i required to reach the failure domain D_f. A large probability p_0 makes it is easier to estimate the intermediate probabilities but leads to a larger number m and thus to a larger simulation budget. A small probability p_0 leads to a faster convergence of the algorithm but the intermediate probabilities are then harder to estimate, therefore the errors made on the estimations may be greater.

Algorithm 1 briefly sums up the logic of the SS method with Markov Chain Monte Carlo (MCMC) and Figure 3.4 illustrates the different steps. In comparison to the crude MC method, the subset simulation algorithm aims to generate observations in more critical regions which are located far from the mean value $\boldsymbol{\mu}_{\mathbf{X}}$ of the inputs distribution. The performance of this subset simulation algorithm heavily relies on the MCMC process which has to generate from the *seeds* additional conditional observations, independent and identically distributed. Markov Chain Monte Carlo [Metropolis et al., 1953] are very powerful techniques for simulating samples according to an arbitrary distribution called the target distribution. The observations are simulated as the states of a Markov chain whose stationary

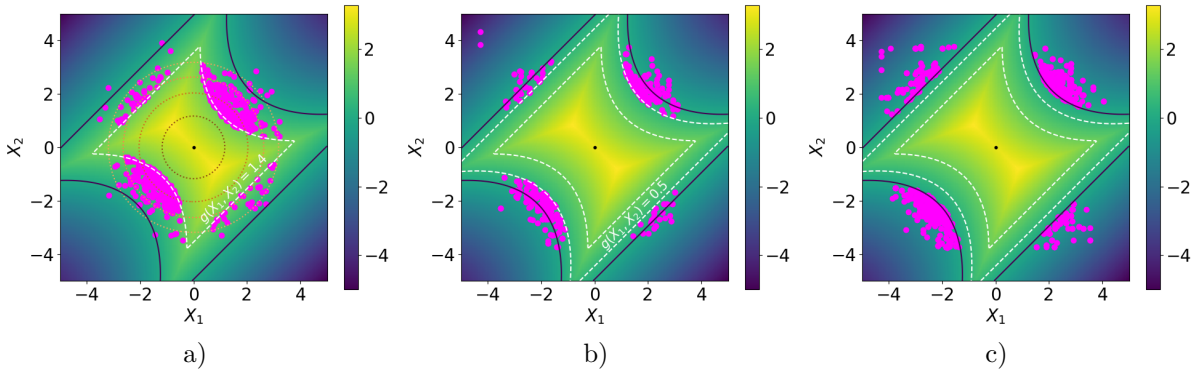


Figure 3.4: Illustration of the main steps of the SS algorithm with $p_0 = 0.1$ a) the first threshold $s_0 = 1.4$ is computed from observations generated from $f_{\mathbf{X}}$, b) the second threshold $s_1 = 0.5$ is computed from observations resulting from the MCMC algorithm, c) the third threshold s_2 computed from observations resulting from the MCMC algorithm is smaller than 0 thus it is set back to 0, the loop stops and the failure probability is estimated with the observations in the failure domain. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dashed white lines represent the intermediate limit state surfaces. The dotted lines represent the isovalues of the joint pdf of the inputs $f_{\mathbf{X}}$. The magenta dots represent the observations located in the intermediate failure domains at each step.

distribution is the target distribution. Here, the target distribution is thus $f_{\mathbf{X}|F_i}$.

The MCMC process selected for the specific needs of the SS algorithm in [Au and Beck, 2001] is the Metropolis algorithm [Bourinet, 2018]. The original Metropolis algorithm, described in the Appendix A.1, is applicable for any joint pdf $f_{\mathbf{X}}$, thus it is adequate for all standard elliptical distributions. It relies on a d-dimensional proposal pdf, set with a spread parameter. However, it does not perform well in high-dimensional standard normal spaces, which is a space of interest in the context of our manuscript. Consequently, the modified Metropolis algorithm as well as other algorithms have been developed for such a space and are presented below.

3.3.2.2 MCMC algorithms for high-dimensional standard normal inputs

Taking into account the special behavior of the standard normal space when the dimension of the system is high, namely the concentration of the probability mass in a narrow important ring (cf Chapter 2.), it was underlined in [Katafygiotis and Zuev, 2008] that the original Metropolis algorithm is inadequate for this particular distribution. In fact, as described in [Au and Beck, 2001], for independent multivariate distributions in high-dimensional spaces, the probability that the next state in the Markov Chain will be equal to the current state is close to 1. Thus the corresponding chain may consist in very few observations repeated numerous times. The SS method is then almost inapplicable.

The modified Metropolis algorithm is then specially tailored for independent distribution in high-dimensional spaces. While the logic remains the same, the way the candidate state is generated changes. Indeed, instead of using a d-dimensional proposal pdf, the independence of the random vector is taken into account and a sequence of univariate proposal densities is employed. Each densities is set with a spread parameter as well. This algorithm allows to overcome the deficiency of the original Metropolis algorithm for independent distributions and is briefly described in the Appendix A.2. In our specific case, the modified Metropolis algorithm is suited for the standard normal distribution.

However, as pointed by Katafygiotis et al. in [Katafygiotis and Zuev, 2008], the modified Metropolis algorithm may be inefficient when the seeds are directed towards *preferred directions*, e.g. $(\sqrt{d}, 0, \ldots, 0)$. If the Markov chain has a length too small, then it might exit the important ring as the chain has not yet reached its underlying stationary distribution. Unfortunately, in practical application, one would use Markov chains of small length as they require smaller simulation budget. Consequences of this phenomenon are displayed in [Katafygiotis and Zuev, 2008], where the resulting SS failure probability estimate $\hat{P}_f^{\mathrm{SS}}$ has a large variance.

To solve this problem, another MCMC algorithm developed by Katafygiotis et al. in [Katafygiotis and Cheung, 2007] explores the stochastic representation of the standard normal space and is briefly described in the Appendix A.3. Using this MCMC algorithm built in the stochastic representation framework solves the *preferred direction* issue. It should be noted that the authors of [Katafygiotis and Cheung, 2007] also used the stochastic representation of the standard normal random variables to another extend, as they proposed a new SS algorithm, called spherical subset simulation, with intermediate failure event constructed as spherical rings of decreasing inner and outer radii, see [Katafygiotis and Cheung, 2007] for more details.

3.3.2.3 Statistical properties of SS

Some of the statistical properties of the Metropolis algorithms are detailed in the Appendix A.4, as the resulting Markov chains heavily relies on the spread parameters of the proposal pdfs previously introduced. Here we present basic statistical properties of the SS algorithm.

We first assume that the intermediate failure events F_i are set *a priori*, independently of the generated samples. As the *seeds* at the start of the Markov chains of step i are used to compute both the intermediate failure probabilities $\mathbb{P}(F_i|F_{i-1})$ of the previous step and $\mathbb{P}(F_{i+1}|F_i)$ of the current step, there is a high correlation between each intermediate failure probability estimate. This source of dependence induces a bias in the final failure probability estimate, and the bias is $\mathcal{O}(1/N)$ (where N is the fixed sample size of each step of the SS algorithm). Nevertheless, the failure probability estimate is asymptotically unbiased (cf [Au and Beck, 2001]). This bias is supposed to become slightly stronger if each F_i depends on the samples, as presented in Algorithm 1.

The variance of the SS estimate is very hard to derive because of the dependence mentioned. However, a Central Limit Theorem (CTL) type of convergence described in [Cérou et al., 2012] allows to derive an asymptotic expression of the variance of the estimate. Since the failure probability estimate is biased, another expression of the coefficient of variation can be used

$$\left(\mathrm{CV}\left(\hat{P}_f^{\mathrm{SS}}\right)\right)^2 = \mathbb{E}\left[\left(\frac{\hat{P}_f^{\mathrm{SS}} - P_f}{P_f}\right)^2\right],$$

which is the CV expressed in terms of deviation with respect to P_f. The following bounds are then often employed to get an approximation of the CV of the failure probability estimate

$$\sum_{i=0}^{m} \delta_i^2 \leq \left(\mathrm{CV}\left(\hat{P}_f^{\mathrm{SS}}\right)\right)^2 \leq \sum_{i=0}^{m}\sum_{l=0}^{m} \delta_i \delta_l,$$

where δ_i are the CV of the intermediate failure probability estimate, which are unbiased. The upper bound formula corresponds to the assumption of fully-correlated intermediate failure probability estimates while the lower bound correspond to the assumption of independent intermediate failure probability estimates. The small induced theoretical bias can be an important drawback of the SS algorithm presented here.

It should be noted that the SS concept, also known as splitting, has been further studied in numerous papers, cf [Cérou et al., 2019] for instance, where the statistical properties presented here are improved. However, SS is not the main subject of this thesis thus these improvements are not derived here.

3.3.2.4 SS algorithm for multiple failure regions

In the particular context of our study, we suppose that there are several failure regions in the failure domain. The following consequence is that the lsf has multiple local minima. As underlined in various studies [Valdebenito et al., 2010, Breitung, 2019, Tabandeh et al., 2022], for particular lsfs, the Metropolis algorithms mentioned can get trapped in specific regions, and do not explore the rest of the inputs space. This phenomenon can occur if there is a change in the topological structure of the intermediate limit state surfaces for instance [Tabandeh et al., 2022], or simply if the spread parameter of the proposal (univariate or multivariate) pdfs is not large enough. If the initial *seeds* are not located towards every part of the failure domain, some failure regions might be left unexplored, as the Markov chains never reach them. Depending on the importance of these unexplored failure

regions, the resulting failure probability estimate could be biased, as its value underrates the true failure probability. A larger simulation budget may then be needed to make sure all failure regions are reached by the Markov chains.

Consequently, the tuning of the spread parameter of the proposal pdfs is crucial. Its setting has influences on the convergence of the algorithm (through the dependence level of the samples, cf Appendix A.4), on the bias of the failure probability estimate and on the simulation budget. Thus the quality of the failure probability estimate and the efficiency of this SS method are not necessarily guaranteed *a priori*.

This SS algorithm will be used for comparison purposes later on in some numerical applications of the manuscript, with the modified Metropolis algorithm as MCMC procedure, as it remains quite efficient in high-dimensional spaces. Next, we introduce the importance sampling framework, which plays a major role in the contributions developed during the thesis. It is therefore presented in detail.

3.3.3 Importance Sampling

The importance sampling method (IS) is part of larger group of so-called *variance-reduction techniques* whose purpose is to lower the simulation budget of the MC estimate by diminishing its variance. The main idea is to introduce an auxiliary density h, whose support includes the support of $\mathbb{I}_{D_f} f_{\mathbf{X}}$, in the integral Eq. (3.1), thus

$$P_f = \int_{\mathbb{R}^d} \mathbb{I}_{D_f}(\mathbf{x}) \frac{f_{\mathbf{X}}(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \mathbb{E}_h \left[\mathbb{I}_{D_f}(\mathbf{X}) \frac{f_{\mathbf{X}}(\mathbf{X})}{h(\mathbf{X})} \right].$$

The IS estimate of P_f is computed with the same logic as the MC estimate, given by the following equation

$$\hat{P}_f^{\mathrm{IS}} = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{D_f}\left(\mathbf{X}^{(j)}\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}^{(j)}\right)}{h\left(\mathbf{X}^{(j)}\right)},$$

where the random variables $\left(\mathbf{X}^{(j)}\right)$ are iid with density h. The density ratio $f_{\mathbf{X}}/h$ is referred to as likelihood ratio and the vector $\left(\mathbb{I}_{D_f}\left(\mathbf{X}^{(j)}\right) f_{\mathbf{X}}\left(\mathbf{X}^{(j)}\right)/h\left(\mathbf{X}^{(j)}\right)\right)_{j=1,\ldots,N}$ are the IS weights of the sample. The IS estimate is theoretically unbiased. The efficiency of the auxiliary density h depends on its ability to generate more observations in the failure domain than $f_{\mathbf{X}}$. The variance of the IS estimate is given by the following formula

$$\begin{aligned} \mathbb{V}\mathrm{ar}\left(\hat{P}_f^{\mathrm{IS}}\right) &= \frac{1}{N} \mathbb{V}\mathrm{ar}_h \left(\mathbb{I}_{D_f}(\mathbf{X}) \frac{f_{\mathbf{X}}(\mathbf{X})}{h(\mathbf{X})} \right) \\ &= \frac{1}{N} \left(\mathbb{E}_h \left[\left(\mathbb{I}_{D_f}(\mathbf{X}) \frac{f_{\mathbf{X}}(\mathbf{X})}{h(\mathbf{X})} \right)^2 \right] - P_f^2 \right). \end{aligned}$$

With an appropriate auxiliary density h, the variance of the IS estimate and hence also its CV can be drastically reduced compared to those of the crude MC. The theoretically optimal IS density h_{opt}, which leads to a zero variance IS probability estimate, is given by [Bucklew, 2004]

$$\forall \mathbf{x} \in \mathbb{R}^d \qquad h_{\mathrm{opt}}(\mathbf{x}) = \frac{\mathbb{I}_{D_f}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{X})}{P_f}$$

This density h_{opt} is unavailable in practice as it depends on the quantity of interest, P_f. The goal of the IS method is then to construct an auxiliary density which is as close as possible to the optimal one, so that the greatest possible variance reduction can be achieved. There are several ways to construct this auxiliary densities [Tabandeh et al., 2022]. Adaptive Importance Sampling (AIS) algorithms aim to get closer to h_{opt} gradually and have been greatly studied. Here in this Section 3.3.3 only the following ones are derived:

- Non parametric Adaptive Importance Sampling (NAIS)
- Cross Entropy-based Adaptive Importance Sampling (CE-AIS)

- Cross Entropy-based Adaptive Multiple Importance Sampling (CE-AMIS)
- improved Cross Entropy-based Adaptive Importance Sampling (iCE-AIS)

The first algorithm presented, NAIS algorithm, is non parametric while the others are all parametric. The last two parametric methods presented, CE-AMIS and iCE-AIS, are improvements of the CE-AIS algorithm. Non adaptive importance sampling algorithms will be mentioned in the Section 3.4 of this chapter. It should be noted that the IS framework is applicable for any inputs distribution, therefore it is adequate for all standard elliptical distributions.

3.3.3.1 Non parametric Adaptive Importance Sampling

The NAIS algorithm was presented in [Zhang, 1996] and aims to sequentially construct an auxiliary density close to h_{opt}, using non parametric densities. Non parametric densities are particularly relevant if no information about the failure domain is available or if the failure domain is supposed to encompass several failure regions. Indeed, non parametric densities are very flexible thus especially pertinent to build multimodal densities as they do not impose a specific pdf family.

Non parametric densities typically refer to kernel densities [Tabandeh et al., 2022] built around a sample. Let $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$ be an iid sample drawn from some multivariate distribution whose pdf is h, the kernel density estimate (kde) of h is

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \hat{h}(\mathbf{x}) = \frac{1}{N\det(B_N)} \sum_{j=1}^{N} K_d\left(B_N^{-1}\left(\mathbf{x} - \mathbf{X}^{(j)}\right)\right),$$

where K_d is a kernel density, a non-negative function that integrates to one, and $B_N = \text{diag}(b_1, \dots, b_N)$ is a diagonal covariance matrix. The diagonal covariance matrix is referred to as bandwidth of the kde. The classical kernel density K_d is the centered normal density (see Chapter 2, Section 2.2.1.3 for the expression of this density) with diagonal covariance matrix $\mathbf{\Sigma} = B_N$ [Morio et al., 2014]. In the specific case of IS, an IS weight is associated to each observation $\mathbf{X}^{(j)}$, therefore, weighted kernel density estimates are used in the NAIS algorithm. Let $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$ be an iid sample drawn from some multivariate distribution h and $(w_j)_{j=1,\dots,N}$ the corresponding weights. The weighted kernel density estimate (wkde) of h is

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \hat{h}(\mathbf{x}) = \frac{1}{\det(B_N)} \sum_{j=1}^{N} \frac{w_j}{\sum_{\ell=1}^{N} w_\ell} K_d\left(B_N^{-1}\left(\mathbf{x} - \mathbf{X}^{(j)}\right)\right), \tag{3.10}$$

with the same notations as above.

The main idea of the NAIS algorithm is to sequentially built with wkdes an estimate of h_{opt}, using ρ-quantiles of the generated sample responses as intermediate failure thresholds s_i, where ρ is in $(0, 1)$. The intermediate threshold s_i defines an intermediate failure domain $D_{f,i} = \{\mathbf{x} \in \mathbb{R}^d | g(\mathbf{x}) \leq s_i\}$. As the iteration goes, the thresholds gradually decrease to reach the true failure threshold equal to zero [Morio et al., 2014], thus it is the same logic than in the subset sampling algorithm; cf Section 3.3.2.1. Algorithm 2 briefly sums up the NAIS algorithm and Figure 3.5 illustrates the different steps.

The parameter ρ greatly influences the convergence of the algorithm and the efficiency of the method. If this parameter is set too small, then the algorithm converges rapidly and the number of iterations to achieve convergence is small. However, since the weighted kernel density estimates are built with only a few points (as the intermediate failure domain indicator function of their associated weight is mostly equal to zero), there can be a greater error in their estimation. In contrast, setting ρ very large enables a very accurate wkde but slows down the convergence of the algorithm and thus the simulation budget gets heavier as the number of iterations gets large as well.

Since all the previously generated samples are reused to construct the new wkde at each iteration, the evaluations of the costly limit state function can be of service for several steps of the algorithm. Several studies have searched for optimal bandwidths parameter $B_{N,i}$, including the initial work of [Zhang, 1996], where an analytical formula is derived by minimizing the asymptotic MSE (mean square error) of the mean value of the IS weights at each iteration. Unfortunately, this formula cannot be used and depends on unknown integrals. Minimizing the AMISE (asymptotic mean integrated

Algorithm 2 NAIS

Require: The joint pdf $f_{\mathbf{X}}$, the quantile level ρ and the sample size N.

Initialization

Generate an iid sample $\left(\mathbf{X}_0^{(j)}\right)_{j=1,\dots,N}$ from the joint pdf $f_{\mathbf{X}}$. Compute the response of the system $\left(Y_0^{(j)} = g\left(\mathbf{X}_0^{(j)}\right)\right)_{j=1,\dots,N}$. Select the initial intermediate threshold s_0 as the maximum between the ρ-quantile of the response values $\left(Y_0^{(j)}\right)_{j=1,\dots,N}$ denoted $Y_0^{(\rho)}$ and zero: $s_0 = \max\left(0, Y_0^{(\rho)}\right)$. Compute the weights with this threshold $\left(w_0^{(j)} = \mathbb{I}_{Y<s_0}\left(Y_0^{(j)}\right)\right)_{j=1,\dots,N}$. Construct the wkde $\hat{h}_1$ with the vector weight $\mathbf{w}_0$, diagonal covariance matrix $B_{N,0}$ and sample $\left(\mathbf{X}_0^{(j)}\right)_{j=1,\dots,N}$, thanks to Eq. (3.10). Set the index i equal to 0.

while $s_i > 0$ **do**

Construction of the wkde for the intermediate failure thresholds

The index i is incremented: $i = i + 1$. Generate an iid sample $\left(\mathbf{X}_i^{(j)}\right)_{j=1,\dots,N}$ from $\hat{h}_i$. Compute the response of the system $\left(Y_i^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)\right)_{j=1,\dots,N}$. Select the intermediate threshold s_i as maximum between the ρ-quantile of the response values $\left(Y_i^{(j)}\right)_{j=1,\dots,N}$ denoted $Y_i^{(\rho)}$ and zero: $s_i = \max\left(0, Y_i^{(\rho)}\right)$. Compute the IS weights with this threshold, for all samples generated so far, thus $\mathbf{w}_i$ is a one-dimensional vector of length $N(i+1)$ where $\left(w_i^{(\ell N+j)} = \mathbb{I}_{Y<s_i}\left(Y_\ell^{(j)}\right) f_{\mathbf{X}}\left(\mathbf{X}_\ell^{(j)}\right) / \hat{h}_\ell\left(\mathbf{X}_\ell^{(j)}\right)\right)_{\ell=0,\dots,i;j=1,\dots,N}$, and $\hat{h}_0 = f_{\mathbf{X}}$. Construct the wkde $\hat{h}_{i+1}$ with the vector weight $\mathbf{w}_i$, diagonal covariance matrix $B_{N,i}$ and sample $\left(\mathbf{X}_\ell^{(j)}\right)_{\ell=0,\dots,i;j=1,\dots,N}$, thanks to Eq. (3.10).

end while

Estimation of the failure probability

Set m equal to i. The probability of failure is estimated with the following formula

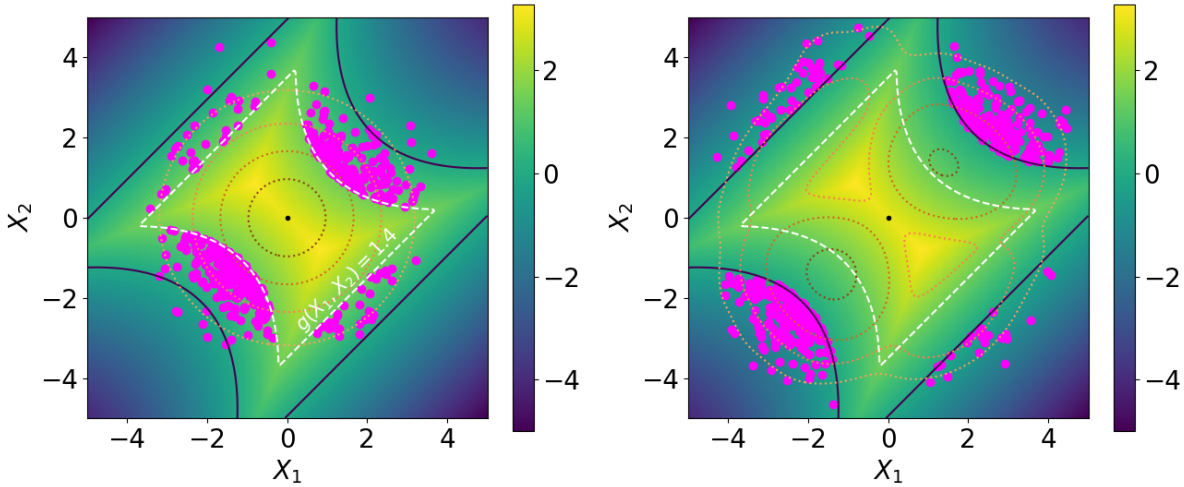
$$\hat{P}_f^{\text{NAIS}} = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}_{Y\leq 0}\left(Y_m^{(j)}\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_m^{(j)}\right)}{\hat{h}_m\left(\mathbf{X}_m^{(j)}\right)}$$

return $\hat{P}_f^{\text{NAIS}}$.

square error) criterion is another possibility and the interested reader can see [Silverman, 2018, Glad et al., 2007]. The NAIS estimate is unbiased and the studies on its convergence rate can be found in [Zhang, 1996]. As other IS estimates, an estimate of its variance is available with the simple formula $\widehat{\mathbb{V}\text{ar}}\left(\hat{P}_f^{\text{NAIS}}\right) = \frac{1}{N}\mathbb{V}\text{ar}_{\hat{h}_m}\left(\mathbb{I}_{Y\leq 0}\left(Y_m\right)\frac{f_{\mathbf{X}}(\mathbf{X}_m)}{\hat{h}_m(\mathbf{X}_m)}\right)$ (using the same notations as in the Algorithm 2) since all the observations are independent.

NAIS for multiple failure regions in high-dimensional spaces

Although the NAIS algorithm is very efficient for multiple failure regions thanks to the kde, it is known to lose its capability as the dimension of the inputs increase. The so-called *curse of the dimensionality* affects the kernel density estimate and the IS weights degenerate: an increasingly large portion of the samples have an associated IS weight close to zero. Moreover, the optimization of the bandwidth parameters gets more complex with the dimension, affecting the efficiency of the algorithm. One idea to overcome this issue is to focus on the most influential or important random variables [Tabandeh et al., 2022], identified with probabilistic sensitivity analysis [Jia and Taflanidis, 2014] for example. The kernel density is then optimized in the lower dimensional space of these inputs.



a) b)

Figure 3.5: Illustration of the main steps of the NAIS algorithm with $\rho = 0.1$ a) the initial threshold $s_0 = 1.4$ is computed from observations generated from $f_{\mathbf{X}}$, b) the next ρ-quantile computed from observations generated from $\hat{h}_1$ is already smaller than 0 thus the loop stops as $s_1 = 0$ and the failure probability is estimated with the observations in the failure domain. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dashed white lines represent the intermediate limit state surfaces. The dotted lines represent the isovalues of the densities $f_{\mathbf{X}}$ and $\hat{h}_1$. The magenta dots represent the observations of non-zero weight at each step.

3.3.3.2 Cross-Entropy based Adaptive Importance Sampling

The concept of the CE-AIS algorithm was first presented by Rubinstein [Rubinstein, 1997, Rubinstein, 1999, Rubinstein and Kroese, 2004] and aims to construct sequentially an auxiliary density close to h_{opt}, with a selected parametric density. Using parametric families as auxiliary densities for the IS estimation can be relevant if prior information about the failure domain is available or if the input distribution $f_{\mathbf{X}}$ is of a specific nature. The parameters $\mathbf{\Theta} \in \mathbb{R}^n$ of the selected family are then gradually optimized to obtain an auxiliary parametric density as close as possible to h_{opt}. In order to define the convergence of the parameters, one must specify a particular distance measure. The typical distance measure is the relative entropy, also referred to as Kullback–Leibler (KL) divergence [Kullback and Leibler, 1951]. The relative entropy between two densities q_1 and q_2, defined both on $\mathbb{R}^d$, is defined as

$$D_{\text{KL}}(q_1 | \, q_2) = \int_{\mathbb{R}^d} q_1(\mathbf{x}) \left(\ln(q_1(\mathbf{x})) - \ln(q_2(\mathbf{x})) \right) \mathrm{d}\mathbf{x}.$$

It should be noted that $D_{\text{KL}}(q_1 | \, q_2) \neq D_{\text{KL}}(q_2 | \, q_1)$ thus D_{KL} is not a proper distance measure as it is not symmetric; furthermore it does not satisfy the triangle inequality. Still, the KL divergence is very practical when optimizing the vector parameter $\mathbf{\Theta}$ of the parametric families which explains its popularity in the literature. Let $h(\cdot\,; \mathbf{\Theta})$ be the density of a specific parametric family. The optimized parameters $\mathbf{\Theta}^*$ that allows $h(\cdot\,; \mathbf{\Theta}^*)$ to be as close as possible to h_{opt} are the result of the following optimization problem

$$\begin{aligned} \mathbf{\Theta}^* &= \underset{\mathbf{\Theta} \in \mathbb{R}^n}{\text{argmin}} \; D_{\text{KL}}(h_{\text{opt}}(\mathbf{x}) | \, h(\mathbf{x}; \mathbf{\Theta})) \\ &= \underset{\mathbf{\Theta} \in \mathbb{R}^n}{\text{argmin}} \int_{\mathbb{R}^d} h_{\text{opt}}(\mathbf{x}) \left(\ln(h_{\text{opt}}(\mathbf{x})) - \ln(h(\mathbf{x}; \mathbf{\Theta})) \right) \mathrm{d}\mathbf{x}, \end{aligned}$$

as h_{opt} does not depend on $\mathbf{\Theta}$, the expression above can be simplified with

$$\mathbf{\Theta}^* = \underset{\mathbf{\Theta} \in \mathbb{R}^n}{\text{argmax}} \int_{\mathbb{R}^d} h_{\text{opt}}(\mathbf{x}) \ln(h(\mathbf{x}; \mathbf{\Theta})) \mathrm{d}\mathbf{x}. \tag{3.11}$$

It is recalled that $h_{\text{opt}} = \mathbb{I}_{D_f} f_{\mathbf{X}}/P_f$. Since P_f is independent of $\boldsymbol{\Theta}$ and $\mathbf{x}$, Eq. (3.11) becomes

$$\begin{aligned}\boldsymbol{\Theta}^* &= \underset{\boldsymbol{\Theta}\in\mathbb{R}^n}{\text{argmax}} \int_{\mathbb{R}^d} \mathbb{I}_{D_f}(\mathbf{x})\ln(h(\mathbf{x};\boldsymbol{\Theta}))f_{\mathbf{X}}(\mathbf{x})\mathrm{d}\mathbf{x} \\ &= \underset{\boldsymbol{\Theta}\in\mathbb{R}^n}{\text{argmax}}\ \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f}(\mathbf{x})\ln(h(\mathbf{x};\boldsymbol{\Theta}))\right].\end{aligned}$$

In order to compute the expected value $\mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f}(\mathbf{x})\ln(h(\mathbf{x};\boldsymbol{\Theta}))\right]$ one could simply use the Monte Carlo method. However since the indicator function of the failure domain is equal to zero most of the time, this method is not relevant. One solution is to compute the integral with the IS method employing another auxiliary density, denoted q. The optimal parameters $\boldsymbol{\Theta}^*$ are then the result of the following optimization

$$\boldsymbol{\Theta}^* \approx \underset{\boldsymbol{\Theta}\in\mathbb{R}^n}{\text{argmax}}\ \frac{1}{N}\sum_{j=1}^{N}\mathbb{I}_{D_f}\left(\mathbf{X}^{(j)}\right)\ln\left(h\left(\mathbf{X}^{(j)};\boldsymbol{\Theta}\right)\right)\frac{f_{\mathbf{X}}\left(\mathbf{X}^{(j)}\right)}{q\left(\mathbf{X}^{(j)}\right)}, \tag{3.12}$$

where the random variables $\left(\mathbf{X}^{(j)}\right)$ are iid with density q.

The logic of the CE-AIS algorithm is quite the same as the NAIS algorithm, as ρ-quantiles of the generated sample responses are also introduced as intermediate failure thresholds, where $\rho \in (0,1)$. However, the way the intermediate auxiliary densities are constructed differs. At a particular step i, the next intermediate auxiliary density built is $h(\cdot\,;\boldsymbol{\Theta}_{i+1})$ where $\boldsymbol{\Theta}_{i+1}$ are the result of the following optimization

$$\boldsymbol{\Theta}_{i+1} \approx \underset{\boldsymbol{\Theta}\in\mathbb{R}^n}{\text{argmax}}\ \frac{1}{N}\sum_{j=1}^{N}\mathbb{I}_{Y<s_i}\left(Y_i^{(j)}\right)\ln\left(h\left(\mathbf{X}_i^{(j)};\boldsymbol{\Theta}\right)\right)\frac{f_{\mathbf{X}}\left(\mathbf{X}_i^{(j)}\right)}{h\left(\mathbf{X}_i^{(j)};\boldsymbol{\Theta}_i\right)}, \tag{3.13}$$

where the random variables $\left(\mathbf{X}_i^{(j)}\right)$ are iid with density $h(\cdot\,;\boldsymbol{\Theta}_i)$ and $Y_i^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)$. This is the same equation as in Eq. (3.12) except for the failure threshold being s_i instead of zero and q being $h(\cdot\,;\boldsymbol{\Theta}_i)$, the previously optimized auxiliary density. As the cross-entropy optimization is performed for intermediate failure threshold, the term *multilevel CE* [Rubinstein and Kroese, 2004] is often used to describe the algorithm. In structural reliability [Rubinstein and Kroese, 2016], the objective functions in Eq. (3.13) are typically convex and differentiable with respect to $\boldsymbol{\Theta}$. Consequently the optimization problem can be efficiently solved by gradient-based algorithms. For distributions belonging to the so-called natural exponential family (NED) [Rubinstein and Kroese, 2016], such as the normal distribution, analytical solutions can be directly derived from Eq. (3.13). For instance, if $h_G(\cdot;\boldsymbol{\Theta})$ is a normal density, with $\boldsymbol{\Theta} = [\boldsymbol{\mu},\boldsymbol{\Sigma}]$, the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ are updated as such [Rubinstein and Kroese, 2016]

$$\boldsymbol{\mu}_{i+1} = \sum_{j=1}^{N}\mathbf{X}_i^{(j)}\bar{w}_i^{(j)} \quad \text{and} \quad \boldsymbol{\Sigma}_{i+1} = \sum_{j=1}^{N}\left(\mathbf{X}_i^{(j)}-\boldsymbol{\mu}_{i+1}\right)\left(\mathbf{X}_i^{(j)}-\boldsymbol{\mu}_{i+1}\right)^{\top}\bar{w}_i^{(j)}, \tag{3.14}$$

where the random variables $\left(\mathbf{X}_i^{(j)}\right)$ are iid with density $h_{\mathrm{G}}\left(\cdot;\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i\right)$ and $\bar{w}_i^{(j)} = w_i^{(j)}/\sum_{\ell=1}^{N} w_i^{(\ell)}$ where $w_i^{(j)} = \mathbb{I}_{Y<s_i}\left(Y_i^{(j)}\right) f_{\mathbf{X}}\left(\mathbf{X}_i^{(j)}\right)/h_{\mathrm{G}}\left(\mathbf{X}_i^{(j)};\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i\right)$ are the sample weights.

The baseline of the method is displayed in Algorithm 3. The initial parameters $\boldsymbol{\Theta}_0$ of the density are usually selected to retrieve the inputs density $f_{\mathbf{X}}$, so that the first intermediate threshold s_0 of the CE-AIS algorithm is in fact the ρ-quantile response value of a crude MC sample. It should be noted that in contrast with the NAIS algorithm, the samples generated in the CE-AIS algorithm are made of use only once.

As for the NAIS algorithm, the parameter ρ greatly influences the convergence of the algorithm and the efficiency of the method. If this parameter is set too low, then the algorithm converges rapidly and the number of iterations is small. However, the optimization of the parameters is performed with an IS estimate of small sample size (as the intermediate failure domain indicator function of their associated weight is mostly equal to zero), thus there can be a greater error in their estimation. In contrast, setting ρ very large enables a very accurate IS estimation of the optimized parameter $\boldsymbol{\Theta}$ but

Algorithm 3 CE-AIS

Require: The joint pdf $f_{\mathbf{X}}$, the quantile level ρ and the sample size N.

Initialization

Choose a distribution family $h(\cdot;\mathbf{\Theta})$ and initial parameters $\mathbf{\Theta}_0$. Generate an iid sample $\left(\mathbf{X}_0^{(j)}\right)_{j=1,\dots,N}$ from $h(\cdot;\mathbf{\Theta}_0)$. Compute the response of the system $\left(Y_0^{(j)} = g\left(\mathbf{X}_0^{(j)}\right)\right)_{j=1,\dots,N}$. Select the initial intermediate threshold s_0 as the maximum between the ρ-quantile of the response values $\left(Y_0^{(j)}\right)_{j=1,\dots,N}$ denoted $Y_0^{(\rho)}$ and zero: $s_0 = \max\left(0, Y_0^{(\rho)}\right)$. Set the index i equal to 0. Compute the new optimized parameters $\mathbf{\Theta}_1$ thanks to Eq. (3.13).

while $s_i > 0$ **do**

Optimization of the parametric density

The index i is incremented: $i = i + 1$. Generate an iid sample $\left(\mathbf{X}_i^{(j)}\right)_{j=1,\dots,N}$ from $h(\cdot;\mathbf{\Theta}_i)$. Compute the response of the system $\left(Y_i^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)\right)_{j=1,\dots,N}$. Select the initial intermediate threshold s_i as the maximum between the ρ-quantile of the response values $\left(Y_i^{(j)}\right)_{j=1,\dots,N}$ denoted $Y_i^{(\rho)}$ and zero: $s_i = \max\left(0, Y_i^{(\rho)}\right)$. If $s_i > 0$, compute the new optimized parameters $\mathbf{\Theta}_{i+1}$ thanks to Eq. (3.13).

end while

Estimation of the failure probability

Set m equal to i. The probability of failure is estimated with the following formula

$$\hat{P}_f^{\mathrm{CE-AIS}} = \frac{1}{N}\sum_{j=1}^{N} \mathbb{I}_{Y\leq 0}\left(Y_m^{(j)}\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_m^{(j)}\right)}{h\left(\mathbf{X}_m^{(j)};\mathbf{\Theta}_m\right)}$$

return $\hat{P}_f^{\mathrm{CE-AIS}}$.

slows down the convergence of the algorithm and thus the simulation budget is heavier as the number of iterations gets large as well. The authors of [Kroese et al., 2013] recommend to select ρ in $[0.01, 0.1]$.

In our specific context, the failure domain encompasses possibly several failure regions and the dimension of the system is high. Special versions of the CE-AIS algorithm were derived for such context and are briefly presented here, starting with the multiple failure regions framework.

CE-AIS for multiple failure regions

The efficiency of the CE-AIS algorithm heavily relies on the selected family density $h(\cdot;\mathbf{\Theta})$. In our specific context, selecting h as a parametric unimodal distribution is consequently not relevant. In order to account for the multiple failure regions, a multimodal distribution as in the mixture importance sampling framework is more appropriate. In mixture importance sampling [Owen, 2013], the observations are generated from the mixture distribution h_{M}, such that

$$\forall\, \mathbf{x} \in \mathbb{R}^d \qquad h_{\mathrm{M}}(\mathbf{x}) = \sum_{k=1}^{K} \pi_k h_k(\mathbf{x}), \tag{3.15}$$

where K is the number of densities in the mixture and π_k are weighting factors such that $\sum_{k=1}^{N} \pi_k = 1$. In the specific case of the CE-AIS algorithm, one typically selects all the h_k from the same family and the weighting factors of dimension K are parameters that have to be optimized as well at each step as they are unknown. Therefore, although these mixture densities are more flexible, the optimization is more complex because of the weighting factors; some approximations are usually made. The introduction of a mixture of densities does not change the unbiasedness of CE-AIS estimate.

The value of K which is the number of densities in the mixture has to be set *a priori*. The setting of the parameter K is crucial, as it highly influences the performance of the algorithm. Choosing K too small results in neglecting failure regions with less importance for the failure probability estimation and may cause bias. In contrast, since the dimension of the parameters of the mixture is proportional

to K, choosing K too large makes the optimization process computationally heavier, which may result in less accurate optimized parameters. If one already knows the number of failure regions in the failure domain, then this value would be given to K. Each weight π_k would give information about the degree of relative impact each failure region has over the whole failure probability. Unfortunately, if the limit state function is a black-box model, which is the case in our study context, then no information about the failure domain is available and the parameter K is consequently difficult to set.

The most commonly used mixture of densities is the Gaussian mixture. Therefore this particular case is derived hereafter, as an example of CE-AIS with a mixture of densities to account for the multiple failure regions.

CE-AIS with Gaussian Mixture (GM)

The Gaussian Mixture (GM) case, introduced in [Kurtz and Song, 2013], results in the following auxiliary density

$$\forall \mathbf{x} \in \mathbb{R}^d \qquad h_{\mathrm{GM}}(\mathbf{x}\,;\boldsymbol{\Theta}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}\,;\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k).$$

Each normal density has a d-vector mean $\boldsymbol{\mu}_k$ and a covariance matrix $\boldsymbol{\Sigma}_k$ (with $d(d+1)/2$ unknown components). The dimension of the parameters $\boldsymbol{\Theta} = [\pi_k, \boldsymbol{\mu_k}, \boldsymbol{\Sigma}_k,\ \text{for } k = 1, \ldots, K]$ is thus equal to $K(d(d+3)/2) + K$. Several studies [Kurtz and Song, 2013, Geyer et al., 2019] have derived updating formulas for the GM parameters. The updating formulas resulting from the gradient-based optimization of Eq. (3.13) with $h = h_{\mathrm{GM}}$, are described in detail in [Kurtz and Song, 2013]. In contrast, the authors in [Geyer et al., 2019] first underline the equivalence between solving the CE optimization problem and obtaining the Maximum Likelihood Estimates (MLE) of the parameters $\boldsymbol{\Theta}$ of the GM density [Rubinstein and Kroese, 2016]. Then, a general Expectation-Maximization (EM) updating scheme for weighted samples is obtained, as the original version of the EM algorithm is suited for unweighted samples [Dempster et al., 1977].

The selection of K is also considered in both studies. In [Kurtz and Song, 2013] a rule of thumb is presented which is to select K as superior to d and superior to the number of components for a system reliability problem. While in [Geyer et al., 2019], several clustering methods are considered to obtain a guess for the number K. These methods include the k-means algorithm [MacQueen, 1967], the k-means++ algorithm [Arthur and Vassilvitskii, 2006] and a technique derived in [Gebru et al., 2016] based on the so-called minimum message length criterion. However the density-based spatial clustering of applications with noise (DBSCAN) algorithm [Ester et al., 1996] is the one selected, with a small modification added in order to take into account the weights of the samples. These methods are not detailed here.

If the number K of Gaussian densities is appropriately set, the CE-AIS algorithm with GM is very efficient for failure domains with multiple failure regions for small to moderate dimension (up until $d = 50$). An illustration of the different steps of the algorithm is displayed in Figure 3.6.

In high-dimensional spaces, as the dimension of $\boldsymbol{\Theta}$ is proportional to d^2, the number of parameters to be optimized drastically increases, especially if K is set large as well. The heavy computational burden of CE-AIS deteriorates its performance. Furthermore, the capability of the clustering algorithms to obtain a guess of the value K, greatly deteriorates with increased dimension because of the *concentration of norm phenomenon* [Klawonn et al., 2012]. Finally, the *curse of dimensionality* affects the IS weight: an increasingly large portion of the samples have an associated IS weight close to zero. One solution is again to reduce the dimension of the IS density. To do so, several techniques have been developed like the screening method [Rubinstein and Glynn, 2009] or projection methods [Uribe et al., 2021, El Masri et al., 2021] for instance. These techniques are mostly suited for failure domains with a unique failure region and need to be adapted to mixture distributions to take into account the multiple failure regions.

In order to deal with systems of greater dimension, other mixture of densities were thus introduced for the CE-AIS algorithm. In the next paragraph we present a mixture of densities specifically tailored for the multiple failure regions of a high-dimensional normal input space, as the important ring of those inputs is taken into account.

CE-AIS with high-dimensional normal inputs

As previously mentioned, the stochastic representation of standard elliptical inputs $\mathbf{X}$ is $\mathbf{X} = R\mathbf{T}$ where R is a positive random variable and $\mathbf{T}$ is a random vector, independent of R and uniformly

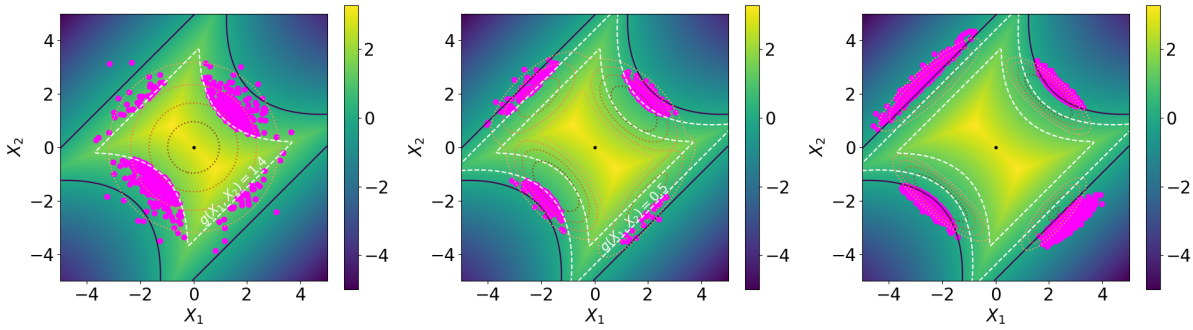


Figure 3.6: Illustration of the main steps of the CE-AIS algorithm with GM, with $K = 4$ and $\rho = 0.1$ a) the initial threshold $s_0 = 1.4$ is computed from observations generated from $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_0) = f_{\mathbf{X}}$, b) the second threshold $s_1 = 0.5$ computed from observations generated from the mixture of 4 Gaussian densities $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_1)$ c) the third ρ-quantile is already smaller than 0 thus the loop stops as $s_2 = 0$ and the failure probability is estimated with the observations in the failure domain generated from the mixture of 4 Gaussian densities $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_2)$. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dashed white lines represent the intermediate limit state surfaces. The dotted lines represent the isovalues of the densities $f_{\mathbf{X}}$, $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_1)$ and $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_2)$. The magenta dots represent the observations of non-zero weight at each step.

distributed on the unit sphere $\mathcal{T}^d = \{(t_1, \ldots, t_d) \in \mathbb{R}^d : t_1^2 + \cdots + t_d^2 = 1\}$. Using this particular writing, the optimal IS auxiliary density gives

$$\forall\, r \in \mathbb{R}^+ \qquad \forall\, \mathbf{t} \in \mathcal{T}^d \qquad h_{\mathrm{opt}}(r\mathbf{t}) = \frac{\mathbb{I}_{D_f}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{T})}{P_f}.$$

The behavior of the standard elliptical inputs in high-dimensional spaces is already underlined in Chapter 2. More precisely, an important ring which encompasses most of the probability mass of the input distribution is defined; it is particularly narrow for standard normal inputs.

Taking into account the stochastic representation of the inputs, a method derived in [Wang and Song, 2016] chooses a parametric family particularly suited for the CE-AIS algorithm in high-dimensional standard normal space. Indeed, the radial variable R is simply considered equal to the constant $\sqrt{d}$, as this value is an accurate approximation of R in high-dimensional standard normal spaces (see Chapter 2). Then, one solely focuses on finding a parametric family for the random vector $\mathbf{T}$ denoted $h_{\mathrm{T}}(\cdot\,;\boldsymbol{\Theta})$. This density is constructed as a mixture of the von Mises–Fisher (vMF) distributions [Banerjee et al., 2005], to take into account possible multiple failure regions. The advantage of the vMF distribution is that the dimension of the parameters is equal to $d+1$, instead of $d(d+3)/2$ for the Gaussian distribution; see Chapter 2 for more detail on the vMF distribution. Thus the parametric density h_{T} is written

$$\forall \mathbf{t} \in \mathcal{T}^d \qquad h_{\mathrm{T}}(\mathbf{t};\boldsymbol{\Theta}) = \sum_{k=1}^{K} \pi_k h_{\mathrm{vMF}}(\mathbf{t};\kappa_k,\boldsymbol{\nu}_k).$$

Each vMF density has a scalar concentration parameter κ_k and a mean direction d-vector $\boldsymbol{\nu}_k$. The dimension of the parameters $\boldsymbol{\Theta} = [\pi_k, \kappa_k, \boldsymbol{\nu}_k \text{ for } k = 1, \ldots, K]$ is thus equal to $K(d+1) + K$ which is considerably smaller than for a GM density.

The updating rules of the parameters, solutions of the optimization Eq. (3.13) are derived in [Wang and Song, 2016], with the additional constraints $\|\boldsymbol{\nu}_k\| = 1$ for $k = 1, \ldots, K$. However one should notice that the optimal density in the KL divergence minimization is not directly h_{opt} but the function defined on $\mathbf{t} \in \mathcal{T}^d$, such that $\mathbf{t} \to h_{\mathrm{opt},\sqrt{\mathrm{d}}}(\mathbf{t}) = \mathbb{I}_{D_f}\left(\sqrt{d}\mathbf{t}\right) f_{\mathbf{T}}(\mathbf{t})/P_{f,\sqrt{d}}$, and $P_{f,\sqrt{d}}$ is the failure probability on the hypersphere of radius $R = \sqrt{d}$. Furthermore, for each iteration, the ρ-quantile of the response values is computed as a convergence criterion, but the threshold in Eq. (3.13) is always set to 0 and not s_i (no intermediate failure thresholds). As for the value of K, a rule of thumb is presented which

is to select $K \geq d/2$. Indeed, if the actual number of failure regions is inferior to K, then the vMF densities in the mixture would merge with each other. Therefore, using a conservative large K does not introduce errors in theory, but it does increase the number of parameters to be optimized. This method is thus particularly relevant in our specific context, as it accounts for the multiple failure regions of the failure domain and for the behavior of the standard normal law in high-dimensional spaces (up to $d = 1000$). However, the setting of the radial variable R could be improved as the approximation $R \approx \sqrt{d}$ is imprecise and can be very inaccurate for moderate dimensions (for $50 \leq d \leq 200$ for instance). This topic is discussed in Section 3.3.3.4.

It should be noted that in contrast to [Wang and Song, 2016], a method derived in [Leng et al., 2022] focuses on finding an IS auxiliary density only for the radial component R of the standard normal inputs, leaving the directional density of $\mathbf{T}$ as the original uniform distribution over $\mathcal{T}^d$. A uniform distribution with bounds r_u and r_l is selected as the radial auxiliary parametric density to be optimized. However another AIS algorithm is developed in [Leng et al., 2022], that does not rely upon the KL divergence minimization to update the parameters. As no important directions are selected in this AIS process, it seems that this method could be irrelevant for some limit state surfaces that do not possess a particular spherical symmetry [Leng et al., 2022].

In the next sections, we present two algorithms derived from the CE-AIS algorithm, the CE-AMIS and the iCE-AIS, which aims at a better use of the limit state function evaluations. Indeed it is recalled here that the limit state function g is numerically expensive, thus each evaluation should be taken advantage of at the maximum (in the same spirit as in the NAIS algorithm for example, but in a parametric context).

3.3.3.3 Cross-Entropy based Adaptive Multiple Importance Sampling

The CE-AMIS algorithm derives from the CE-AIS algorithm, with the introduction of the Adaptive Multiple IS (AMIS) structure. In the AMIS [Cornuet et al., 2012, Marin et al., 2012] framework, a recycling procedure is introduced to improve the lsf evaluations profitability. As mentioned in the previous section, the introduction of intermediate failure thresholds is not mandatory for AIS algorithms; see [Kurtz and Song, 2013, Wang and Song, 2016, Leng et al., 2022]. For the CE-AMIS algorithm presented here, the convergence criteria is thus related to the theoretical coefficient of variation of the failure probability estimate rather than the ρ-quantile of the response values [Tabandeh et al., 2022]. A CV level δ is thus selected in the initialization of the algorithm.

The optimization equation of parameter $\mathbf{\Theta}_{i+1}$ at step i of the CE-AMIS algorithm is the following

$$\mathbf{\Theta}_{i+1} \approx \underset{\mathbf{\Theta}\in\mathbb{R}^n}{\operatorname{argmax}} \frac{1}{(i+1)N} \sum_{\ell=0}^{i} \sum_{j=1}^{N} \mathbb{I}_{Y<0}\left(Y_\ell^{(j)}\right) \ln\left(h\left(\mathbf{X}_\ell^{(j)};\mathbf{\Theta}\right)\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_\ell^{(j)}\right)}{\sum_{p=0}^{i} h\left(\mathbf{X}_\ell^{(j)};\mathbf{\Theta}_p\right)/(i+1)}, \quad (3.16)$$

where the random variables $\left(\mathbf{X}_\ell^{(j)}\right)_{\ell=0,\dots,i;j=1,\dots,N}$ are generated with density $\sum_{p=0}^{i} h\left(\cdot,\mathbf{\Theta}_p\right)/(i+1)$ and $Y_\ell^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)$. Thus all generated samples are taken into account in the updating formula of $\mathbf{\Theta}$, which is an idea close to the NAIS procedure. However, all the observations are assumed to be generated from the mixture $\sum_{p=0}^{i} h\left(\cdot,\mathbf{\Theta}_p\right)/(i+1)$ in a multiple IS (MIS) context, whereas the observations in the NAIS algorithm are associated to one density only. The failure probability is estimated at each step, taking advantage of all the previously generated samples as well with the MIS formula

$$\hat{P}_{f,i}^{\mathrm{CE-AMIS}} = \frac{1}{(i+1)N} \sum_{\ell=0}^{i} \sum_{j=1}^{N} \mathbb{I}_{Y<0}\left(Y_\ell^{(j)}\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_\ell^{(j)}\right)}{\sum_{p=0}^{i} h\left(\mathbf{X}_\ell^{(j)};\mathbf{\Theta}_p\right)/(i+1)}, \quad (3.17)$$

with the same notation as in Eq. (3.16). It should be noted that the IS weight of each observation $\mathbf{X}_\ell^{(j)}$ is modified at each step, since another density is added to the mixture; even for observations generated at previous iterations. This differs significantly from the NAIS algorithm. The theoretical variance is estimated along, with all the samples, with the approximation

$$\hat{\mathbb{V}}\mathrm{ar}\left(\hat{P}_{f,i}^{\mathrm{CE-AMIS}}\right) \approx \frac{1}{(i+1)N} \mathbb{V}\mathrm{ar}\left(\mathbb{I}_{Y<0}\left(Y\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}\right)}{\sum_{p=0}^{i} h\left(\mathbf{X};\mathbf{\Theta}_p\right)/(i+1)}\right), \quad (3.18)$$

Algorithm 4 CE-AMIS

Require: The joint pdf $f_{\mathbf{X}}$, the CV level δ and the sample size N.

Initialization

Choose a distribution family $h(\cdot;\boldsymbol{\Theta})$ and initial parameters $\boldsymbol{\Theta}_0$. Generate an iid sample $\left(\mathbf{X}_0^{(j)}\right)_{j=1,\dots,N}$ from $h(\cdot;\boldsymbol{\Theta}_0)$. Compute the response of the system $\left(Y_0^{(j)} = g\left(\mathbf{X}_0^{(j)}\right)\right)_{j=1,\dots,N}$. Compute the theoretical CV of the estimate, $\mathrm{CV}_0 = \sqrt{\hat{\mathbb{V}}\mathrm{ar}\left(\hat{P}_{f,0}^{\mathrm{CE-AMIS}}\right)}/\hat{P}_{f,0}^{\mathrm{CE-AMIS}}$ with Eqs. (3.17) and (3.18). Set the index i equal to 0. Compute the new optimized parameters $\boldsymbol{\Theta}_1$ thanks to Eq. (3.16).

while $\mathrm{CV}_i > \delta$ **do**

Optimization of the parametric density

The index i is incremented: $i = i + 1$. Generate an iid sample $\left(\mathbf{X}_i^{(j)}\right)_{j=1,\dots,N}$ from $h(\cdot;\boldsymbol{\Theta}_i)$. Compute the response of the system $\left(Y_i^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)\right)_{j=1,\dots,N}$. Compute the theoretical CV of the estimate, $\mathrm{CV}_i = \sqrt{\hat{\mathbb{V}}\mathrm{ar}\left(\hat{P}_{f,i}^{\mathrm{CE-AMIS}}\right)}/\hat{P}_{f,i}^{\mathrm{CE-AMIS}}$ with Eqs. (3.17) and (3.18). Compute the new optimized parameters $\boldsymbol{\Theta}_{i+1}$ thanks to Eq. (3.16).

end while

Estimation of the failure probability

Set m equal to i. The probability of failure is estimated with $\hat{P}_{f,m}^{\mathrm{CE-AMIS}}$ of Eq. (3.17).

return $\hat{P}_{f,m}^{\mathrm{CE-AMIS}}$.

assuming all the observations are iid. Algorithm 4 describes the main structure of the CE-AMIS method. As no intermediate failure threshold is introduced, the quality of the initial parameter vector $\boldsymbol{\theta}_0$ is crucial. Some observations $\mathbf{X}_0^{(j)}$ must belong to the failure domain in step $i = 0$, or else the algorithm cannot run.

Eqs. (3.16), (3.17) and (3.18) can be adapted to the case where the size of the sample N_i vary for each iteration [Owen, 2013]. Instead of dividing by $(i+1)N$, the terms are divided by $N_{\mathrm{Total}} = \sum_{\ell=0}^{i} N_\ell$ and the weights of the densities are not all equal to $1/(i+1)$ but $\alpha_p = N_p/N_{\mathrm{Total}}$. If the simulation budget has to be very small, then one could want to have higher values of N_i as the iteration goes, to sample more observations with densities further optimized. Recently, methods presented in [Sbert and Havran, 2017] and [Sbert et al., 2018] allow to compute optimal weights α_p for the multiple IS framework.

The choice of the target CV level δ greatly influences the performance of the algorithm. If the value of δ is too small, then the simulation budget will be large as the number of iterations gets large as well. On the contrary, if the value of δ is too large, then the resulting probability estimate $\hat{P}_{f,m}^{\mathrm{CE-AMIS}}$ will have a large variance and may be inaccurate as the algorithm converges rapidly. Therefore a trade-off must be studied between the simulation budget and the quality of the failure probability estimate.

At the end of the algorithm, the last density $h(\cdot\,;\boldsymbol{\Theta}_m)$ is the best density with this parametric setting to sample in the failure region. Nevertheless, the intermediate densities $h(\cdot\,;\boldsymbol{\Theta}_m)$ and their associated samples X_i are also reused in the final MIS estimate. This choice improves the stability of the estimate and speeds up the convergence process [Cornuet et al., 2012].

Statistical properties of the CE-AMIS algorithm

The statistical properties of the CE-AMIS algorithm have been studied in numerous papers [Douc et al., 2007, Cornuet et al., 2012, Marin et al., 2012] as they are particularly challenging to derive. One of the main result of these studies is that the CE-AMIS estimate, and more generally the AMIS estimate, is biased [Cornuet et al., 2012]. The sample $\left(\mathbf{X}_i^{(j)}\right)_{j=1,\dots,N}$ is dependent of all the previously generated samples, since $\boldsymbol{\Theta}_i$ is. Therefore, the samples of iterations $\ell = 0,\dots,i$ are all generated with the MIS density $\sum_{p=0}^{i} h(\cdot;\boldsymbol{\Theta}_p)/(i+1)$ but they are not independent. This intricate dependence induces a bias and $\hat{P}_{f,i}^{\mathrm{CE-AMIS}}$ is biased at each step of the algorithm, except at the initial step $i = 0$. Consequently, the variance formula Eq. (3.18) is only an approximation, assuming all the samples are

independent. The convergence properties of the AMIS algorithm are also difficult to address. The convergence in probability of the AMIS estimate under weak conditions is presented in [Cornuet et al., 2012].

In [Marin et al., 2012], a slight modification of the updating formulas is made, allowing the consistency of the resulting algorithm to be proven. Indeed, only the samples generated at iteration i are considered in the optimization of parameters $\mathbf{\Theta}_{i+1}$. However, all of the samples of the previous iterations $\ell = 0, \ldots, m-1$ are still reused to compute the final failure probability estimate $\hat{P}_{f,m}^{\text{CE-AMIS}}$.

3.3.3.4 Improved Cross Entropy-based Adaptive Importance sampling

The improved CE-AIS was presented in [Papaioannou et al., 2019] and introduces a modification in the CE-AIS which allows a better use of the limit state function evaluations $g(\mathbf{X})$, in a different manner than CE-AMIS. Indeed, in the classical CE-AIS algorithm, only ρ percent of the sample of each step contributes to the optimization of the parameters $\mathbf{\Theta}_{i+1}$ in Eq. (3.13), because of the intermediate failure domain indicator function $\mathbb{I}_{Y<s_i}$. The solution proposed by [Papaioannou et al., 2019] is to sequentially get closer to h_{opt} with a smooth approximation of the real failure domain indicator function $\mathbb{I}_{D_f}$, rather than with intermediate thresholds. The smooth approximation selected by the authors is the standard normal cdf Ψ, which verifies

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \mathbb{I}_{D_f}(\mathbf{x}) = \lim_{\sigma \to 0} \Psi\left(\frac{-g(\mathbf{x})}{\sigma}\right),$$

where $\sigma > 0$. However one should notice that other smoothing functions could be appropriate as well [Lacaze et al., 2015]. With this approximation, the optimal density h_{opt} is gradually approached with the following sequence of densities

$$\forall \mathbf{x} \in \mathbb{R}^d \quad h_i(\mathbf{x}) = \frac{\Psi\left(\frac{-g(\mathbf{x})}{\sigma_i}\right) f_{\mathbf{X}}(\mathbf{x})}{P_i}, \tag{3.19}$$

where P_i is the normalizing constant and $\sigma_i > 0$ is the smoothing parameter. It should be noted that the distribution sequence in Eq. (3.19) can be used for deriving others AIS approaches, as in [Beaurepaire et al., 2013, Dubourg and Sudret, 2014], or in the Sequential IS (SIS) framework [Papaioannou et al., 2016, Papaioannou et al., 2018] for instance. Using this density instead of h_{opt} in the computation of the KL divergence gives the following optimization equations for the parameters $\mathbf{\Theta}_{i+1}$

$$\mathbf{\Theta}_{i+1} \approx \underset{\mathbf{\Theta} \in \mathbb{R}^n}{\text{argmax}} \frac{1}{N} \sum_{j=1}^{N} \Psi\left(\frac{-Y_i^{(j)}}{\sigma_i}\right) \ln\left(h\left(\mathbf{X}_i^{(j)}; \mathbf{\Theta}\right)\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_i^{(j)}\right)}{h\left(\mathbf{X}_i^{(j)}; \mathbf{\Theta}_i\right)}. \tag{3.20}$$

With this equation, all of the limit state function evaluations $\left(Y_i^{(j)}\right)_{j=1,\ldots,N}$ are taken into account. The parameter σ_{i+1} is computed with the following optimization

$$\sigma_{i+1} = \underset{\sigma \in [0, \sigma_i]}{\text{argmin}} \left(\text{CV}(\mathbf{W}_i) - \delta\right)^2, \tag{3.21}$$

where $\left(W_i^{(j)} = \Psi\left(\frac{-Y_i^{(j)}}{\sigma_i}\right) \frac{f_{\mathbf{X}}\left(\mathbf{X}_i^{(j)}\right)}{h\left(\mathbf{X}_i^{(j)}; \mathbf{\Theta}_i\right)}\right)_{j=1,\ldots,N}$ and δ is the CV level used as the convergence criterion of the algorithm. The sequence of σ_i is thus decreasing. The structure of the iCE-AIS method is displayed in Algorithm 5.

The initial parameters $\mathbf{\Theta}_0$ of the density are selected to retrieve the inputs density $f_{\mathbf{X}}$. The stopping criterion is related to the CV of the ratio between the failure domain indicator function $\mathbb{I}_{D_f}$ and the smooth approximation $\Psi(\cdot/\sigma)$ to make sure the approximation is close enough to the indicator function. This criterion ensures that the CV of the estimate $\hat{P}_f^{\text{iCE-AIS}}$ is low; see [Papaioannou et al., 2019] for more detail.

The choice of the target CV level δ greatly influences the performance of the algorithm. It is discussed in [Papaioannou et al., 2016, Papaioannou et al., 2018]. As for the CE-AMIS algorithm, selecting a value of δ too small makes the simulation budget heavier as it leads to a larger number of iterations. In contrast, if the value of δ is too large, the smooth function $\Psi(\cdot/\sigma)$ can give an

Algorithm 5 iCE-AIS

Require: The joint pdf $f_{\mathbf{X}}$, the CV level δ and the sample size N.

Initialization

Set $\sigma_0 = +\infty$. Choose a distribution family $h(\cdot;\mathbf{\Theta})$ and initial parameters $\mathbf{\Theta}_0$. Generate an iid sample $\left(\mathbf{X}_0^{(j)}\right)_{j=1,\dots,N}$ from $h(\cdot;\mathbf{\Theta}_0)$. Compute the response of the system $\left(Y_0^{(j)} = g\left(\mathbf{X}_0^{(j)}\right)\right)_{j=1,\dots,N}$. Compute the CV of the ratio $\left(\mathbb{I}_{D_f}\left(Y_0^{(j)}\right)/\Psi\left(Y_0^{(j)}/\sigma_0\right)\right)_{j=1,\dots,N}$ denoted CV_0. Set the index i equal to 0. Compute the new smoothing parameter σ_1 with Eq. (3.21). Compute the new optimized parameters $\mathbf{\Theta}_1$ thanks to Eq. (3.20).

while $\mathrm{CV}_i > \delta$ **do**

Optimization of the parametric density

The index i is incremented: $i = i+1$. Generate an iid sample $\left(\mathbf{X}_i^{(j)}\right)_{j=1,\dots,N}$ from $h(\cdot;\mathbf{\Theta}_i)$. Compute the response of the system $\left(Y_i^{(j)} = g\left(\mathbf{X}_i^{(j)}\right)\right)_{j=1,\dots,N}$. Compute the CV of the ratio $\left(\mathbb{I}_{D_f}\left(Y_i^{(j)}\right)/\Psi\left(Y_i^{(j)}/\sigma_i\right)\right)_{j=1,\dots,N}$ denoted CV_i. Compute the new smoothing parameter σ_{i+1} with Eq. (3.21). Compute the new optimized parameters $\mathbf{\Theta}_{i+1}$ thanks to Eq. (3.20).

end while

Estimation of the failure probability

Set m equal to i. The probability of failure is estimated with the following formula

$$\hat{P}_f^{\mathrm{iCE-AIS}} = \frac{1}{N}\sum_{j=1}^{N}\mathbb{I}_{Y\leq 0}\left(Y_m^{(j)}\right)\frac{f_{\mathbf{X}}\left(\mathbf{X}_m^{(j)}\right)}{h\left(\mathbf{X}_m^{(j)};\mathbf{\Theta}_m\right)}$$

return $\hat{P}_f^{\mathrm{iCE-AIS}}$.

inaccurate approximation of the failure domain indicator function, resulting in a poor sampling scheme.

iCE-AIS for multiple failure regions with high-dimensional normal inputs

In [Papaioannou et al., 2019] a parametric family particularly suited for failure domain with multiple failure regions is derived, in the high-dimensional standard normal space. The paper follows the work of [Wang and Song, 2016] described in Section 3.3.3.2 as the stochastic representation of the standard normal inputs is taken advantage of as well. The density for the directional vector $\mathbf{T}$ is also selected as the von Mises–Fisher density, but an improvement for the radial density is made. Indeed, instead of considering the radial component R equal to $\sqrt{d}$, a Nakagami density is selected as the auxiliary radial density. The Nakagami distribution is defined by the following pdf

$$\forall r \in \mathbb{R}^+ \qquad h_{\mathrm{N}}(r;\omega,\Omega) = \frac{2\omega^\omega}{\Gamma(\omega)\Omega^\omega}r^{2\omega-1}\exp\left(-\frac{\omega}{\Omega}r^2\right),$$

where $\omega \geq 0.5$ is the shape parameter and $\Omega > 0$ is the spread parameter. This density can be viewed as a generalization of the χ distribution. For standard normal inputs $R \sim \chi(d)$, therefore this particular choice for the radial density is relevant. The density mixture, denoted vMFNM for von Mises–Fisher Nakagami mixture, is thus the following

$$\forall r \in \mathbb{R}^+ \qquad \forall \mathbf{t} \in \mathcal{T}^d \qquad h_{\mathrm{vMFNM}}(r,\mathbf{t};\mathbf{\Theta}) = \sum_{k=1}^{K}\pi_k h_{\mathrm{N}}(r;\omega_k,\Omega_k)h_{\mathrm{vMF}}(\mathbf{t};\kappa_k,\boldsymbol{\nu_k}).$$

Each Nakagami density has a scalar shape parameter ω_k and a scalar spread parameter Ω_k. Each vMF density has a scalar concentration parameter κ_k and a mean direction d-vector $\boldsymbol{\nu}_k$. The dimension of the parameter $\mathbf{\Theta} = [\pi_k,\omega_k,\Omega_k,\kappa_k,\boldsymbol{\nu}_k,\ \text{for } k=1,\dots,N]$ is thus equal to $K(d+3)+K$.

The updating rules of the parameters, solution of Eq. (3.20), are derived with a weighted Expectation Maximization (EM) algorithm, using the same equivalence as in [Geyer et al., 2019] (cf Section 3.3.3.2) between the optimization problem of the CE algorithm and the weighted MLE of the distribution parameters. The algorithm is illustrated in Figure 3.7. The selection of the value of K is

discussed as well, but the clustering algorithms are inappropriate in high-dimensional space, as already mentioned in Section 3.3.3.2. Therefore, there is still a trade-off between a reasonable dimension of the parameters $\mathbf{\Theta}$, which is linear with K, and an accurate representation of the multiple failure regions.

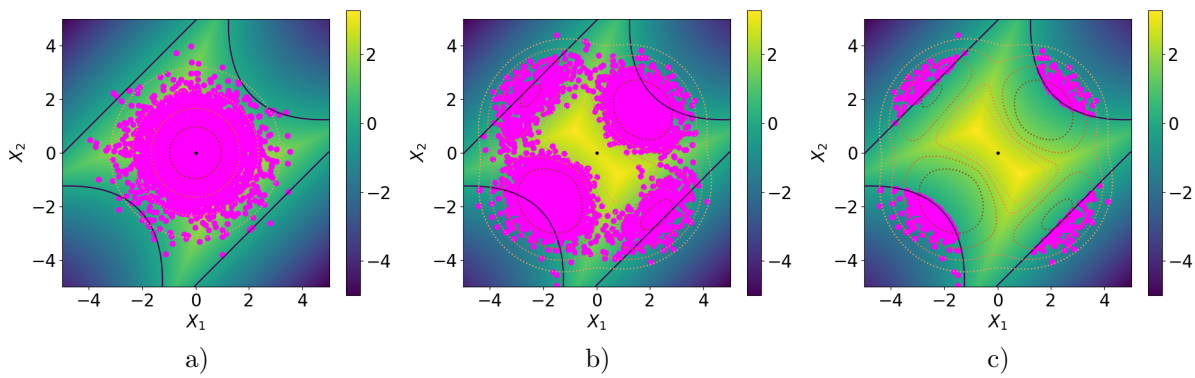


Figure 3.7: Illustration of the main steps of the iCE-vMFNM algorithm, with $K = 4$ and $\delta = 4$ a) the initial CV: $\mathrm{CV}_0 = 28.9$ is computed from observations generated from $h_{\mathrm{vMFNM}}(\cdot\,;\mathbf{\Theta}_0) = f_{\mathbf{X}}$ b) the second CV computed from the observations generated from the mixture density $h_{\mathrm{vMFNM}}(\cdot\,;\mathbf{\Theta}_1)$ is already smaller than δ thus the loop stops c) the failure probability is estimated with the observations in the failure domain generated from $h_{\mathrm{vMFNM}}(\cdot\,;\mathbf{\Theta}_1)$. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dotted lines represent the isovalues of the densities $f_{\mathbf{X}}$ and $h_{\mathrm{vMFNM}}(\cdot\,;\mathbf{\Theta}_1)$. The magenta dots represent the observations generated at each step in a) and b) and only the observations lying in the failure domain in c).

This algorithm, denoted iCE-vMFNM, is thus particularly well suited for the context of this thesis, for normal inputs. It will be used later on for comparison purposes in some numerical applications of the manuscript. We present next a sampling method which relies on the stochastic representation of the standard elliptical inputs. As it is assumed in this manuscript that the inputs of the system can be modeled with standard elliptical distributions, this sampling method is particularly interesting.

3.3.4 Directional Sampling

3.3.4.1 DS in the standard elliptical space

Directional Sampling (DS) is a sampling technique first introduced in [Ditlevsen et al., 1986] and derived for structural reliability problems in [Bjerager, 1988,Ditlevsen et al., 1990] for standard normal space, but we generalize here the concept for all standard elliptical distributions, as the idea remains the same. This particular sampling technique requires the inputs to follow a standard elliptical distribution, in order to take advantage of their stochastic representation. The inputs are thus written $\mathbf{X} = R\mathbf{T}$, where $R \sim f_R$ and $\mathbf{T} \sim \mathcal{U}\left(\mathcal{T}^d\right)$ and R and $\mathbf{T}$ are independent. Conditioning on $\mathbf{T} = \mathbf{t}$, the failure probability is written with the stochastic representation of the inputs

$$P_f = \mathbb{P}(g(R\mathbf{T}) \leq 0) = \int_{\mathcal{T}^d} \mathbb{P}(g(R\mathbf{t}) \leq 0 | \mathbf{T} = \mathbf{t}) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}\mathbf{t} = \int_{\mathcal{T}^d} \mathbb{P}(g(R\mathbf{t}) \leq 0) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}\mathbf{t}, \tag{3.22}$$

where the last equality comes from the fact that R an $\mathbf{T}$ are independent. The DS estimate of P_f is then obtain with the following MC estimate

$$\hat{P}_f^{\mathrm{DS}} = \frac{1}{N} \sum_{j=1}^{N} \mathbb{P}\left(g\left(R\mathbf{T}^{(j)}\right) \leq 0\right),$$

where the $\mathbf{T}^{(1)}, \mathbf{T}^{(2)}, \ldots$ are iid according to $\mathcal{U}(\mathcal{T}^d)$. The DS estimate is unbiased and its variance is equal to

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(\hat{P}_f^{\mathrm{DS}}\right) &= \frac{1}{N}\mathbb{V}\mathrm{ar}_{f_{\mathbf{T}}}\left(\mathbb{P}\left(g\left(R\mathbf{T}\right) \leq 0\right)\right) \\ &= \frac{1}{N}\left(\mathbb{E}_{f_{\mathbf{T}}}\left[\left(\mathbb{P}\left(g\left(R\mathbf{T}\right) \leq 0\right)\right)^2\right] - P_f^2\right).\end{aligned}$$

If the analytical evaluation of the conditional probability for each $\mathbf{T}^{(j)}$ is possible, the DS estimate has a smaller variance than the basic MC estimate [Bjerager, 1988]. More specifically, for spherical limit state surface, only one single observation is needed in order to obtain the exact failure probability. Thus this sampling technique is particularly suited for *almost spherical* limit state surfaces, and its efficiency decreases for non-spherical limit state surfaces for an increasing dimension d.

Introducing the IS technique in Eq. (3.22) leads to the following Directional IS (DIS) estimate

$$\hat{P}_f^{\mathrm{DIS}} = \frac{1}{N}\sum_{j=1}^{N}\mathbb{P}\left(g\left(R\mathbf{T}^{(j)}\right) \leq 0\right)\frac{f_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)}{h_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)},$$

where $\mathbf{T}^{(1)}, \mathbf{T}^{(2)}, \ldots$ are iid according to $h_{\mathbf{T}}$, the auxiliary DIS density. The variance of this unbiased estimate is

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(\hat{P}_f^{\mathrm{DIS}}\right) &= \frac{1}{N}\mathbb{V}\mathrm{ar}_{h_{\mathbf{T}}}\left(\mathbb{P}\left(g\left(R\mathbf{T}\right) \leq 0\right)\frac{f_{\mathbf{T}}\left(\mathbf{T}\right)}{h_{\mathbf{T}}\left(\mathbf{T}\right)}\right) \\ &= \frac{1}{N}\left(\mathbb{E}_{h_{\mathbf{T}}}\left[\left(\mathbb{P}\left(g\left(R\mathbf{T}\right) \leq 0\right)\frac{f_{\mathbf{T}}\left(\mathbf{T}\right)}{h_{\mathbf{T}}\left(\mathbf{T}\right)}\right)^2\right] - P_f^2\right).\end{aligned}$$

With an appropriate auxiliary density $h_{\mathbf{T}}$, the variance of the DIS estimate can be greatly reduced. The optimal DIS density, which leads to a variance equal to zero, is $h_{opt,\mathbf{T}}(\mathbf{t}) = \mathbb{P}(g(R\mathbf{t}) \leq 0)f_{\mathbf{T}}(\mathbf{t})/P_f$. The goal of DIS is to generate more observations in important directions, in contrast to the original distribution $\mathcal{U}\left(\mathcal{T}^d\right)$.

3.3.4.2 DS/DIS for star shaped limit state surfaces

In the specific case where the limit state surfaces are star-shaped, then the following property holds

$$\mathbb{P}\left(g\left(R\mathbf{T}^{(j)}\right) \leq 0\right) = \mathbb{P}(R \geq r_j) = 1 - F_R(r_j) = H_R(r_j), \tag{3.23}$$

where H_R is the survival function of the radial distribution (χ distribution for standard normal inputs for example) and r_j is called the root of the lsf g in the direction $\mathbf{T}^{(j)}$. The root r_j is the smallest value of R such that $g\left(r_j\mathbf{T}^{(j)}\right) = 0$ and it is thus assumed that $g\left(r\mathbf{T}^{(j)}\right) \leq 0$ for all $r \in [r_j, +\infty)$. Consequently, in order to compute the conditional probability, one only has to find the root r_j for each direction, as illustrated in Figure 3.8.

Assuming the gradient of the lsf g is available, the root of a particular direction is obtained with a gradient-based line-search method. For more complex limit state functions, i.e that possess several roots for the same direction $\mathbf{T}^{(j)}$, the formula Eq. (3.23) no longer holds. The conditional probability formula then depends on the number n_r of roots for each direction [Cheng et al., 2023].

The numerical accuracy of each root influences a lot the resulting DS/DIS estimate. If the value of r_j is underrated, then the probability $\mathbb{P}(R \geq r_j)$ is overrated and vice versa. The DS/DIS estimate is then biased. Therefore particular care must be taken when performing the line-search for each direction.

3.3.4.3 Variants of the DIS method

Although the DS/DIS framework is suited for all standard elliptical laws, it has been mainly developed for standard normal inputs only. The following methods are thus tailored in the standard normal space, but they could be generalized for other elliptical spaces.

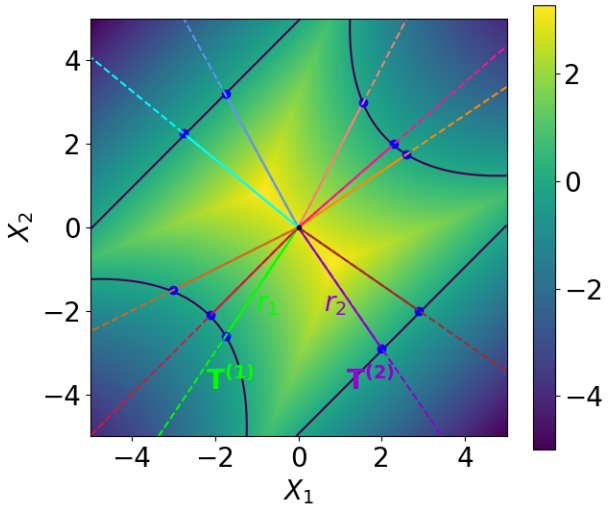


Figure 3.8: Illustration on a two-dimensional example of the DS main idea. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The colored lines represent different directions $\mathbf{T}^{(j)}$, the blue dots represent the intersection between the directions and the limit state surface. The associated roots r_j are the distance between the blue dots and the origin for each $\mathbf{T}^{(j)}$.

The DIS method has recently been combined with the subset sampling logic in [Cheng et al., 2023] to introduce a new algorithm named Sequential Directional Importance Sampling (SDIS). The failure probability is computed as a product of intermediate failure probabilities as in SS; the first one is estimated with a basic MC simulation but the other intermediate failure probabilities are estimated with DIS. Two MCMC algorithms are specifically tailored to generate the samples in this context. The SDIS algorithm is efficient for multiple failure regions in both small and moderate dimension. Its performance deteriorates with the dimension and the authors suggest to favor the SS algorithm for higher dimensions ($d \geq 100$).

The CE-AIS method presented in Section 3.3.3.2 has been adapted to the DIS framework in [Shayanfar et al., 2018] (CE-ADIS). However, it is not applicable for the reliability problems with multiple failure domains. Following this study, a recent method presented in [Zhang et al., 2022] (CE-DIS) allows to take into account the multiple failure regions, by selecting the DIS auxiliary directional density as a mixture of vMF densities, as in [Wang and Song, 2016]; see Section 3.3.3.2. The updating formulas for the parameters of the vMF mixture are then derived with the EM algorithm for weighted MLE as in [Papaioannou et al., 2019]. The DBSCAN algorithm is used to choose the value of the number of failure regions K. This CE-DIS algorithm is very efficient for small to moderate dimension (up to $d = 50$). Nevertheless, in high-dimensional standard normal spaces, finding the roots along the direction vector can become meaningless, as most of the probability mass of the radial component is gathered in the important ring. Furthermore, the clustering algorithms become inefficient and it is difficult to set the initial value of K, as already mentioned in Section 3.3.3.2. Consequently, although this method is interesting in our context, as it takes into account the stochastic representation of the elliptical inputs, the DS idea needs adaptation to high-dimensional spaces.

3.4 Simulation techniques based on the design point

In this section we present several sampling methods that rely on the design points of the system, as some may be efficient in our specific context. Therefore, we suppose that the design points $(P_k^*)_{k=1,\dots,K}$ are available with optimizations as derived in Section 3.2 for example, in the standard elliptical space. The goal of these sampling techniques is thus to take advantage of the information concerning the design points locations, to generate observations in the failure domain more easily. As in the previous

section, for each method presented, its efficiency regarding failure domains encompassing multiple failure regions is underlined, as well as its performance in high-dimensional standard elliptical spaces.

3.4.1 Radial-Based Importance Sampling

3.4.1.1 RBIS in the standard elliptical space

The Radial-Based Importance Sampling (RBIS) method is a non adaptive method first introduced by Harritz in [Harbitz, 1986] for standard normal space, but we generalize here the concept for all standard elliptical distributions, as the idea remains the same. Let β denote the norm of the closest design point written P^* in the standard elliptical space: $\beta = \min\{\|P_k^*\| \text{ for } k = 1, \dots, K\} = \|P^*\|$. The main idea of RBIS is to exclude from the sampling domain a centered hypersphere of radius β called the β-sphere, which belongs to the safety domain. Indeed, as P^* is the closest failing point of the failure domain, the centered hypersphere of radius β cannot contain any failure region and is completely safe.

Conditioning on $R \geq \beta$, the failure probability of Eq. (3.2) is written with the stochastic representation

$$\begin{aligned} P_f = \mathbb{P}(g(R\mathbf{T}) \leq 0) &= \mathbb{P}(g(R\mathbf{T}) \leq 0 | R \geq \beta)\mathbb{P}(R \geq \beta) \\ &= \mathbb{P}(g(R\mathbf{T}) \leq 0 | R \geq \beta) H_R(\beta) \\ &= H_R(\beta) \int_\beta^{+\infty} \int_{\mathcal{T}^d} \mathbb{I}_{D_f}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{t}) \, \mathrm{d}r \, \mathrm{d}\mathbf{t}, \end{aligned}$$

where H_R is the survival function of the radial distribution. This integral is the result of an IS technique with the radial auxiliary density $f_{R|R\geq\beta} = \mathbb{I}_{R\geq\beta} f_R / H_R(\beta)$. The RDIS estimate of P_f is then obtain with the following MC estimate

$$\hat{P}_f^{\text{RBIS}} = \frac{H_R(\beta)}{N} \sum_{j=1}^{N} \mathbb{I}_{D_f}\left(R^{(j)}\mathbf{T}^{(j)}\right),$$

where $\mathbf{T}^{(1)}, \mathbf{T}^{(2)}, \dots$ are iid according to $f_{\mathbf{T}} \sim \mathcal{U}^d$, and $R^{(1)}, R^{(2)}, \dots$ are iid according to $f_{R|R\geq\beta}$. In crude MC, most of the observations generated with $f_{\mathbf{T}}$ and f_R would be located in the β-sphere. Thus RBIS is very efficient compared to crude MC. Nevertheless, since the direction density $f_{\mathbf{T}}$ is left as the original uniform distribution over the hypersphere $\mathcal{T}^d$, it seems the RBIS framework is particularly suited for *almost spherical* limit state surfaces, like the DS method. An illustration of the RBIS method is displayed in Figure 3.9.

It should be noted that if the norm of P^* is incorrectly evaluated, the quality of the RBIS estimate decreases. Indeed, if the value of β is overrated, then $\hat{P}_f^{\text{RBIS}}$ could be biased with an underrated value. On the contrary, if β is underrated, the failure probability estimate should be unbiased, but the method is then less efficient as the estimate has a larger variance.

3.4.1.2 RBIS without the design point

In the particular case where the design points are unknown, an algorithm presented in [Grooteman, 2008] describes an adaptive scheme to derive an optimal radius β (ARBIS). Recently, the Multisphere-based Important Sampling method (MBIS) [Thedy and Liao, 2021] focuses on excluding an even larger part of the sampling domain. Indeed, multiple hyperspheres with various centers and radii are identified and excluded from the sampling domain. Consequently, the sampling scheme is much improved and MBIS is more appropriate for failure domains with multiple failure regions. However, the RBIS method and its variant are known to be specially designed for use in low-dimension problems [Thedy and Liao, 2021], and their performance deteriorates greatly with an increasing dimension.

3.4.2 Line Sampling

3.4.2.1 LS in the standard normal space

Line sampling was first introduced in [Hohenbichler and Rackwitz, 1988] and aims to correct the FORM/SORM approximations with sampling. Suppose the design point is unique and written P^*.

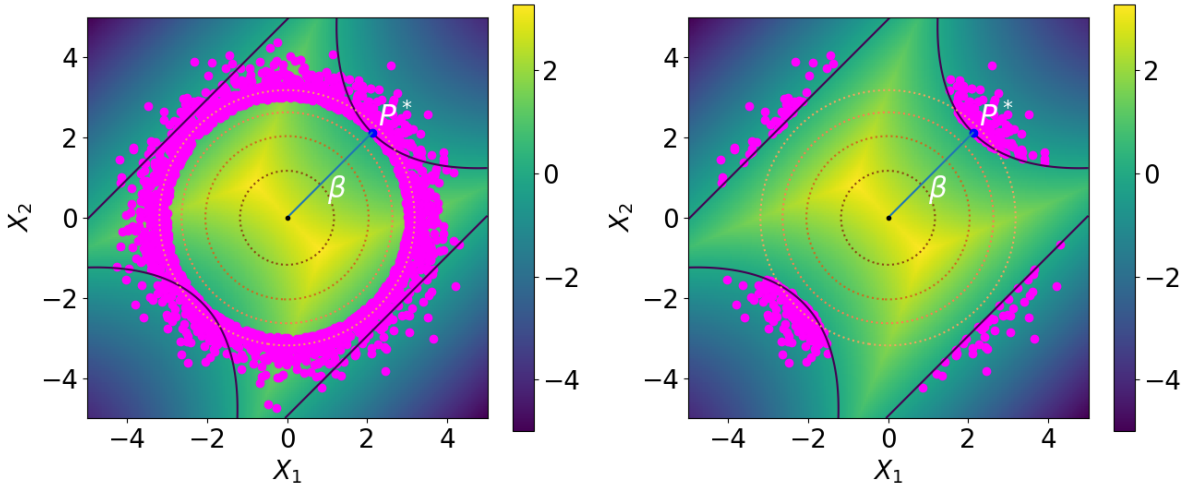


Figure 3.9: Illustration on a two-dimensional example of the RBIS main idea a) the observations are generated from $f_{R|R\geq\beta}$ where $\beta = \|P^*\|$ b) the failure probability is estimated with the observations in the failure domain generated from $f_{R|R\geq\beta}$. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dotted lines represent the isovalues of the inputs density $f_{\mathbf{X}}$. The magenta dots represent observations generated from $f_{R|R\geq\beta}$ in a) and only those belonging to the failure domain in b).

The idea is to sample on the hyperplane denoted $\mathbf{V}_1$ perpendicular to the direction pointing towards the design point, in the standard normal space $\mathbf{U}$. In order to do so, the sampling space is rotated and reduced by one dimension: $\mathbf{V} = \mathbf{R}\mathbf{U}$, with $\mathbf{V} = \left[\mathbf{V}_1^\top, V_d\right]^\top$, where $\mathbf{V}_1$ contains the $(d-1)$-dimensional subspace, V_d is a coordinate parallel to the direction of the design point P^*, and $\mathbf{R}$ is a rotation matrix. The rotation matrix $\mathbf{R}$ verifies $\mathbf{R}^\top\mathbf{R} = \mathbf{I}_d$ and its d-th row is the unit direction $P^*/\|P^*\|$. Since the standard normal distribution is invariant by orthogonal transformation [Lebrun and Dutfoy, 2009], $\mathbf{V}$ is also standard normal. The probability integral of Eq. (3.1) in then written in the rotated space as

$$P_f = \int_{\mathbb{R}^{d-1}} \int_{D_f(\mathbf{v}_1)} f_{V_d}(v_d) f_{\mathbf{V}_1}(\mathbf{v}_1) \mathrm{d}v_d \, \mathrm{d}\mathbf{v}_1, \tag{3.24}$$

where f_{V_d} is the univariate standard normal pdf, $f_{\mathbf{V}_1}$ is the $(d-1)$-variate standard normal pdf on the hyperplane $v_d = 0$ and $D_f(\mathbf{v}_1) = \left\{v_d \in \mathbb{R} |\, g\left(\mathbf{R}^\top \left[\mathbf{v}_1^\top, v_d\right]^\top\right) \leq 0\right\}$ is the one-dimensional failure domain for a set vector $\mathbf{v}_1$. The smallest value v_d which verifies $g\left(\mathbf{R}^\top \left[\mathbf{v}_1^\top, v_d\right]^\top\right) = 0$ is referred to as the root of g for the vector $\mathbf{v}_1$ and denoted $b(\mathbf{v}_1)$ and it is assumed to be unique [Papaioannou and Straub, 2021a]. The root $b(\mathbf{v}_1)$ represents the distance to the failure surface in a direction orthogonal to the hyperplane $v_d = 0$ at $\mathbf{v}_1$. An illustration is displayed in Figure 3.10.

From the uniqueness assumption, if $v_d \geq b(\mathbf{v}_1)$ then $v_d \in D_f(\mathbf{v}_1)$; therefore Eq. (3.24) is written

$$P_f = \int_{\mathbb{R}^{d-1}} \int_{b(\mathbf{v}_1)}^{+\infty} f_{V_d}(v_d) f_{\mathbf{V}_1}(\mathbf{v}_1) \mathrm{d}v_d \, \mathrm{d}\mathbf{v}_1. \tag{3.25}$$

Using the symmetry property of the standard normal distribution, Eq. (3.25) becomes

$$P_f = \int_{\mathbb{R}^{d-1}} \Psi(-b(\mathbf{v}_1)) f_{\mathbf{V}_1}(\mathbf{v}_1) \mathrm{d}\mathbf{v}_1,$$

where Ψ denotes the univariate standard normal cdf. With a MC estimation of the previous integral, the LS estimate is written

$$\hat{P}_f^{\mathrm{LS}} = \frac{1}{N} \sum_{j=1}^{N} \Psi\left(-b\left(\mathbf{V}_1^{(j)}\right)\right), \tag{3.26}$$

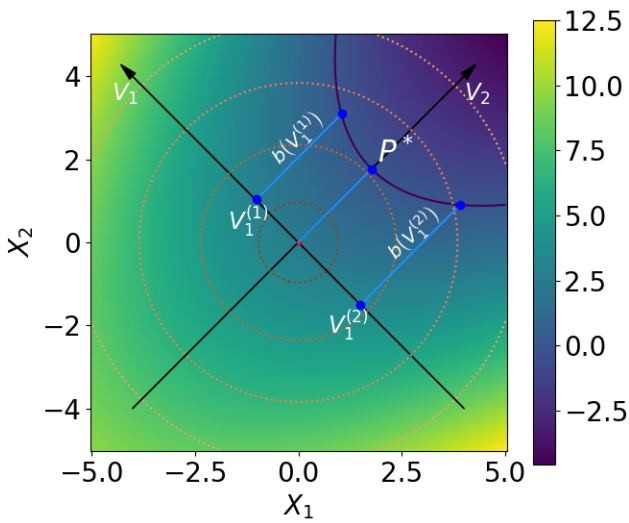


Figure 3.10: Illustration on a two-dimensional example of the LS main idea. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The black arrows are the axes of the rotated space $\mathbf{V}$. The dotted lines represent the isovalues of the standard normal inputs density $f_{\mathbf{X}}$.

where $\mathbf{V}_1^{(j)}$ are iid according to $f_{\mathbf{V}_1}$, the $(d-1)$-variate standard normal pdf on the hyperplane $v_d = 0$. Hence, for each sample $\mathbf{V}_1^{(j)}$, only the computation of the root $b\left(\mathbf{V}_1^{(j)}\right)$ is required to estimate the failure probability. This computation is typically achieved with a gradient-based line-search method. If the uniqueness of the root is not verified, then Eq. (3.25) no longer holds and the LS estimate does not have the expression as in Eq.(3.26).

If the roots are accurately computed, the LS estimate is unbiased and its variance is always smaller or equal to the one of the crude MC estimate, cf [Koutsourelakis et al., 2004]. More specifically, if the limit state surface is an hyperplane normal to the design point, only one single observation is needed in order to obtain the exact failure probability. Therefore, the LS framework allows to evaluate how much does the lsf deviate from a hyperplane normal to the design point, which is the FORM hypothesis. And if the lsf is nonlinear, the FORM/SORM estimates will be biased but not the LS estimate.

The independence of the standard normal distribution is crucial in the decomposition $\mathbf{V} = \left[\mathbf{V}_1^\top, V_d\right]^\top$ since it allows to separate the pdfs of V_d and $\mathbf{V}_1$. As stated in [Koutsourelakis et al., 2004], the LS method can be potentially applied for any d-variate distribution, as long as at least one input is independent of the others. Since all the other standard elliptical distributions are dependent, LS can hardly be generalized to other standard elliptical distributions than the normal one.

3.4.2.2 LS for multiple design points

As derived in [Schueller et al., 2004, Lu et al., 2008], line sampling can also be applied to an arbitrary number of important directions, provided that a small modification in the basic procedure is made. Each point of the space is then associated in a disjunct manner to one of the given important directions, according to an Euclidean distance criterion. As the failure regions are then not overlapping (thanks to the small modification), the failure probability estimate is thus the sum of the individual failure probability estimates. It should be noted however that the distance criterion mentioned may not be efficient in high-dimensional spaces because of the *concentration of norm phenomenon*, which states that the Euclidean norm tends to be identical for high-dimensional data [Klawonn et al., 2012].

3.4.2.3 LS for high-dimensional standard normal inputs

One of the main advantage of the LS method is that the high-dimensional problem is reduced to a conditional one-dimensional problem, since only the roots $b(\mathbf{v}_1)$ have to be computed to estimate the failure probability. Therefore, it is particularly efficient in high-dimensional spaces [Valdebenito et al., 2010]. However, its efficiency is based on the importance of the design points' directions. As mentioned in Section 3.2.3, the design points rarely belong to the standard normal important ring. Therefore, if the design points' directions are, in fact, not the main failure directions in the important ring, then the LS estimate performs quite poorly [Valdebenito et al., 2010]. A larger simulation budget is required as the selected important directions are suboptimal. In other words, the LS method with design points assumes that the directions of the design points stay relevant inside the important ring, but this hypothesis can be wrong for nonlinear lsf. If there exists no important direction at all, then the LS framework may not be appropriate [Pradlwarter et al., 2007]. Even if the directions of the design points are suboptimal, the LS estimate remains unbiased, provided the roots are accurately computed.

3.4.2.4 LS without the design point

The line sampling method without the design point information has been derived in [Koutsourelakis et al., 2004], where a Markov chain-based technique allows to find the important direction for example. Other simulation methods have been developed since. The Advanced Line Sampling (ALS) method [de Angelis et al., 2015], derives an adaptive simulation scheme to improve a poor initial choice of the important direction. More recently, the Combination Line Sampling (CLS) framework [Papaioannou and Straub, 2021a] allows to generalize the ALS method and introduce the CLS estimate, which is a weighted combination of LS estimates. The weights are then optimized to enhance the performance of the LS method. Both ALS and CLS are efficient in small to moderate dimensions, however as the dimension increases, the chances of finding a better important direction throughout simulations decrease. Thus these simulation methods are not appropriate for high-dimensional spaces. Furthermore, they need to be adapted to the case where multiple important directions exist.

If the gradient of the lsf is available, on can simply select the important direction as the negative of the gradient of g at the origin of the standard normal space [Valdebenito et al., 2018], or at another chosen point [Pradlwarter et al., 2005]. However, this choice supposes that there is only one important direction, and that the gradient is somewhat constant which may be inaccurate.

3.4.3 Importance Sampling using design points

3.4.3.1 IS using design points in the standard elliptical space

Using the positions of the design points to build the auxiliary density of the IS estimate was first derived in [Schuëller and Stix, 1987] in the standard normal space. In this non adaptive method, for each design point P_k^*, an auxiliary density is built as an independent normal density centered on the design point location. Since it is difficult to give some criteria to select proper variance values for the covariance matrices, unit variances are usually chosen [Fujita and Rackwitz, 1988]. This method can be generalized to the other standard elliptical inputs distributed as $\mathcal{E}(\mathbf{0}, \mathbf{I}_d, h_{\mathbf{X}})$, by selecting for each design point P_k^* the auxiliary density $\mathcal{E}(P_k^*, \mathbf{I}_d, h_{\mathbf{X}})$ centered on P_k^*. The IS density is then constructed as a mixture of those auxiliary densities, where the weight π_k of each density $\mathcal{E}(P_k^*, \mathbf{I}_d, h_{\mathbf{X}})$ reflects the contribution of the failure region associated to P_k^* to the failure probability; i.e the design point P_k^* with small $\beta_k = \|P_k^*\|$ value is more significant in the failure probability estimation than other design points with larger β values. For example, the ratio

$$\pi_k = \frac{F_{X_1}(-\beta_k)}{\sum_{j=1}^{K} F_{X_1}(-\beta_j)}$$

allows to take into account this criterion, where F_{X_1} is the cdf of the standard elliptical density. The IS density is thus written

$$\forall\, \mathbf{x} \in \mathbb{R}^d \qquad h(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{E}(P_k^*, \mathbf{I}_d, h_{\mathbf{X}}). \tag{3.27}$$

In contrast to the FORM/SORM approximations, where the positions of the design points P_k^* have to be determined with great accuracy, this method is not sensitive to the exact position since observations are sampled in the vicinity of P_k^* [Schueller et al., 2004]. The resulting estimate is unbiased. An illustration of this density in the standard normal case is displayed in Figure 3.11.

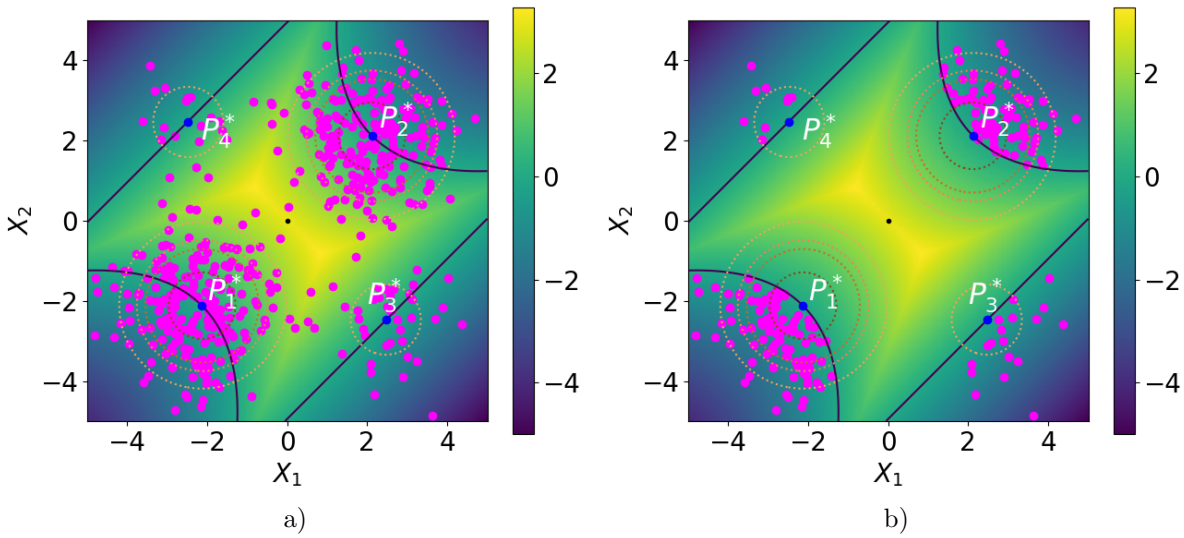


Figure 3.11: Illustration on a two-dimensional example of the IS auxiliary mixture built with design points a) the observations are generated from h of Eq. (3.27), b) the failure probability is estimated with the observations in the failure domain generated from h. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dotted lines represent the isovalues of the standard normal mixture density h. The magenta dots represent observations generated from h in a) and only those in the failure domain in b).

If the dimension d increases, the important ring of the standard normal inputs is particularly narrow. Therefore the behavior of the IS mixture must be studied for this particular elliptical distribution.

3.4.3.2 IS using design points in the high-dimensional standard normal space

IS with a Gaussian mixture built as Eq. (3.27), with $h_{\mathbf{X}}$ the normal generator density, is still relevant provided the variances are all equal to one [Au and Beck, 2003]. In fact, it is proven in [Au and Beck, 2003] that the Gaussian mixture IS density stays pertinent as long as the covariance matrices of each auxiliary density k is the same than the one of the inputs. The resulting IS estimate is then empirically unbiased and its variance is smaller than the one of a crude Monte Carlo, for dimension up to $d = 1000$, as long as the lsf is not strongly nonlinear. In the strongly nonlinear case however, the method is not as efficient [Katafygiotis and Zuev, 2008]. The main reason is because the positions of the design points do not necessarily reflect the locations of the failure regions in the important ring, as already discussed in Section 3.4.2.3 for the LS method. Therefore, if the limit state function is a black-box model, which is the context of this manuscript, the efficiency of this IS estimate is not certain.

3.4.3.3 IS using random failing points in the high-dimensional standard normal space

If the Gaussian mixture is built with independent normal densities centered on random failing points $(P_k)_{k=1,\dots,K}$ obtained with a pre-sampling step, then the relevance of a Gaussian mixture as IS density no longer holds [Au and Beck, 2003]. Because of the pre-sampling step, these failing points P_k have high chances of belonging to the important ring, whereas the design points usually do not [Katafygiotis and Zuev, 2008]. Suppose P_1 belongs to the important ring, thus $\|P_1\| \approx \sqrt{d}$. Then, as described in [Katafygiotis and Zuev, 2008] the observations generated from the normal density centered on P_1 are all located at a distance from P_1 close to $\sqrt{d}$ and belong to the important ring of P_1. Therefore,

it is very unlikely that an observation belonging to the important ring centered on P_1, belongs as well to the important ring of the original input distribution centered on 0. Such an observation O should verify $\|O\| \approx \sqrt{d}$ and $\|O - P_1\| \approx \sqrt{d}$, which is very rare [Katafygiotis and Zuev, 2008]. Consequently, the IS weight of most of the observations will be extremely small; it is the *curse of dimensionality* previously mentioned, the IS weights degenerate making the probability estimation impracticable, no matter the linearity of the lsf.

3.5 Conclusion

In this chapter, a non-exhaustive review of failure probability estimation methods has been derived. We favored methods that perform well for the specific context of the manuscript: high-dimensional inputs following a standard elliptical distribution, and a failure domain which encompasses several failure regions. The limit state function is a black-box model, thus the number of failure regions is unknown, but its gradient is available.

The approximation methods based on the design points require a negligible simulation budget, however it is challenging to evaluate the error made on the failure probability estimation, as the approximations of the limit state function are hardly ascertainable. A scheme is presented to find all the failure regions of the failure domain, which is particularly relevant in our context. Nevertheless, the estimation error is then expected to grow with multiple design points. Furthermore, the approximation of the lsf at the design points location could be erroneous inside the important ring of the inputs, especially in the standard normal case, resulting in a poor failure probability estimate.

The simulation methods require a larger simulation budget. Nonetheless, it is then possible to assess the estimation error with the computation of their coefficient of variation. Several parametric adaptive importance sampling algorithms are particularly efficient in the special context of this manuscript. Indeed, the stochastic representation of the inputs is taken advantage of, to present an auxiliary density constructed as a mixture of densities built as a product of a density for the radial component R and a density for the directional vector $\mathbf{T}$. However, the number of densities in the mixture is difficult to set, since no information about the number of failure regions is available.

Combining the location information of the design points with the sampling methods allows then to take the advantages of both processes. However, the efficiency of the resulting methods is not guaranteed in the specific context of high-dimensional standard elliptical inputs with multiple failure regions, especially if the limit state function is nonlinear.

In order to meet this need, a new CE-AMIS algorithm has been developed during the thesis and is presented in Chapter 5. The failure regions are searched for as in the approximation methods framework. The CE-AMIS context allows a better use of the limit state function evaluations and several IS methods based on the design points inspired the construction of the IS auxiliary density, like RBIS for instance. Furthermore, this algorithm is relevant for other elliptical distributions than the normal distribution.

Finally, several simulation methods are presented in this chapter while only the FORM/SORM approximations are described. It should be noted that a large literature concerning surrogate-assisted approximation methods [Echard et al., 2011] is available but is not mentioned here. The main idea is to built a numerically inexpensive limit state function approximation with a surrogate model constructed from training points. However, the *curse of dimensionality* is known to affect such methods in high-dimensional spaces [Tabandeh et al., 2022]. Consequently, we chose not to derive surrogate-assisted methods in this manuscript.

Analyse de sensibilité locale dans un contexte d'étude de fiabilité du système

Contexte d'étude

Dans un but d'analyse de fiabilité du système, la probabilité de défaillance P_f du système est estimée, avec des méthodes décrites au Chapitre 3. Cette probabilité dépend des valeurs données aux entrées déterministes $\mathbf{s}$ du système. Ainsi, nous nous intéressons aussi dans ce manuscrit à la dérivée de cette probabilité P_f par rapport au vecteur $\mathbf{s}$. L'estimation de cette dérivée fait partie de l'analyse de sensibilité du système [Iooss and Saltelli, 2017]. Le calcul de cette dérivée se faisant à un point précis de l'espace $\mathbf{s} \in \mathbb{R}^d$, on parle alors de d'analyse de sensibilité locale [Chabridon, 2018].

Selon la nature des variables déterministes $\mathbf{s}$, cette dérivée ne s'exprime pas de la même manière et n'a pas la même utilité. Nous rappelons ici que le vecteur d'entrée aléatoire $\mathbf{X}$ suit une loi elliptique standard, qui peut être le résultat d'une transformation isoprobabiliste, avec $\mathbf{Z}$ le vecteur aléatoire de départ. Le vecteur d'entrée déterministe $\mathbf{s}$ peut alors contenir les différents paramètres de distribution, notés $\boldsymbol{\theta}$, des entrées de départ $\mathbf{Z}$. Les dérivées de P_f selon $\boldsymbol{\theta}$ apportent une information essentielle quant à la dépendance de la défaillance du système par rapport au modèle probabiliste choisi. Ces dérivées permettent aussi de déterminer les variables aléatoires les plus importantes [Sues and Cesare, 2005, Wu, 1994b], dont le rôle est prépondérant dans la défaillance du système. Selon le choix de l'espace aléatoire dans lequel est mené l'analyse de sensibilité, les dérivées de P_f par rapport à $\boldsymbol{\theta}$ s'expriment alors avec une intégrale de domaine (dans l'espace de départ $\mathbf{Z}$) ou avec une intégrale surfacique (dans l'espace standard $\mathbf{X}$).

Le vecteur déterministe $\mathbf{s}$ peut aussi contenir les variables physiques du système, appelées variables de conception, notées $\boldsymbol{\delta}$. L'estimation de la dérivée de P_f par rapport à $\boldsymbol{\delta}$ est alors fondamentale en optimisation fiabiliste (Reliability-Based Design Optimization (RBDO)) [Moustapha and Sudret, 2019]. Ce domaine de recherche consiste par exemple à minimiser le coût de construction d'une structure avec des contraintes liées à la fiabilité de la structure. La dérivée de P_f par rapport à $\boldsymbol{\delta}$ s'exprime nécessairement avec une intégrale surfacique.

Dans ce chapitre nous présentons plusieurs méthodes d'estimation des dérivées de P_f par rapport au vecteur $\mathbf{s}$. Tout d'abord dans l'espace de départ $\mathbf{Z}$, lorsque $\mathbf{s} = \boldsymbol{\theta}$ uniquement, puis dans l'espace standard $\mathbf{X}$ pour tout type de variable $\mathbf{s}$. La dimension du système est toujours supposée grande $d > 50$, et il y a plusieurs régions de défaillance dans le domaine de défaillance, peu importe l'espace aléatoire. Nous supposons que l'estimateur de la probabilité de défaillance a déjà été calculé; il s'agit alors d'augmenter le moins possible le budget de simulation pour l'estimation de la dérivée.

Estimation de la dérivée de la probabilité de défaillance selon l'espace aléatoire sélectionné

Dérivation par rapport aux paramètres de distribution dans l'espace aléatoire de départ

Notons $f_{\mathbf{Z}}$ la densité du vecteur aléatoire $\mathbf{Z}$, $g_{\mathbf{Z}}$ la fonction d'état limite dans l'espace de départ et $D_{f_{\mathbf{Z}}}$ le domaine de défaillance associé. La dérivée de P_f par rapport à $\boldsymbol{\theta}$ s'écrit avec l'intégrale de domaine suivante

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{D_{f_{\mathbf{Z}}}} \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell}\,\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell}\,\mathrm{d}\mathbf{z}.$$

Afin d'estimer cette intégrale, la méthode dite de "score function" [Rubinstein, 1986] consiste à introduire une densité auxiliaire, tout comme en échantillonnage préférentiel. Si cette densité auxiliaire est $f_{\mathbf{Z}}$, la dérivée s'écrit alors

$$\begin{aligned}\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} &= \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{1}{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})} \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell} f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\,\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{\partial \ln\left(f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\right)}{\partial \theta_\ell} f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\,\mathrm{d}\mathbf{z} \\ &= \mathbb{E}_{f_{\mathbf{Z}}}\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) \frac{\partial \ln\left(f_{\mathbf{Z}}(\mathbf{Z};\boldsymbol{\theta})\right)}{\partial \theta_\ell}\right].\end{aligned}$$

Une méthode de Monte-Carlo classique permet alors d'estimer cette espérance, en utilisant les mêmes échantillons qui ont servis pour l'estimation de la probabilité de défaillance P_f. Le budget de simulation est donc constant, et l'estimateur de la dérivée est alors non biaisé. Sa variance a un taux de convergence de $1/N$.

Cette méthode de "score function" a été adaptée à d'autres densités auxiliaires plus efficace que $f_{\mathbf{Z}}$ [Wu, 1994a], pour obtenir un estimateur de variance plus faible. Elle a aussi été adaptée à la technique de "subset simulation" présentée au Chapitre 3 [Song et al., 2009].

Dérivation valable pour toutes variables déterministes dans l'espace standard elliptique

Lorsque l'analyse de sensibilité est menée dans l'espace standard elliptique, la dérivée de la probabilité P_f s'écrit nécessairement avec une intégrale surfacique, peu importe la nature du vecteur $\mathbf{s}$. En supposant que le gradient $\nabla_{\mathbf{x}} g(\mathbf{s},\mathbf{x}) \neq \mathbf{0}$ pour tous $\mathbf{x}$ et $\mathbf{s}$ sur la surface d'état limite $\{g(\mathbf{s},\mathbf{x}) = 0\}$, la dérivée de $P_f(\mathbf{s})$ par rapport à s_ℓ pour $\ell = 1,\ldots,p$ est égale à l'intégrale surfacique suivante [Breitung, 2006]

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = -\int_{g(\mathbf{s},\mathbf{x})=0} \frac{1}{\parallel \nabla_{\mathbf{x}} g(\mathbf{s},\mathbf{x}) \parallel} \frac{\partial g(\mathbf{s},\mathbf{x})}{\partial s_\ell} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}s(\mathbf{x}),$$

où $\mathrm{d}s(\mathbf{x})$ dénote l'intégration surfacique sur la surface d'état limite $\{g(\mathbf{s},\mathbf{x}) = 0\}$. Cette intégrale surfacique peut être estimée de différentes manières. Nous nous concentrons sur les techniques dérivant des méthodes d'estimation de probabilité de défaillance présentées au Chapitre 3, appropriées au contexte du manuscrit, afin de pouvoir réutiliser au maximum les évaluations de la fonction d'état limite nécessaires au calcul de P_f.

Lorsque la probabilité de défaillance est calculée avec les méthodes d'approximation, comme FORM et SORM, la fonction d'état limite est alors approchée par une version simplifiée. La dérivée de la probabilité s'obtient alors en dérivant cette version simplifiée [Ditlevsen and Madsen, 1996], et le coût de simulation supplémentaire est très faible. Néanmoins il est toujours difficile d'évaluer le biais et l'erreur de l'estimateur obtenu de la dérivée.

Dans certain cas, l'expression de la dérivée de la probabilité dans l'espace standard elliptique est toujours une intégrale de domaine, et non de surface. En effet, dans le contexte d'échantillonnage directionnel et d'échantillonnage de ligne "line sampling", la fonction indicatrice du domaine de défaillance est implicite. Ainsi la dérivation de l'intégrale de domaine P_f donne une intégrale de domaine, qui peut être estimée avec des méthodes de Monte-Carlo classiques [Papaioannou et al., 2013]. Cependant, pour chaque observation, l'évaluation de la dérivée de g par rapport à $\mathbf{s}$ ainsi que le gradient de g

par rapport à $\mathbf{x}$ sont nécessaires. En conséquence, le budget de simulation est considérablement alourdi.

Une alternative à l'intégrale surfacique peut s'obtenir en remplaçant la fonction indicatrice du domaine de défaillance par une fonction plus lisse. Cette méthode, appelée approche Faible, "Weak approach" [Torii, 2020], consiste à utiliser une fonction de répartition F_σ univariée, dérivable, paramétrée avec un scalaire σ, qui vérifie la propriété suivante

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \mathbb{I}_{D_f}(\mathbf{x}) = \mathbb{I}_{g(\mathbf{s},\cdot)\leq 0}(\mathbf{x}) = \lim_{\sigma \to 0} F_\sigma\left(-g(\mathbf{s},\mathbf{x})\right).$$

En remplaçant la fonction indicatrice par la fonction de répartition pour un $\tilde{\sigma}$ faible fixé, la dérivée de la probabilité de défaillance peut être approchée par

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} \approx -\int_{\mathbb{R}^d} \frac{\partial g(\mathbf{s},\mathbf{x})}{\partial s_\ell} f_{\tilde{\sigma}}(-g(\mathbf{s},\mathbf{x})) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x},$$

où $f_{\tilde{\sigma}}$ est la densité de probabilité de la fonction de répartition $F_{\tilde{\sigma}}$. Cette intégrale de domaine est ensuite estimée avec les méthodes de Monte-Carlo classiques. La dérivée de g par rapport à s_ℓ est donc évaluée pour chaque observation, ce qui alourdit le budget de simulation.

Puisque $\tilde{\sigma} \neq 0$, l'estimateur obtenu est forcément biaisé. Son erreur quadratique moyenne (mean square error) est alors fonction de σ. Selon la valeur de σ, deux comportements distincts sont observés. Pour les σ élevés, l'erreur de l'estimateur est principalement due à son biais au carré qui est proportionnel à σ^4. L'erreur est donc fortement croissante. Pour les σ faibles, l'erreur de l'estimateur est principalement due à sa variance qui est proportionnelle à $1/\sigma$. L'erreur est donc décroissante. La valeur de $\tilde{\sigma}$ idéale est donc celle à l'intersection de ces deux courbes, qui correspond au minimum de l'erreur. Malheureusement cette valeur n'est pas disponible en pratique puisqu'elle dépend de la dérivée qui est la quantité d'intérêt.

Enfin il est possible d'estimer la dérivée de la probabilité par différence finie [Atkinson, 1991]. Les formules de différences finies sont alors paramétrés avec un scalaire $h > 0$ et l'estimateur est biaisé. La dérivée de la fonction d'état limite g n'est alors pas nécessaire. Cependant, le budget de simulation est doublé ou triplé.

Conclusion

Plusieurs méthodes permettant d'estimer la dérivée de P_f par rapport aux variables déterministes $\mathbf{s}$ du système ont été présentées, dans le contexte du manuscrit. Pour chaque méthode, le budget de simulation supplémentaire, par rapport à celui nécessaire pour l'estimation de la probabilité de défaillance, a été souligné.

Si les variables déterministes sont des paramètres de distributions, deux options sont possibles pour estimer la dérivée. La première option consiste à utiliser la méthode de "score function" dans l'espace aléatoire de départ $\mathbf{Z}$. Cette méthode ne nécessite aucune évaluation de la fonction d'état limite supplémentaire et les estimateurs sont non biaisés. La seconde option consiste à estimer la dérivée dans l'espace standard elliptique. Dans cet espace, le nombre de méthodes disponibles pour estimer la probabilité est bien plus conséquent. Néanmoins, le budget de simulation augmente, et les estimateurs sont alors majoritairement biaisés, avec un biais difficilement contrôlable. Si les variables déterministes sont des paramètres de conception, le budget de simulation augmente obligatoirement. La dérivée est estimée dans l'espace standard elliptique et les estimateurs sont biaisés ou bien requièrent un budget de simulation significatif.

Nous proposons alors un nouvel algorithme d'estimation de dérivée de P_f par rapport à n'importe quelle variable déterministe $\mathbf{s}$, présenté au Chapitre 6. Un estimateur de la dérivée s'obtient grâce à une régression polynomiale hétéroscédastique, dans l'espace standard elliptique. Cette nouvelle approche nécessite seulement l'évaluation de la dérivée de la fonction d'état limite g par rapport à $\mathbf{s}$ en plus du budget de simulation nécessaire pour estimer P_f et s'inspire de l'approche Faible. Cependant, le biais de l'estimateur est alors contrôlable avec les paramètres de la régression polynomiale.

Chapter 4

Local reliability-based sensitivity analysis

Contents

4.1 Introduction **76**
4.1.1 Brief presentation of the reliability-based sensitivity analysis 76
4.1.1.1 Sensitivity analysis of the system 76
4.1.1.2 Introduction of the reliability-based sensitivity analysis of the system 76
4.1.1.3 RSA of the system with the quantity of interest being the failure probability 76
4.1.2 Local RSA with respect to deterministic inputs 77
4.1.2.1 Expression of the local RSA with respect to distribution parameters 77
4.1.2.2 Expression of the local RSA with respect to design parameters . . 78
4.1.2.3 Estimation of the local sensitivities 78
4.2 Local RSA in the original input space **79**
4.2.1 Score function method for sensitivity analysis 79
4.2.2 Statistical properties of the score function estimate with crude MC 79
4.2.3 Score function method with advanced sampling techniques 80
4.2.3.1 Score function method with subset sampling 80
4.2.3.2 Score function method with importance sampling 81
4.3 Local RSA in the standard elliptical space **82**
4.3.1 Sensitivity through the design point of the system 82
4.3.1.1 Expression in the FORM context 82
4.3.1.2 Sensitivity with multiple design points in the standard normal space 83
4.3.1.3 Sensitivity with other approximation methods inspired by FORM 84
4.3.2 Sensitivity with direct simulation method 84
4.3.2.1 Sensitivity estimation with directional sampling 84
4.3.2.2 Sensitivity estimation with line sampling 85
4.3.3 Sensitivity through an approximation of the failure domain indicator function 86
4.3.3.1 Weak approach: approximation of the indicator function 87
4.3.3.2 Statistical properties of the Weak approach with crude MC 88
4.3.3.3 Weak approach with advanced sampling techniques 89
4.3.4 Sensitivity with finite difference schemes 90
4.3.4.1 Expression of the finite difference schemes 90
4.3.4.2 Statistical properties of the finite difference scheme estimates . . . 90
4.4 Conclusion **91**

4.1 Introduction

The estimation of the failure probability P_f gives crucial information about the reliability of the system but it greatly depends on the settings of the different inputs. We recall here that the inputs of the system are divided into two categories: deterministic inputs $\mathbf{s} \in \mathbb{R}^p$ and random inputs $\mathbf{X} \in \mathbb{R}^d$. The study of the influences of the inputs on P_f is the purpose of the Reliability-based Sensitivity Analysis (RSA) of the system.

4.1.1 Brief presentation of the reliability-based sensitivity analysis

4.1.1.1 Sensitivity analysis of the system

Reliability-based sensitivity analysis is part of the sensitivity analysis of the system [Chabridon, 2018]. The sensitivity analysis (SA) consists in understanding how the uncertainty of the model random inputs $\mathbf{X}$ influences the uncertainty of the model output [Saltelli et al., 2004, Iooss and Saltelli, 2017]. Thus, a large literature focuses on studying the dependence of the output $Y = \mathcal{M}(\mathbf{s}, \mathbf{X})$ on each random input or group of random inputs.

Two different SA are then defined [Shekhar and Xiong, 2007]: local SA and global SA. In local SA, one investigates the local impact of a small variation of the inputs $\mathbf{X}$ on the output Y. For instance, the value of one input can be slightly changed while the others inputs remain constant. Local SA typically requires the evaluation of gradients or partial derivatives. In global SA, one investigates the impact of large variations of possibly several inputs simultaneously, on the output Y. Therefore both individual random inputs and groups of random inputs are considered, and the variations cover the inputs' entire domain of definition. Several importance measures for global SA have been defined, such as the variance-based Sobol' indices [Sobol', 1990] for example.

4.1.1.2 Introduction of the reliability-based sensitivity analysis of the system

In reliability-based SA, instead of directly studying the output $Y = \mathcal{M}(\mathbf{s}, \mathbf{X})$, one focuses on two other quantities of interest, depending on this output. The first one is the state of the system (with the value of the failure domain indicator function $\mathbb{I}_{g(\mathbf{s},\mathbf{X})\leq 0}$) and the second one is the failure probability P_f of the system. One is thus interested in the influence of the inputs on one of these quantities, which allow to describe the reliability of the system. If the quantity of interest is the state of the system, global RSA can be conducted with respect to the random inputs $\mathbf{X}$ by adapting Sobol' indices to the failure domain indicator function [Saltelli et al., 2008, Perrin and Defaux, 2019] for instance. If the quantity of interest is the failure probability P_f of the system, both local and global RSA can be conducted, with respect to the random variables $\mathbf{X}$ as well as the deterministic variables $\mathbf{s}$ of the system. Depending on the nature of the inputs, those local and global RSA with quantity of interest P_f are further detailed next.

4.1.1.3 RSA of the system with the quantity of interest being the failure probability

There are two categories of inputs: the random input variables $\mathbf{X}$ and the deterministic inputs $\mathbf{s}$ of the system. Furthermore, the deterministic inputs $\mathbf{s}$ of the system can be of different nature and, for the rest of this chapter, depending on their nature, a specific notation will be used. The deterministic inputs can represent physical variables like design parameters of the system, denoted $\boldsymbol{\delta} \in \mathbb{R}^n$. The deterministic inputs can also represent the different distribution parameters of the random inputs, denoted $\boldsymbol{\theta} \in \mathbb{R}^m$. Thus the deterministic input vector is written from now on $\mathbf{s} = [\boldsymbol{\theta}, \boldsymbol{\delta}]$, and $m+n = p$.

RSA with respect to deterministic inputs. The local methods consist in computing the derivatives of $P_f(\mathbf{s})$ according to $\mathbf{s}$ for a fixed value of $\mathbf{s}$. The local sensitivity of P_f with regard to the distribution parameters $\boldsymbol{\theta}$ gives valuable information about the influence of the probabilistic model selected for the random inputs on the system's failure occurrence. Whereas the local sensitivity of P_f with regard to the design parameters $\boldsymbol{\delta}$ can be of great use for reliability-based design optimization (RBDO) [Moustapha and Sudret, 2019].

In the global context, several options can be tested to extend the local derivative information. One could simply repeat the local RSA to multiple values of $\mathbf{s}_j$. Otherwise, one could measure global sensitivity through the integration of the local RSA [Wang et al., 2013]. Another solution is to adapt

global sensitivity analysis to the deterministic inputs; for instance a variance-based important measure has been developed for distribution parameters $\boldsymbol{\theta}$ in [Morio, 2011].

RSA with respect to random inputs. The local methods consist in measuring the importance of each input variables X_i on the failure probability at a particular location of the random input space, for instance in the vicinity of the design point [Ditlevsen and Madsen, 1996, Papaioannou and Straub, 2021b]. While in the global context, one is interested in importance measures that reflect the variation of the failure probability induced by the random inputs, either considered isolated or gathered in groups of different sizes. For instance, one can compute the conditional failure probability given that a subset of the random inputs is fixed [Li et al., 2019].

In this thesis, we uniquely focus on local RSA with respect to deterministic inputs, with the quantity of interest being the failure probability of the system. In the next section, the expression of these local sensitivities are presented in detail.

4.1.2 Local RSA with respect to deterministic inputs

In this manuscript, it is assumed that the random inputs $\mathbf{X}$ follow a standard elliptical distribution. However, this may be the result of an isoprobalistic transformation: let $\mathbf{Z}$ be the original inputs of the system, then $\mathbf{X} = T(\mathbf{Z})$ where T is a diffeomorphism described in Chapter 2 Section 2.2.2.2. We denote $g_{\mathbf{Z}}$ the limit state function in the original input space and g the transformed limit state function in the standard elliptical space. The joint pdf of $\mathbf{Z}$ is written $f_{\mathbf{Z}}$, while the joint pdf of $\mathbf{X}$ is written $f_{\mathbf{X}}$.

4.1.2.1 Expression of the local RSA with respect to distribution parameters

We describe here the framework of the local RSA with respect to $\boldsymbol{\theta}$, the distribution parameters of the original inputs $\mathbf{Z}$. The derivatives of P_f with regard to $\boldsymbol{\theta}$ give valuable information about the influence of the probabilistic model selected for the random inputs on the system's failure occurrence. This sensitivity also allows to identify which random inputs should be focused on, for quality control improvement programs for instance [Wu, 1994b, Sues and Cesare, 2005]. In order to compute the derivatives of $P_f(\boldsymbol{\theta})$ two options are possible.

The first option is to evaluate both the failure probability P_f and its derivatives in the original input space. The dependence of P_f on $\boldsymbol{\theta}$ is then contained in the joint distribution $f_{\mathbf{Z}}$ of the original inputs $\mathbf{Z}$. It is assumed here that $f_{\mathbf{Z}}$ is continuously differentiable with respect to $\boldsymbol{\theta}$. Furthermore, we assume that the integration domain of $\mathbf{Z}$ does not depend on $\boldsymbol{\theta}$. The failure probability is then written

$$P_f(\boldsymbol{\theta}) = \int_{D_{f_{\mathbf{Z}}}} f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\,\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\,\mathrm{d}\mathbf{z},$$

where $D_{f_{\mathbf{Z}}} = \left\{\mathbf{z} \in \mathbb{R}^d \,|\, g_{\mathbf{Z}}(\mathbf{z}) \leq 0\right\}$ is the failure domain in the original space. The derivatives of P_f with respect to θ_ℓ for $\ell \in [1,\dots,m]$ are then straightforward

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{D_{f_{\mathbf{Z}}}} \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell}\,\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell}\,\mathrm{d}\mathbf{z}. \tag{4.1}$$

The resulting expression of the derivative of P_f in Eq. (4.1) is thus a domain integral which depends on the derivatives of $f_{\mathbf{Z}}$ with respect to θ_ℓ, and is defined on $D_{f_{\mathbf{Z}}}$.

The second option is to evaluate both the failure probability P_f and its derivatives in the standard elliptical space. The dependence of P_f on $\boldsymbol{\theta}$ is then contained in the transformed limit state function g while the standardized joint distribution $f_{\mathbf{X}}$ is parameter-free. The failure probability is then written

$$P_f(\boldsymbol{\theta}) = \int_{D_f(\boldsymbol{\theta})} f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} \mathbb{I}_{D_f(\boldsymbol{\theta})}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

where $D_f(\boldsymbol{\theta}) = \left\{\mathbf{x} \in \mathbb{R}^d \,|\, g(\boldsymbol{\theta},\mathbf{x}) \leq 0\right\}$ is the failure domain in the standard space. Assuming the gradient $\nabla_{\mathbf{x}} g(\boldsymbol{\theta},\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x}$ and $\boldsymbol{\theta}$ on the limit state surface $\{g(\boldsymbol{\theta},\mathbf{x}) = 0\}$, the derivatives of

$P_f(\boldsymbol{\theta})$ with respect to θ_ℓ for $\ell \in [1, \dots, m]$ are defined by the following surface integral [Breitung, 2006]

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = -\int_{g(\boldsymbol{\theta},\mathbf{x})=0} \frac{1}{\parallel \nabla_{\mathbf{x}} g(\boldsymbol{\theta},\mathbf{x}) \parallel} \frac{\partial g(\boldsymbol{\theta},\mathbf{x})}{\partial \theta_\ell} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}s(\mathbf{x}), \tag{4.2}$$

where $\mathrm{d}s(\mathbf{x})$ stands for surface integration over the limit state surface $\{g(\boldsymbol{\theta},\mathbf{x}) = 0\}$. The resulting expression of the derivative of P_f in Eq. (4.2) is thus a surface integral which depends on the derivative of g with respect to θ_ℓ and its gradient $\nabla_{\mathbf{x}} g$.

4.1.2.2 Expression of the local RSA with respect to design parameters

Knowing the influence of the design parameters $\boldsymbol{\delta}$ on the failure of system can be of great use for reliability-based design optimization (RBDO) [Moustapha and Sudret, 2019], as previously mentioned. RBDO is a research domain useful to engineers for the design of structures under uncertainties. The formulation of the RBDO problem typically attempts to balance the cost of the structure with its reliability. Therefore, the optimization problem can be written with the minimization of an objective function describing the cost of the structure [Hilton and Feigen, 1960], under some probabilistic constraints as such

$$\boldsymbol{\delta}^* = \underset{\boldsymbol{\delta}\in\mathbb{R}^n}{\arg\min}\ c(\boldsymbol{\delta}) \quad \text{subject to} \quad \begin{cases} l_j(\boldsymbol{\delta}) \leq 0 & \text{for } j = 1, \dots, n_s, \\ P_f(\boldsymbol{\delta}) \leq \bar{P}_f, \end{cases}$$

where c is the cost function and the constraints are classified into two categories: the deterministic functions l_j, which define the feasible domain of the design parameters, and the failure probability function which has to remain lower than an upper threshold set $\bar{P}_f$, cf [Dubourg et al., 2011]. The derivatives of P_f with respect to $\boldsymbol{\delta}$ are then needed for the gradient-based algorithm used to solve this RBDO problem.

The expression of the dependence of P_f on the design parameters $\boldsymbol{\delta}$ is necessarily contained in the limit state function of the input space. Therefore, no matter the input space selected, the expression of the derivatives of P_f is a surface integral as in Eq. (4.2) where one adapts the integral to $f_{\mathbf{Z}}$ or $f_{\mathbf{X}}$. In the rest of this chapter we assume that the sensitivities with regard to design parameters are computed in the standard elliptical space of the inputs $\mathbf{X}$. The surface integral is thus given by the following equation

$$\frac{\partial P_f(\boldsymbol{\delta})}{\partial \delta_\ell} = -\int_{g(\boldsymbol{\delta},\mathbf{x})=0} \frac{1}{\parallel \nabla_{\mathbf{x}} g(\boldsymbol{\delta},\mathbf{x}) \parallel} \frac{\partial g(\boldsymbol{\delta},\mathbf{x})}{\partial \delta_\ell} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}s(\mathbf{x}), \tag{4.3}$$

with the same assumptions and notations as in Eq. (4.2), for $\ell \in [1, \dots, n]$. This sensitivity depends then necessarily on the derivative of g with respect to δ_ℓ and its gradient $\nabla_{\mathbf{x}} g$.

4.1.2.3 Estimation of the local sensitivities

The main challenge of the estimation of the derivatives with respect to $\mathbf{s}$, no matter its nature, is to increase as little as possible the simulation budget needed for the estimation of the probability of failure $P_f(\mathbf{s})$. Consequently, it is assumed that the computation of this sensitivity reuses as much as possible the evaluations of the limit state function needed for the computation of P_f. In other words, it is uncommon to estimate P_f and its derivatives in different input spaces, as the limit state functions differ.

According to the input space, the derivatives of P_f with regard to distribution parameters $\boldsymbol{\theta}$ are either domain integrals Eq. (4.1) or surface integrals Eq. (4.2). Domain integrals are much easier to handle than surface integrals. However, the literature on failure probability estimation in standardized spaces is more luxuriant than in the original input space. For instance, the methods presented in the Chapter 3 which depend on the location of the design points of the system require standardized input space (FORM/SORM approximations, radial-based importance sampling, line sampling, etc...). Therefore, both options present advantages and drawbacks.

The rest of this chapter is organized as follows. Section 4.2 describes local RSA methods with respect to the distribution parameters $\boldsymbol{\theta}$ of $\mathbf{Z}$ estimated in the original input space. Section 4.3 derives local RSA methods with respect to any deterministic inputs $\mathbf{s}$, estimated in the standard elliptical space. We recall here that we assume the dimension of the system d to be quite large, $d > 50$ and that the failure domains of both spaces encompasses several failure regions. Therefore, we focus here on methods that allow the estimation of the failure probability as well as its sensitivity in such a context.

4.2 Local RSA in the original input space

In this section, we suppose the vector $\mathbf{s}$ contains only distribution parameters $\boldsymbol{\theta}$. We denote $\mathbf{Z}$ the original random inputs of the system and $f_{\mathbf{Z}}$ its joint pdf. The dependence of P_f on $\boldsymbol{\theta}$ is then expressed in the joint pdf $f_{\mathbf{Z}}$, as previously mentioned. The joint pdf $f_{\mathbf{Z}}$ is supposed to be continuously differentiable with respect to $\boldsymbol{\theta}$. We further assume that the integration domain of $\mathbf{Z}$ does not depend on $\boldsymbol{\theta}$. The derivatives of P_f are written as a domain integral Eq. (4.1), and can be estimated with various reliability methods. We present here some of these methods, based on the score function method.

4.2.1 Score function method for sensitivity analysis

In order to estimate the derivatives of P_f, the score function method was first introduced in [Rubinstein, 1986] and later in [Wu, 1994a, Wu, 1994b]. It consists in introducing a density function in the domain integral, like in the important sampling framework; see Chapter 3 Section 3.3.3. The density is simply taken as the joint pdf of the inputs $\mathbf{Z}$, and the sensitivity is then written

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell} \,\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{1}{f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})} \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell} f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) \,\mathrm{d}\mathbf{z},$$

for $\ell = 1, \dots, m$. Noticing the following equality $\partial \ln\left(f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\right)/\partial\theta_\ell = \left(1/f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\right) \partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})/\partial\theta_\ell$ holds for all $\mathbf{z} \in \mathbb{R}^d$, the derivative is thus equal to

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{\partial \ln\left(f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})\right)}{\partial \theta_\ell} f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta}) \,\mathrm{d}\mathbf{z} = \mathbb{E}_{f_{\mathbf{Z}}}\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) \frac{\partial \ln\left(f_{\mathbf{Z}}(\mathbf{Z};\boldsymbol{\theta})\right)}{\partial \theta_\ell}\right].$$

The derivative $Q(\cdot) = \partial \ln\left(f_{\mathbf{Z}}(\cdot;\boldsymbol{\theta})\right)/\partial\theta_\ell$ is then called *score function*. In order to compute the sensitivity with regard to the distribution parameters of $\mathbf{Z}$, the following expected value must thus be evaluated: $\mathbb{E}_{f_{\mathbf{Z}}}\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) Q(\mathbf{Z})\right]$. This expected value can be estimated with a crude MC method, and the MC estimate of the derivatives of P_f is then

$$\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{MC}} = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{D_{f_{\mathbf{Z}}}}\left(\mathbf{Z}^{(j)}\right) Q\left(\mathbf{Z}^{(j)}\right) \tag{4.4}$$

where the observations $\mathbf{Z}^{(j)}$ are iid generated from $f_{\mathbf{Z}}$. Therefore, the observations generated to obtain the crude MC estimate of the failure probability $\hat{P}_f^{\mathrm{MC}}$ can be entirely reused here and the simulation budget is constant as the score function does not depend on the costly limit state function g.

4.2.2 Statistical properties of the score function estimate with crude MC

In this section, the bias, variance and mean square error of the MC sensitivity estimate are derived. As noted in [Torii and Novotny, 2021], the crude MC estimate of Eq. (4.4) is unbiased, and its variance is written

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{MC}}\right) &= \frac{1}{N} \mathbb{V}\mathrm{ar}\left(\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) Q(\mathbf{Z})\right) \\ &= \frac{1}{N}\left(\mathbb{E}_{f_{\mathbf{Z}}}\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) Q(\mathbf{Z})^2\right] - \left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}\right)^2\right),\end{aligned}$$

since $\left(\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z})\right)^2 = \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z})$ for all $\mathbf{z} \in \mathbb{R}^d$. As the MC sensitivity estimate is unbiased, the expression of its mean square error (mse) $\mathrm{E}_{\mathrm{mse}}^{\mathrm{MC}}$ is equal to its variance.

The error estimate is then easy to derive as the quantity $\mathbb{E}_{f_{\mathbf{Z}}}\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) Q(\mathbf{Z})^2\right]$ can be evaluated with a crude MC estimate using the same observations generated to compute both P_f and $\partial P_f(\boldsymbol{\theta})/\partial\theta_\ell$ with the Monte Carlo method. In the literature, several studies have focused on the expression of

the mean square error of estimates obtained with the score function method [Rubinstein and Kroese, 2016, Rubinstein and Shapiro, 1990], not necessarily in the failure probability framework. Here we derive the expression presented in [Torii and Novotny, 2021] which is fitted to our particular case of $\mathrm{E}_{\mathrm{mse}}^{\mathrm{MC}}$. The mean square error is bounded with the following inequalities

$$\frac{1}{N}\left(\frac{1}{P_f}-1\right)\left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}\right)^2 \leq \mathrm{E}_{\mathrm{mse}}^{\mathrm{MC}} \leq \frac{1}{N}\left(\sqrt{P_f}\sqrt{\mathbb{E}\left[Q\left(\mathbf{Z}\right)^4\right]}-\left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}\right)^2\right). \tag{4.5}$$

The lower bound of Eq. (4.5) is an expression depending on the failure probability and its derivatives thus it can be easily derived in practice. The upper bound depends on another quantity: $\mathbb{E}\left[Q\left(\mathbf{Z}\right)^4\right]$, in addition to P_f and the sensitivity. This quantity can have an analytical expression in very simple cases. Suppose $\mathbf{Z}$ is a univariate random variables which follows a normal distribution of mean $\mu_{\mathbf{Z}}$ and variance $\sigma_{\mathbf{Z}}^2$. We are interested in the derivative of P_f with respect to the mean value $\mu_{\mathbf{Z}}$. In this scenario, $\mathbb{E}\left[Q\left(\mathbf{Z}\right)^4\right] = 3/\sigma^4$ [Torii and Novotny, 2021] and the right bounds become

$$\mathrm{E}_{\mathrm{mse}}^{\mathrm{MC}} \leq \frac{1}{N}\left(\frac{\sqrt{3P_f}}{\sigma^2}-\left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \mu_{\mathbf{Z}}}\right)^2\right).$$

In the other cases, the analytical expression of $\mathbb{E}\left[Q\left(\mathbf{Z}\right)^4\right]$ can be more cumbersome, and the quantity is thus estimated with the Monte Carlo method, reusing the same sample once more.

From the inequality Eq. (4.5), we can conclude that the error, and therefore the variance of the estimate, is proportional to $1/N$. Consequently, the same variance convergence applies for both the estimation with the Monte Carlo method of the failure probability $\hat{P}_f^{\mathrm{MC}}$ (cf Chapter 3 Section 3.3.1) and its derivatives with respect to distribution parameters $\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{MC}}$.

4.2.3 Score function method with advanced sampling techniques

As presented in Chapter 3, various advanced sampling techniques have been developed, based on the MC method, to estimate the failure probability more efficiently. Some of these techniques can be applied to the score function method context, to estimate the expected value $\mathbb{E}_{f_{\mathbf{Z}}}\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})Q(\mathbf{Z})\right]$ in addition to P_f, making it possible to evaluate the derivatives of P_f according to the distribution parameters $\boldsymbol{\theta}$ of $\mathbf{Z}$. We derive some of these methods here.

4.2.3.1 Score function method with subset sampling

We recall here that the failure probability estimate with subset sampling has the following expression, with underlined dependence on $\boldsymbol{\theta}$

$$P_f(\boldsymbol{\theta}) = \prod_{i=0}^{m} P_i(\boldsymbol{\theta}) \quad \text{and} \quad \hat{P}_f^{\mathrm{SS}}(\boldsymbol{\theta}) = \prod_{i=0}^{m} \hat{P}_i^{\mathrm{MC}}(\boldsymbol{\theta}), \tag{4.6}$$

where $P_0(\boldsymbol{\theta}) = \mathbb{P}(F_0)$ and $P_i(\boldsymbol{\theta}) = \mathbb{P}(F_i|F_{i-1})$ for $i > 0$; see Chapter 3 Section 3.3.2 for more detail. We assume here that intermediate failure domains $F_0, F_1, \ldots, F_{m-1}$ are independent of $\boldsymbol{\theta}$ and are chosen *a priori*. As underlined in [Song et al., 2009], taking the derivatives of Eq. (4.6) with respect to θ_ℓ, for $\ell = 1, \ldots, m$ gives the following equation

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \sum_{i=0}^{m} \frac{P_f(\boldsymbol{\theta})}{P_i(\boldsymbol{\theta})} \frac{\partial P_i(\boldsymbol{\theta})}{\partial \theta_\ell}, \tag{4.7}$$

where the terms of the sum are defined with the following equations

$$\frac{\partial P_0(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{\mathbb{R}^d} \mathbb{I}_{F_0}(\mathbf{z}) \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell}\, \mathrm{d}\mathbf{z}, \tag{4.8}$$

and for $i > 0$

$$\frac{\partial P_i(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{\mathbb{R}^d} \mathbb{I}_{F_i}(\mathbf{z}) \frac{\partial f_{\mathbf{Z}|F_{i-1}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell} \, \mathrm{d}\mathbf{z} \quad \text{where} \quad \forall \mathbf{z} \in \mathbb{R}^d \quad f_{\mathbf{Z}|F_{i-1}}(\mathbf{z};\boldsymbol{\theta}) = \frac{\mathbb{I}_{F_{i-1}}(\mathbf{z}) f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\mathbb{P}(F_{i-1})}. \quad (4.9)$$

It should be noted that each $\mathbb{P}(F_{i-1})$ is dependent of $\boldsymbol{\theta}$ as well, even if the notation is not explicit. Applying the score function method to the Eqs. (4.8) and (4.9) leads to the final sensitivity estimate

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \frac{P_f(\boldsymbol{\theta})}{P_0(\boldsymbol{\theta})} \mathbb{E}_{f_{\mathbf{Z}}} \left[\mathbb{I}_{F_0}(\mathbf{Z}) Q(\mathbf{Z}) \right] + \sum_{i=1}^{m} \frac{P_f(\boldsymbol{\theta})}{P_i(\boldsymbol{\theta})} \mathbb{E}_{f_{\mathbf{Z}|F_{i-1}}} \left[\mathbb{I}_{F_i}(\mathbf{Z}) Q_i(\mathbf{Z}) \right],$$

where $Q_i = \partial \ln \left(f_{\mathbf{Z}|F_{i-1}}(\cdot\,;\boldsymbol{\theta}) \right) / \partial \theta_\ell$ for $i > 0$. Those expectations can be evaluated with crude Monte Carlo estimates, reusing the exact same samples generated during the SS procedure

$$\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{SS}} = \frac{\hat{P}_f^{\mathrm{SS}}(\boldsymbol{\theta})}{\hat{P}_0^{\mathrm{MC}}(\boldsymbol{\theta})} \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{F_0} \left(\mathbf{Z}_0^{(j)} \right) Q \left(\mathbf{Z}_0^{(j)} \right) + \sum_{i=1}^{m} \frac{\hat{P}_f^{\mathrm{SS}}(\boldsymbol{\theta})}{\hat{P}_i^{\mathrm{MC}}(\boldsymbol{\theta})} \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{F_i} \left(\mathbf{Z}_i^{(j)} \right) Q_i \left(\mathbf{Z}_i^{(j)} \right),$$

where the random variables $\mathbf{Z}_0^{(j)}$ are generated from f_Z and $\mathbf{Z}_i^{(j)}$ are generated from $f_{\mathbf{Z}}(\mathbf{s}, \cdot\,|F_{i-1})$, with $i > 0$, see [Song et al., 2009] for more details. The simulation budget is therefore constant as the score functions Q_i do not depend on the costly limit state function g.

The statistical properties of $\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{SS}}$ are quite similar as those of the failure probability estimate $\hat{P}_f^{\mathrm{SS}}$ (see Chapter 3 Section 3.3.2.3). The MC estimates of $\mathbb{E}_{f_{\mathbf{Z}}} \left[\mathbb{I}_{F_0}(\mathbf{Z}) Q(\mathbf{Z}) \right]$ and $\mathbb{E}_{f_{\mathbf{Z}|F_{i-1}}} \left[\mathbb{I}_{F_i}(\mathbf{Z}) Q_i(\mathbf{Z}) \right]$ for $i > 0$ are each unbiased. However, there is a small bias induced by the dependence between the samples generated at each step, if one uses the Metropolis algorithms for the Markov chain MC procedure. This bias affects $\hat{P}_f^{\mathrm{SS}}$ and the SS sensitivity estimate. The dependence makes the variance of the sensitivity estimate along with the mean square error hard to derive. Consequently these statistical quantities are not displayed here.

4.2.3.2 Score function method with importance sampling

The score function method is, by definition, already an importance sampling technique, since a density is added in the sensitivity integral. Nevertheless, if this density is selected as $f_{\mathbf{Z}}$, then the resulting MC estimate is not very efficient, as most of the observations will not be in the failure domain. Suppose the failure probability is estimated in the IS context with an auxiliary density h. This auxiliary density is supposed to be capable of generating more observations in the failure domain; see Chapter 3 Section 3.3.3. Using the same density h, the sensitivity of the failure probability can be estimated as

$$\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} = \int_{\mathbb{R}^d} \mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{z}) \frac{1}{h(\mathbf{z})} \frac{\partial f_{\mathbf{Z}}(\mathbf{z};\boldsymbol{\theta})}{\partial \theta_\ell} h(\mathbf{z}) \, \mathrm{d}\mathbf{z} = \mathbb{E}_h \left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) \frac{1}{h(\mathbf{Z})} \frac{\partial f_{\mathbf{Z}}(\mathbf{Z};\boldsymbol{\theta})}{\partial \theta_\ell} \right] = \mathbb{E}_h \left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) Q_h(\mathbf{Z}) \right],$$

where $Q_h(\cdot) = (1/h) \times (\partial f_{\mathbf{Z}}(\cdot;\boldsymbol{\theta}) / \partial \theta_\ell) = (f_{\mathbf{Z}}/h) \times (\partial \ln(f_{\mathbf{Z}}(\cdot;\boldsymbol{\theta})) / \partial \theta_\ell)$. The IS sensitivity estimate is then written

$$\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{IS}} = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}_{D_{f_{\mathbf{Z}}}} \left(\mathbf{Z}^{(j)} \right) Q_h \left(\mathbf{Z}^{(j)} \right),$$

where the random variables $\mathbf{Z}^{(j)}$ are iid with density h and are the same observations used for the estimations of $\hat{P}_f^{\mathrm{IS}}$. The simulation budget is constant as the score function Q_h does not depend on the costly limit state function g. The IS sensitivity estimate is then unbiased, and its variance, as well as the mean square error, is written

$$\mathbb{V}\mathrm{ar} \left(\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\mathrm{IS}} \right) = \mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}} = \frac{1}{N} \left(\mathbb{E}_h \left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z}) \, Q_h(\mathbf{Z})^2 \right] - \left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} \right)^2 \right).$$

Using the same computations as in [Torii and Novotny, 2021], it is possible to derive an upper for the $\mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}}$ as such (see Appendix B)

$$\mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}} \leq \frac{1}{N} \left(\sqrt{P_f} \sqrt{\mathbb{E}_h \left[\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})} \right)^3 Q(\mathbf{Z})^4 \right]} - \left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell} \right)^2 \right).$$

Therefore, we notice that compared to the formula Eq. (4.5) the likelihood ratio $f_{\mathbf{Z}}/h$ to the power 3 is introduced in the expected value. The convergence of $\mathrm{E}_{\mathrm{mse}}^{\mathrm{MC}}$ is thus greatly affected by the choice of the auxiliary density h.

The use of another IS auxiliary density than $f_{\mathbf{z}}$ in the score function framework was initiated by Wu in [Wu, 1994a, Wu, 1994b] where an adaptive importance sampling scheme is presented and the resulting auxiliary density h is close to $h_{\mathrm{opt}} = \mathbb{I}_{D_{f_{\mathbf{Z}}}} f_{\mathbf{Z}}/P_f$. Later in [Song et al., 2009], the authors developed a simulation method combining the SS framework and the IS technique to build the failure probability estimate, along with the sensitivity estimate with the score function method. Indeed, instead of estimating each intermediate failure probability $\mathbb{E}_{f_{\mathbf{Z}|F_{i-1}}}[\mathbb{I}_{F_i}(\mathbf{Z})]$ and intermediate sensitivity $\mathbb{E}_{f_{\mathbf{Z}|F_{i-1}}}[\mathbb{I}_{F_i}(\mathbf{Z})Q_i(\mathbf{Z})]$ of the SS framework with crude MC estimates, they employ the IS technique. The samples drawn by IS are then iid and the variance of both the final failure probability estimate and its sensitivity can be derived.

In short, local RSA in the original input space is performed without any additional evaluation of the lsf. The simulation budget is kept constant, no matter the simulation method selected. Furthermore, the sensitivity estimates are unbiased (except in the SS framework with Metropolis algorithms). In the next section, we present local RSA performed in the standard elliptical space.

4.3 Local RSA in the standard elliptical space

In this section we assume that after an isoprobabilistic transformation T, the random inputs $\mathbf{Z}$ have been transformed into standard elliptical inputs $\mathbf{X}$, such as $\mathbf{X} = T(\mathbf{Z})$; see Chapter 3 Section 3.2.2.2. We denote $g_{\mathbf{Z}}$ the limit state function in the original input space and g the transformed limit state function in the standard elliptical space. The dependence of P_f on $\mathbf{s}$ is then necessarily expressed in the transformed limit state function g. According to the nature of the parameters $\mathbf{s}$, this dependence can be further derived.

Suppose the vector $\mathbf{s}$ contains distribution parameters only, $\mathbf{s} = \boldsymbol{\theta}$. Consequently, these parameters have no influence on the limit state function in the $\mathbf{Z}$-space and the influence on the limit state function in the $\mathbf{X}$-space is actually solely contains in the transformation T, denoted $T_{\boldsymbol{\theta}}$ such as $\mathbf{x} = T_{\boldsymbol{\theta}}(\mathbf{z})$, therefore

$$\forall \mathbf{x} \in \mathbb{R}^d \quad g(\boldsymbol{\theta}, \mathbf{x}) = g_{\mathbf{Z}}(\mathbf{z}) \quad \text{where} \quad \mathbf{z} = T_{\boldsymbol{\theta}}^{-1}(\mathbf{x}),$$

with $T_{\boldsymbol{\theta}}^{-1}$ the inverse isoprobabilistic transformation with fixed $\boldsymbol{\theta}$.

Suppose the vector $\mathbf{s}$ contains design parameters only $\mathbf{s} = \boldsymbol{\delta}$. Consequently, the parameters $\boldsymbol{\delta}$ have influence on the definition of the limit state function in both spaces. The relationship between the two limit state functions g and $g_{\mathbf{Z}}$ can then be written

$$\forall \mathbf{x} \in \mathbb{R}^d \quad g(\boldsymbol{\delta}, \mathbf{x}) = g(\boldsymbol{\delta}, T(\mathbf{z})) = g_{\mathbf{Z}}(\boldsymbol{\delta}, \mathbf{z}) \quad \text{where} \quad \mathbf{z} = T^{-1}(\mathbf{x}).$$

Independently of the nature of $\mathbf{s}$, the sensitivity of P_f is expressed with a surface integral: Eq. (4.2) for $\boldsymbol{\theta}$ and Eq. (4.3) for $\boldsymbol{\delta}$. This surface integral can be estimated with various methods [Royset and Polak, 2004, Jensen et al., 2009] for RBDO purposes. However in this chapter we mainly focus on methods that directly come from frameworks employed to compute the failure probability estimates presented in Chapter 3, in the specific context of high-dimensional standard elliptical spaces with multiple failure regions.

4.3.1 Sensitivity through the design point of the system

4.3.1.1 Expression in the FORM context

In the FORM and SORM context, the limit state function is approximated with Taylor series expansions in the neighborhood of the design points in the standard elliptical space; see Chapter 3 Section 3.2.1. We derive here the sensitivities of P_f for the FORM approximation, first in a context where there is a unique design point P^*. The failure probability is estimated with the equation

$$\hat{P}_f^{\mathrm{FORM}} = 1 - F_{X_1}(\beta(\mathbf{s})),$$

where F_{X_1} is the cdf of the univariate standard elliptical variable X_1. The quantity β is equal to

$$\beta(\mathbf{s}) = \boldsymbol{\alpha}^\top P^* \quad \text{where} \quad \boldsymbol{\alpha} = -\frac{\nabla_{\mathbf{x}} g(\mathbf{s}, P^*)}{\|\nabla_{\mathbf{x}} g(\mathbf{s}, P^*)\|}.$$

Therefore, the value of β is a function of $\mathbf{s}$. The derivative of P_f according to s_ℓ in the FORM context is then simply

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\text{FORM}} = -\frac{\partial \beta(\mathbf{s})}{\partial s_\ell} f_{\mathbf{X}_1}(\beta(\mathbf{s})),$$

where f_{X_1} is the pdf of the random variable X_1 and $\partial\beta(\mathbf{s})/\partial s_\ell$ is the derivatives of β with respect to $\mathbf{s}$. After some calculus, see [Ditlevsen and Madsen, 1996], the derivative is equal to

$$\frac{\partial \beta(\mathbf{s})}{\partial s_\ell} = \frac{1}{\|\nabla_{\mathbf{x}} g(\mathbf{s}, P^*)\|} \frac{\partial g(\mathbf{s}, P^*)}{\partial s_\ell}. \tag{4.10}$$

Consequently the sensitivity of P_f with respect to s_ℓ, with $\ell \in [1, \ldots, p]$ is equal to

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\text{FORM}} = -\frac{1}{\|\nabla_{\mathbf{x}} g(\mathbf{s}, P^*)\|} \frac{\partial g(\mathbf{s}, P^*)}{\partial s_\ell} f_{\mathbf{X}_1}(\beta(\mathbf{s})), \tag{4.11}$$

with the FORM approximation context. Therefore, one needs the value of the gradient $\nabla_{\mathbf{x}} g$ evaluated in $[\mathbf{s}, P^*]$, as well as the derivative of g with respect to s_ℓ evaluated in $[\mathbf{s}, P^*]$. Apart from these two quantities, no additional evaluation of the costly lsf is required to estimate the sensitivities. Depending on the nature of the deterministic inputs $\mathbf{s}$, this derivative can be furthered detailed [Ditlevsen and Madsen, 1996]. If the vector $\mathbf{s}$ represents distribution parameters only, $\mathbf{s} = \boldsymbol{\theta}$, then

$$\frac{\partial \beta(\boldsymbol{\theta})}{\partial \theta_\ell} = \boldsymbol{\alpha}^\top \frac{\partial T(\boldsymbol{\theta}, \mathbf{z}^*)}{\partial \theta_\ell},$$

where $\mathbf{z}^* = T_{\boldsymbol{\theta}}^{-1}(P^*)$. Therefore the derivative of P_f is approximated with

$$\widehat{\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}}^{\text{FORM}} = -\boldsymbol{\alpha}^\top \frac{\partial T(\boldsymbol{\theta}, \mathbf{z}^*)}{\partial \theta_\ell} f_{\mathbf{X}_1}(\beta(\boldsymbol{\theta})).$$

If the vector $\mathbf{s}$ represents design parameters only, $\mathbf{s} = \boldsymbol{\delta}$, it follows that $\partial g(\boldsymbol{\delta}, P^*)/\partial \delta_\ell = \partial g_{\mathbf{Z}}(\boldsymbol{\delta}, \mathbf{z}^*)/\partial \delta_\ell$, where $\mathbf{z}^* = T^{-1}(P^*)$.

It should be noted that the study of the bias and the error of the sensitivity estimate is particularly challenging, in a black-box model context. As for the failure probability estimate in the FORM framework, the method requires very small simulation budget but the statistical properties of the resulting estimates are arduous to obtain. This remark is also valid when there are multiple design points. The computation of the sensitivities in this context is detailed next.

4.3.1.2 Sensitivity with multiple design points in the standard normal space

For systems with multiple design points, an accurate formula of P_f is difficult to derive; cf Chapter 3 Section 3.2.2. However, in the standard normal space, with the approximation introduced in [Hohenbichler, 1984], the derivatives can be easily derived in the FORM context. Suppose K design points have been identified $P_1^*, P_2^*, \ldots, P_K^*$; we recall here that the failure probability is approximated with the formula

$$\hat{P}_f^{\text{FORM}} = 1 - F_K(\boldsymbol{\beta}(\mathbf{s}); \mathbf{C}(\mathbf{s})),$$

where F_K is the cdf of a centered normal vector of dimension K with covariance matrix $\mathbf{C}(\mathbf{s})$ and $\boldsymbol{\beta}(\mathbf{s})$ is the vector of the norms of all the design points $(\beta_k = \|P_k^*\|)_{k=1,\ldots,K}$. The covariance matrix $\mathbf{C}(\mathbf{s})$ is defined such as $C_{i,j} = \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_j$ where $\boldsymbol{\alpha}_i = -\nabla_{\mathbf{x}} g(\mathbf{s}, P_i^*)/\|\nabla_{\mathbf{x}} g(\mathbf{s}, P_i^*)\|$. The derivative of the failure probability with respect to s_ℓ is written

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\text{FORM}} = -\sum_{k=1}^{K} \left(-\frac{\partial \beta_k(\mathbf{s})}{\partial s_\ell} \frac{\partial F_K(\boldsymbol{\beta}(\mathbf{s}); \mathbf{C}(\mathbf{s}))}{\partial x_k} + \sum_{j=1}^{k-1} \frac{\partial C_{k,j}(\mathbf{s})}{\partial s_\ell} \frac{\partial F_K(\boldsymbol{\beta}(\mathbf{s}); \mathbf{C}(\mathbf{s}))}{\partial C_{k,j}} \right),$$

where the derivative of each β_k can be derived from Eq. (4.10). The interested reader can find the derivatives of $C_{k,j}$ in [Ditlevsen and Madsen, 1996]. Another approach to compute the sensitivity of P_f in a series system is presented in [Sues and Cesare, 2005]. First, a MC sampling is performed using the various approximated lsfs built at the different design points, in a FORM/SORM context, to estimate P_f. Then, the derivatives are obtained with the score function method with IS, using, in the $\mathbf{Z}$-space, the failing observations obtained from the computation of P_f; cf Section 4.2.3.2. This special approach is restricted to distribution parameters $\boldsymbol{\theta}$ only.

4.3.1.3 Sensitivity with other approximation methods inspired by FORM

As previously mentioned in Chapter 3 Section 3.2, the FORM approximation is accurate for linear lsfs only. Sensitivities derived in a SORM context are proposed in [Yoo et al., 2014], when $\mathbf{s}$ are distribution parameters, $\mathbf{s} = \boldsymbol{\theta}$. However, the context slightly differs from the SORM presented in Chapter 3 and is named *novel* SORM. This particular context takes advantage of the radial distribution of the standard normal inputs ($R^2 \sim \chi^2(d)$, see Chapter 2, Section 2.2.1) to derive another quadratic approximation of the limit state surface. The sensitivities are then estimated with the differentiation of this novel approximation. This method is therefore only applicable for inputs transformed in standard normal spaces, and for $\mathbf{s} = \boldsymbol{\theta}$. This method is tested for systems of various dimensions (up to $d = 20$).

Although not presented in Chapter 3, other approximation methods inspired by the FORM/SORM framework have been derived in the reliability literature. These approximation methods aim to give better results than FORM and SORM in a nonlinear lsf context. We mention here the multi-hyperplane combination method (MHCM), first introduced in [Feng, 1990]. In the MHCM framework, several hyperplanes are employed to approximate the lsf at the vicinity of a design point, in the standard normal space. The sensitivities with respect to the distribution parameters $\boldsymbol{\theta}$ of the original inputs $\mathbf{Z}$ in the MHCM context are presented in [Dong et al., 2014]. Another approximation method presented in [Melchers and Ahammed, 2004] is completely independent of the design points' position, as the approximation of the lsf is a hyperplane constructed from a Monte Carlo sample. The sensitivities with respect to the distribution parameters are then computed with this hyperplane.

In a broader context, the sensitivities computed with these approximation methods generally require very few to none additional lsf evaluation (or its gradient). However, the error of the resulting sensitivity estimate is usually difficult to measure without any reference value [Yoo et al., 2014, Dong et al., 2014].

4.3.2 Sensitivity with direct simulation method

Surface integrals cannot be directly estimated with Monte Carlo methods. Nevertheless, there are special contexts in which the derivative of the domain integral P_f is still a domain integral, in the standard elliptical space. Indeed, if the failure probability is estimated in the directional sampling framework or in the line sampling framework, the failure domain indicator function is no longer explicit in the definition of the failure probability integral. Consequently, the differentiation of the probability results in a domain integral as well, which can be estimated with MC methods.

The sensitivity analysis in both sampling frameworks is detailed here, where no assumption on the nature of $\mathbf{s}$ is made. In both frameworks, the observations generated to compute the failure probability estimate can be reused to compute the failure probability sensitivity estimate. As derived in Chapter 3, line sampling and directional sampling are two methods that can be efficient in high-dimensional spaces, for multiple failure regions. Line sampling is tailored to the standard normal space while directional sampling is suited for all standard elliptical spaces.

4.3.2.1 Sensitivity estimation with directional sampling

Directional sampling is presented in Chapter 3 Section 3.3.4 and we focus here on the case where the limit state function is star-shaped. The probability of failure, with the dependency on $\mathbf{s}$ included, is written

$$P_f(\mathbf{s}) = \mathbb{P}(g(\mathbf{s}, R\mathbf{T}) \leq 0) = \int_{\mathcal{T}^d} \mathbb{P}(g(\mathbf{s}, R\mathbf{t}) \leq 0) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}t = \int_{\mathcal{T}^d} (1 - F_R(r(\mathbf{s}, \mathbf{t})) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}t,$$

where $\mathbf{t} \mapsto r(\mathbf{s}, \mathbf{t})$ is the function which assigns to every direction $\mathbf{t}$ its root, where the dependency of $\mathbf{s}$ is underlined, and f_R is the pdf of the radial component of the standard elliptical inputs $\mathbf{X}$; see

Section 3.3.4.2 of Chapter 3. The sensitivity of P_f can be simply derived with the equation

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = \int_{\mathcal{T}^d} \left(-\frac{\partial r(\mathbf{s},\mathbf{t})}{\partial s_\ell} f_R(r(\mathbf{s},\mathbf{t})) \right) f_{\mathbf{T}}(\mathbf{t}) \mathrm{d}t,$$

and the derivatives of r is equal to [Papaioannou et al., 2013]

$$\forall \mathbf{t} \in \mathcal{T}^d \qquad \frac{\partial r(\mathbf{s},\mathbf{t})}{\partial s_\ell} = -\frac{1}{\nabla_{\mathbf{x}} g\left(\mathbf{s}, r(\mathbf{s},\mathbf{t})\mathbf{t}\right)^\top \mathbf{t}} \frac{\partial g\left(\mathbf{s}, r(\mathbf{s},\mathbf{t})\mathbf{t}\right)}{\partial s_\ell}.$$

Therefore, the failure probability sensitivity can be estimated with

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\mathrm{DS}} = \frac{1}{N} \sum_{j=1}^{N} \frac{f_R\left(r\left(\mathbf{s}, \mathbf{T}^{(j)}\right)\right)}{\nabla_{\mathbf{x}} g\left(\mathbf{s}, r\left(\mathbf{s}, \mathbf{T}^{(j)}\right) \mathbf{T}^{(j)}\right)^\top \mathbf{T}^{(j)}} \frac{\partial g\left(\mathbf{s}, r\left(\mathbf{s}, \mathbf{T}^{(j)}\right) \mathbf{T}^{(j)}\right)}{\partial s_\ell}, \tag{4.12}$$

where the observations $\mathbf{T}^{(j)}$ are iid from $f_{\mathbf{T}}$. Introducing the IS framework, the sensitivity estimate in the DIS context is written

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\mathrm{DIS}} = \frac{1}{N} \sum_{j=1}^{N} \frac{f_R\left(r\left(\mathbf{s}, \mathbf{T}^{(j)}\right)\right)}{\nabla_{\mathbf{x}} g\left(\mathbf{s}, r\left(\mathbf{s}, \mathbf{T}^{(j)}\right) \mathbf{T}^{(j)}\right)^\top \mathbf{T}^{(j)}} \frac{\partial g\left(\mathbf{s}, r\left(\mathbf{s}, \mathbf{T}^{(j)}\right) \mathbf{T}^{(j)}\right)}{\partial s_\ell} \frac{f_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)}{h_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)}, \tag{4.13}$$

where the observations $\mathbf{T}^{(j)}$ are iid from $h_{\mathbf{T}}$. The estimators Eq. (4.12) and Eq. (4.13) are unbiased. Since all the observations are iid, the variance and mean square error are derived with the classical formulas. The observations are the same as the ones used to compute the failure probability estimate. However, it should be noted that the denominator of the first term of both estimates is the cosine between the sampling direction and the gradient of the lsf evaluated at the intersection between the sampling direction and the limit state surface. Consequently, an evaluation of the gradient $\nabla_{\mathbf{x}} g$ is required for each observation $\mathbf{T}^{(j)}$, in addition to the evaluation of the derivative of g with respect to s_ℓ. Therefore, the simulation budget becomes substantially heavier.

In [Song et al., 2011], the sensitivities with respect to distribution parameters only, $\mathbf{s} = \boldsymbol{\theta}$, are furthered detailed. Furthermore, a link between the root function r, and the norm β of the design point in the FORM framework is underlined. Indeed, for each observation $r\left(\mathbf{s}, \mathbf{T}^{(j)}\right) \mathbf{T}^{(j)}$ at the intersection of the sampling direction and the limit state surface, a hyperplane can be built as in the FORM context to locally approximate the limit state function. There is consequently an equivalence between the derivatives of r and those of β; cf Section 4.3.1.1. The resulting sensitivity estimate can then be seen as the mean value, with a correction factor, of all the local sensitivities computed on the approximated lsfs.

4.3.2.2 Sensitivity estimation with line sampling

Line sampling is presented in Chapter 3 Section 3.4.2 and we focus here on the case where the roots of the limit state function for the vector $\mathbf{v}_1$ are unique. Suppose the design point P^* is unique. The probability of failure with the dependency on $\mathbf{s}$ is written

$$P_f(\mathbf{s}) = \int_{\mathbb{R}^{d-1}} \Psi(-b(\mathbf{s},\mathbf{v}_1)) f_{\mathbf{V}_1}(\mathbf{v}_1) \mathrm{d}\mathbf{v}_1,$$

where $\mathbf{v}_1 \mapsto b(\mathbf{s},\mathbf{v}_1)$ is the function which assigns to every vector $\mathbf{v}_1$ its root, where the dependency of $\mathbf{s}$ is underlined, Ψ denotes the cdf of an univariate standard normal variable and $f_{\mathbf{V}_1}$ is the pdf of the $(d-1)$-variate standard normal random vector on the hyperplane $v_d = 0$; see Chapter 3 Section 3.4.2.1. The sensitivity of P_f can simply be derived with the equation

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = \int_{\mathbb{R}^{d-1}} \left(-\frac{\partial b(\mathbf{s},\mathbf{v}_1)}{\partial s_\ell} \phi\left(b(\mathbf{s},\mathbf{v}_1)\right) \right) f_{\mathbf{V}_1}(\mathbf{v}_1) \mathrm{d}\mathbf{v}_1,$$

where ϕ denotes the pdf of a univariate standard normal variable (which is symmetric), and the derivatives of b are equal to [Papaioannou et al., 2013]

$$\forall \mathbf{v}_1 \in \mathbb{R}^{d-1} \qquad \frac{\partial b(\mathbf{s}, \mathbf{v}_1)}{\partial s_\ell} = -\frac{1}{\nabla_{\mathbf{x}} g\left(\mathbf{s}, \mathbf{R}^\top \left[\mathbf{v}_1^\top, b(\mathbf{s}, \mathbf{v}_1)\right]^\top\right)^\top \boldsymbol{a}} \frac{\partial g\left(\mathbf{s}, \mathbf{R}^\top \left[\mathbf{v}_1^\top, b(\mathbf{s}, \mathbf{v}_1)\right]^\top\right)}{\partial s_\ell},$$

where $\boldsymbol{a} = P^*/\|P^*\|$. Consequently, the failure probability sensitivity can be estimated with

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\mathrm{LS}} = \frac{1}{N} \sum_{j=1}^{N} \frac{\phi\left(b\left(\mathbf{s}, \mathbf{V}_1^{(j)}\right)\right)}{\nabla_{\mathbf{x}} g\left(\mathbf{s}, \mathbf{R}^\top \left[\mathbf{V}_1^{(j)\top}, b\left(\mathbf{s}, \mathbf{V}_1^{(j)}\right)\right]^\top\right)^\top \boldsymbol{a}} \frac{\partial g\left(\mathbf{s}, \mathbf{R}^\top \left[\mathbf{V}_1^{(j)\top}, b\left(\mathbf{s}, \mathbf{V}_1^{(j)}\right)\right]^\top\right)}{\partial s_\ell}, \tag{4.14}$$

where the observations $\mathbf{V}_1^{(j)}$ are iid from $f_{\mathbf{V}_1}$. The sensitivity estimate is unbiased. Since all the observations are iid, the variance and mean square error are derived with the classical formulas; the CV is displayed in [Valdebenito et al., 2018] for the interested reader. The observations are the same as the ones used to compute the failure probability estimate. However, it should be noted that the denominator of the first term is the cosine between the direction $\boldsymbol{a}$ towards the design point and the gradient of the lsf evaluated at the intersection between the vector parallel to $\boldsymbol{a}$ at $\mathbf{V}_1^{(j)}$ and the limit state surface. Consequently, as in the DS framework, an evaluation of the gradient $\nabla_{\mathbf{x}} g$ is required for each observation $\mathbf{V}_1^{(j)}$, in addition to the evaluation of the derivative of g with respect to s_ℓ. Therefore the simulation budget also becomes substantially heavier with this sampling method.

The case of the systems with multiple design points is studied in [Lu et al., 2008], where the sensitivities with regard to distribution parameters only, $\mathbf{s} = \boldsymbol{\theta}$, are further detailed. As previously mentioned in Chapter 3 Section 3.4.2.2, a small modification in the line sampling scheme is required to obtain non overlapping failure regions. Once the failure regions are non-overlapping, the sensitivity of the failure probability with regard to a deterministic parameter is simply equal to the sum of the failure probability sensitivities associated to each design point. Here we derive the formula for any deterministic input s_ℓ, as the logic presented in [Lu et al., 2008] is the same for design parameters. Suppose there are K design points, the derivative of P_f is written

$$\widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}^{\mathrm{LS}} = \sum_{k=1}^{K} \widehat{\frac{\partial P_f(\mathbf{s})}{\partial s_\ell}}_k^{\mathrm{LS}} \tag{4.15}$$

where each term in the sum is equal to Eq. (4.14) adapted to the design point P_k^*. The resulting sensitivity estimate is still unbiased and since all the samples are independent, its variance can be computed with the classical formula.

As in the DS framework, a link between the derivation of the root function b and the derivation of the β in the FORM framework can be derived [Lu et al., 2008, Valdebenito et al., 2018, Shufang et al., 2009]. In [Valdebenito et al., 2018] a second approach is presented to compute the sensitivity with regard to distribution parameters only, $\mathbf{s} = \boldsymbol{\theta}$, in the LS framework. Indeed, in the particular case where the original inputs $\mathbf{Z}$ are independent, the expression of the sensitivity with regard to the distribution parameters is first derived in the original input space $\mathbf{Z}$ with the score function method (cf Section 4.2.1). Next, this domain integral is recast in the standard normal space $\mathbf{X}$ as the transformation T is quite simple since $\mathbf{Z}$ is already an independent vector. The LS method is then applied to the resulting domain integral in the standard normal space; see [Valdebenito et al., 2018] for more details.

4.3.3 Sensitivity through an approximation of the failure domain indicator function

As previously mentioned, surface integrals cannot be directly computed with Monte Carlo methods. However, it is possible to have a close estimate of the surface integrals Eqs. (4.2) and (4.3) with a domain integral, through the approximation of the failure domain indicator function. This approach

is called the Weak approach by Torii in [Torii, 2020] and [Torii and Novotny, 2021], as the sensitivity is evaluated in the sense of distribution. The Weak approach is presented here and the statistical properties of the sensitivity estimate are derived, for any deterministic input s_ℓ, for $\ell = 1, \dots, p$.

4.3.3.1 Weak approach: approximation of the indicator function

The failure domain indicator function makes the direct differentiation of the failure probability integral not possible. By replacing the failure domain indicator function by a smoother function, the differentiation can be performed and results in a domain integral [Papaioannou et al., 2013]. Several smooth approximations of the indicator function have been derived in the literature [Lacaze et al., 2015], which are typically cdfs of continuous univariate variables [Torii, 2020]. Here we derive the approximation chosen in [Papaioannou et al., 2013, Papaioannou et al., 2018] which comes from the following limit

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \mathbb{I}_{D_f}(\mathbf{x}) = \lim_{\sigma \to 0} \Psi\left(-\frac{g(\mathbf{s}, \mathbf{x})}{\sigma}\right),$$

where Ψ is the standard normal univariate cdf, and $\sigma > 0$. An illustration of this approximation of the indicator function is displayed in Figure 4.1. It should be noted however that the uniform distribution defined on $[-\sigma/2, \sigma/2]$, with $\sigma > 0$, is another univariate cdf often considered for the approximation of the indicator function [Torii and Novotny, 2021, Torii, 2020].

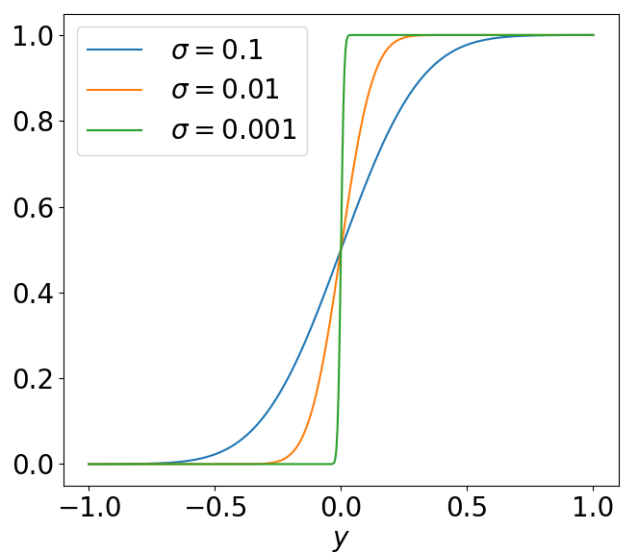


Figure 4.1: Approximation of the indicator function $\mathbb{I}_{y>0}$ with the function $y \to \Psi(y/\sigma)$ for three different values of σ, where Ψ is the cdf of a standard normal univariate variable.

In the sense of distributions, the derivative of P_f with respect to s_ℓ can then be defined as [Torii, 2020]

$$\begin{aligned}\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} &= \lim_{\sigma \to 0} \int_{\mathbb{R}^d} \frac{\partial \Psi\left(-g(\mathbf{s}, \mathbf{x})/\sigma\right)}{\partial s_\ell} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} \\ &= -\lim_{\sigma \to 0} \int_{\mathbb{R}^d} \frac{1}{\sigma} \frac{\partial g(\mathbf{s}, \mathbf{x})}{\partial s_\ell} \phi\left(-\frac{g(\mathbf{s}, \mathbf{x})}{\sigma}\right) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x},\end{aligned}$$

where ϕ is the univariate standard normal pdf. The convergence is proven in [Papaioannou et al., 2013] under some regularity conditions. Consequently, an approximation of the failure probability sensitivity can be obtained with the following domain integral

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} \approx -\int_{\mathbb{R}^d} \frac{1}{\widetilde{\sigma}} \frac{\partial g(\mathbf{s}, \mathbf{x})}{\partial s_\ell} \phi\left(-\frac{g(\mathbf{s}, \mathbf{x})}{\widetilde{\sigma}}\right) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \frac{\partial \widetilde{P_f}(\mathbf{s}, \widetilde{\sigma})}{\partial s_\ell}$$

$$\text{where} \quad \widetilde{P_f}(\mathbf{s}, \widetilde{\sigma}) = \int_{\mathbb{R}^d} \Psi\left(-\frac{g(\mathbf{s}, \mathbf{x})}{\widetilde{\sigma}}\right) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x},$$

for $\widetilde{\sigma}$ a set positive value. One should notice that the following convergence also holds: $\widetilde{P_f}(\mathbf{s}, \sigma)$ converges to $P_f(\mathbf{s})$ as $\sigma \to 0$. The domain integral $\partial \widetilde{P_f}(\mathbf{s}, \widetilde{\sigma})/\partial s_\ell$ can be estimated with the Monte Carlo method with the following estimate

$$\widehat{\frac{\partial \widetilde{P_f}(\mathbf{s}, \widetilde{\sigma})}{\partial s_\ell}}^{\mathrm{MC}} = -\frac{1}{N} \sum_{j=1}^{N} \frac{1}{\widetilde{\sigma}} \frac{\partial g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)}{\partial s_\ell} \phi\left(-\frac{g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)}{\widetilde{\sigma}}\right),$$

where the $\mathbf{X}^{(j)}$ are iid from $f_{\mathbf{X}}$. The same samples can be reused to compute both the failure probability estimate with MC and this sensitivity estimate, however, for each observation $\mathbf{X}^{(j)}$, the derivative of the lsf with respect to s_ℓ has to be evaluated, which increases the simulation budget. This estimate is biased, since $\widetilde{\sigma} \neq 0$ and its statistical properties are detailed in the next section.

4.3.3.2 Statistical properties of the Weak approach with crude MC

The statistical properties of this sensitivity estimate have been studied in numerous papers [Papaioannou et al., 2013, Torii, 2020, Torii and Novotny, 2021] as the choice of $\widetilde{\sigma}$ is crucial and has influence on both the bias and the variance of the estimate. As the estimate is biased, it is interesting to use the expression of the mean square error denoted $\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}(\sigma)$, derived in [Torii, 2020], which is a function of $\sigma > 0$

$$\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}(\sigma) = \left(\mathbb{E}\left[\widehat{\frac{\partial \widetilde{P_f}(\mathbf{s}, \sigma)}{\partial s_\ell}}^{\mathrm{MC}}\right] - \frac{\partial P_f(\mathbf{s})}{\partial s_\ell}\right)^2 + \mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial \widetilde{P_f}(\mathbf{s}, \sigma)}{\partial s_\ell}}^{\mathrm{MC}}\right)$$

$$\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}(\sigma) = e^2 + \mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial \widetilde{P_f}(\mathbf{s}, \sigma)}{\partial s_\ell}}^{\mathrm{MC}}\right),$$

where e is the bias. Intuitively, one would want the value of $\widetilde{\sigma}$ to be as small as possible to reduce the bias. However, for very small σ, the domain integral of $\partial \widetilde{P_f}(\mathbf{s}, \sigma)/\partial s_\ell$ converges to a surface integral over the limit state surface $g(\mathbf{s}, \mathbf{x}) = 0$. Consequently, most observations of the sensitivity MC estimate would then have a negligible weight in the mean sample value, as only few of them are close to the limit state surface.

In [Torii, 2020] analytical formulas are derived when the uniform cdf defined on the interval $[-\sigma/2, -\sigma/2]$ is chosen as the failure domain indicator function approximation, for the one-dimensional case ($d = 1$). The resulting bias e is proportional to σ^2, thus $e^2 \propto \sigma^4$ and the variance is proportional to $1/N\sigma$. The proportional property of the bias is extended to the multidimensional case in [Torii and Novotny, 2021]. Consequently, decreasing the parameter σ greatly reduces the bias. Nevertheless, past a certain point, the variance of the estimate increases for smaller σ values, if N is kept constant. This is schematically illustrated in Figure 4.2. The horizontal axis is separated into two stages: the bias stage where the error is mainly affected by the bias of the estimate and the variance stage where the error is mainly affected by the variance of the estimate.

Therefore, the optimal value of $\widetilde{\sigma}$ should be selected at the intersection between the bias stage and the variance stage. This intersection value is challenging to obtain in practice and is problem dependent [Lacaze et al., 2015]. An analytical formula of the optimal $\widetilde{\sigma}$ is detailed in [Torii and Novotny, 2021] for the one-dimensional case, with the uniform cdf approximation mentioned above. This formula is proportional to $1/N^{1/5}$ and depends on the quantity of interest $\partial P_f(\mathbf{s})/\partial s_\ell$ as well as the derivative of order 3 of P_f with respect to s_ℓ; thus its computation is intractable in practice. With this optimal $\widetilde{\sigma}$ value, the value of the mean square error is computed and $\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}(\widetilde{\sigma}) \propto 1/N^{4/5}$. This convergence rate is slower than the error convergence rate of the MC estimate obtained with the score function method, which is $1/N$, see Section 4.2.2. Consequently, even with the best $\widetilde{\sigma}$ value, the score function method is more efficient, but it is only applicable when $\mathbf{s}$ are distribution parameters only.

Finally, it should be underlined that the bias and variance of the sensitivity estimate are independent of $P_f(\mathbf{s})$, as noticed in [Torii, 2020]. They are dependent on the derivatives of P_f of order 1 and

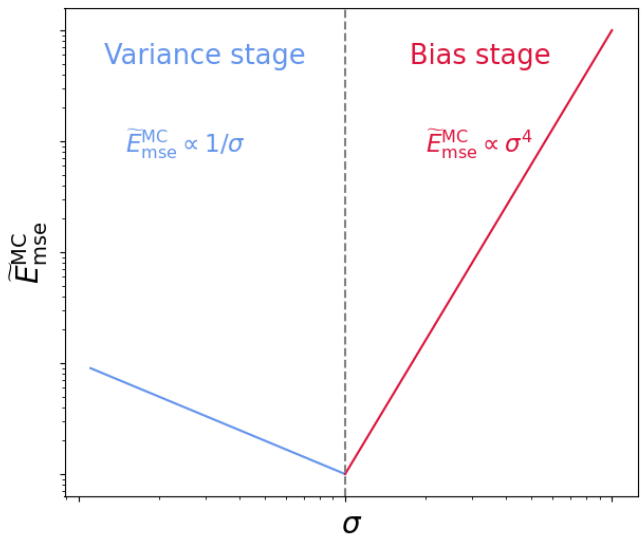


Figure 4.2: Illustration of the relation between $\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}$ and σ in log-log scale, for a fixed sample size N, with the approximations derived in [Torii, 2020] in a one-dimensional case.

above. In comparison, the mean square error of the MC estimate obtained with the score function method depends on P_f, as shown in Eq. (4.5). Consequently, the sensitivities of $P_f(\mathbf{s})$ computed with the Weak approach and the failure probability estimate do not depend on the same quantities. Improvements on the failure probability estimation will not necessarily lead to improvements on the failure probability sensitivity estimation and vice versa.

4.3.3.3 Weak approach with advanced sampling techniques

The Weak approach has been associated to the sequential importance sampling (SIS) framework in [Papaioannou et al., 2018], to obtain a sensitivity estimate more efficient than the MC one. The selected value of $\widetilde{\sigma}$ is the smallest value of σ such as the coefficient of variation of the sensitivity estimate is close to a CV target

$$\widetilde{\sigma} = \underset{\sigma \in (0, \sigma_L)}{\operatorname{argmin}} \left(\mathrm{CV} \left(\widehat{\frac{\partial \widetilde{P_f}(\mathbf{s}, \sigma)}{\partial s_\ell}}^{\mathrm{SIS}} \right) - \mathrm{CV}_{\mathrm{target}} \right)^2 ,$$

where σ_L is the final value of the decreasing sequence σ_i of the SIS framework (see [Papaioannou et al., 2016]). This selection does not require any additional evaluation of the lsf. The resulting sensitivity estimate has a controlled CV however it is not possible to measure its bias, other than with the approximation $e \propto \widetilde{\sigma}^2$. This method is efficient for multiple failure regions [Papaioannou et al., 2018]. The SIS framework, only briefly mentioned in Chapter 3, is also quite robust in high-dimensional spaces [Papaioannou et al., 2016]. Consequently the method presented in [Papaioannou et al., 2018] is particularly appropriate in our context to derive the sensitivities, but the bias is not controlled.

The Weak approach has been associated to the subset sampling framework in [Lacaze et al., 2015], for sensitivities with respect to design parameters. Eq. (4.7) presented in Section 4.2.3.1 is still valid (with $\boldsymbol{\delta}$ instead of $\boldsymbol{\theta}$), however the dependence on $\boldsymbol{\delta}$ in each $P_i(\boldsymbol{\delta})$ is then contained in the intermediate failure domain indicator function, as well as in $\mathbb{P}(F_{i-1})$. Each intermediate sensitivity $\partial P_i(\boldsymbol{\delta})/\partial \delta_\ell$ is then estimated with the Weak approach. The selected value of $\widetilde{\sigma}$ is obtained with a formula that depends on the number of failing observations in the sampling procedure, on $P_f(\mathbf{s})$, and on the order of magnitude of the lsf g. More precisely, $\widetilde{\sigma}$ is chosen as a quantile of the $\left\|Y^{(j)}\right\| = \left\|g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)\right\|$ ordered values, for $j = 1, \ldots, N$. This choice results in a number N_r of response values within $\pm\widetilde{\sigma}$, where N_r

depends on N and $P_f(\mathbf{s})$ and is set *a priori*. As previously mentioned, the bias of this estimate cannot be controlled other than with the approximation $e \propto \widetilde{\sigma}^2$.

4.3.4 Sensitivity with finite difference schemes

Although the gradient of the limit state function is supposed to be available in this manuscript, we briefly present here approaches that do not rely on the derivatives of the limit state function g with respect to $\mathbf{s}$ to compute the sensitivity of P_f. These approaches are applicable for high-dimensional systems whose failure domain encompasses several failure regions, and will be used as reference values later in the manuscript.

4.3.4.1 Expression of the finite difference schemes

The derivatives of P_f with respect to $\mathbf{s}$ can be directly based on finite difference schemes [Atkinson, 1991, Quarteroni et al., 2010]. Denoting $\mathbf{h}_\ell$ the p-vector whose components equal zero except the ℓ-th component equals to h, with $h > 0$, we recall here the forward formula

$$\frac{\partial \widehat{P_f(\mathbf{s}, \mathbf{h}_\ell)}^{\mathrm{F}}}{\partial s_\ell} = \frac{P_f(\mathbf{s}+\mathbf{h}_\ell) - P_f(\mathbf{s})}{h} = \frac{\mathbb{E}_{f_\mathbf{X}}\left[\mathbb{I}_{D_f(\mathbf{s}+\mathbf{h}_\ell)}(\mathbf{X})\right] - \mathbb{E}_{f_\mathbf{X}}\left[\mathbb{I}_{D_f(\mathbf{s})}(\mathbf{X})\right]}{h},$$

and the central formula

$$\frac{\partial \widehat{P_f(\mathbf{s}, \mathbf{h}_\ell)}^{\mathrm{C}}}{\partial s_\ell} = \frac{P_f(\mathbf{s}+\mathbf{h}_\ell) - P_f(\mathbf{s}-\mathbf{h}_\ell)}{2h} = \frac{\mathbb{E}_{f_\mathbf{X}}\left[\mathbb{I}_{D_f(\mathbf{s}+\mathbf{h}_\ell)}(\mathbf{X})\right] - \mathbb{E}_{f_\mathbf{X}}\left[\mathbb{I}_{D_f(\mathbf{s}-\mathbf{h}_\ell)}(\mathbf{X})\right]}{2h}.$$

The expected values of these formulas can then be estimated with MC methods. Here we use the same formulations to present the different approaches as in [Torii, 2020]. The Direct approach consists in employing independent samples for the estimations of $P_f(\mathbf{s}+\mathbf{h}_\ell)$, $P_f(\mathbf{s})$ and $P_f(\mathbf{s}-\mathbf{h}_\ell)$. Therefore, in order to obtain the failure probability sensitivity, the simulation budget is multiplied by two with the forward formula and by three with the central formula. In contrast, the Common Random Variable (CRV) approach employs the same sample for the estimations of $P_f(\mathbf{s}+\mathbf{h}_\ell)$, $P_f(\mathbf{s})$ and $P_f(\mathbf{s}-\mathbf{h}_\ell)$. However, the limit state functions with parameter $\mathbf{s}+\mathbf{h}_\ell$ and $\mathbf{s}-\mathbf{h}_\ell$ differ from $g(\mathbf{s}, \cdot)$. Consequently, the simulation budget is still multiplied by two with the forward formula and by three with the central formula. Using the same sample allows to cancel the noise in the expected values MC estimates, which justifies the CRV approach.

4.3.4.2 Statistical properties of the finite difference scheme estimates

Both approaches are biased since $h \neq 0$. The expression of the bias and variance of both approaches, for both finite difference schemes can be found in [Torii, 2020]. The bias e of both approaches is proportional to h with the forward formula and h^2 for the central formula. The variance of the Direct approach is proportional to $1/Nh^2$. Whereas the variance of the CRV approach is proportional to $1/Nh$. Consequently, the CRV is theoretically superior to the Direct approach. Furthermore, amongst the two formulas, the central formula is the best theoretically as the bias converges more rapidly, for small h. The evolution of the mean square error as a function of h can be separated in two stages for both methods: the variance stage for small values of h and the bias stage for greater values of h, as presented for the Weak approach in Section 4.3.3.2.

An optimal value of h is displayed in [Torii and Novotny, 2021] for the central formula in the CRV approach. This value is proportional to $1/N^{1/5}$ and depends on the quantity of interest $\partial P_f(\mathbf{s})/\partial s_\ell$ as well as the derivative of order 3 of P_f with respect to s_ℓ; thus its computation is intractable in practice. With this optimal h, the value of the mean square error is computed and $\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}(\widetilde{\sigma}) \propto 1/N^{4/5}$, which is the same convergence rate than the Weak approach with the optimal $\widetilde{\sigma}$. The same conclusion can be drawn: this convergence rate is slower than the error convergence rate of the MC estimate obtained with the score function method, which is $1/N$, see Section 4.2.2. Consequently, for sensitivities with respect to distribution parameters, the score function method is more efficient.

4.4 Conclusion

In this chapter we presented several methods to compute the derivatives of the failure probability with respect to the deterministic inputs $\mathbf{s}$ of the system, for a local reliability-based sensitivity analysis purpose. We focused on methods efficient in the special context of the manuscript. The dimension of the system is thus large, $d > 50$, and the failure domain encompasses possibly several failure regions. For each method, the additional simulation budget needed is underlined, compared to the simulation budget necessary to the estimation of the failure probability.

If the deterministic inputs are distribution parameters, denoted $\boldsymbol{\theta}$, two different frameworks are possible. The first framework relies on the score function methods and the estimation is performed in the original input space. The simulation budget is then constant and the sensitivity estimates are typically unbiased. However, there are few efficient techniques available to compute the failure probability in the original space. In the second framework, the original inputs are transformed into standard elliptical inputs with an isoprobabilistic transformation. In this input space, there is a greater number of efficient methods to compute the failure probability. Nevertheless, the sensitivity estimation causes a heavier simulation budget, as the derivatives of the limit state function with respect to $\boldsymbol{\theta}$ must be evaluated. Some approaches result in unbiased estimates, as in the DS and LS context, but require an even greater simulation budget as the gradient of g with respect to the random inputs is needed. While the other approaches lead to biased estimates, whose bias depend on a scalar parameter (σ or h), or remains unknown (in the FORM context). The resulting bias cannot not be properly controlled.

If the deterministic parameters are design parameters, denoted $\boldsymbol{\delta}$, the evaluation of the derivatives of g with respect to $\boldsymbol{\delta}$ is mandatory, except in the finite difference schemes context (where the simulation budget is then doubled or tripled). The sensitivity analysis is performed in the standard elliptical space and the same conclusions can be drawn concerning the simulation budget and the bias than with distribution parameters. Thus, the simulation budget substantially increases with LS or DS and the bias cannot be properly managed with the other approaches.

In Chapter 6 we present a new approach to compute the derivatives of the failure probability with respect to any deterministic input $\mathbf{s}$, as a byproduct of a heteroscedastic polynomial regression, in the standard elliptical space. This new approach only requires the evaluation of the derivatives of the limit state function with respect to $\mathbf{s}$ in addition to the simulation budget necessary for the estimation of P_f. It is inspired by the Weak approach. However, the resulting bias can be controlled in an innovative way with the parameters of the polynomial regression.

It should be noted that a large literature concerning the derivatives of P_f with respect to deterministic inputs in the meta-model framework exists [Torii et al., 2017, Dubourg and Sudret, 2014, Moustapha and Sudret, 2019, Yun et al., 2020], but as mentioned in Chapter 3, these methods are not presented here.

Estimation de probabilité de défaillance avec échantillonnage préférentiel multiple dans un espace standard elliptique de grande dimension

Nous présentons ici un algorithme d'estimation de probabilité de défaillance d'un système de grande dimension, dont le domaine de défaillance possède plusieurs régions de défaillance. Les variables aléatoires du système sont modélisées avec une distribution elliptique standard. La dépendance de la probabilité de défaillance aux variables déterministes $\mathbf{s}$ du système est supposée implicite dans ce chapitre.

Présentation du schéma adaptatif de l'algorithme

Toutes les régions de défaillance doivent être identifiées dans l'espace elliptique de grande dimension afin d'obtenir un estimateur précis de la probabilité de défaillance. Au lieu de construire un mélange de densités paramétriques, optimisées ensemble, ayant pour but d'échantillonner dans le domaine de défaillance entier, nous proposons ici de recourir au paradigme "diviser pour mieux régner". Ainsi, chaque région de défaillance est identifiée et échantillonnée seule, avec une unique densité optimisée. La probabilité de défaillance est ensuite estimée par échantillonnage préférentiel multiple, avec un mélange de toutes les densités construites au fur et à mesure.

L'estimation de la probabilité de défaillance est alors décomposée en deux problèmes distincts. Tout d'abord, il y a le problème d'identification, qui consiste à trouver dans l'espace standard elliptique les différentes régions de défaillance. Ensuite, il y a le problème d'échantillonnage, qui consiste à construire une densité auxiliaire optimisée pour chaque région de défaillance. Ces problèmes sont bien distincts mais ils ne peuvent pas être résolus indépendamment : dès qu'une région de défaillance est identifiée, l'échantillonnage dans cette zone commence. La recherche des autres régions de défaillance s'ensuit, en prenant en compte la région de défaillance déjà identifiée : c'est l'étape d'adaptation. Ce schéma est répété tant que toutes les régions de défaillance ne sont pas identifiées et échantillonnées. Nous détaillons à présent chaque étape de la méthode.

Le problème d'identification des régions de défaillance

La recherche des régions de défaillance est inspirée de la méthode FORM/SORM dans le cas de points de conception multiples [Der Kiureghian and Dakessian, 1998]. Les régions de défaillance sont alors identifiées grâce à des points défaillants représentatifs. Le point représentatif d'une région est le point défaillant de la région de norme euclidienne minimale, dans l'espace standard elliptique. Il est appelé point de conception. En effet, ce point défaillant le plus proche de l'origine est celui qui maximise la vraisemblance i.e. c'est le point défaillant de densité la plus élevée de toute la région en question. Néanmoins, il est souligné dans [Katafygiotis and Zuev, 2008] que les points de conception n'appartiennent généralement pas à l'anneau d'importance dans l'espace standard normal. Pour pren-

dre en compte cette situation, la solution considérée consiste alors à chercher les points défaillants les plus proches de l'origine dans l'anneau d'importance.

Selon la localisation de la région de défaillance et de l'anneau d'importance, deux situations, appelées Cas 1 et Cas 2 sont considérées. Dans le Cas 1, le point défaillant le plus proche de l'origine se trouve dans l'anneau d'importance. Ainsi ce point est le point sélectionné pour être représentatif de la région de défaillance. Il est trouvé par optimisation sous contrainte, en minimisant la norme des points $\mathbf{X}$ appartenant à l'anneau d'importance et à la région de défaillance. Dans le Cas 2, le point défaillant le plus proche de l'origine ne se trouve pas dans l'anneau d'importance et se situe plus près de l'origine que la borne inférieure LB de l'intervalle d'importance. Pour trouver un point représentatif de la région de défaillance, deux optimisations sont alors nécessaires, la première étant la même que dans le Cas 1. Puisque le résultat de cette première optimisation est trouvé sur la borne inférieure, la deuxième optimisation consiste à chercher sur cette borne inférieure le point le plus représentatif de la région de défaillance. Le résultat de ces deux optimisations est alors un point défaillant qui minimise la fonction d'état limite g sur la borne inférieure de l'anneau d'importance.

Puisque plusieurs régions de défaillance doivent être identifiées, ces optimisations sont effectuées avec une technique de points de départ multiples aléatoires. Une liste L de minima locaux est ainsi obtenue. Cette liste est ensuite triée pour favoriser les points représentants les régions de défaillance les plus proches ou bien ceux dont la valeur de la fonction d'état limite est la plus négative. Le premier point P_1 de cette liste représente la première région de défaillance identifiée.

Le problème d'échantillonnage de la région de défaillance

Une fois qu'une région de défaillance a été identifiée via son point représentatif P_1, il s'agit d'échantillonner dans cette région avec une densité auxiliaire optimale. En prenant en compte la représentation stochastique des variables elliptiques, la densité paramétrique utilisée pour échantillonner la région est construite comme un produit d'une densité h_R sur la variable aléatoire R et d'une densité $h_{\mathbf{T}}$ sur le vecteur aléatoire $\mathbf{T}$.

La densité radial h_R sélectionnée est la densité radiale elliptique de départ f_R conditionnée à ce que R soit plus grand qu'une certaine valeur scalaire appelée $r_{\text{opt},1}$. La densité directionnelle choisie est une densité de von Mises–Fisher [Banerjee et al., 2005], qui dépend d'un paramètre de concentration κ_1 scalaire et d'une direction moyenne $\boldsymbol{\nu}_1 \in \mathcal{T}^d$. Plus le paramètre de concentration κ_1 est élevé, plus la direction $\boldsymbol{\nu}_1$ est importante. Il y a donc $d+2$ paramètres à optimiser pour la densité auxiliaire choisie $h_R \times h_{\mathbf{T}}$. Au lieu d'optimiser ces paramètres tous ensemble avec un algorithme d'entropie croisée comme dans [Wang and Song, 2016, Papaioannou et al., 2019], les coordonnées du point représentatif P_1 sont utilisées pour définir la direction moyenne $\boldsymbol{\nu}_1$ et le scalaire $r_{\text{opt},1}$.

Pour le scalaire $r_{\text{opt},1}$, la localisation de la région de défaillance dans l'anneau d'importance influence le paramétrage. Ainsi le scalaire $r_{\text{opt},1}$ est égal à la norme euclidienne de P_1 dans le Cas 1 et à 0 dans le Cas 2. Pour les autres paramètres, les configurations sont les mêmes peu importe le cas. La direction moyenne $\boldsymbol{\nu}_1$ est égale à la direction normée du point P_1. Le paramètre scalaire de concentration κ_1 est optimisé avec un algorithme d'entropie croisée avec échantillonnage préférentiel multiple adaptatif, où la densité optimale à approcher est celle de la région de défaillance et non pas celle du domaine de défaillance entier. La convergence de l'algorithme est défini avec un critère de coefficient de variation théorique.

L'utilisation de cet algorithme d'entropie croisée avec échantillonnage préférentiel multiple adaptatif introduit de la dépendance entre les échantillons générés. Néanmoins, nous avons remarqué que cette dépendance était négligeable lorsque le paramètre à optimiser était seulement le paramètre de concentration κ scalaire. L'échantillonnage préférentiel multiple permet alors de rentabiliser au mieux les évaluations coûteuses de la fonction d'état limite. En effet, les évaluations sont recyclées à chaque itération de l'algorithme. Cela augmente alors le nombre total d'observations défaillantes et l'optimisation du κ est plus stable. À la fin de cet algorithme, la probabilité de défaillance de cette région est estimée.

Étape d'adaptation

Lorsqu'une région de défaillance a été identifiée et échantillonnée, la recherche des autres régions de défaillance commence. Pour se faire, nous nous inspirons toujours de la méthode FORM avec plusieurs points de conception [Der Kiureghian and Dakessian, 1998], en adaptant la méthode à notre contexte

spécifique. Une fois que la première région de défaillance est trouvée, un bulbe est construit autour de P_1 et est ajouté à la fonction d'état limite g. La nouvelle fonction d'état limite est appelée fonction d'état limite modifiée et est notée $\tilde{g}$. Cette région de défaillance représentée par P_1 n'est alors plus considérée dans la recherche des autres régions de défaillance, puisqu'elle appartient désormais au domaine de sûreté du système.

Ce bulbe est défini à l'aide d'un rayon et d'un paramètre d'échelle "scale". Après avoir construit ce bulbe, une première mise à jour de la liste L contenant tous les points représentatifs est effectuée. Tous les points se situant dans le bulbe de P_1 sont enlevés car ils représentent la même région de défaillance, qui a déjà été échantillonnée.

Ce bulbe peut ne pas englober toute la zone de défaillance associée à P_1, tout particulièrement dans le Cas 2 mentionné plus haut, où P_1 se situe sur la borne inférieure de l'anneau d'importance alors que la zone de défaillance s'étale sur tout l'anneau d'importance. Ainsi, un cône de défaillance associé à P_1 est défini grâce à l'angle maximal entre P_1 et les points défaillants générés dans la région de défaillance lors du problème d'échantillonnage. Tous les points défaillants de la liste L se trouvant dans ce cône de défaillance sont alors retirés de la liste, puisqu'ils représentent aussi la même région de défaillance. C'est la deuxième mise à jour de la liste

À la suite de ces mises à jour, les points défaillants se trouvant dans la liste L sont représentatifs de régions de défaillance qui n'ont pas encore été échantillonnées. Le premier point de cette liste mise à jour, P_2, sert à identifier la seconde région de défaillance. Le problème d'échantillonnage est alors effectué pour cette seconde région, avec la construction d'une nouvelle densité auxiliaire. La fonction d'état limite utilisée pour le problème d'échantillonnage est toujours celle d'origine g et non pas celle modifiée. Un bulbe autour de P_2 est ensuite ajoutée à la fonction d'état limite modifiée pour la mettre à jour et un cône est aussi associé au point P_2. La liste L est mise à jour à nouveau. Ce schéma est répété tant que la liste L n'est pas vide.

Recherche des régions de défaillance négligées

Il peut arriver que la liste L obtenue suite aux multiples optimisations effectuées lors du problème d'identification ne contiennent pas de points représentatifs de toutes les régions de défaillance du système. En effet, une région de défaillance peut être particulièrement dominante, au point que les optimisations locales convergent toutes vers le même point représentatif. Il peut alors y avoir plusieurs régions de défaillance qui n'ont pas été identifiées. Afin de remédier à ce problème, un autre ensemble d'optimisation est considéré, c'est un autre problème d'identification qui s'amorce.

Ces optimisations sont semblables à celles présentées dans le problème d'identification, à deux exceptions près. Tout d'abord, la fonction d'état limite utilisée n'est pas celle d'origine g, mais celle modifiée $\tilde{g}$ prenant en compte tous les bulbes construits. Le domaine de défaillance n'est donc plus celui de départ. Ensuite, la technique de points de départ multiples aléatoires des optimisations est centrée autour de la position opposée à la somme des points défaillants déjà trouvés.

Une fois ces optimisations effectuées, une nouvelle liste L de points représentatifs de région de défaillance est obtenue. Une mise à jour de cette liste est effectuée en enlevant les point défaillants contenus dans le cône de défaillance associé à chaque région de défaillance déjà trouvée. La liste est ensuite triée, le premier point de cette liste est représentatif d'une nouvelle région de défaillance à échantillonner. Le schéma précédemment décrit est donc réitéré pour cette nouvelle liste, et ainsi de suite.

Ainsi, l'algorithme s'arrête lorsqu'une nouvelle liste L ne continent plus aucune nouvelle région de défaillance. Il semble alors que toutes les régions de défaillance du système ont été identifiées et échantillonnées.

Estimation de la probabilité de défaillance

La probabilité de défaillance globale est estimée avec un estimateur d'échantillonnage préférentiel multiple (MIS), en utilisant toutes les densités construites au fur et à mesure et tous les échantillons générés. Supposons que le nombre de région de défaillance final soit égal à K, et que pour chaque itération de l'algorithme d'échantillonnage, N observations soient générées. Pour chaque région de défaillance, on note n_k l'indice final du paramètre de concentration κ optimisé. L'équation final de

l'estimateur s'écrit alors

$$\hat{P}_f^{\mathrm{MIS}} = \frac{1}{n_{\mathrm{total}} \times N} \sum_{j=1}^{n_{\mathrm{total}} \times N} \mathbb{I}_{D_f}\left(R^{(j)}\mathbf{T}^{(j)}\right) \frac{f_R\left(R^{(j)}\right) f_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)}{h_{\mathrm{MIS}}\left(R^{(j)}, \mathbf{T}^{(j)}\right)},$$

où les paires aléatoires $\left(R^{(1)}, \mathbf{T}^{(1)}\right), \left(R^{(2)}, \mathbf{T}^{(2)}\right), \ldots$ sont générées par la densité h_{MIS}

$$\forall r \in \mathbb{R}^+ \quad \forall \mathbf{T} \in \mathcal{T}^d \quad h_{\mathrm{MIS}}(r\mathbf{t}) = \frac{1}{n_{\mathrm{total}}} \sum_{k=1}^{K} h_R(r; r_{\mathrm{opt},k}) \left\{ \sum_{\ell=0}^{n_k} h_{\mathbf{T}}(\mathbf{t}; \kappa_{k,\ell}, \boldsymbol{\nu}_k) \right\}.$$

Le nombre n_{total} de densités auxiliaires est donc égal à $n_1 + \cdots + n_K + K$. Ce nombre de densité est plus élevé que dans d'autres algorithmes [Papaioannou et al., 2019, Geyer et al., 2019]. Néanmoins, ce grand nombre de densités et d'échantillons associés rendent l'estimateur d'échantillonnage préférentiel plus précis en termes de coefficient de variation.

Le budget de simulation de la méthode est alors réparti entre les différents problèmes d'identification et d'échantillonnage nécessaires au fur et à mesure de l'algorithme. Le gradient de la fonction d'état limite est nécessaire pour résoudre les problèmes d'identification. Selon la précision des optimisations menées lors des différents problèmes d'identification, le budget de simulation peut être grandement impacté : si les coordonnées de chaque P_k ne sont pas précises, des itérations supplémentaires sont nécessaires dans l'algorithme d'échantillonnage.

Un estimateur théorique de la variance de $\hat{P}_f^{\mathrm{MIS}}$, supposant que toutes les observations sont indépendantes et identiquement distribuées, est disponible. Il permet alors de déterminer la qualité de l'estimateur en une seule simulation.

Applications numériques

La performance de cet algorithme est mise à l'épreuve avec quatre applications numériques. Les trois premières sont des applications fréquemment utilisées dans la littérature, la dernière est une application industrielle plus réaliste, dans un contexte de fonction d'état limite boîte noire.

La technique de points de départ multiples aléatoires utilisée pour les optimisations des problèmes d'identification est celle d'échantillonnage par hypercube latin [Shields et al., 2015], de taille 10. Les différent points de départ sont alors répartis sur l'espace des entrées en globalité.

Système en série de quatre fonctions d'état limite linéaire

Cette première application permet de tester la robustesse de l'algorithme à trouver toutes les régions de défaillance (ici quatre) du système, en fonction de plusieurs paramètres. Les variables aléatoires sont normales standards. La probabilité de défaillance est alors indépendante de la dimension du système, sa valeur a un ordre de grandeur de 10^{-6}. Cet exemple est représentatif du Cas 2 mentionné plus tôt.

Tout d'abord, une étude sur la valeur du paramètre de concentration $\kappa_{k,0}$ initial est menée. Ce paramètre influence la convergence de l'algorithme d'échantillonnage d'entropie croisée. Trois dimensions du système sont alors étudiées : $d = 200$, $d = 300$ et $d = 400$, pour des valeurs de $\kappa_{k,0}$ allant de 50 à 150. Plus la dimension du système est grande, plus les valeurs finales convergées de κ sont élevées. Dans de très rares cas, l'algorithme trouve trois ou cinq régions de défaillance au lieu de quatre. Ces cas correspondent à des valeurs de $\kappa_{k,0}$ particulièrement basses ou bien très élevées par rapport à la dimension du système. Cependant, l'algorithme est globalement très robuste et permet d'estimer correctement la probabilité de défaillance du système, cf Figure 5.11.

La performance de l'algorithme est ensuite comparée à d'autres méthodes de simulation adaptées à la grande dimension : un algorithme de "subset simulation" [Au and Beck, 2001] ainsi qu'un algorithme d'entropie croisée améliorée (iCE) [Papaioannou et al., 2019]. Pour les trois dimensions concernées, l'algorithme proposé requiert un budget de simulation plus faible, cf Table 5.1. De plus, le coefficient de variation de l'estimateur de la probabilité de défaillance de notre méthode est aussi le plus bas. En augmentant la valeur de la probabilité de défaillance à estimer (10^{-4}), la méthode présentée est toujours la plus performante, en termes de budget de simulation ainsi que de coefficient de variation,

cf Table 5.2. Il est toutefois rappelé que le gradient de la fonction d'état limite, supposé disponible, doit être évalué dans l'algorithme proposé, alors que les deux autres méthodes n'en n'ont pas besoin.

L'algorithme est enfin expérimenté en très grande dimension : $d = 500$ et $d = 1000$. Les résultats sont également concluant, la probabilité de défaillance est correctement estimée. Le budget de simulation nécessaire est toujours bas (inférieur à 12500) et le coefficient de variation reste faible aussi (inférieur à 5%).

Oscillateur de Duffing

La fonction d'état limite de cette seconde application est non linéaire et représente le déplacement maximal d'un oscillateur de Duffing [Zuev, 2009] soumis à une excitation aléatoire. Les variables aléatoires sont normales standards. La probabilité de défaillance est indépendante de la dimension du système, sa valeur a un ordre de grandeur de 10^{-4}. L'algorithme a trouvé pour cette application, deux régions de défaillance. Cette application est aussi représentative du Cas 2 mentionné plus haut.

La performance de l'algorithme est testée en dimension $d = 100$, $d = 200$ et $d = 300$ et elle est comparée avec un algorithme de "subset simulation" ainsi qu'un algorithme d'entropie croisée améliorée (iCE). Les mêmes conclusions que pour l'exemple précédent peuvent être observées, cf Table 5.3. Pour les trois dimensions concernées, l'algorithme proposé requiert un budget de simulation plus faible et le coefficient de variation de l'estimateur est le plus faible aussi. Alors que la fonction d'état limite est non linéaire, le budget de simulation nécessaire aux multiples optimisations des différents problèmes d'identification reste relativement bas.

Pour cette application non linéaire, l'influence de la largeur de l'anneau d'importance est aussi étudiée en dimension $d = 100$. Les directions des points défaillants représentatifs ne varient pas avec un anneau d'importance plus étroit. Ainsi l'estimation de la probabilité de défaillance ne semble pas être affectée par ce paramètre.

Perte d'un portefeuille

Ce troisième exemple est fréquemment utilisé en finance [Chan and Kroese, 2012]. Il s'agit d'un large portefeuille de prêt de dimension $d = 250$ où chaque débiteur a une probabilité non nulle de faire défaut. Les variables aléatoires suivent une loi de Student multivariée. La probabilité de défaillance correspond à une perte excessive du portefeuille. La valeur de cette probabilité a un ordre de grandeur de 10^{-3}. Avant de pouvoir utiliser l'algorithme proposé, les variables d'entrées doivent être réduites, afin de travailler dans l'espace de Student multivarié standard. De même, la fonction d'état limite, qui dépend de plusieurs fonctions indicatrice, n'est pas directement dérivable. Une approximation dérivable de la fonction d'état limite est alors utilisée pour effectuer les optimisations nécessaires aux problèmes d'identification de la méthode. L'algorithme a trouvé pour cette application, une seule région de défaillance. Cette application est représentative du Cas 1 mentionné plus haut. Il est rappelé que l'anneau d'importance des lois de Student est très large, ce qui favorise l'apparition de situations de type Cas 1.

La performance de l'algorithme est comparée à celle d'une méthode de minimisation de variance [Chan and Kroese, 2012]. Contrairement aux deux applications précédentes, un budget de simulation assez conséquent est nécessaire pour l'algorithme afin d'obtenir un estimateur dont le coefficient de variation est faible (inférieur à 4%), cf Table 5.4. De plus, ce coefficient de variation est supérieur à celui de la méthode de minimisation de variance [Chan and Kroese, 2012]. Ainsi, pour cet application, l'algorithme ne semble pas être aussi performant que précédemment.

Une étude sur la distribution auxiliaire radiale permet d'expliquer ce résultat moins satisfaisant. Après avoir effectué un échantillonnage Monte-Carlo de large dimension, il semble que la valeur r_{opt} calculée grâce aux optimisations ne soit pas assez élevée. La densité auxiliaire radiale construite par l'algorithme n'est alors particulièrement pas optimale, ce qui explique le budget de simulation important et le coefficient de variation élevé. En augmentant cette valeur r_{opt}, l'efficacité de l'algorithme augmente aussi. De même, le budget de simulation diminue si la valeur du paramètre de concentration $\kappa_{1,0}$ initiale est abaissée. En effet, le domaine de défaillance semble avoir une symétrie sphérique.

Cette mauvaise estimation de r_{opt} peut être due à l'approximation de la fonction d'état limite mentionnée plus haut. Cette exemple prouve l'importance de la précision des optimisations effectuées dans les problèmes d'identification de la méthode.

Profil aérodynamique dans un flux non visqueux transsonique

Le dernier exemple examine le coefficient de traînée d'un profil aérodynamique dans un flux non visqueux transsonique soumis à une déformation aléatoire de forme. Cet exemple est représentatif d'un problème industriel puisque la fonction d'état limite est analytiquement non disponible et se comporte comme une boîte noire [Economon et al., 2016]. Les variables aléatoires représentent l'amplitude des déformations et suivent une loi normale centrée d'écart type 5×10^{-4}. La probabilité de référence est obtenue avec un échantillon Monte-Carlo de taille 106770 et sa valeur a un d'ordre de grandeur de 10^{-4}. Plus de 8 jours ont été nécessaires pour effectuer cette simulation Monte-Carlo puisque la boîte noire est effectivement très coûteuse en termes de temps de calcul (résolution de l'équation d'Euler par la méthode des volumes finis, parallélisée sur 24 CPU). L'algorithme a trouvé pour cette application trois régions de défaillance. Cet exemple est représentatif du Cas 2 mentionné plus haut.

La performance de l'algorithme est comparée au résultat de la méthode Monte-Carlo, pour trois simulations indépendantes seulement. Le budget de simulation nécessaire est alors beaucoup plus faible pour chaque simulation, comparé à l'échantillonnage Monte-Carlo classique, cf Table 5.5. Pour les deux premières simulations, la probabilité trouvée est proche de celle de la méthode Monte-Carlo, avec un coefficient de variation théorique plus bas. Pour la dernière simulation, la probabilité trouvée est plus faible que la valeur de référence. Cette valeur pour la troisième simulation est le résultat de coordonnées imprécises du point représentatif de la deuxième région de défaillance. En effet, l'optimiseur ne trouve pas le "bon" point représentatif, et l'échantillonnage de cette zone est alors sous-optimal.

Comme l'illustre les deux premières simulations, l'algorithme présenté peut donc proposer une estimation correcte de la probabilité de défaillance, en réduisant le budget de simulation par 7 comparé à un échantillonnage Monte-Carlo. Ce résultat est vraiment intéressant pour cet exemple où chaque évaluation de la fonction d'état limite est particulièrement coûteuse. Aussi, la probabilité de défaillance de chacune de ces trois régions est évaluée. La contribution à la probabilité de défaillance globale de chaque région de défaillance dans l'espace des entrées est alors soulignée. Cette information est primordiale pour mieux comprendre quelles configurations du profil aérodynamique mènent à la défaillance du système.

Conclusion

Dans ce chapitre, un nouvel algorithme d'estimation de probabilité de défaillance d'un système a été présenté. Les variables aléatoires du système sont modélisées par une loi elliptique standard de grande dimension. Nous supposons qu'il y a plusieurs régions de défaillance dans le domaine de défaillance. Nous abordons ici quelques remarques finales ainsi que des perspectives.

Nécessité de l'espace standard elliptique

Nous pouvons nous demander si la modélisation des variables aléatoires du système par une loi elliptique standard est obligatoire pour le bon fonctionnement de l'algorithme. Il semblerait que la réponse soit positive, pour plusieurs raisons.

Tout d'abord, l'indépendance entre la variable radiale R et le vecteur directionnel $\mathbf{T}$ des variables elliptiques standards permet de définir un anneau d'importance avec deux hypersphères. Sans cette indépendance, il est difficile de définir un anneau d'importance puisqu'il pourrait ne pas exister. Si les optimisations nécessaires aux problèmes d'identification ne sont pas contraintes à un espace restreint comme l'anneau d'importance, l'algorithme pourrait ne pas s'arrêter. De même, des régions de défaillance dont la masse probabiliste est extrêmement faible pourraient être considérées dans le problème d'échantillonnage, ce qui serait un gâchis d'évaluations de la fonction d'état limite coûteuse.

De plus, si les entrées ne sont pas elliptiques, il est aussi difficile de sélectionner une densité auxiliaire radiale efficace. Dans l'algorithme proposé, cette densité dépend de la loi de la variable R des entrées. Si une telle loi n'existe pas, plusieurs solutions sont alors possibles pour construire une densité auxiliaire radiale (comme avec des densités de noyau par exemple). Ces solutions nécessiteraient cependant une étape d'échantillonnage supplémentaire.

Enfin, si les entrées ne sont pas elliptiques, les poids de l'estimateur d'échantillonnage préférentiel multiple pourraient dégénérer avec une dimension d élevée. En effet, avec des entrées elliptiques, le choix de la densité radiale mentionnée au-dessus permet une simplification des poids de l'estimateur,

qui ne serait pas possible avec des entrées d'autre nature. Les poids pourraient alors dégénérer en grande dimension, ce qui est un phénomène très courant en échantillonnage préférentiel.

Influence des densités auxiliaires de l'échantillonnage préférentiel

Les applications numériques présentées permettent de rendre compte de l'importance des densités auxiliaires sélectionnées pour l'échantillonnage préférentiel. En effet, dans l'espace standard gaussien, il semble que la densité auxiliaire directionnelle soit la composante la plus importante du produit de densités, pour les trois applications considérées. De bonnes valeurs des paramètres de cette densité rend l'échantillonnage préférentiel efficace. Au contraire, avec l'application avec les lois de Student, puisque le domaine de défaillance possède une symétrie sphérique, c'est la bonne valeur du paramètre de la densité auxiliaire radiale qui est primordiale. Prendre en compte la représentation stochastique des entrées elliptiques est donc pertinent dans les deux espaces aléatoires considérés.

Algorithme proposé sans l'utilisation du gradient de la fonction d'état limite

Dans ce manuscrit, il est supposé que le gradient de la fonction d'état limite est disponible pour les optimisations nécessaires aux problèmes d'identification. Dans le cas où ce gradient n'est pas aisément calculable, ces optimisations sont alors beaucoup plus difficiles à effectuer, en particulier avec une dimension du système élevée. Des travaux récents permettent de trouver les points de conception du système sans avoir recours au gradient de la fonction d'état limite [Zhong et al., 2020,Zhu et al., 2022]. En combinant ces techniques avec la restriction à l'anneau d'importance lors de la recherche des points représentatifs des régions de défaillance, une version alternative de l'algorithme ne nécessitant pas le gradient pourrait être envisageable.

Chapter 5

New failure probability estimation method with MIS

Contents

5.1 Introduction . . . 102
5.2 Adaptive method to estimate the failure probability . . . 103
5.2.1 Presentation of the proposed method . . . 103
5.2.1.1 Presentation of the scheme for sampling in the entire failure domain 103
5.2.1.2 Presentation of the MIS estimate . . . 104
5.2.2 Identification problem . . . 104
5.2.2.1 Search for the multiple failure regions . . . 104
5.2.2.2 Optimization problems in the important ring . . . 105
5.2.2.3 Selection of a first failure region . . . 106
5.2.3 Sampling problem . . . 106
5.2.3.1 Selection of a parametric family for the standard elliptical inputs . 106
5.2.3.2 Setting of the parameters . . . 107
5.2.3.3 Remarks on the CE-AMIS algorithm for κ . . . 110
5.2.4 Adaptation step . . . 111
5.2.4.1 Construction of a bulge around the failure region . . . 111
5.2.4.2 Definition of a failure cone associated with the failure region . . . 111
5.2.4.3 Identification of the other failure regions . . . 112
5.2.5 Search for missed failure regions . . . 112
5.2.6 Estimation of the global failure probability . . . 114
5.2.6.1 Final MIS estimate expression . . . 114
5.2.6.2 Discussion on the simulation budget allocation . . . 114
5.2.6.3 Example in dimension 2: A 4-branch system . . . 115
5.3 Numerical investigations with high-dimensional input space . . . 117
5.3.1 Series system of four linear limit state functions . . . 118
5.3.1.1 Presentation of the application . . . 118
5.3.1.2 Definition of the important ring . . . 119
5.3.1.3 Settings of the proposed algorithm . . . 120
5.3.1.4 Study of the influence of κ_0 on the failure probability estimate . . 120
5.3.1.5 Comparison with the iCE-vMFNM method and the SS algorithm 121
5.3.1.6 Influence of the value of P_f . . . 123
5.3.1.7 Test on higher dimensions . . . 124
5.3.2 Duffing oscillator . . . 124
5.3.2.1 Presentation of the application . . . 124
5.3.2.2 Definition of the important ring . . . 125
5.3.2.3 Comparison with the iCE-vMFNM method and the SS algorithm 125

5.3.2.4 Influence of the width of the important ring 127
5.3.3 Portfolio loss 128
5.3.3.1 Presentation of the application 128
5.3.3.2 Adaptation to the proposed method 128
5.3.3.3 Definition of the important ring 129
5.3.3.4 Performance of the proposed method 129
5.3.3.5 Study on radial distribution of the failing inputs 131
5.3.4 Airfoil in inviscid transonic flow 132
5.3.4.1 Design problem and optimal airfoil 132
5.3.4.2 Presentation of the reliability problem 133
5.3.4.3 Definition of the important ring 134
5.3.4.4 Performance of the proposed method 134
5.4 Conclusion 138
5.4.1 Necessity of standard elliptical inputs 138
5.4.1.1 Definition of an important ring 138
5.4.1.2 Selection of the radial auxiliary density 138
5.4.1.3 Degeneration of the MIS weights 139
5.4.2 Influence of IS auxiliary density in high-dimensional spaces 139
5.4.3 Proposed method without the gradient of the limit state function 139

5.1 Introduction

In this chapter, we present a new algorithm developed to estimate the failure probability of a high-dimensional system with standard elliptical inputs and a failure domain which encompasses several failure regions. The main highlights of the proposed algorithm are as follows:

- the failure probability is estimated in the importance sampling framework, with parametric adaptively optimized auxiliary densities;
- the auxiliary IS parametric densities are sequentially constructed for each failure region identified as part of the failure domain with a CE-AMIS method (cf Chapter 3 Section 3.3.3.3);
- the search for failure regions is achieved through optimization and is inspired by the FORM/SORM multiple design points method (cf Chapter 3 Section 3.2.2). However, an original adaptation of the optimization is proposed to take into account the important ring of the standard elliptical inputs (see Chapter 2 Section 2.2.2.1);
- the stochastic representation of the elliptically distributed inputs is exploited in the structure of the auxiliary parametric densities, which are expressed as the product of a parametric conditional distribution for the radial component, and a parametric von Mises–Fisher distribution for the directional vector;
- the failure probability of the entire failure domain is then estimated by multiple importance sampling (MIS) with a mixture of the auxiliary densities.

This chapter is organized as follows. A complete description of the proposed algorithm is first detailed in Section 5.2. Each step is described with illustrations. Section 5.3 presents four numerical examples with elliptical inputs: three with the Gaussian distribution and one with the Student distribution. The results are summarized and some conclusions are drawn in Section 5.4.

As the random inputs $\mathbf{X}$ are assumed to follow a standard elliptical distribution, the different formulas are written with the stochastic representation of the inputs: $\mathbf{X} = R\mathbf{T}$. The important ring defined in Chapter 2 Section 2.2.2.1 is then the region in the standard elliptical space encompassing most of the probability mass of $\mathbf{X}$. Furthermore, the expression of the dependence on the deterministic inputs of the system $\mathbf{s}$ is dropped in this chapter for the sake of clarity, for both P_f and the limit state function g.

5.2 Adaptive method to estimate the failure probability

As previously mentioned, several failure regions in the high-dimensional input space have to be identified in order to get an accurate estimate of the failure probability. In the specific case of parametric densities for IS, several methods [Wang and Song, 2016, Papaioannou et al., 2019, Kurtz and Song, 2013, Geyer et al., 2019] have relied on a mixture structure to account for the multiple failure regions. The IS density is then built as a weighted mixture of K parametric densities, with weight π_k. However, determining the number of failure regions K *a priori* is challenging; cf Chapter 3 Section 3.3.3.2. Furthermore, the weights have to be computed as well, which increases the number of unknown parameters.

Instead of trying to sample in the failure domain as a whole with a mixture of parametric densities optimized together, the method proposed here sequentially identifies and generates samples in each of the failure regions. In the spirit of the *divide and conquer* paradigm, the global IS density is thus sought through a local IS density for each failure region. The final failure probability estimate is then computed with a mixture of the optimized single IS auxiliary densities.

5.2.1 Presentation of the proposed method

5.2.1.1 Presentation of the scheme for sampling in the entire failure domain

In order to estimate the failure probability, it is proposed to decompose the problem in two parts. First, an identification problem is considered which consists of identifying the different regions of the failure domain in the input space. Next, a sampling problem is solved by constructing an optimized auxiliary parametric density for each failure region. Although these two problems are distinct, they are intertwined and cannot be solved separately. As soon as a failure region is found, the construction of its IS auxiliary density is considered. Next, the search for a new failure region is initiated, taking into account the failure region previously found; this is the adaptation step. This scheme is repeated until all failure regions have been found and sampled as illustrated in Figure 5.1.

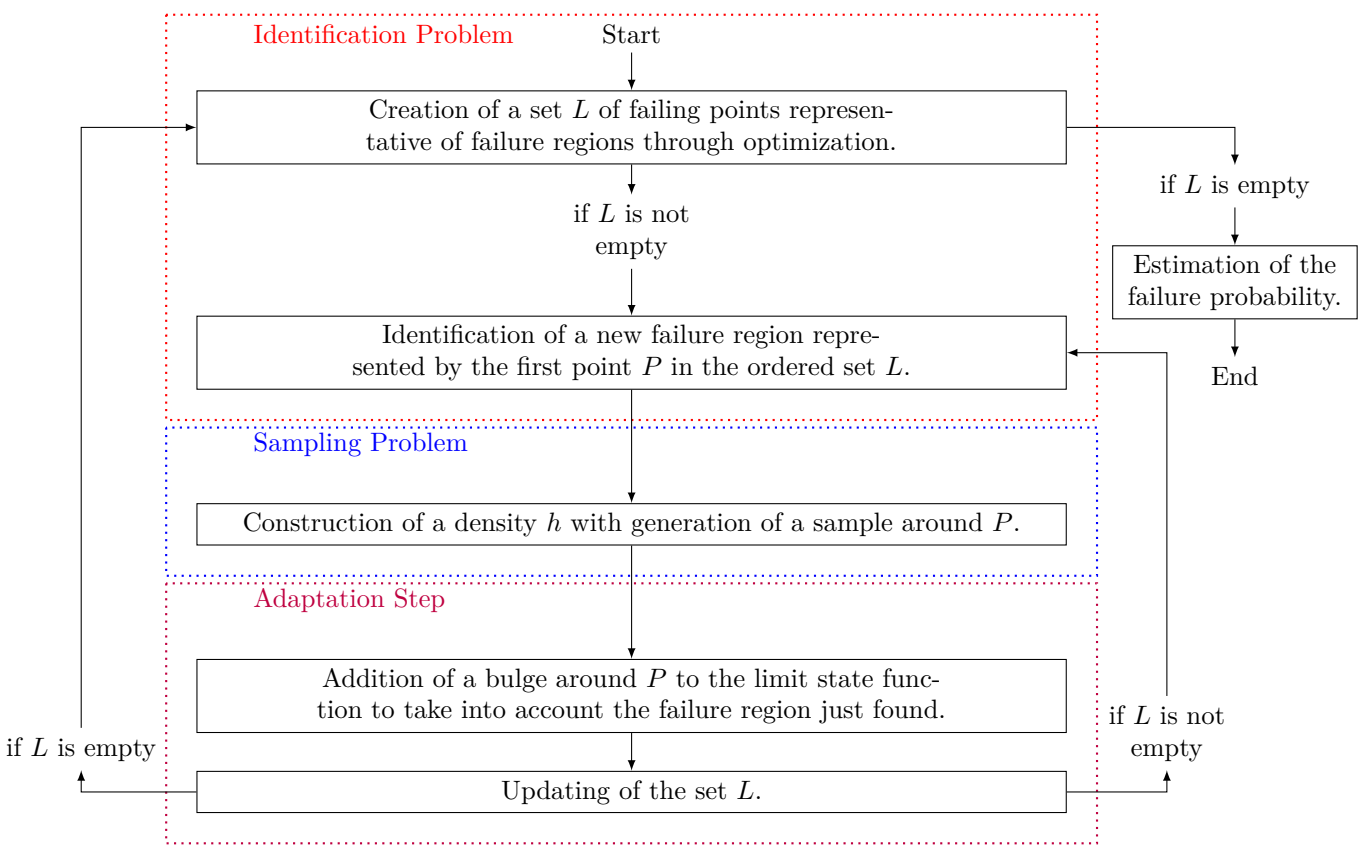


Figure 5.1: Scheme of the proposed method to find and model every failure region.

Section 5.2.2 describes the identification problem: the search for a set L of failing points which are representative of the failure regions, and the selection of a first failure region to be sampled. The sampling problem for this failure region is then considered in Section 5.2.3. The adaptation step in described in Section 5.2.4, as well as the loop which must be initiated when the set L is not empty after the adaptation step. The procedure to be followed when the set L is empty after the adaptation step is laid out in Section 5.2.5.

5.2.1.2 Presentation of the MIS estimate

At the end of this scheme, every failure region has been identified and sampled, and the failure probability can be estimated by the MIS framework; see Chapter 3 Section 3.3.3.3. For every failure region j, a number $n_j + 1$ of densities $h_{j,0}, \ldots, h_{j,n_j}$ are created successively in the construction of the optimal density for the sampling problem derived in Section 5.2.3. They result from the parameter optimization process with $h_{j,0}$ being the initial density and h_{j,n_j} the optimized density that best fits the failure region. Instead of using only the last density h_{j,n_j} and its sample to represent the failure region j, all the intermediates densities and samples are included to improve the stability of the estimate, as in AMIS framework [Cornuet et al., 2012]. As previously noted in Chapter 3, the AMIS method induces dependence amongst the different samples, which lead to a bias estimate of the failure probability. However, in the rest of this chapter, we assume that the dependence of the samples is negligible. This assumption will be discussed in Section 5.2 and Section 5.3.

Let K denote the number of failure regions. The stochastic representation of the inputs is taken advantage of, to propose the following MIS probability estimate

$$\hat{P}_f^{\mathrm{MIS}} = \frac{1}{N_{\mathrm{samp}}} \sum_{j=1}^{N_{\mathrm{samp}}} \mathbb{I}_{D_f}\left(R^{(j)}\mathbf{T}^{(j)}\right) \frac{f_R\left(R^{(j)}\right) f_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)}{h_{\mathrm{MIS}}\left(R^{(j)}\mathbf{T}^{(j)}\right)},$$

where the random pairs $\left(R^{(1)}, \mathbf{T}^{(1)}\right), \left(R^{(2)}, \mathbf{T}^{(2)}\right), \ldots$ are generated from h_{MIS}

$$\forall r \in \mathbb{R}^+ \quad \forall \mathbf{t} \in \mathcal{T}^d \quad h_{\mathrm{MIS}}(r\mathbf{t}) = \frac{1}{n_{\mathrm{total}}} \sum_{j=1}^{K} \sum_{\ell=0}^{n_j} h_{j,\ell}(r, \mathbf{t}).$$

Here, N_{samp} is the total number of observations generated in the various sampling problems performed for the estimation, and the total number of densities in the mixture h_{MIS} is n_{total}. The number n_{total} of IS auxiliary densities is equal to $n_1 + \cdots + n_K + K$. It should be noted that in this mixture, the weighting factors π_k are all equal to $1/n_{\mathrm{total}}$. Therefore all of the constructed densities have the same contribution to the mixture h_{MIS}. This topic will be discussed in Section 5.2.3.

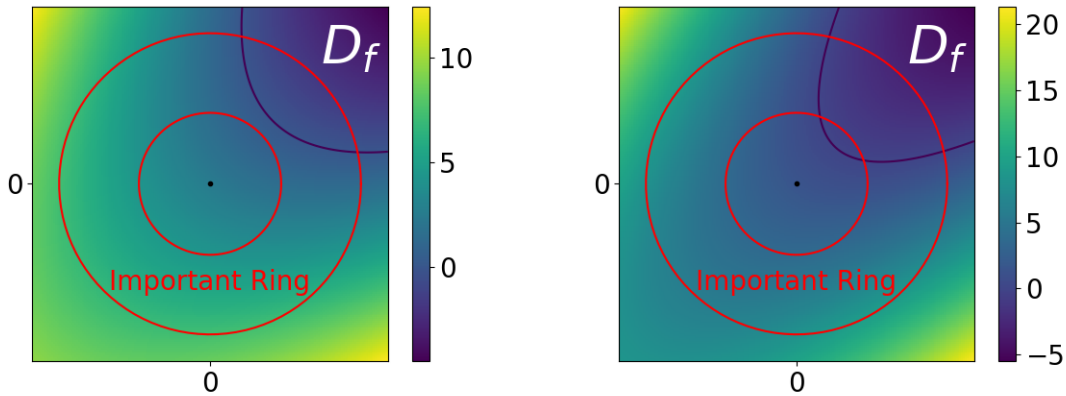
The following sections describe the successive steps of the method summarized in Figure 5.1 to determine a valuable IS mixture density h_{MIS}.

5.2.2 Identification problem

5.2.2.1 Search for the multiple failure regions

The search for the failure regions in the standard elliptical space is inspired by what is achieved in the approximation methods FORM/SORM for multiple design points detailed in Chapter 3 Section 3.2.2. The main idea is recalled here. In those methods, each failure region is identified with the failing point closest to the origin in the standard elliptical space [Der Kiureghian and Dakessian, 1998,Lebrun and Dutfoy, 2009]. Indeed the closest failing point, the so-called design point or most probable point [Grooteman, 2008], of a failure region is the most likely single realization of the input random variables that causes failure. Nevertheless, as reported in [Katafygiotis and Zuev, 2008], in high-dimensional standard normal space those design points seldom belong to the important ring. Hence the design points are of little interest in the search for the main failure regions in such a space, as the probability mass associated with their vicinity is most likely negligible. To take this issue into account, the solution proposed here is to look for the closest failing points inside the important ring.

For each failure region, depending on its location with respect to the important ring, two different situations must be considered. In Case 1, the closest failing point of the failure region is inside the



a) b)

Figure 5.2: Illustration of the location of the failure region and the important ring in a) Case 1 where the closest failing point of the failure region is inside the important ring, and in b) Case 2 where the closest failing point of the failure region is outside the important ring. The red circles delineate the lower bound and the upper bound of the important ring. In both examples, there is only one failure region in the failure domain, located in the upper right corner delimited in dark blue. The color legend represents the value of the limit state function. The representation is in dimension two, however the aim of the algorithm is high-dimensional spaces.

important ring. In Case 2, the closest failing point of the failure region does not belong to the important ring, as it is closer than the lower bound LB of the important interval, and the failure region spreads across the important ring. Case 2 is typical of contexts where the distribution of the radial component R leads to a particularly narrow important ring, as in the high-dimensional standard normal space.

Figure 5.2 provides schematic illustrations of the two cases. It should be noted that other configurations of the failure region position regarding the important ring exist, however, these two cases seem to be the most common situations encountered when testing the algorithm.

5.2.2.2 Optimization problems in the important ring

In each one of these situations, the failure region is identified with a failing point found as follows:

Case 1: The failure region is identified with the closest failing point. The local optimization problem is written in the form

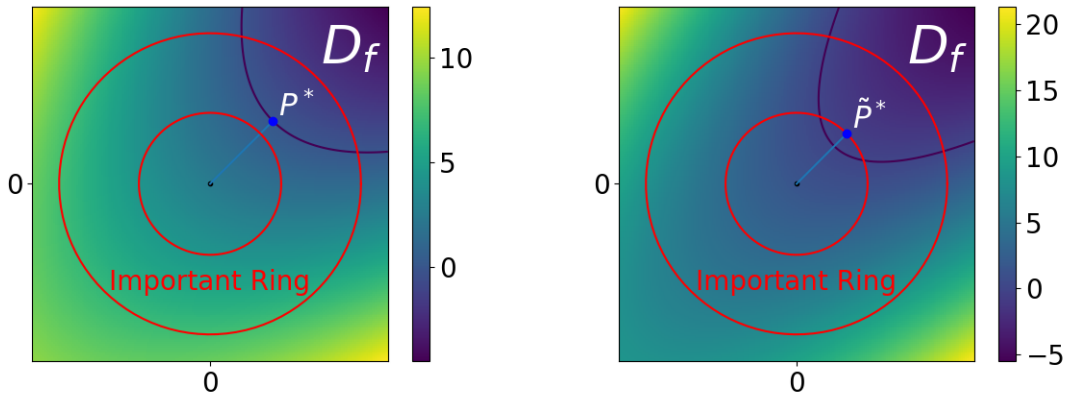
$$P^* = \underset{\mathbf{x} \in \text{Imp.Ring} \cap D_f}{\text{argmin}} \|\mathbf{x}\|, \tag{5.1}$$

where P^* is a point belonging to both the failure domain D_f and the important ring (Imp.Ring), which minimizes the Euclidean norm $\|\cdot\|$; see Figure 5.3 a). This optimization is exactly the same as performed in the FORM/SORM approximations and P^* is the design point of the system.

Case 2: The failure region is identified with a failing point resulting from two optimizations, as optimization Eq. (5.1) then has an infinite number of solutions: any point in the failure region on the lower bound of the important ring is a solution to problem Eq. (5.1). It is assumed that the optimizer then randomly picks up one of them, say P^*. A second optimization is performed to find a point that would better represent this particular failure region. The proposed solution consists of finding the point $\tilde{P}^*$ that minimizes the limit state function in the lower bound of radius LB of the important ring, viz.

$$\tilde{P}^* = \underset{\mathbf{x} \in D_f,\, \|\mathbf{x}\| = \text{LB}}{\text{argmin}} g(\mathbf{x}). \tag{5.2}$$

See Figure 5.3 b). The starting point of the second optimization is P^*. This local optimization is performed at fixed radius value with a gradient algorithm and hence relatively cheap in terms of limit state function evaluation.



a) b)

Figure 5.3: Illustration of the failing point representative of the failure region P^* in a) Case 1 and $\tilde{P}^*$ in b) Case 2. The blue line represents the Euclidean distance between the representative point and the origin. The red circles delineate the lower bound and upper bound of the important ring. In both examples, there is only one failure region in the failure domain, located in the upper right corner delimited in dark blue. The color legend represents the value of the limit state function.

Depending on the shape of the failure region and the important ring, the representative point is thus either the starting point P^* of this failure region in the important ring (Case 1) or the most negative point $\tilde{P}^*$ at the lower bound of the important ring for failure regions starting before the important ring (Case 2), as illustrated in Figure 5.3. The result of the first optimization Eq. (5.1) indicates which Case it is: if $\|P^*\|$ =LB then it is Case 2, else it is Case 1.

5.2.2.3 Selection of a first failure region

To find all the failure regions, these optimizations are repeated with a random multi-start technique which is detailed in Section 5.3. The resulting local minima P^* or $\tilde{P}^*$ representing the failure regions of the systems are gathered in a set L. It may happen that some points represent the same failure region. To select a first failure region for sampling, the representative points of the set L are ranked and sorted. In Case 1, representative points P^* of failure regions are ranked according to their norm $\|P^*\| >$ LB. In Case 2, representative points $\tilde{P}^*$ of failure regions are ranked according to the value of their lsf $g(\tilde{P}^*) < 0$, as their norms are equal to LB. The set L is then sorted in ascending order with regard to the ranking value of each representative point.

It should be noted that the representative points of Case 2 whose ranking values are negative are necessarily ranked before those of Case 1 whose ranking values are positive. This numerical trick allows to favor points representative of Case 2, as we expect them to be associated with larger failure regions than in Case 1.

A first failure region is then identified with the first point P_1 of the ordered set L. This point is thus the closest to the origin and its lsf evaluation is the most negative if its norm is equal to LB. The identification problem is summarized in Algorithm 6.

5.2.3 Sampling problem

5.2.3.1 Selection of a parametric family for the standard elliptical inputs

Taking advantage of the stochastic decomposition of the elliptical inputs, the parametric IS density used to sample in the failure region is constructed as the product of a density h_R on the random variable R and a density $h_{\mathbf{T}}$ on the random vector $\mathbf{T}$.

The selected radial distribution h_R is the original law of R, conditional on being greater than a

Algorithm 6 Identification Problem

Require: The lsf g and its gradient ∇g.
Initialization
The set L is empty. Generate a random vector $\mathbf{V}$ of starting points $\mathbf{X}_0$.
Multi-start local optimizations
for $\mathbf{X}_0 \in \mathbf{V}$ **do**
Search for the closest failing point with Eq. (5.1) with starting point $\mathbf{X}_0$. The result is P^*.
if $\|P^*\| = \text{LB}$ **then**
The failure region starts before the important ring: This is Case 2. Perform the optimization given in Eq. (5.2), with starting point P^*, to find the most negative point in the lower bound. The result is $\tilde{P}^*$. Add the representative point $\tilde{P}^*$ to the set L, with associated ranking value $g(\tilde{P}^*)$.
else
The closest failing point is inside the important ring: This is Case 1. Add the representative point P^* to the set L, with associated ranking value $\|P^*\|$.
end if
end for
Identification of a first failure region
Sort the set L in ascending order with respect to the ranking value of each representative point. Identify P_1 as the first point of the ordered set L.
return The ordered set L and a first identified failure region with P_1.

scalar parameter $r_{\text{opt}} \in \mathbb{R}^+$

$$\forall r \in \mathbb{R}^+ \qquad h_R(r; r_{\text{opt}}) = f_{R|R>r_{\text{opt}}}(r) = \frac{\mathbb{I}_{R>r_{\text{opt}}}(r) f_R(r)}{\mathbb{P}(R > r_{\text{opt}})}. \tag{5.3}$$

This choice of distribution for the radial component, which excludes part of the original distribution, is inspired by the RBIS framework presented in Chapter 3 Section 3.4.1.

The selected directional distribution $h_{\mathbf{T}}$ is a von Mises–Fisher distribution, a choice already proven to be efficient in [Zhang et al., 2022, Wang and Song, 2016, Papaioannou et al., 2019]. It depends on two parameters, namely the concentration parameter $\kappa \in \mathbb{R}^+$ and the mean direction parameter $\boldsymbol{\nu} \in \mathcal{T}^d$, cf Chapter 2 Section 2.2.1.3. It is recalled here that as κ grows, the vMF distribution is more and more concentrated around the mean direction ν; if $\kappa = 0$ then the vMF distribution is equivalent to the uniform distribution on the unit hypersphere $\mathcal{U}(\mathcal{T}^d)$.

The choice of these parametric densities leads to a total of $d+2$ scalar parameters to be set. For the first failure region identified above with P_1, the goal is thus to find the optimal parameters $r_{\text{opt},1}, \kappa_1$ and $\boldsymbol{\nu}_1$ such that the density $r, \mathbf{t} \mapsto h_R(r; r_{\text{opt},1}) \times h_{\mathbf{T}}(\mathbf{t}; \kappa_1, \boldsymbol{\nu}_1)$ is close to the optimal density of this failure region defined by

$$\forall r \in \mathbb{R}^+ \qquad \forall \mathbf{t} \in \mathcal{T}^d \qquad h_{\text{opt},1}(r, \mathbf{t}) = \frac{\mathbb{I}_{D_{f,1}}(r\mathbf{t}) f_R(r) f_{\mathbf{T}}(\mathbf{t})}{P_{f,1}},$$

where $D_{f,1}$ denotes the failure domain of this particular failure region and $P_{f,1}$ its probability. Instead of optimizing the parameters $r_{\text{opt},1}$, κ_1, and $\boldsymbol{\nu}_1$ of dimension $d+2$ with a Cross-Entropy (CE) algorithm as in [Wang and Song, 2016, Papaioannou et al., 2019, Kurtz and Song, 2013, Geyer et al., 2019], one can take advantage of the coordinates of the optimized point P_1 representative of the failure region as derived below. This simplifies the optimization of the parameters.

5.2.3.2 Setting of the parameters

Scalar parameter r_{opt} of the radial conditional distribution
The parameter $r_{\text{opt},1}$ defines the minimum value of R such that $R\mathbf{T}$ belongs to $D_{f,1}$, for all $\mathbf{T}$. Depending on the location of the failure region compared to the important ring, $r_{\text{opt},1}$ is set differently. Using the same notation as in Section 5.2.2 with Cases 1 and 2, the settings of $r_{\text{opt},1}$ are as follows:

Case 1: Set $r_{\text{opt},1} = \|P_1\|$. Indeed, P_1 is the closest failing point in the important ring. No point whose norm is below $r_{\text{opt},1}$ can be in the failure region of the important ring. This choice is the same as in the RBIS context, when there is only one failure region.

Case 2: The norm of P_1 equals the distance between the origin and the lower bound of the important ring. Thus, $r_{\text{opt},1}$ is set to 0 to avoid any bias, given that additional information on the radial component of the failure region is not available. Therefore, the distribution of the variable R is left unchanged in this particular case.

Mean direction $\boldsymbol{\nu}$ of the vMF distribution

The mean direction $\boldsymbol{\nu}_1$ is set to be the direction of P_1, viz. $\boldsymbol{\nu}_1 = P_1/\|P_1\|$. This decision implies that the direction of the point P_1 is where the likelihood of the density $h_{\mathbf{T}}$ is largest. It makes it possible to set the value of the vector $\boldsymbol{\nu}_1$ of dimension d directly with the coordinates of P_1, which are available after the optimizations of the identification problem described in Section 5.2.2.

Selecting the direction of P_1 as the most important direction for this failure region is relevant both in Cases 1 and 2. Indeed in Case 1, P_1 is the design point of the system; thus, selecting its direction as the most important of the failure region is the same idea as in the line sampling framework, described in Chapter 3 Section 3.4.2.1. And since P_1 belongs to the important ring, the issue discussed in Section 3.4.2.3 of Chapter 3 is solved: this direction is indeed relevant for the failure probability estimation of this particular failure region.

In Case 2, given that the evaluation of P_1 with the lsf has the most negative value in the lower bound of the important ring, the direction of P_1 is the most likely to be well centered with regard to the failure region location. In this case P_1 is not the *true* design point of the system. The true design point is located closer to the origin. In the line sampling framework, one would have chosen the direction of the true design point of the system as the most important direction. But as previously underlined in Section 3.4.2.3 of Chapter 3, this direction may be irrelevant in the important ring for nonlinear lsf. Therefore, selecting the most important direction with the direction of P_1 is more relevant, and allows the method to be more robust to nonlinear lsf.

Concentration parameter κ of the vMF distribution

Finally, the scalar κ_1 is set as the result of a Cross-Entropy based Adaptive MIS method (CE-AMIS) [Cornuet et al., 2012], presented in Section 3.3.3.3 of Chapter 3. Here, the parameter κ has to minimize the KL distance between the optimal IS density $r, \mathbf{t} \to h_{\text{opt},1}(r, \mathbf{t})$ of this particular failure region and the product density $r, \mathbf{t} \to h_R(r; r_{\text{opt},1}) \times h_{\mathbf{T}}(\mathbf{t}; \kappa, \boldsymbol{\nu}_1)$. The resulting updating formula for $\kappa_{1,i+1}$ in the stochastic representation of the inputs is written

$$\kappa_{1,i+1} \approx \underset{\kappa \in \mathbb{R}^+}{\operatorname{argmax}} \frac{1}{(i+1)N} \sum_{\ell=0}^{i} \sum_{j=1}^{N} \mathbb{I}_{D_{f,1}} \left(R_\ell^{(j)} \mathbf{T}_\ell^{(j)} \right) \ln \left(h_{\mathbf{T}} \left(\mathbf{T}_\ell^{(j)}; \kappa, \boldsymbol{\nu}_1 \right) \right) \frac{f_R \left(R_\ell^{(j)} \right) f_{\mathbf{T}} \left(\mathbf{T}_\ell^{(j)} \right)}{h_R \left(R_\ell^{(j)}; r_{\text{opt},1} \right) \sum_{p=0}^{i} \frac{1}{i+1} h_{\mathbf{T}} \left(\mathbf{T}_\ell^{(j)}; \kappa_{1,p}, \boldsymbol{\nu}_1 \right)}, \tag{5.4}$$

where the random variables $\left(R_\ell^{(j)}, \mathbf{T}_\ell^{(j)} \right)_{\ell=0,\dots,i; j=1,\dots,N}$ are generated with the multiple IS density $h_R(\cdot\,; r_{\text{opt},1}) \sum_{p=0}^{i} \frac{1}{i+1} h_{\mathbf{T}}(\cdot\,; \kappa_{1,p}, \boldsymbol{\nu}_1)$. Replacing the expression of h_R of Eq. (5.3) in Eq. (5.4) allows a simplification in the MIS weights

$$\kappa_{1,i+1} \approx \underset{\kappa \in \mathbb{R}^+}{\operatorname{argmax}} \frac{1}{(i+1)N} \sum_{\ell=0}^{i} \sum_{j=1}^{N} \mathbb{I}_{D_{f,1}} \left(R_\ell^{(j)} \mathbf{T}_\ell^{(j)} \right) \ln \left(h_{\mathbf{T}} \left(\mathbf{T}_\ell^{(j)}; \kappa, \boldsymbol{\nu}_1 \right) \right) \frac{\mathbb{P}(R > r_{\text{opt},1}) f_{\mathbf{T}} \left(\mathbf{T}_\ell^{(j)} \right)}{\sum_{p=0}^{i} h_{\mathbf{T}} \left(\mathbf{T}_\ell^{(j)}; \kappa_{1,p}, \boldsymbol{\nu}_1 \right) / (i+1)}. \tag{5.5}$$

If the indicator function of the failure domain and the IS weights were removed, this optimization would be exactly the same as a maximum likelihood estimation (MLE) of the vMF distribution parameters, when $\boldsymbol{\nu}$ is known. Therefore, as in [Papaioannou et al., 2019, Geyer et al., 2019], the equivalence between the CE optimization and the MLE is underlined here. With the indicator function and the IS weights, the optimization (5.5) becomes a weighted MLE problem whose solution is known. After

some calculations using formulas from [Banerjee et al., 2005, Wang and Song, 2016], the optimal κ_1 is found to be the solution of the equation

$$\frac{I_{d/2}(\kappa_{1,i+1})}{I_{d/2-1}(\kappa_{1,i+1})} = \frac{\mathbf{r}^\top \boldsymbol{\nu}_1}{\sum_{\ell=0}^{i}\sum_{j=1}^{N} \mathbb{I}_{D_{f,1}}\left(R_\ell^{(j)}\mathbf{T}_\ell^{(j)}\right) \frac{f_{\mathbf{T}}\left(\mathbf{T}_\ell^{(j)}\right)}{\sum_{p=0}^{i} h_{\mathbf{T}}\left(\mathbf{T}_\ell^{(j)};\kappa_{1,p},\boldsymbol{\nu}_1\right)/(i+1)}}, \tag{5.6}$$

where

$$\mathbf{r} = \sum_{\ell=0}^{i}\sum_{j=1}^{N} \mathbb{I}_{D_{f,1}}\left(R_\ell^{(j)}\mathbf{T}_\ell^{(j)}\right) \frac{f_{\mathbf{T}}\left(\mathbf{T}_\ell^{(j)}\right)}{\sum_{p=0}^{i} h_{\mathbf{T}}\left(\mathbf{T}_\ell^{(j)};\kappa_{1,p},\boldsymbol{\nu}_1\right)/(i+1)} \mathbf{T}_\ell^{(j)}.$$

the term $\mathbb{P}(R > r_{\text{opt},1})$ simplifies in the ratio. As one cannot solve Eq. (5.6) analytically, it is proposed in [Banerjee et al., 2005] to approximate $\kappa_{1,i+1}$ with

$$\kappa_{1,i+1} \approx \frac{\xi d - \xi^3}{1-\xi^2}, \tag{5.7}$$

where $\xi = I_{d/2}(\kappa_{1,i+1})/I_{d/2-1}(\kappa_{1,i+1})$. If $\kappa_{1,i+1}$ happens to be negative, then it is set to 0. In the CE-AMIS procedure, the failure probability is evaluated at each step of the algorithm. Here the formula presented in Chapter 3 Section 3.3.3.3 becomes

$$\hat{P}_{f,1,i} = \frac{1}{(i+1)N}\sum_{\ell=0}^{i}\sum_{j=1}^{N} \mathbb{I}_{D_{f,1}}\left(R_\ell^{(j)}\mathbf{T}_\ell^{(j)}\right) \frac{\mathbb{P}(R > r_{\text{opt},1}) f_{\mathbf{T}}\left(\mathbf{T}_\ell^{(j)}\right)}{\sum_{p=0}^{i} h_{\mathbf{T}}\left(\mathbf{T}_\ell^{(j)};\kappa_{1,p},\boldsymbol{\nu}_1\right)/(i+1)}, \tag{5.8}$$

And the theoretical variance is estimated along, with the following approximation

$$\hat{\mathbb{V}}\text{ar}\left(\hat{P}_{f,1,i}\right) \approx \frac{1}{(i+1)N}\mathbb{V}\text{ar}\left(\mathbb{I}_{D_{f,1}}(R\mathbf{T}) \frac{\mathbb{P}(R > r_{\text{opt},1}) f_{\mathbf{T}}(\mathbf{T})}{\sum_{p=0}^{i} h_{\mathbf{T}}(\mathbf{T};\kappa_{1,p},\boldsymbol{\nu}_1)/(i+1)}\right), \tag{5.9}$$

assuming all the samples are iid (we suppose the dependence of the samples to be negligible). With these equations, the value of $\kappa_{1,i+1}$ is gradually updated with the convergence of the CE-AMIS algorithm as follows.

1. Initialization. Set the index i equal to 0. Starting from a value $\kappa_{1,0} \neq 0$, an iid sample S_0 of size N is generated from $r, \mathbf{t} \to h_R(r, r_{\text{opt},1}) \times h_{\mathbf{T}}(\mathbf{t}, \kappa_{1,0}, \boldsymbol{\nu}_1)$. With this sample, the failure probability of this region $\hat{P}_{f,1,0}$ is estimated with the formula (5.8). The theoretical CV of $\hat{P}_{f,1,0}$ denoted CV_0, is given by $\text{CV}_0 = \sqrt{\hat{\mathbb{V}}\text{ar}\left(\hat{P}_{f,1,0}\right)}/\hat{P}_{f,1,0}$ where $\hat{\mathbb{V}}\text{ar}\left(\hat{P}_{f,1,0}\right)$ is given by E. (5.9). Then $\kappa_{1,1}$ is computed with population S_0 and Eqs. (5.7).

It is assumed that $\mathbb{I}_{D_{f,1}}\left(R_\ell^{(j)}\mathbf{T}_\ell^{(j)}\right) = \mathbb{I}_{D_f}\left(R_\ell^{(j)}\mathbf{T}_\ell^{(j)}\right)$ while the observations are generated with $\kappa_{1,p} \neq 0$ around $\boldsymbol{\nu}_1$. Thus in Eqs. (5.6), (5.7), (5.8) and (5.9) no distinction is made between those two failure domains. The situation where not a single $R_\ell^{(j)}\mathbf{T}_\ell^{(j)}$ from S_0 is failing may occur in two different settings. First, if the initial $\kappa_{1,0}$ is too small, then the observations are not concentrated enough around $\boldsymbol{\nu}_1$ and with a particularly narrow failure region, $\mathbb{I}_{D_{f,1}}\left(R_\ell^{(j)}\mathbf{T}_\ell^{(j)}\right) = 0$ for all S_0. Second, if the failure region associated with P_1 is extremely small or strongly nonlinear, then even with a large $\kappa_{1,0}$ all points of S_0 may belong to the safety domain. Indeed, it is recalled here that P_1 is located on the lower bound of the important ring in Case 2. Furthermore in Case 2 the distribution of the radial component R is left untouched: most of the observations generated will have a norm close to the $\mathbb{E}[R]$. Depending on the important ring setting, the lower bound can be very distant from $\mathbb{E}[R]$. Therefore, the region in the same direction than P_1 but with higher radial values may be in the safety domain. For both of these situations the Cross-Entropy optimization is stopped and the failure region associated with P_1 is considered negligible for the estimation of the failure probability.

2. Updating of the concentration parameter. While $\text{CV}_i \geq 10\%$, the index i is incremented: $i = i+1$. An iid sample S_i of size N is generated from $r, \mathbf{t} \to h_R(r; r_{\text{opt},1}) \times h_{\mathbf{T}}(\mathbf{t}; \kappa_{1,i}, \boldsymbol{\nu}_1)$. The failure

probability of this region $\hat{P}_{f,1,i}$ is estimated from the $i+1$ samples $S_0, \ldots, S_i$ generated, using the MIS formula (5.8). The theoretical CV of $\hat{P}_{f,1,i}$, denoted CV_i, is given by $\mathrm{CV}_i = \sqrt{\hat{\mathbb{V}}\mathrm{ar}\left(\hat{P}_{f,1,i}\right)}/\hat{P}_{f,1,i}$ where $\hat{\mathbb{V}}\mathrm{ar}\left(\hat{P}_{f,1,i}\right)$ is given by E. (5.9). Then $\kappa_{1,i+1}$ is updated with populations $S_0, \ldots, S_i$ and Eqs. (5.7).

3. End of the algorithm. Once the CV convergence criterion is reached, n_1 is set to $n_1 = i+1$. A last iid sample S_{n_1} of size N is generated from $r, \mathbf{t} \to h_R(r; r_{\mathrm{opt},1}) \times h_{\mathbf{T}}(\mathbf{t}; \kappa_{1,n_1}, \boldsymbol{\nu}_1)$ and the final probability of this failure region is computed $\hat{P}_{f,1} = \hat{P}_{f,1,n_1}$. The matrix M_1 is defined as the vector $M_1 = (S_0, \ldots, S_{n_1})$ of all the generated samples.

5.2.3.3 Remarks on the CE-AMIS algorithm for κ

Settings of the CE-AMIS structure

Compared to the CE-AMIS algorithm introduced in Chapter 3 Section 3.3.3.3, the last step differs. Indeed, once the CV criterion is reached, a final sample is generated with the final optimized value of κ_1. This choice was made on purpose, as we noticed that for some examples, the CV criterion was already reached in the initial step. This can occur if the position of the failing point P_1 is truly well located inside the failure region and for a particularly good choice of $\kappa_{1,0}$. Consequently, in the original CE-AMIS algorithm, the concentration parameter would not have been updated. And the resulting algorithm would be very dependent on the initial concentration parameter $\kappa_{1,0}$. Therefore, we added this extra sampling to prevent a strong dependence on the parameter setting of the CE-AMIS algorithm. Since this extra sampling was added, the CV criterion is set to a value quite large: 10%, as it is not the final theoretical CV of the probability of this failure region.

Here, the sample size of each step is the same and equal to N. This choice results in equal weights for each density $h_R(\cdot\,; r_{\mathrm{opt},1}) \times h_{\mathbf{T}}(\cdot\,; \kappa_{1,p}, \boldsymbol{\nu}_1)$ in the MIS mixture. As mentioned in the presentation of the CE-AMIS algorithm in Chapter 3 Section 3.3.3.3, a better setting of the size of the samples $S_0, \ldots, S_{n_1}$ could be considered, as several recent studies make it possible to compute optimal weights for each density; cf [Sbert and Havran, 2017, Sbert et al., 2018]. For simplicity, we chose to let the sizes equal. Moreover, it should be underlined that a recent paper [El-Laham et al., 2019] focus on selecting amongst the (n_1+1) densities generated in the AMIS procedure, a smaller number of components (a reduced mixture) for a better efficiency; it is the efficient AMIS (EAMIS). Nevertheless, for simplicity purposes again, we chose to keep all of the (n_1+1) densities for each step of the CE-AMIS algorithm.

Choice of the CE-AMIS method over the CE-AIS method

Updating $\kappa_{1,i+1}$ with a CE-AMIS algorithm rather than a CE-AIS algorithm as done in [Wang and Song, 2016, Papaioannou et al., 2019, Zhang et al., 2022] is a non-trivial choice. Indeed, it is well known that the AMIS structure does induce dependence between the different samples $S_0, \ldots, S_{n_1}$, which leads to a bias in the failure probability estimate [Cornuet et al., 2012], whereas the CE-AIS method does not. In fact, we observed during the investigation that if the CE-AMIS method is used to update both the scalar parameter κ and the mean direction $\boldsymbol{\nu}$ of a vMF density, a bias is then observable. Nevertheless, the parameter optimized here is only the scalar concentration κ, since $\boldsymbol{\nu}$ is already set. We noticed that in this situation, the bias of the failure probability estimate becomes then negligible, and the approximation of the theoretical variance, assuming all the samples are independent, is valid. Moreover, the concentration parameter κ of the vMF is known to be quite unstable in the CE-AIS algorithm. Indeed, it converges fast to high values and several techniques are used to slow down its convergence. In [Wang and Song, 2016, Zhang et al., 2022] a bisection technique is employed to modify the concentration parameter after its updating. Whereas in [Papaioannou et al., 2019], an upper bound is set on the value of ξ in Eq. (5.7) to prevent numerical instability. We did not encounter such instability with the CE-AMIS algorithm. This is due to the higher number of observations with non-zero weights in Eqs. (5.6) and (5.7), which is a property of AMIS: all of the generated observations are recycled instead of using only the ones generated at the current step.

It should also be noted that in classical adaptive Cross-Entropy methods, the initial parameters are often selected to retrieve the original input density cf Chapter 3 Section 3.3.3.2 and Section 3.3.3.4, as no information about the failure region is available. In contrast here, a search for the failure regions has led to the coordinates of the failing point P_1: an important direction has already been identified and a minimum value of R as well. Therefore, the first sample is generated directly towards the failure

region and has higher chances of belonging to the failure domain than any sample generated from $f_{\mathbf{X}}$. Consequently, the sampling procedure is already optimized and all the samples can be relevant for the CE procedure. Keeping all of the samples in an AMIS structure to update the concentration parameter and to estimate the failure probability of this region is thus justified. For the different reasons listed above, the CE-AMIS algorithm is more suited to our method than others CE-AIS algorithm; furthermore, the AMIS framework allows a better use of the costly lsf evaluations.

The accurate computation of the coordinates of P_1 with the local optimization problems of Section 5.2.2 is thus crucial in the efficiency of the CE-AMIS algorithm. Indeed, a greater number of iterations could be required to reach the CV convergence criterion if the mean direction $\boldsymbol{\nu}_1$ or $r_{\text{opt},1}$ are inaccurate. This would lead to a larger simulation budget. Therefore, the optimizations described in Section 5.2.2.2 must be performed with care.

5.2.4 Adaptation step

5.2.4.1 Construction of a bulge around the failure region

Once a first failure region has been identified and sampled, the search for the remaining failure regions can be initiated. As mentioned at the beginning of Section 5.2.2.1, the search for the failure regions is greatly inspired by the FORM method for multiple failure modes [Der Kiureghian and Dakessian, 1998] described in Section 3.2.2 of Chapter 3. This method is adjusted to our specific context and is part of the Adaptation step mentioned in Section 5.2.1.1. Once the first failure region is found, a bulge, denoted by Bulge_1, is built around the location of its associated representative failing point P_1 and added to the lsf g. The modified limit state function, written $\tilde{g}$, is then

$$\tilde{g}(\mathbf{x}) = g(\mathbf{x}) + \text{Bulge}_1(\mathbf{x}),$$

and the modified failure domain is written $\tilde{D}_f = \{r\mathbf{t} : \tilde{g}(r\mathbf{t}) \leq 0\}$. This failure region is thus not considered again in the search for the remaining failure regions, as it belongs to the safety domain.

In the proposed method, two different situations have been defined as Case 1 and Case 2 depending on the location of the failure region within the important ring. Given that the value of the lsf at the failing point $g(P_1)$ is not equal to zero in Case 2, a small modification has been added to the classical bulge equation presented in [Der Kiureghian and Dakessian, 1998] to introduce a bulge equation that suits both situations:

$$\text{Bulge}_1(\mathbf{x}) = \begin{cases} s_1 \left(\alpha_1^2 - \|\mathbf{x} - P_1\|^2\right)^2 - \mathbb{I}_{g(P_1)<0} \dfrac{g(P_1)}{\alpha_1} \left(\alpha_1 - \|\mathbf{x} - P_1\|\right) & \text{if } \|\mathbf{x} - P_1\| \leq \alpha_1, \\ 0 & \text{elsewhere,} \end{cases}$$

where $\mathbb{I}_{g(P_1)<0} = 1$ when $g(P_1) < 0$ and 0 otherwise. This bulge is continuous and differentiable. The resulting modified failure regions in those two situations are illustrated in Figure 5.4.

The scalar α_1 is the radius of the bulge defined by $\alpha_1 = \gamma \times \beta_1$ and s_1 is the scale of the bulge, defined by

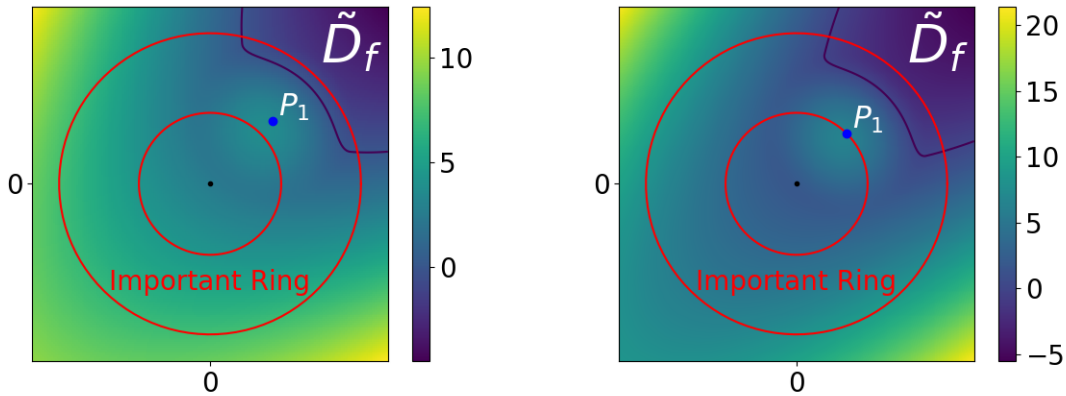
$$s_1 = \frac{\delta\beta_1 \|\nabla g(P_1)\|}{\{(\gamma\beta_1)^2 - (\delta\beta_1)^2\}^2},$$

where $\delta = 0.75$ and $\gamma = 1.1$ as in [Der Kiureghian and Dakessian, 1998], and $\beta_1 = \|P_1\|$. Any point in the bulge is thus no longer in the updated failure domain $\tilde{D}_f$. A **first updating** of the ordered set L is made to remove any of those points.

5.2.4.2 Definition of a failure cone associated with the failure region

In Case 2, P_1 is located on the lower bound of the important ring and hence $\beta_1 = \text{LB}$. However, most of the observations generated are not located around P_1. Indeed, as the distribution of the radial component R is left unchanged (see Section 5.2.3.2), most observations have a radial value closer to the mean value $\mathbb{E}[R]$ of R than LB. Therefore, the bulge built around P_1 is not representative of the samples generated in this failure region. A failing point belonging to the same failure region as P_1 may be found at a distance from P_1 superior to α_1.

In Case 1, the same situation may happen if the radius α_1 of the bulge is not large enough. Thus another criterion, inspired by the angle criterion in [Der Kiureghian and Dakessian, 1998], is introduced



a) b)

Figure 5.4: Illustration of the bulge built around the representative point P_1 of the failure region and the resulting modified failure domain $\tilde{D}_f$ in a) Case 1 and b) Case 2. This illustration should be compared to Figure 5.3 for a better understanding. The red circles represent the lower bound and upper bound of the important ring. In both examples, there is only one failure region in the failure domain, located in the upper right corner delimited in dark blue. The color legend represents the value of the modified limit state function $\tilde{g}$.

for both Case 1 and Case 2 in order to prevent this phenomenon: the maximum angle ϕ_1 between P_1 and the failing points of the population M_1 is estimated and the failure cone C_1 associated with P_1 is defined as $C_1 = \{\mathbf{x} : \widehat{\mathbf{x}P_1} \leq \gamma \times \phi_1\}$. Any failing point belonging to this cone is representative of the same failure region and must also be removed from the ordered set L to avoid redundancy, it is the **second updating**.

Adding an angle criterion is relevant in high-dimensional spaces, as it provides a measure that is independent of the dimension d of the system. Relying only on the distance criterion α_1, defined with the Euclidean norm that does depend on d, becomes pointless in very high dimensions, as the distances tend to be identical for all data; it is the *concentration of norm phenomenon* [Klawonn et al., 2012], already mentioned in Chapter 3 Section 3.3.3.2 in the clustering algorithms context and in Section 4.2.2 in the context of LS for multiple design points.

5.2.4.3 Identification of the other failure regions

After this second updating, if the ordered set L is not empty, then the remaining points represent some failure regions that are different from the first one and need to be considered. The next failure region is identified with the first point of the updated ordered set L, denoted as P_2. The sampling problem (Section 5.2.3) is then repeated with this failure region represented by P_2. The failure domain used in the sampling problem is still $D_f = \{r\mathbf{t} : g(r\mathbf{t}) \leq 0\}$ without the modification $\tilde{g}$ defined above. It prevents introducing a bias in the probability estimation. At the end of the sampling problem, the second failure region represented by P_2 has been properly sampled and another auxiliary density has been iteratively constructed. A second bulge Bulge_2 is created in the same way as Bulge_1, so that $\tilde{g}$ is updated again. The ordered set L is also updated with the distance criterion α_2 and the cone criterion C_2.

This scheme goes on until the set L is empty, with the number of failure regions found equal to k. Algorithm 7 sums up the process of sampling in all failure regions represented by points in the set L. The modified lsf is initialized with $\tilde{g} = g$ and the number of failure regions k is initialized with $k = 0$ as there is no identified failure region at the start of the algorithm.

5.2.5 Search for missed failure regions

The set L found in the identification problem presented in Section 5.2.2.3 may not include all the existing failure regions. For example, one failure region may be so dominant that all the optimized points

Algorithm 7 Sampling in all failure regions of the ordered set L

Require: The ordered set L, the lsf g, the modified lsf $\tilde{g}$ and the number k of failure regions.
Construction of densities and sampling
while $L \neq \emptyset$ **do**
Set $k = k + 1$. Select P_k as the first element of the ordered set L.
Sampling Problem: Construct an optimal auxiliary density for the failure region represented by P_k and generate a population M_k in this failure region as in Section 5.2.3.
Adaptation Step: Add the Bulge$_k$ to the modified lsf $\tilde{g}$ and update the set L by removing points closer to α_k from P_k and points belonging to the cone C_k.
end while
All the failure regions represented by failing points in the set L have been sampled.
return The number k of failure regions and the modified lsf $\tilde{g}$.

of the set L have converged to this particular region. Therefore, some failure regions that are slightly harder to find by optimization may be missed. To solve this problem, another set of optimizations is performed, to find the failure regions that could have been overshadowed. The optimizations are the same as the one presented in Section 5.2.2.2 with Eqs. (5.1) and (5.2) except for two updates. First, the failure domain is no longer D_f but the modified one $\tilde{D}_f = \{r\mathbf{t} : \tilde{g}(r\mathbf{t}) \leq 0\}$. Thereby, all the bulges constructed earlier prevent the optimization from converging back to the same points, because they are no longer in the failure domain. Second, the multi-start technique of the first optimization Eq. (5.1) is centered around the opposite direction $\mathbf{x}_0$ of the failing points previously found, as suggested in the FORM method with multiple design points [Der Kiureghian and Dakessian, 1998]

$$\mathbf{x}_0 = -\epsilon(P_1 + \cdots + P_k), \tag{5.10}$$

where $\epsilon = 0.5$. After the optimizations, a new set L is found. The failing points of this set are all located at least α_i-away from each point P_i with $i \in \{1, \ldots, k\}$. After removing the failing points of the set belonging to the cone C_i of each P_i, the set L contains only failing points representative of failure regions that have not been sampled yet. It is sorted as described in Section 5.2.2.3. Algorithm 7 is then repeated for this ordered set L until is it empty. Algorithm 8 sums up the sampling of all failure regions of the failure domain.

Algorithm 8 Search and sampling of all failure regions of the failure domain

Require: The lsf g and its gradient ∇g.
Initialization
Identification Problem: Create an ordered set L as output of the Algorithm 6 with inputs $(g, \nabla g)$.
Set the number of failure region $k = 0$. Initialize the modified lsf with $\tilde{g} = g$
Sampling all failure regions represented by failing points in the set L
while $L \neq \emptyset$ **do**
Sampling Problem and Adaptation Step: Generate a population in every failure region represented by failing points in the ordered set L as derived in Algorithm 7 with inputs $(L, g, \tilde{g}, k)$ and outputs $(k, \tilde{g})$.
Identification Problem: Create a new set L with the optimizations presented in Section 5.2.2 with the modified lsf $\tilde{g}$ and the modified failure domain $\tilde{D}_f = \{r\mathbf{t} : \tilde{g}(r\mathbf{t}) \leq 0\}$ and $\mathbf{x}_0 = -\epsilon(P_1 + \cdots + P_k)$. Update this set L by removing points located within the cones C_i for each P_i with $i \in \{1, \ldots, k\}$ and sort the set L.
end while
All failure regions of the failure domain have been identified and sampled.
Set K equal to the last value of k, such that K is the final number of failure regions identified and sampled, in the failure domain.
return The total number of failure regions K.

The algorithm ends when the new set L is found to be the empty set. It means that no other failure region has been found, i.e., the entire failure domain has (presumably) been covered. It should be noted that the bulges keep getting added to the modified lsf $\tilde{g}$ through the different created sets L. It is recalled here that the modified limit state function is used only for the optimizations performed

in the identification problems, and not for the construction of the auxiliary densities in the sampling problems.

5.2.6 Estimation of the global failure probability

5.2.6.1 Final MIS estimate expression

The global failure probability is estimated at the end of Algorithm 8 with the MIS estimate presented before. The global IS auxiliary density is a mixture of a number n_{total} of densities structured as $h_R \times h_{\mathbf{T}}$. The final equation of the estimate is

$$\hat{P}_f^{\text{MIS}} = \frac{1}{n_{\text{total}} \times N} \sum_{j=1}^{n_{\text{total}} \times N} \mathbb{I}_{D_f}\left(R^{(j)}\mathbf{T}^{(j)}\right) \frac{f_R\left(R^{(j)}\right) f_{\mathbf{T}}\left(\mathbf{T}^{(j)}\right)}{h_{\text{MIS}}\left(R^{(j)}, \mathbf{T}^{(j)}\right)},$$

where the random pairs $\left(R^{(1)}, \mathbf{T}^{(1)}\right), \left(R^{(2)}, \mathbf{T}^{(2)}\right), \ldots$ are generated from h_{MIS}

$$\forall r \in \mathbb{R}^+ \quad \forall \mathbf{T} \in \mathcal{T}^d \quad h_{\text{MIS}}(r\mathbf{t}) = \frac{1}{n_{\text{total}}} \sum_{k=1}^{K} h_R(r; r_{\text{opt},k}) \left\{ \sum_{\ell=0}^{n_k} h_{\mathbf{T}}(\mathbf{t}; \kappa_{k,\ell}, \boldsymbol{\nu}_k) \right\}.$$

Thus $N_{\text{samp}} = n_{\text{total}} \times N$. As previously noted, the number n_{total} of IS auxiliary densities is equal to $n_1 + \cdots + n_K + K$. The number n_{total} is thus quite larger than in other algorithms where only one density is associated with a particular failure region k [Papaioannou et al., 2019, Geyer et al., 2019]. We compared the proposed MIS estimate with a *lighter* version, where only the last density of the CE-AMIS algorithm and its sample were kept in a MIS estimate. The *full* MIS estimate has proven to be the more stable in terms of empirical CV than the *lighter* version, therefore this framework was selected.

This estimate does not require any prior knowledge of the number K of failure regions, as it is computed within the algorithm. The failure probability associated with each failure region $\hat{P}_{f,k}$ is computed as well and it can be an interesting quantity to understand the contribution of each failure region to the global failure probability. If the failure regions are disjoint, then the crude sum of these failure probabilities gives the same estimate than the proposed MIS estimate. However if they are not disjoint, the proposed MIS estimate takes into account the union of these regions while the crude sum of the $\hat{P}_{f,k}$ does not and returns a higher biased failure probability value.

It should be noted that if there exist some errors in the gradient estimation or in the local optimizations performed during the identification steps, the identification of the different failure regions and the coordinates of their representative point can be impacted. The resulting MIS probability estimate may be biased if the parameters $r_{\text{opt},k}$ are not identified well. Therefore, the quality of the proposed MIS estimate highly depends on the accuracy of the optimizations presented in Section 5.2.2.2.

The samples associated with different failure regions are independent, however there is a small dependence amongst the samples associated with the same failure region, as previously mentioned in the CE-AMIS algorithm in Section 5.2.3.3. We assume this dependence to be negligible to approximate the theoretical variance of the estimate as such

$$\hat{\mathbb{V}}\text{ar}\left(\hat{P}_f^{\text{MIS}}\right) \approx \frac{1}{n_{\text{total}} \times N} \mathbb{V}\text{ar}\left(\mathbb{I}_{D_f}\left(R\mathbf{T}\right) \frac{f_R\left(R\right) f_{\mathbf{T}}\left(\mathbf{T}\right)}{h_{\text{MIS}}\left(R\mathbf{T}\right)}\right). \tag{5.11}$$

5.2.6.2 Discussion on the simulation budget allocation

The number $(n_k + 1)$ of densities associated with the failure region k depends on the CV criterion of the CE-AMIS algorithm presented in Section 5.2.2.2. The harder it is to sample in the failure region k and have a good estimation of $\hat{P}_{f,k}$, the higher this number $(n_k + 1)$ is and the larger the generated population $M_k = (S_0, \ldots, S_{n_k})$ gets, as N is fixed. Thus in the final estimate, a large population could be associated with a failure region particularly hard to reach but not necessarily important in the failure domain, i.e with a small contribution to the failure probability. In other algorithms, which consider the failure domain in its entirety, this issue would not happen, as the algorithms would stop as long as enough observations are in the failure domain, no matter the region (for algorithms with ρ-quantile criterion for instance). Therefore, this could be a drawback of the proposed method. The

global simulation budget may not be pertinently distributed amongst the different failure regions. However, as a consequence, smaller failure regions are theoretically not left out or overshadowed by other larger failure regions. This situation may occur with the other algorithms mentioned.

To keep track of the simulation budget, the evaluations of the lsf and its gradient needed in the proposed method is now further detailed. The lsf g is evaluated through several repetitions of the identification problem (Section 5.2.2) and the sampling problem (Section 5.2.3). The numbers $N_{\text{opt},1}$ and $N_{\text{opt},2}$ denote the global numbers of evaluations of g and ∇g, respectively, required for all the optimizations performed in the identification problems, and N_{samp} is the final number of lsf evaluations required in the sampling problems. The total number of lsf evaluations is equal to $N_{\text{opt},1} + N_{\text{samp}}$. It is reminded here that a greater simulation budget N_{samp} may be required to reach the CV convergence criterion of the CE-AMIS algorithm if the sampling density is sub-optimal because of inaccurate mean direction $\boldsymbol{\nu}_k$ or $r_{\text{opt},k}$. Therefore the simulation budget is strongly dependent on the quality of the optimizations presented in Section 5.2.2.2.

5.2.6.3 Example in dimension 2: A 4-branch system

To clarify the procedure, an example in dimension 2 is presented. This example is meant to be for illustrative purposes only, as the goal of the algorithm is to estimate the failure probability of systems involving a large number of input variables. The following example is frequently found in the literature and was used as an illustration example for the different failure probability algorithms of Chapter 3. The inputs are assumed to be Gaussian $\mathbf{X} \sim \mathcal{N}(\mathbf{0}_2, \mathbf{I}_2)$, and the lsf equation is given by

$$g(\mathbf{X}) = \min \left\{ \begin{array}{c} (X_1 - X_2)^2/10 - (X_1 + X_2)/\sqrt{2} + 3 \\ (X_1 - X_2)^2/10 + (X_1 + X_2)/\sqrt{2} + 3 \\ (X_1 - X_2) + 7/\sqrt{2} \\ (X_2 - X_1) + 7/\sqrt{2} \end{array} \right\}.$$

The reference value for the failure probability is 2.22×10^{-3}. Figure 5.5 displays the function g and the contour of the failure domain.

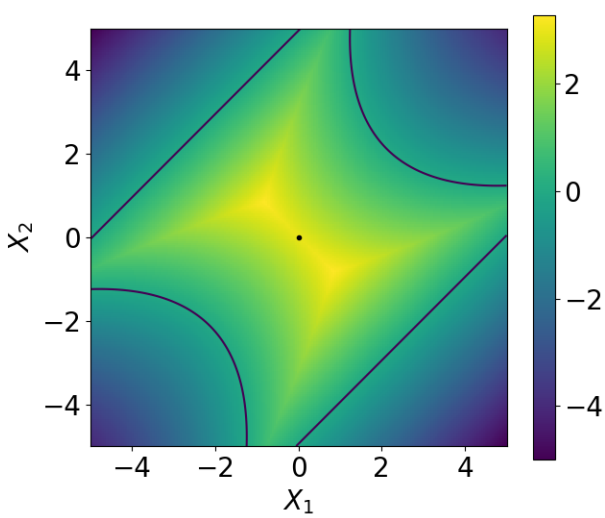


Figure 5.5: Limit state function for the 4-branch system. This system has four failure regions located at the four corners of the input space. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface.

The first step of the algorithm is to find P_1 through optimization (Section 5.2.2); this is the identification problem. Here the dimension is too low for the important ring to be any restrictive; it encompasses the whole input space. Hence, all the failure regions start inside the important ring.

This application is thus an example of the situation referred to as Case 1 in the previous sections. After the identification problem (with a random multi-start of size 10), the set L contains only three different representative points, $(-2.12, -2.12)$, $(2.12, 2.12)$ and $(2.47, -2.47)$. The Euclidean norm of both $(2.12, 2.12)$ and $(2.12, -2.12)$ is very close to 3 while the Euclidean norm of $(2.47, -2.47)$ is close to 3.5; thus it is last in the ordered set. This first identification problem has missed the fourth failure region.

The first point of the set is $P_1 = (-2.12, -2.12)$ with $r_{\text{opt},1} = 2.99$ and $\boldsymbol{\nu}_1 = (-0.71, -0.71)$. The sampling of the failure region around P_1 is performed (Section 5.2.3) with $\kappa_{1,0} = 20$ and $N = 1000$. The first CV_0 computed is equal to 2.6%, thus κ is only updated once and its final value is 33.5. The population M_1 of size $2N$ is illustrated in Figure 5.6.

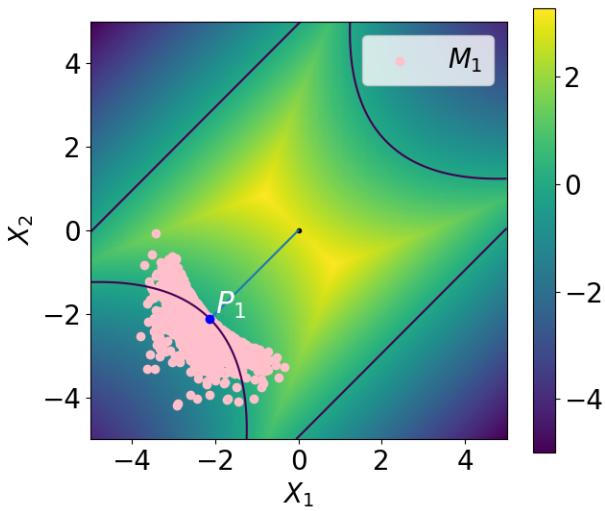


Figure 5.6: Sampling of the first failure region of the 4-branch system. One has $P_1 = (-2.12, -2.12)$, $r_{\text{opt},1} = 2.99$ and $\boldsymbol{\nu}_1 = (-0.71, -0.71)$. The population M_1 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{1,0} = 20$ and $\kappa_{1,1} = 33.5$ and is represented with pink dots. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface.

The Bulge_1 is added to the limit state function g to create the modified lsf $\tilde{g}$ and the set L is updated (adaptation step). The same scheme is repeated for $P_2 = (2.12, 2.12)$ and $P_3 = (2.47, -2.47)$ (Section 5.2.4). Populations M_2 and M_3 are displayed in Figure 5.7. The set L is empty at this point, and the search for missed failure regions begins (identification problem, Section 5.2.5) with a new set L resulting of the optimization on the modified limit state function $\tilde{g}$, displayed in Figure 5.8.

The new set L results in the singleton $\{(-2.47, 2.47)\}$. The sampling of the fourth failure region is performed (sampling problem), Bulge_4 is added to $\tilde{g}$ and once this is done, the set L is empty (adaptation step). The resulting population M_4 is displayed in Figure 5.9 as well as the new modified lsf.

The new set L resulting from the optimization fails to detect any failure region different from the ones previously found (identification problem) and thus the algorithm stops. The failure probability is estimated with a MIS estimate of eight densities, two for each failure region. The estimated probability is equal to 2.23×10^{-3}, with a theoretical CV equal to 1.3%. The total number of evaluations of the gradient of g for the optimizations of the identification problems is equal to $N_{\text{opt},2} = 330$. The number of evaluations of g comprises $N_{\text{opt},1} = 476$ evaluations for the optimizations of the identification problems, and $N_{\text{samp}} = 4 \times 2 \times 1000 = 8000$ evaluations for the sampling problems of the four failure regions.

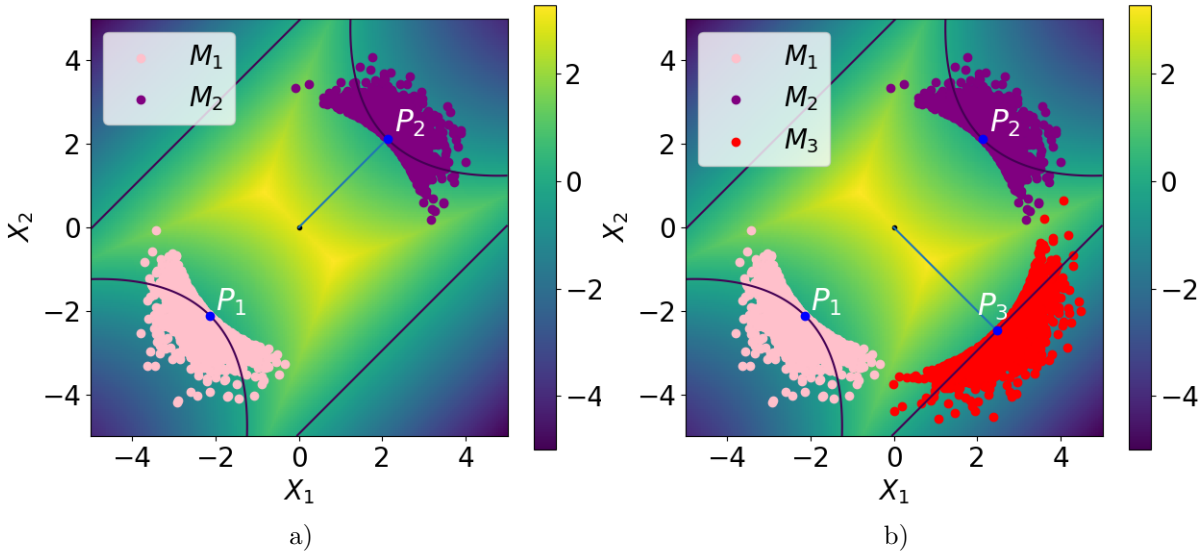


Figure 5.7: Sampling of a) the second and b) third failure regions of the 4-branch system. One has $P_2 = (2.12, 2.12)$, $r_{\text{opt},2} = 2.99$, and $\boldsymbol{\nu}_2 = (0.71, 0.71)$. Population M_2 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{2,0} = 20$ and $\kappa_{2,1} = 27.6$. Moreover, $P_3 = (2.74, -2.47)$, $r_{\text{opt},3} = 3.49$, and $\boldsymbol{\nu}_3 = (0.71, -0.71)$. Population M_3 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{3,0} = 20$ and $\kappa_{3,1} = 18.1$. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface.

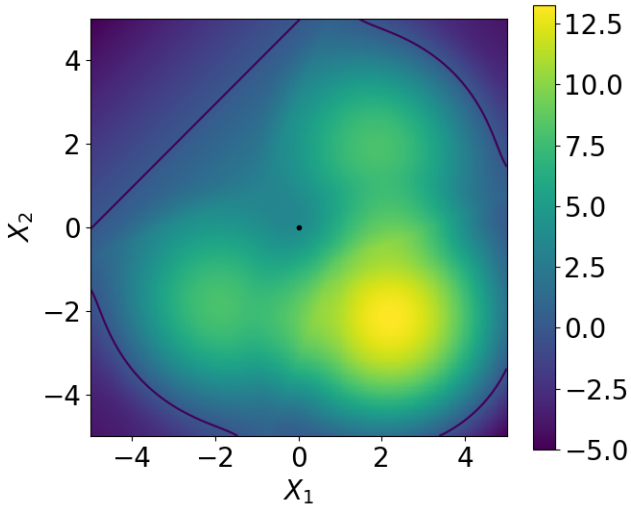


Figure 5.8: The modified limit state function $\tilde{g}$ after the additions of Bulge$_1$, Bulge$_2$, and Bulge$_3$. The color legend represents the value of the modified limit state function $\tilde{g}$. The dark blue line is the modified limit state surface.

5.3 Numerical investigations with high-dimensional input space

The performance of the proposed algorithm is investigated with three numerical examples taken from the IS literature and with a realistic aerodynamic engineering application. The first two applications are compared with the improved Cross-Entropy vMF-Nakagami Mixture algorithm (iCE-vMFNM) of [Papaioannou et al., 2019] presented in Section 3.3.3.4 of Chapter 3 and with a subset simulation

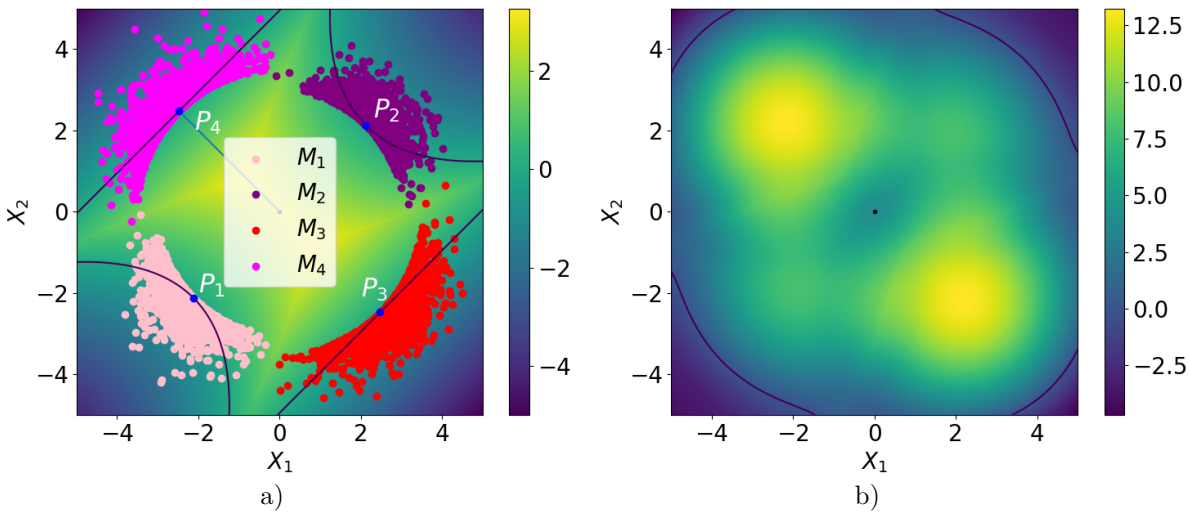


Figure 5.9: a) Sampling of the fourth failure region of the 4-branch system. One has $P_4 = (-2.74, 2.47)$, $r_{\text{opt},4} = 3.49$, and $\boldsymbol{\nu}_4 = (-0.71, 0.71)$. Population M_4 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{4,0} = 20$ and $\kappa_{4,1} = 16.4$. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface. b) The modified limit state function after the addition of Bulge$_4$ to the previous modified limit state function. The color legend represents the value of the modified limit state function $\tilde{g}$. The dark blue line is the modified limit state surface.

(SS) algorithm [Au and Beck, 2001] presented in Section 3.3.2 of Chapter 3. The Monte Carlo method is used for reference value. The inputs of the first two examples and the realistic engineering application follow the standard Gaussian distribution, while the inputs of the third example follow a multivariate Student distribution.

As previously noted, all the examples involve inputs of dimension 100 and above. Nevertheless, as shown at the end of the previous section, the algorithm can also perform well in small dimension, even though it is not the primary objective of the proposed method.

To compare the different methods, 500 independent simulation runs were performed to calculate the statistics of the probability estimates and the other quantities of interest, except for the last example where the algorithm ran only once.

5.3.1 Series system of four linear limit state functions

5.3.1.1 Presentation of the application

The first example is a series system problem with four components for which the number of random variables can be chosen without affecting the probability of failure; it was inspired by a numerical application from [Wang and Song, 2016] and [Papaioannou et al., 2019]. This first example demonstrates the robustness of the proposed algorithm to find all the failure regions in the high-dimensional input

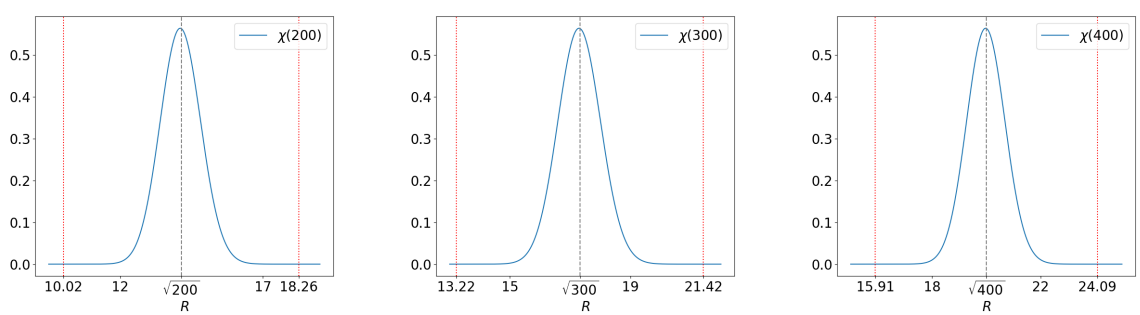
space. It is defined standard Gaussian space $\mathbf{X} \sim (\mathbf{0}_d, \mathbf{I}_d)$ and the lsf is given by

$$g_1(\mathbf{X}) = \min \left\{ \begin{array}{c} \beta + \frac{1}{\sqrt{d}} \sum_{i=1}^{d} X_i \\ \beta - \frac{1}{\sqrt{d}} \sum_{i=1}^{d} X_i \\ \beta + \frac{1}{\sqrt{d}} \left(\sum_{i=1}^{d/2} X_i - \sum_{i=d/2+1}^{d} X_i \right) \\ \beta + \frac{1}{\sqrt{d}} \left(-\sum_{i=1}^{d/2} X_i + \sum_{i=d/2+1}^{d} X_i \right) \end{array} \right\}. \tag{5.12}$$

The failure probability is then independent of the number d of random variables. In order to have a low failure probability, the β parameter is set to $\beta = 5$. The reference probability value computed with the Monte Carlo method is equal to $P_f = 1.15 \times 10^{-6}$ with an empirical CV of 4.1% and a sample of size 5×10^8. There are four failure regions in this failure domain, each associated with a limit state function.

5.3.1.2 Definition of the important ring

The performance of the proposed method is studied in dimensions 200, 300, and 400. As the inputs are standard normal, one has $R \sim \chi(d)$. The important ring is defined as in [Katafygiotis and Zuev, 2008]: the important interval [LB, UB] is centered around $\sqrt{d}$ and the smallest value of ϵ such that $\mathbb{P}(\sqrt{d} - \epsilon \leq R \leq \sqrt{d} + \epsilon) \geq 1 - 10^{-8}$ is selected. It should be noted that another definitions of the important ring is possible, cf Section 2.2.2.1 of Chapter 2. The interval is $[10.02, 18.26]$ for $d = 200$, $[13.22, 21.42]$ for $d = 300$ and $[15.91, 24.09]$ for $d = 400$, and is displayed in Figure 5.10.



a) b) c)

Figure 5.10: Illustration of the important interval of the $\chi(d)$ distribution for the three dimensions tested a) $d = 200$, b) $d = 300$ c) $d = 400$. The dashed line represents the value of $\sqrt{d}$, the red dotted lines are the two important interval bounds.

The four design points of the system are straightforward. For the first two linear lsfs, the design points are $\left\{-(\beta/\sqrt{d})\mathbf{1}_d, (\beta/\sqrt{d})\mathbf{1}_d\right\}$ where $\mathbf{1}_d = [1, \ldots, 1]$ and for the other two linear lsfs they are $\left\{-(\beta/\sqrt{d})\bar{\mathbf{1}}_d, (\beta/\sqrt{d})\bar{\mathbf{1}}_d\right\}$ where $\bar{\mathbf{1}}_d$ is the vector $\bar{\mathbf{1}}_d = [1, \ldots, 1, -1, \ldots, -1]$ in which the change of sign occurs at the $(d/2+1)$-th component. Each of these design points has a norm equal to β, thus they do not belong to any of the important rings defined above. Therefore each of the four failure regions starts before the beginning of the important ring and spreads across it. Thus this application is an example of the situation referred to as Case 2 in the previous sections.

It should be noted that although the failure regions are not all disjoints, the failure probability value is very close to the crude sum of the four FORM estimates of each failure regions $P_f \approx 4 \times \Psi(-\beta) = 1.15 \times 10^{-6}$, where Ψ is the cdf of an univariate standard normal law; cf Section 3.2.1 Chapter 3. This results from the linearity of each limit state function.

5.3.1.3 Settings of the proposed algorithm

The optimizations presented in Section 5.2.2.2 are performed with SLSQP algorithm [Sahin, 2019], where the precision goal for the value of the objective function in the stopping criterion is set to 10^{-3} for both of the optimizations Eq. (5.1) and Eq. (5.2). The maximum number of iterations in the algorithm is set to 1000. Although the optimizations must be very accurate to improve the efficiency of the CE-AMIS algorithm, the simulation budget associated with these optimizations ($N_{\text{opt},1}$ and $N_{\text{opt},2}$) must not be too large either.

The random vector for the multi-start technique mentioned in Section 5.2.2.3 and Section 5.2.5 in the search for the failure regions is generated with Latin Hypercube Sampling (LHS) [Shields et al., 2015] of size 10. LHS is a sampling method developed for independent inputs which aims to cover the inputs space with a stratified sampling strategy. In the proposed method, LHS is thus performed with standard normal inputs in the first identification problem of Section 5.2.2.3 and then with independent normal inputs of unit variance centered on $\mathbf{x}_0$ of Eq. (5.10) in the following identification problems of Section 5.2.5. Although the choice of 10 for the sampling size is arbitrary, it proved to be sufficient for the example considered here.

In the CE-AMIS algorithm, N is set to 1000 and the maximum number of iterations, denoted $p_{\max}$ is set to $p_{\max} = 10$. The number $p_{\max}$ is rather low, but it is adequate as the only updated parameter is the scalar concentration parameter κ of the vMF density.

One should remark that the limit state function g_1 is not differentiable at all points: the gradient is not defined for all $\mathbf{x}$ where two of the four linear lsfs evaluations are equal. This occurs for all $\mathbf{x} = [0, \ldots, 0, x_{d/2+1}, \ldots, x_d]$ and $\mathbf{x} = [x_1, \ldots, x_{d/2}, 0, \ldots, 0]$. In practice, the optimizer did not evaluate the gradient of g_1 at those particular points during our investigation.

5.3.1.4 Study of the influence of κ_0 on the failure probability estimate

For each failure region k, an initial value of $\kappa_{k,0}$ has to be set *a priori* in the CE-AMIS algorithm of Section 5.2.3.2. A study of the influence of this $\kappa_{k,0}$, simply denoted κ_0, on the failure probability estimate has been conducted for this application and is displayed in Figure 5.11. The failing points representative of the four failure regions are found in the same direction as the actual design points, but with a norm equal to LB instead of β. For each of the four failure regions, the final converged κ value of the CE-AMIS algorithm is the same. This value grows with the dimension. For $d = 200$, the converged κ value is around 79. For $d = 300$, the converged κ value is around 94. Finally for $d = 400$, the converged κ value is around 107.

The number K of failure regions found by the algorithm is displayed in Figure 5.11 c). In dimension 300, where the converged κ value is equal to 94, setting $\kappa_0 = 50$ or $\kappa_0 = 150$ does not prevent the algorithm from finding the four failure regions each time. However, in dimension 200, if κ_0 is greater than the converged value 79, the algorithm finds on rare occasions five failure regions instead of four (once every 500 independent simulations, for $\kappa_0 = 120$ and $\kappa_0 = 150$). This means that the samples generated in one particular failure region are too concentrated and do not cover the whole failure region; as a result, the remaining part is identified as another new failure region (as both the bulge and the cone associated with the representative failing point are not wide enough, see Section 5.2.4). This is not a problem in terms of failure probability estimation, but it results in a heavier budget N_{samp} for these particular simulation runs.

In contrast in dimension 400, if κ_0 is a lot smaller than the optimal κ equal to 107, the algorithm sometimes fails to find the four failure regions and stops at three (18 times out of 500 independent simulations, for $\kappa_0 = 50$). Indeed, if κ_0 is too small, then the sample generated in the first step of the CE-AIS algorithm is too wide, and the situation occurs where not a single $R^{(j)}\mathbf{T}^{(j)}$ is failing. This particular failure region is thus considered as negligible (see Section 5.2.3) and therefore $K = 3$. The resulting probability estimate is then biased (see Figure 5.11 a)) and the variance of the estimate is higher (see Figure 5.11 d)). However, it should be noted that finding three failure regions or five remains rare instances, and that the algorithm is quite robust regarding the number of failure regions found depending on the value of κ_0.

The probability estimate is very close to the reference value and is not dependent on κ_0, except in dimension 400 when $\kappa_0 = 50$. It seems that the value of κ_0 that minimizes the number N_{samp} of evaluations of the limit state function is when κ_0 is really close to the optimal κ, which is expected, (see Figure 5.11 b)). Then, as κ_0 is set greater or smaller, the budget N_{samp} increases as more steps

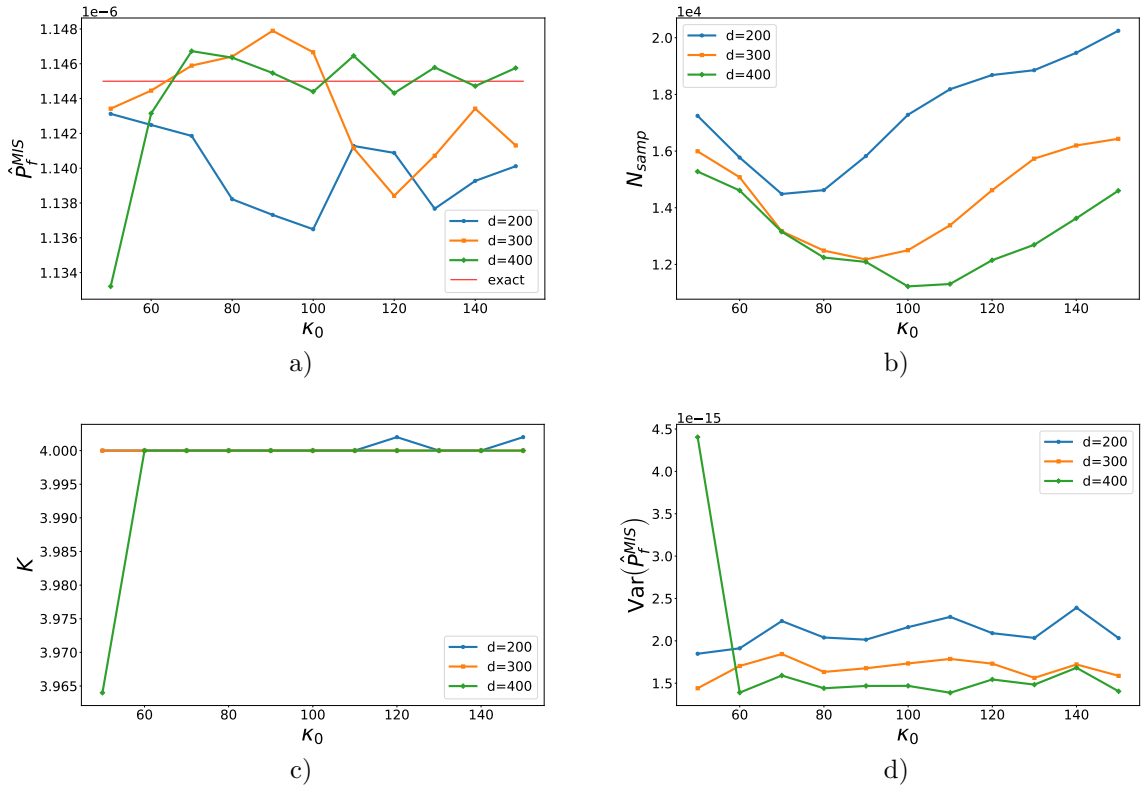


Figure 5.11: a) Evolution of the failure probability estimate with respect to κ_0, for dimensions 200, 300, and 400. b) Evolution of the number of lsf evaluations N_{samp} with respect to κ_0. c) Evolution of the number K of failure regions found with respect to κ_0. d) Evolution of variance of the failure probability estimate with respect to κ_0.

are necessary for the CE-AMIS algorithm to converge.

The global numbers $N_{\text{opt},1}$ and $N_{\text{opt},2}$ of evaluations of g and ∇g required for all the optimizations performed in the identification problems did not vary throughout the study. Indeed, as the initial LHS is of size 10, if the algorithm finds three or five failure regions instead of four, the number of performed optimizations remains the same in the end. Thus κ_0 does not influence the numerical cost of the optimization part.

The conclusion of this study is to choose κ_0 relatively high when no information concerning the failure regions is available and the higher the dimension, the higher κ_0.

5.3.1.5 Comparison with the iCE-vMFNM method and the SS algorithm

Settings of the Subset Sampling algorithm

The subset sampling algorithm used for this application employs the modified Metropolis algorithm described in Appendix A.2 as it is suited for high-dimensional standard normal inputs. The proposal univariate pdfs of the modified Metropolis algorithm are chosen as uniform pdfs of width $w = 1$, centered at the coordinate of the current sample: for component j, the proposal univariate pdf is $p_j^*\left(\cdot\,|\theta_j^{(k)}\right) = \mathbb{I}_{\left[\theta_j^{(k)}-w,\theta_j^{(k)}+w\right]}/2w$, where $\boldsymbol{\theta^{(k)}}$ is the current state (cf Appendix A.2). The p_0 probability is set equal to 0.1. For this particular application, the spread parameter w was set to 1 as it gives better result than if $w = 2$ (which is the value suggested by [Au and Beck, 2001]). The sample size at each iteration is set to $N = 15000$, which is quite large but necessary as there are four failure regions. A maximum number of iteration is set to $p_{\max} = 10$ for comparison purposes.

Settings of the iCE-vMFNM algorithm

The parameters of the iCE-vMFNM algorithm are set as follow : $\delta = 10$ and $N = 100000$ (cf Chapter 3

Table 5.1: Comparison of the proposed algorithm with the iCE-vMFNM algorithm and with a SS algorithm for g_1. The reference probability is equal to 1.15×10^{-6} for the three dimensions considered. The symbol [$**$] indicates that fewer than 500 simulations converged.

Series system of four linear limit state functions

	Proposed algorithm		iCE-vMFNM		Subset sampling	
$d = 200$						
	1.14×10^{-6}		1.00×10^{-6}		1.15×10^{-6}	
	CV	4.1% (3.9%)	CV	56.4%	CV	10.6%
	N_{samp}	14486	N_{samp}	449400	N_{samp}	91530
	$N_{\text{opt},1}$	261				
	$N_{\text{opt},2}$	204				
$d = 300$						
	1.15×10^{-6}		1.03×10^{-6}		1.15×10^{-6}	
	CV	3.6% (3.6%)	CV	50.6%	CV	10.9%
	N_{samp}	12174	N_{samp}	456800	N_{samp}	91860
	$N_{\text{opt},1}$	231				
	$N_{\text{opt},2}$	184				
$d = 400$						
	1.15×10^{-6}		1.05×10^{-6}		1.15×10^{-6}	
	CV	3.3% (3.4%)	CV	33.1% [**]	CV	10.3%
	N_{samp}	11308	N_{samp}	400000	N_{samp}	91470
	$N_{\text{opt},1}$	255				
	$N_{\text{opt},2}$	213				

Section 3.3.3.4). This CV target value $\delta = 10$ resulted in being the best value after several tests. The number N is very large as well but it is necessary as the failure probability is very small. It is recalled here that in the iCE-vMFNM algorithm, the first sample is generated from $f_{\mathbf{X}}$ and there are no intermediate failure thresholds. The number of densities in the mixture, $K = 4$ is a parameter to be set; therefore, its previous knowledge is required unlike in the proposed algorithm where no assumption is made on K. A maximum number of iteration is set to $p_{\max} = 10$ for comparison purposes.

Comparison of the methods

To compare the algorithm with the other available methods, κ_0 is selected as follows: 70 when $d = 200$, 90 when $d = 300$, and 110 when $d = 400$. The performance of the algorithms is summarized in Table 5.1 where the CV displayed for all the methods is the empirical one, estimated over the 500 independent simulation runs. The numbers in parentheses are the mean of the theoretical CV estimates of the proposed method, computed with the approximation of the variance of Eq. (5.11).

The estimates of the proposed method and the subset simulation method are quite close to the reference value 1.15×10^{-6}. However, the proposed algorithm has the smallest CV with the smallest global budget. Indeed, the total number of evaluations of the lsf $N_{\text{opt},1} + N_{\text{samp}}$ is very low: with less than 15,000 evaluations, the CV is below 4% whatever the dimension d. For every simulation, the number of failure regions found by the proposed method is equal to 4. The CE-AMIS algorithm converges in 3 steps for each failure region on average, no matter the dimension. Consequently, the number N_{samp} seems to be independent of the dimension, which is expected since only the scalar parameter κ is optimized with the CE-AMIS method. The failure probability of each region $\hat{P}_{f,k}$ is close to $\Psi(-5) = 2.87 \times 10^{-7}$, its FORM estimate. Therefore the regions contribute equally to the global failure probability. The optimization budgets $N_{\text{opt},1}$ and $N_{\text{opt},2}$ remain negligible compared to the number of lsf evaluations required by the other two methods. The estimates of the theoretical CVs are very close to the empirical CVs, which indicates that the precision of the algorithm can be correctly determined with only one simulation run and that the independence assumption is relevant.

Table 5.2: Comparison of the proposed algorithm with the iCE-vMFNM algorithm and with a SS algorithm for g_1, with $\beta = 3.5$. The reference probability is equal to 9.30×10^{-6} for the $d = 200$.

Series system of four linear limit state functions

	Proposed algorithm		iCE-vMFNM		Subset sampling	
$d = 200$	9.29×10^{-4}		9.32×10^{-4}		9.33×10^{-4}	
	CV	2.6% (2.6%)	CV	6.5%	CV	7.7%
	N_{samp}	12312	N_{samp}	41080	N_{samp}	38240
	$N_{\text{opt},1}$	229				
	$N_{\text{opt},2}$	184				

The iCE-vMFNM algorithm performs poorly compared to the other two algorithms. Indeed, the mean estimate is a bit biased compared to the reference value, the CV is very high and the simulation budget is huge compared to the other two methods. In dimension 400, the displayed statistics result from less than 500 independent simulations, hence the symbol [**], given that the algorithm did not converge most of the time.

The number of evaluations of g_1 for the subset sampling method is independent of the dimension of the inputs. The simulation budget is quite high: at least six times higher than the proposed algorithm global budget. This method leads to large CVs as well: it is two to three times superior to the proposed estimate CVs.

5.3.1.6 Influence of the value of P_f

In order to study the influence of the value of the probability on the efficiency of the algorithms, a study of same application is conducted with a value of $\beta = 3.5$ in dimension 200. The reference probability value computed with the Monte Carlo method is equal to $P_f = 9.30 \times 10^{-4}$ with an empirical CV of 3.2% and a sample of size 10^6. The probability value is once more very close to the crude sum of the four FORM estimate: $P_f \approx 4 \times \Psi(-\beta)$. This probability is a lot larger than when $\beta = 5$. To compare the methods, a few changes considering the settings of the algorithms are made. In the proposed algorithm, since the probability is quite large, the value of the parameter γ, which influences the size of the bulges of Section 5.2.4.1 and the cones of Section 5.2.4.2 are lowered from $\gamma = 1.1$ to $\gamma = 1$ (only for this particular application). In the SS algorithm, the size of the sample are lowered to $N = 10000$. In the iCE-vMFNM algorithm, the target CV is lowered to $\delta = 4$ and the sample size is lowered to $N = 20000$.

The value of κ_0 is set the same as in the previous study, thus $\kappa_0 = 70$. The performance of the algorithms is summarized in Table 5.2, the CV displayed for all the methods is the empirical one, estimated over the 500 independent simulation runs. The numbers in parentheses are the mean of the theoretical CV estimates of the proposed method, computed with the approximation of the variance of Eq. (5.11).

The estimates of the three methods are very close to the reference value 9.30×10^{-4}. Once more, the proposed algorithm has the smallest CV with the smallest global budget. The total number of evaluations of the lsf $N_{\text{opt},1} + N_{\text{samp}}$ is still very low: with less than 13,000 evaluations, the CV is below 3%. For every simulation, the number of failure regions found by the proposed method is equal to 4. As previously, the CE-AMIS algorithm converges in 3 steps for each failure region on average. Consequently, the number N_{samp} seems to be independent of the rarity of the failure event. The final converged κ value of the four failure regions is close to 55. The failure probability of each region $\hat{P}_{f,k}$ is close to $\Psi(-3.5) = 2.33 \times 10^{-4}$, its FORM estimate. The optimization budgets $N_{\text{opt},1}$ and $N_{\text{opt},2}$ remain negligible compared to the number of lsf evaluations required by the other two methods. The estimates of the theoretical CVs are still very close to the empirical CVs. Selecting the value of κ_0 as 70, even though the failure region is much larger here (the final converged κ value is equal to 55 instead of 79 when $\beta = 5$), underlines the robustness of the algorithm concerning this κ_0 value.

The iCE-vMFNM algorithm performs significantly better compared to the previous study. The

simulation budget as well as the CV are divided by 10. Consequently, the efficiency of this algorithm depends on the rarity of the failure event. The simulation budget is still three times higher than the proposed method.

The SS algorithm performs slightly better in this configuration, as the simulation budget is lower along with the CV. Consequently, the efficiency of this algorithm depends as well on the rarity of the failure event. The simulation budget is still three times higher than the proposed method.

Remarks on the setting of the γ parameter

It should be noted that if the value of γ is set so $\gamma = 1.1$ instead of $\gamma = 1$ in this configuration, the number of failure regions found by the algorithm is not always equal to 4. Indeed over 500 simulations, 79 simulations resulted in only 3 failure regions found, due to the width of the failure regions. Indeed, as the failure event is not very rare, the failure regions are quite large and *relatively* close to one another. It may happen that in the CE-AMIS algorithm, an observation generated in the sampling process of the failure region k reaches another failure region k'. The maximum angle ϕ_k associated with the failure region k is then very large, and multiplying it by 1.1 makes it even larger. The resulting failure cone C_k encompasses then the failing point $P_{k'}$ associated with the other failure region k' found by the optimizations. Consequently, this failure region k' is not accounted for, or only very little as just a small part of it has been sampled. At the end of the algorithm, the failure probability estimate is then biased. Nevertheless, this phenomenon occurs on rare occasions and lowering the value of γ solves the problem. The value of γ is kept to 1.1 for the rest of the applications.

5.3.1.7 Test on higher dimensions

The performance of proposed method is investigated for this application in dimensions 500 and 1000 as well, to test the algorithm in very high dimensions. For this test, the first setting of the application is used ($\beta = 5$, $\gamma = 1.1$) and 100 independent simulations are run.

Dimension 500

The initial value of the concentration parameter κ_0 is set to $\kappa_0 = 130$. The number of failure regions found by the proposed method is equal to 4. As previously, the CE-AMIS algorithm converges in 3 steps for each failure region on average, $N_{\text{samp}} = 11520$. The final converged value of κ is close to 119. The failure probability estimate equals $\hat{P}_f = 1.15 \times 10^{-6}$, with an empirical CV of 3.3% and a theoretical CV of 3.2% (assuming all the samples are independent). The optimization budgets are $N_{\text{opt},1} = 248$ and $N_{\text{opt},2} = 205$.

Dimension 1000

The initial value of the concentration parameter κ_0 is set to $\kappa_0 = 200$. The number of failure regions found by the proposed method is equal to 3.98, thus for two runs only 3 failure regions are found. For these two runs, it seems the same problem mentioned in the previous section occurred: the failure cone is a little too wide. The CE-AMIS algorithm converges in 3 steps for each failure region on average, $N_{\text{samp}} = 11930$. The final converged value of κ is close to 167. The failure probability estimate equals $\hat{P}_f = 1.14 \times 10^{-6}$, with an empirical CV of 4.6% and a theoretical CV of 2.8% (assuming all the samples are independent). The optimization budgets are $N_{\text{opt},1} = 273$ and $N_{\text{opt},2} = 213$.

The proposed method stays relevant in very high-dimensional spaces. In dimension 1000, it seems a smaller setting of the γ parameter could lead to better results, however even with $\gamma = 1.1$, the proposed method is already very efficient. This value is kept for the others applications.

5.3.2 Duffing oscillator

5.3.2.1 Presentation of the application

The second example is a nonlinear elastic system as described in [Zuev, 2009]. The Duffing oscillator is modeled by the equation

$$m\ddot{z}(t) + c\dot{z}(t) + k\left\{z(t) + \gamma z(t)^3\right\} = f(t),$$

where $z(t)$, $\dot{z}(t)$, and $\ddot{z}(t)$ are the displacement, velocity, and acceleration of the oscillator at time t, respectively, and the constants are taken as in [Zuev, 2009], with $m = 1000$ kg, $c = 200\pi$ Ns/m, $k = 1000(2\pi)^2$ N/m and $\gamma = 1$ m^{-2}. The oscillator is subjected to the random loading $f(t, \mathbf{X})$, discretized in the frequency domain as in [Papaioannou et al., 2019]

$$f(t, \mathbf{X}) = -m\sigma \sum_{i=1}^{d/2} \left\{ X_i \cos(\omega_i t) + X_{d/2+i} \sin(\omega_i t) \right\},$$

where $\omega_i = i\Delta\omega$ with $\Delta\omega = 30\pi/d$ and $\sigma = \sqrt{2S\Delta\omega}$, where $S = 0.005$ m^2/s^3 is the intensity of the loading. Here, $\mathbf{X}$ are independent standard Gaussian random variables of dimension d. The lsf g_2 is the maximum displacement of the oscillator at $t = 2$ s defined as

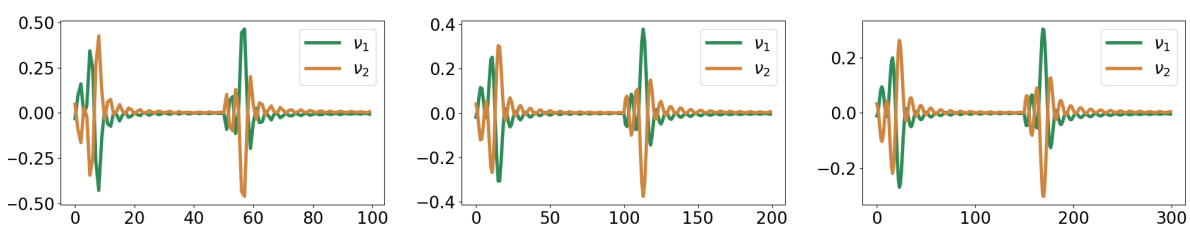
$$g_2(\mathbf{X}) = \min\{z_{\text{crit},1} - z(2s), z(2s) - z_{\text{crit},2}\},$$

where $z_{\text{crit},1} = 0.1$ and $z_{\text{crit},2} = -0.06$. The initial values are set to $z(0) = 0$ and $\dot{z}(0) = 1.5$. The limit state function along with its gradient are computed with an Euler method. The reference value is computed with a Monte Carlo method and is equal to 4.28×10^{-4} with an empirical CV of 4.9% with a sample of size of 10^6, whatever the dimension d.

5.3.2.2 Definition of the important ring

The performance of the proposed method is investigated in dimensions 100, 200, and 300 and the important ring is defined as in the previous application Section 5.3.1.2. Thus for $d = 100$, the importance interval equals $[5.83, 14.17]$ and for the other two dimensions the interval is the same as before.

Compared to the previous example, the search for design points of the system is not as straightforward. The proposed method found each time 2 failure regions. Both failure regions start before the beginning of the important ring and spread across it; thus this application is another example of the situation referred to as Case 2 in the previous sections. The direction $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ of the associated failing point P_1 and P_2 are displayed in Figure 5.12. They seem to be opposite to each other and the plots are the same no matter the dimension, which is expected. Therefore the failure event is linked to extreme values of the amplitudes associated with the lowest frequencies in the harmonic discretisation of the random loading. It is reminded here that the random loading formula associates the random vector's components 1 to $d/2$ with a cosine and $d/2 + 1$ to d with a sine. As the failure regions are opposite to each other, the sum of the failure probability associated with each failure region should be equal to the final MIS estimate.



a) b) c)

Figure 5.12: Illustration of the direction of the failing points P_1 and P_2 representative of the two failure regions) in dimension 100 b) in dimension 200 c) in dimension 300.

5.3.2.3 Comparison with the iCE-vMFNM method and the SS algorithm

The setting of the proposed method is the same as the one presented in Section 5.3.1.3. The setting of the SS algorithm is the same as the one presented in Section 5.3.1.5, except for the size of the sample at each iteration which is set to $N = 10000$. For the iCE-vMFNM algorithm, the number of densities in the mixture is set to $K = 2$ and the CV target value is set to $\delta = 2$ for all dimensions. However, the sample size is set depending on the dimension. It is recalled here that the iCE-vMFNM algorithm requires a number of $2(d + 3) + 2$ parameters to be optimized; cf Chapter 3 Section 3.3.3.4.

Table 5.3: Comparison of the proposed algorithm with the iCE-vMFNM algorithm and with a SS algorithm for g_2. The reference probability is equal to 4.28×10^{-4} for the three dimensions considered.

Duffing oscillator						
	Proposed algorithm		iCE-vMFNM		Subset sampling	
$d = 100$						
	4.28×10^{-4}		4.28×10^{-4}		4.29×10^{-4}	
	CV	4.7% (4.6%)	CV	5.0%	CV	8.7%
	N_{samp}	5818	N_{samp}	9312	N_{samp}	40000
	$N_{\text{opt},1}$	386				
	$N_{\text{opt},2}$	278				
$d = 200$						
	4.26×10^{-4}		4.26×10^{-4}		4.28×10^{-4}	
	CV	4.0% (4.1%)	CV	5.0%	CV	9.5%
	N_{samp}	5624	N_{samp}	16170	N_{samp}	40000
	$N_{\text{opt},1}$	244				
	$N_{\text{opt},2}$	187				
$d = 300$						
	4.26×10^{-4}		4.25×10^{-4}		4.28×10^{-4}	
	CV	4.5% (4.3%)	CV	4.6%	CV	8.9%
	N_{samp}	5978	N_{samp}	23660	N_{samp}	40000
	$N_{\text{opt},1}$	268				
	$N_{\text{opt},2}$	201				

Consequently the higher the dimension the larger the simulation budget for comparable precision: $N = 3000$ for $d = 100$, $N = 5000$ for $d = 200$ and $N = 7000$ for $d = 300$. In the proposed method, this issue does not occur since only the concentration parameter κ is optimized in the CE-AMIS algorithm.

To compare the algorithm with the other available methods, the initial κ_0 are selected as follows: $\kappa_0 = 50$ when $d = 100$, $\kappa_0 = 70$ when $d = 200$ and $\kappa_0 = 90$ when $d = 300$. These particular values are inspired by the study of the previous example. The performance of the algorithms is summarized in Table 5.3, where the CV displayed for all the methods is the empirical one, estimated over the 500 independent simulation runs. The numbers in parentheses are the mean of the theoretical CV estimates of the proposed method, computed with the approximation of the variance of Eq. (5.11).

The estimates of the three methods are close to the reference value 4.28×10^{-4}. As in the previous example, the proposed algorithm has the smallest CV with the smallest global budget. The total number of evaluations of the lsf $N_{\text{opt},1} + N_{\text{samp}}$ is extremely low: with less than 7000 evaluations, the CV is below 5% whatever the dimension d. For every simulation, the number of failure regions found by the proposed method is equal to 2. The CE-AMIS algorithm converges in 3 steps for both failure regions on average, no matter the dimension. The final converged value of κ depends on the failure regions and the dimension: $\kappa_{1,n_1} = 39$ and $\kappa_{2,n_2} = 43$ for $d = 100$, $\kappa_{1,n_1} = 54$ and $\kappa_{2,n_2} = 57$ for $d = 200$ and $\kappa_{1,n_1} = 64$ and $\kappa_{2,n_2} = 69$ for $d = 300$. Therefore, the first failure region is wider than the second. The failure probability of the failure regions are given by $\hat{P}_{f,1} = 3.14 \times 10^{-4}$ and $\hat{P}_{f,2} = 1.14 \times 10^{-4}$, for every dimension. The contribution of the first failure region to the global failure probability is the strongest.

Given that the number of failure regions found is twice smaller than in the previous example, it is not surprising that N_{samp} is also twice smaller in the proposed method. It seems that N_{samp} is thus strongly linked to the number of failure regions of the system and completely independent of the dimension. The numbers $N_{\text{opt},1}$ and $N_{\text{opt},2}$ have the same magnitude as in the previous application even though the lsf is more complex; they remain below 400, which is very low compared to N_{samp}. Once more, the estimates of the theoretical CV are very close to the empirical CVs.

For this application, the iCE-vMFNM algorithm performs better than in the previous example,

even with the low β setting of Section 5.3.1.6. The number of lsf evaluations gradually increases, for a CV close to 5%. Consequently, the iCE-vMFNM performance depends on the dimension of the system, as well as the number of failure regions. This is not surprising considering the number of parameters in the CE optimization which depends both on d and K. In dimension $d = 300$, the number of lsf evaluations required is still four times higher than with the proposed algorithm.

The number of evaluations of g_2 for the subset sampling method is the highest with also the highest CV and it is still independent of the dimension. In contrast to the other two algorithms, the number N_{samp} is not lower than in the previous application with $\beta = 3.5$ although there are only two failure regions. Compared to the proposed method, the simulation budget is six to seven times superior.

5.3.2.4 Influence of the width of the important ring

In this application, two failure regions start before the beginning of the important ring and spreads across it; a situation denoted as Case 2 in the previous sections. One can wonder about the influence of the width of the important ring, on the coordinates of the important direction $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$. Indeed, in Case 2, the distribution of the radial component is left unchanged. Moreover, in this particular application, since $\mathbf{X}$ are standard normal inputs then $R \sim \chi(d)$ and $\mathbb{E}[R] \approx \sqrt{d}$. Therefore, the narrower the important ring is defined, the closer the lower bound LB is to $\sqrt{d}$. As the representative failing points P_1 and P_2 are on the hypersphere of radius LB, the closer LB is to $\sqrt{d}$, the closer P_1 and P_2 are to the observations generated in the various CE-AMIS procedures. In other words, searching for the important directions on a hypersphere of radius closer $\sqrt{d}$ gives a better auxiliary density $h_R \times h_{\mathbf{T}}$, considering the chosen radial auxiliary density h_R; see the Appendix C for more details.

In the previous example, as the lsf is linear, the important directions would be the same no matter the radius LB of the important ring. In this application however it may differ. Consequently, the performance of proposed method is investigated in dimension 100 with an important ring defined as such: the interval [LB, UB] is centered around $\sqrt{d}$ and the smallest value of ϵ such that $\mathbb{P}(\sqrt{d}-\epsilon \leq R \leq \sqrt{d}+\epsilon) \geq 1-10^{-4}$ is selected. The resulting important interval is $[7.25, 12.75]$, instead of $[5.83, 14.17]$ and is displayed in Figure 5.13.

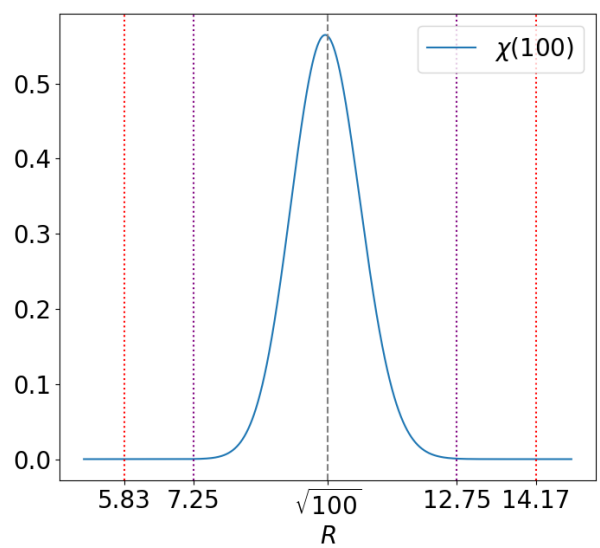


Figure 5.13: Illustration of the important intervals of the $\chi(100)$ distribution for the two different settings. The dashed line represents the value $\sqrt{100}$, the red dotted lines are the two important interval bounds of the first setting, the purple dotted lines are the two important interval bounds of the second setting.

The important directions found are the same as with the previous important interval. Consequently, the CE-AMIS algorithm performs in the same manner and the simulation budget N_{samp} remains the same. However the optimization budgets decrease a little: $N_{\text{opt},1} = 354$ instead of 386 and $N_{\text{opt},2} = 261$

instead of 278. The conclusion of this study is that for different values of LB, the most negative directions found by the optimization of Eq. (5.2) on the hypersphere of radius LB remain the same, therefore the limit state function g_2 does not have a strong nonlinear behavior.

5.3.3 Portfolio loss

5.3.3.1 Presentation of the application

The third example is often used in financial studies [Bassamboo et al., 2008, Chan and Kroese, 2012, El Masri et al., 2021]. It consists of a large portfolio of loans with 250 obligors, each of whom having a non-zero probability of default. It is assumed that a d-vector of underlying latent variables $\mathbf{Z}$ represents the obligors as such: when $Z_i > 0.5\sqrt{d}$, then the ith obligor defaults. The portfolio loss is expressed as

$$L(\mathbf{Z}) = \mathbb{I}_{Z_1 > 0.5\sqrt{250}} + \cdots + \mathbb{I}_{Z_{250} > 0.5\sqrt{250}}. \tag{5.13}$$

The value of interest is the probability that the portfolio loss exceeds 0.25×250; thus, the limit state function is given by

$$g_3(\mathbf{Z}) = 62.5 - L(\mathbf{Z}).$$

In order to model the underlying correlations between the obligors, this simple single factor model is employed [Bassamboo et al., 2008]

$$Z_i = \frac{\rho L + \sqrt{1-\rho}\,\eta_i}{W} \qquad \text{for} \qquad i = 1, \ldots, 250,$$

where the factor L is a standard normal variable which measures the effects impacting all the obligors, η_i is a centered normal variable of variance 9 independent of the factor L which captures the idiosyncratic risk, W is a random variable independent of L and η_i, which represents a common shock affecting the obligors, of distribution

$$W \sim \sqrt{\chi^2(\nu)/\nu},$$

and ρ is a correlation parameter. The resulting inputs $\mathbf{Z}$ follow thus a centered multivariate Student distribution with ν degrees of freedom (see Chapter 2 Section 2.2.1.3) and $d \times d$ dispersion matrix

$$\boldsymbol{D}_{\mathbf{Z}} = \begin{pmatrix} 9-8\rho^2 & \rho^2 & . & . & . & \rho^2 \\ \rho^2 & 9-8\rho^2 & \rho^2 & . & . & . \\ . & & & & & . \\ . & & & & & . \\ \rho^2 & . & . & . & \rho^2 & 9-8\rho^2 \end{pmatrix}.$$

For application purposes, ρ is set to 0.25 and $\nu = 4$. The reference probability value computed with a Monte Carlo method is equal to 8.12×10^{-3}.

5.3.3.2 Adaptation to the proposed method

To use the proposed method, the first step is to apply a linear transformation to remove the correlation of the inputs (but they remain dependent). To do so, one has to pre-multiply the inputs $\mathbf{Z}$ by the matrix A^{-1} such that $AA^\top$ is the Cholesky decomposition of $\boldsymbol{D}_{\mathbf{Z}}$. The limit state function with the standard variable $\mathbf{X} = A^{-1}\mathbf{Z}$ is then

$$\bar{g}_3(\mathbf{X}) = 62.5 - L(A\mathbf{X}).$$

The optimizations of Section 5.2.2.2 performed in the various identification problems require the gradient of $\bar{g}_3$. Nevertheless, this particular limit state function is not differentiable, as it involves the indicator function in the portfolio loss function L. Thus, only for the optimizations performed in the identification problems, a substitute $\hat{g}_3$ of $\bar{g}_3$ is used

$$\hat{g}_3(\mathbf{X}) = 62.5 - \hat{L}(A\mathbf{X}),$$

where

$$\hat{L}(A\mathbf{X}) = \Psi\left([A\mathbf{X}]_1 - 0.5\sqrt{250}\right) + \cdots + \Psi\left([A\mathbf{X}]_{250} - 0.5\sqrt{250}\right),$$

with Ψ being the cdf of the standard normal distribution. This regularization for the indicator function is quite common (as derived in [Lacaze et al., 2015]) and makes differentiation possible.

5.3.3.3 Definition of the important ring

Given that $\mathbf{X}$ is a standard multivariate Student random vector with $\nu = 4$ degrees of freedom and dimension $d = 250$, then $R^2/250 \sim \mathcal{F}(250, 4)$ with $\mathcal{F}$ a Fisher–Snedecor distribution. This distribution is heavy-tailed and is not symmetric around its mean; cf Chapter 2 Section 2.2.2.1. The important ring is defined with the two hyperspheres of radii LB and UB chosen as quantiles of the Fisher–Snedecor distribution such that $\text{LB} = \sqrt{250 \times q_{10^{-5}/2}(\mathcal{F}(250, 4))}$ and $\text{UB} = \sqrt{250 \times q_{1-10^{-5}/2}(\mathcal{F}(250, 4))}$. The important interval is thus equal to $[5.62, 398.22]$ and is a lot wider than for the χ distribution of the previous applications; it is displayed in Figure 5.14.

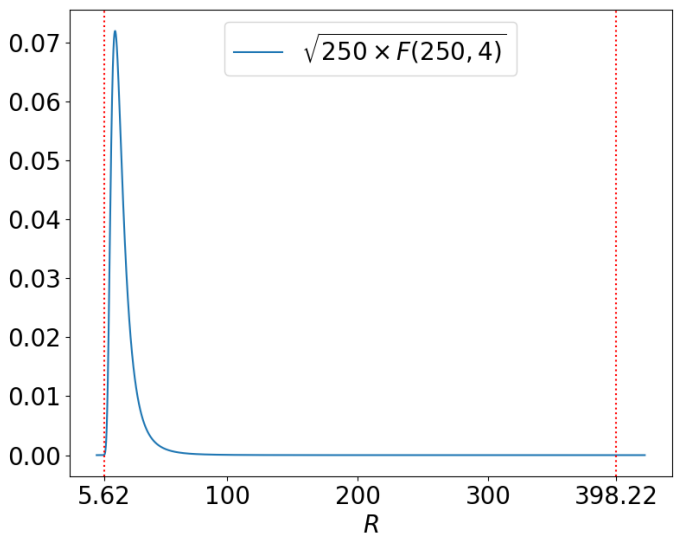


Figure 5.14: Illustration of the important interval of the $\sqrt{250 \times \mathcal{F}(250, 4)}$ distribution. The red dotted lines are the two important interval bounds.

The proposed method found each time one failure region. The failure region starts inside the important ring; therefore, this application is an example of the situation referred to as Case 1 in the previous sections. The failing point P_1 is thus the design point of the system. Its norm is equal to $\|P_1\| = 21.19$ and its direction ν_1 is displayed in Figure 5.15, in both the space of $\mathbf{X}$ and the space of $\mathbf{Z}$. This important direction is quite difficult to interpret as there are many jumps in the value of the components. In the $\mathbf{Z}$ space, each Z_i plays the same part in the loss function of Eq. (5.13), however the value of the components oscillates between two constants.

5.3.3.4 Performance of the proposed method

In this application, the setting of the LHS differs from Section 5.3.1.3 as the inputs are no longer independent; the LHS framework requires independent inputs. Therefore, the LHS is performed with an independent mixture of d univariate Student distributions (which is not a multivariate Student law) centered either on 0 or $\mathbf{x}_0$ of Eq. (5.10) according to the identification problem. The rest of the settings are the same. The proposed method is also tested with a higher size of sample $N = 13000$ for comparison purposes.

The proposed method is compared with the results of a variance minimization (VM) algorithm presented in [Chan and Kroese, 2012] which is not detailed here, to have a comparative benchmark. The basic idea of this VM algorithm is to optimize the parameters of the IS density to obtain an IS estimate with minimum variance, instead of the IS density with the minimum CE divergence with $h_{\text{opt}} = \mathbb{I}_{D_f} f_{\mathbf{X}}/P_f$.

The performances of the algorithms are displayed in Table 5.4 with $\kappa_0 = 80$ for $N = 1000$ and $N = 13000$. This particular value of κ_0 is inspired by the study of Section 5.3.1.4. The numbers in parentheses are the theoretical coefficients of variation of the proposed method, computed with

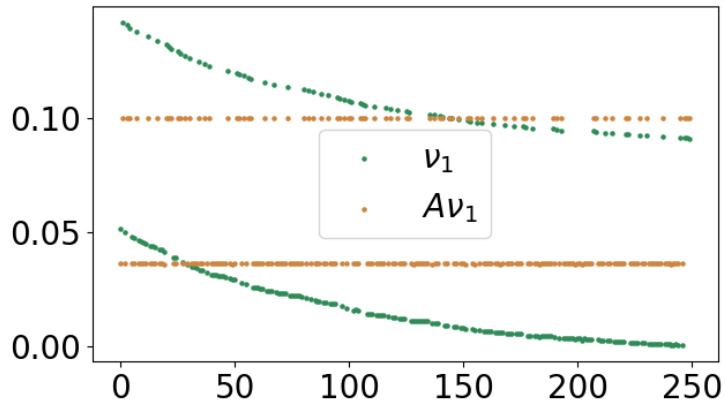


Figure 5.15: Illustration of the direction of the failing points P_1, in the space of $\mathbf{X}$ in green and the space of $\mathbf{Z}$ in brown.

the approximation of the variance of Eq. (5.11), over 500 independent simulations. The results are compared with the classical Monte Carlo algorithm as well.

Table 5.4: Comparison of the proposed algorithm with the VM algorithm of [Chan and Kroese, 2012] and with a classical Monte Carlo for $\bar{g}_3$. The reference probability is equal to 8.12×10^{-3}.

Portfolio Loss							
Proposed algorithm ($N = 1000$)		Proposed algorithm ($N = 13000$)		VM		MC	
8.12×10^{-3}		8.15×10^{-3}		8.14×10^{-3}		8.12×10^{-3}	
CV	8.4% (8.4%)	CV	3.9% (3.7%)	CV	0.5%	CV	1.1%
N_{samp}	7452	N_{samp}	50258	N_{samp}	55000	N_{samp}	10^6
$N_{\text{opt},1}$	1354	$N_{\text{opt},1}$	1340				
$N_{\text{opt},2}$	1119	$N_{\text{opt},2}$	1108				

The failure probability estimate of the proposed method is very close to the reference value 8.12×10^{-3}. For every simulation, the number of failure regions found by the proposed method is equal to 1. The CE-AMIS algorithm converges in 7 steps on average for $N = 1000$ and on 4 steps on average for $N = 13000$. The final converged value of κ_{1,n_1} is equal to $\kappa_{1,n_1} = 7$. Therefore, the failure region is very wide; it appears that the failure domain has a spherical symmetry, as the resulting vMF density $h_{\mathbf{T}}$ is close to the original uniform distribution on the hypersphere $f_{\mathbf{T}}$. Consequently, the important direction displayed in Figure 5.15 might not be very important. The simulation budget N_{samp} is quite large, even though there is only one failure region. Compared to the two previous examples, the proposed algorithm here is not as precise in terms of CV. Indeed, a very large budget is needed to reach a CV value lower to 4%. The numbers $N_{\text{opt},1}$ and $N_{\text{opt},2}$ are a lot larger than in the previous applications as well: they are no longer negligible.

The VM algorithm, in contrast, is very precise. However, to achieve such results, it is assumed that it is possible to generate observations from h_{opt} with a Markov Chain Monte Carlo algorithm specifically set for the multivariate Student distribution: the Gibbs sampler [Chan and Kroese, 2012]. This assumption is possible in this particular case as there is only one failure region. The VM algorithm would take a very long time to converge if there were several failure regions, resulting to a larger simulation budget.

As a conclusion, the proposed algorithm remains relevant, as it allows to compute the probability with a CV lower than 10% with a budget lower than 10000. Furthermore, the estimates of the theoretical CV are still very close to the empirical CV. This example with the Student distribution

shows the flexibility of the proposed algorithm in terms of elliptically distributed inputs.

5.3.3.5 Study on radial distribution of the failing inputs

As the results of the proposed method are not as good as for the two previous applications, a study on the radial distribution of the failing inputs is conducted. With a large Monte Carlo, a histogram of the norm of the points in the failure domain D_f of the system is plotted, against the chosen radial auxiliary density $h_R = f_{R|R>r_{\text{opt},1}}$ with $r_{\text{opt},1} = 21.19$. The result is illustrated in Figure 5.16. The conclusion of this study is that the value of $r_{\text{opt},1}$, obtained with the optimization of Section 5.2.2.2 is underestimated. Indeed, it seems that the failure domain starts at the hypersphere of radius 40, which is a lot larger than 21.19. This explains why the CE-AMIS takes a lot of iterations to converge. As there is no particular important direction, an efficient radial auxiliary density is crucial. Unfortunately, the settings of h_R are not optimal because $r_{\text{opt},1}$ is too low; therefore the algorithms requires a larger simulation budget to reach the CV criterion.

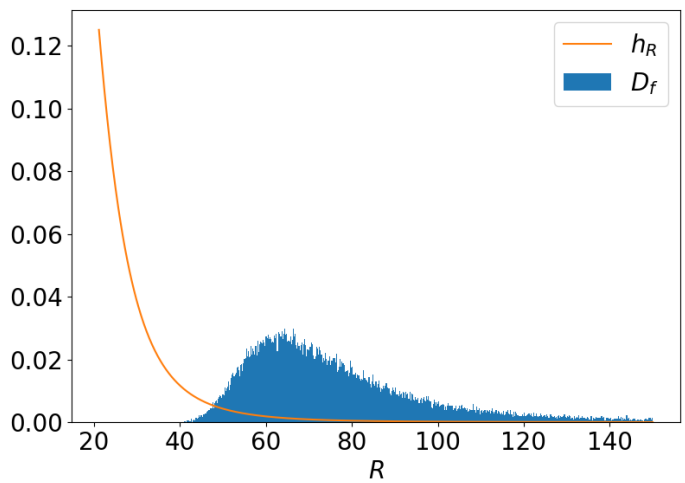


Figure 5.16: Density of $h_R = f_{R|R>r_{\text{opt},1}}$ with $r_{\text{opt},1} = 21.19$ plotted with an histogram of the norm of the points in D_f obtained with a MC sampling.

This poor estimation of $r_{\text{opt},1}$ can be related to the approximation of the lsf $\hat{g}_3$, which may not be the most appropriate to accurately find the design point. To analyze if this poor $r_{\text{opt},1}$ is the only reason for the results of the proposed method, the following test is conducted: the CE-AMIS algorithm is run independently 500 times in three different settings. For all the settings, the size of the sample is set to $N = 1000$ and the important direction $\boldsymbol{\nu_1}$ is set to $\boldsymbol{\nu_1} = \mathbf{1}_d/\sqrt{d}$ (since there is no important direction). In the first setting, denoted Setting 1, the initial $\kappa_{1,0}$ value is kept to 80, but $r_{\text{opt},1} = 38$ (a little lower than 40, in case the previous MC did not sample the entire failure region D_f, to avoid any bias). In the second setting, denoted Setting 2, the value of $r_{\text{opt},1}$ is kept to 21.19, but the initial $\kappa_{1,0}$ value is lowered to 20. In the third setting, denoted Setting 3, the initial $\kappa_{1,0}$ value is set to 20 and $r_{\text{opt},1} = 38$.

Results of Setting 1

The failure probability estimate is equal to 8.12×10^{-3}, with an empirical CV of 4.8% and a theoretical CV (assuming the samples are independent) of 5.0%. The simulation budget is $N_{\text{samp}} = 4022$, which indicates that the CE-AMIS algorithm converges in 4 steps in average, instead of 7 in Table 5.4. Therefore, the CV is almost divided by two, along with the simulation budget, for $N = 1000$. The influence of $r_{\text{opt},1}$ is thus very strong.

Results of Setting 2

The failure probability estimate is equal to 8.08×10^{-3}, with an empirical CV of 8.3% and a theoretical CV (assuming the samples are independent) of 8.4%. The simulation budget is $N_{\text{samp}} = 5544$, which indicates that the CE-AMIS algorithm converges in 5 to 6 steps in average, instead of 7 in Table 5.4.

Therefore, the simulation budget is lower, but the CV remains the same. The influence of $\kappa_{1,0}$ is thus moderate.

Results of Setting 3
The failure probability estimate is equal to 8.07×10^{-3}, with an empirical CV of 5.0% and a theoretical CV (assuming the samples are independent) of 4.9%. The simulation budget is $N_{\text{samp}} = 2610$, which indicates that the CE-AMIS algorithm converges in 2 to 3 steps in average. Compared to Setting 1, the simulation budget is lower but the CV is the same. Compared to Setting 2, the simulation budget is lower along with the CV.

The conclusion of these tests is that the results of the proposed method displayed in Table 5.4 are mostly due to a poor estimation of $r_{\text{opt},1}$ which has a great influence on the CV of the failure probability estimate and the simulation budget. The initial setting of $\kappa_{1,0}$ plays a minor part in the resulting CV but affects the simulation budget.

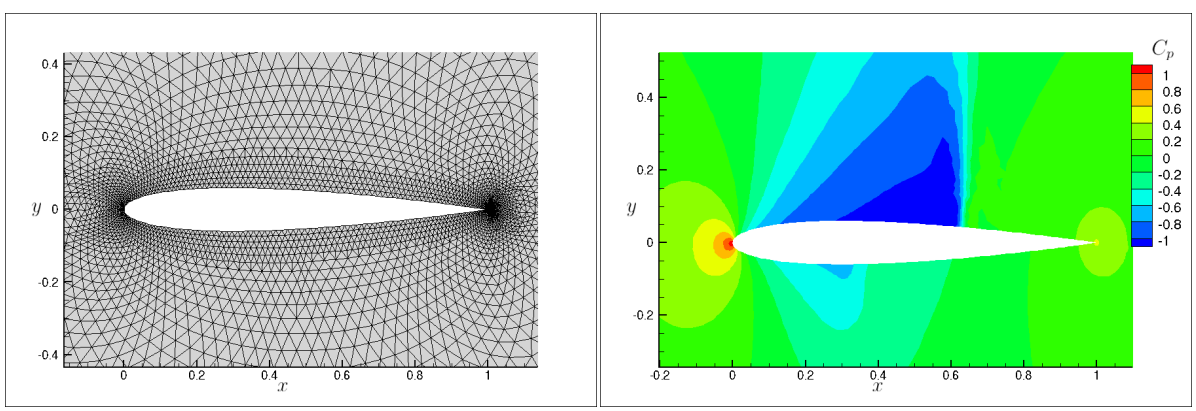
5.3.4 Airfoil in inviscid transonic flow

The last example studies the drag coefficient of an airfoil in inviscid transonic flow subject to random shape distortion. This example is representative of an industrial problem as the limit state function is analytically unavailable and acts like a black-box. Therefore this example illustrates better the context of the manuscript.

5.3.4.1 Design problem and optimal airfoil

This example is inspired by an optimization test case provided by the Computational Fluid Dynamic (CFD) solver SU2 [Economon et al., 2016]. For the sake of completeness, the main features of this problem are now detailed. The objective of this test case is to design the shape of an airfoil that minimizes the drag coefficient in transonic inviscid flow conditions. The baseline shape is a NACA 0012 airfoil, the freestream pressure is set to 101,325 Pa, its temperature is set to 273.15 K and the Mach number is equal to 0.8. The 2D Euler fluid model is solved by the finite volume method on a computational mesh of 5233 points and 10,216 triangular elements. Figure 5.17 a) presents the mesh around the airfoil. Under such conditions a transonic shock appears on the upper surface of the airfoil leading to a high drag coefficient equals to $C_D^{\text{NACA } 0012} = 2.13 \times 10^{-2}$, as presented in Figure 5.17 b).

The shape of the airfoil is parameterized through Hicks-Henne bump functions defined around the airfoil. These functions include two parameters, the location of their center and the amplitude of the bump. For this example 10 Hicks–Henne bump functions are used, five on the upper surface of the airfoil and five on the lower surface of the airfoil, respectively centered at 5%, 25%, 50%, 75%, 95% of the airfoil chord. The optimization problem is then to minimize the drag coefficient with respect to the 10 amplitudes of the Hicks–Henne bump functions, under a constraint of lift coefficient equals



a) b)

Figure 5.17: a) Close view on the initial computational mesh of the NACA 0012 airfoil. b) Pressure coefficient field showing the transonic shock on the NACA 0012 airfoil.

to $C_L = 0.326$ (value obtained with the initial NACA 0012 airfoil). This problem is solved using a gradient based optimizer and the gradient of the objective function is obtained through adjoint approach. Readers interested in the details of this resolution are referred to the SU2 tutorials and to the many articles about the adjoint approach in CFD computation; see [Kenway et al., 2019] for example.

The solution of the optimization problem is illustrated in Figure 5.18. One can note that the optimal solution (Figure 5.18 a)) reaches some characteristics of a supercritical airfoil, like a flattened upper surface near the leading edge. The obtained pressure coefficient field (Figure 5.18 b)) shows no transonic shock and leads to an overall drag coefficient equal to $C_D^{\text{opt}} = 1.02 \times 10^{-3}$.

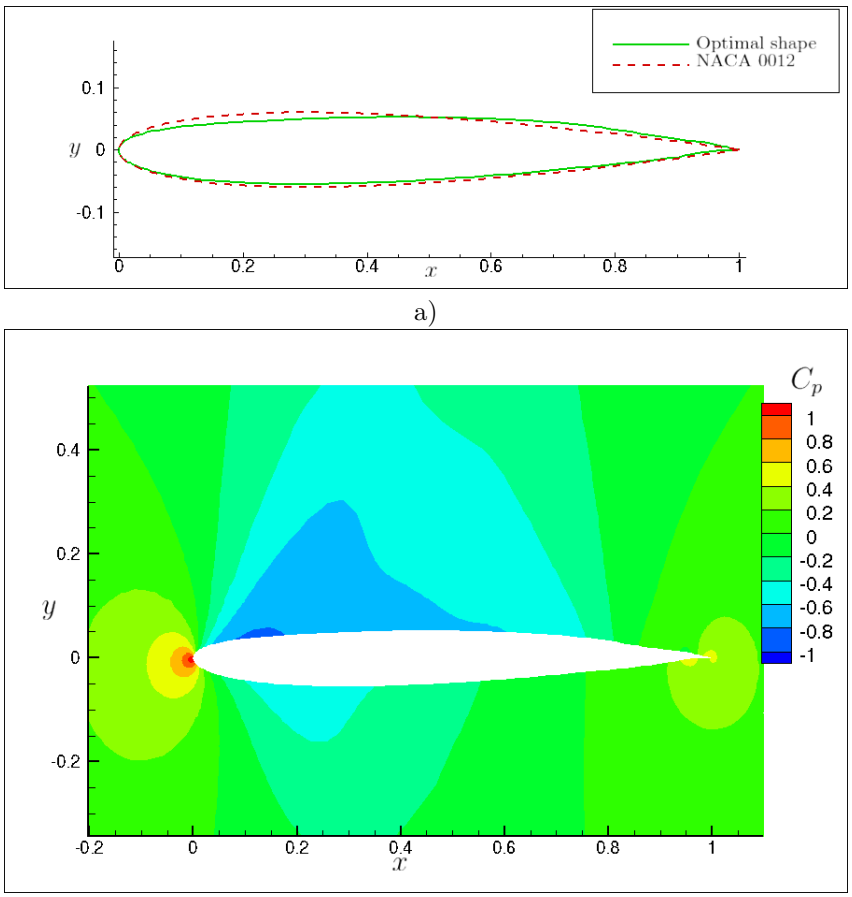


Figure 5.18: a) Comparison between the optimal airfoil solution and the baseline NACA 0012 airfoil. b) Pressure coefficient field around the optimal solution.

5.3.4.2 Presentation of the reliability problem

It is now assumed that the optimal airfoil shape previously defined is subjected to random distortion. This distortion is represented by 100 Hicks–Henne bump functions, 50 at the lower airfoil surface and 50 at the upper airfoil surface. The centers of the 50 functions are linearly spaced between 5% and 95% of the airfoil chord. It is assumed that the amplitude of the Hicks–Henne functions are independent and normally distributed. The mean value is null and the variance is chosen to create reasonable deformation of the airfoil, modeling manufacturing uncertainties for example. Using previous notations this leads to $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, 2.5 \times 10^{-7} \times \mathbf{I}_d)$. The failure of the system occurs when the distortion causes a value of the drag coefficient that is above a certain threshold equal to 0.01, thus the limit state function is written

$$g_4(\mathbf{X}) = 0.01 - C_D(\mathbf{X}).$$

First a Monte Carlo estimation is performed in order to get a reference value for this problem. A total of 106,770 simulations were run leading to a failure probability estimation equal to $P_f^{\mathrm{MC}} = 5.61 \times 10^{-4}$ with a theoretical coefficient of variation of 12.9%. Each CFD simulation is run on a cluster using 24 CPU and takes approximately $7s$. The Monte Carlo estimation thus lasted approximately 8.5 days on 24 CPU. The equations between the inputs $\mathbf{X}$ and the resulting drag coefficient C_D are the Euler and adjoint Euler equations. The lsf for this engineering problem is a black-box function as the analytical form of the drag coefficient C_D is not available but is numerically computed by finite volume CFD code. One can note that it is thus not possible to know *a priori* the number of failure regions.

5.3.4.3 Definition of the important ring

The definition of the important ring is the same as in Section 5.3.2.2 for dimension $d = 100$, but given that the standard deviation of the inputs is equal to 5×10^{-4}, the bounds of the importance interval are multiplied by 5×10^{-4} as well, and $\mathrm{IR} = [2.92 \times 10^{-3}, 7.08 \times 10^{-3}]$. Since the inputs do not have a variance equal to 1, a standard deviation factor is simply added in the algorithm.

The proposed algorithm finds three different failure regions, $K = 3$. For each failure region, the representative point P_i is found on the lower bound of the important ring, thus the failure regions start before the beginning of the important ring and spread across it; this application is another example of the situation referred to as Case 2 in the previous sections. The vector values of these three points represent the worst configurations for the shape of the airfoil. Indeed they lead to shapes that produce particularly high drag. The shapes induced by the three representative failing point P_1, P_2 and P_3 are displayed in Figure 5.19. The pressure fields resulting from these three shapes are illustrated in Figure 5.20. Shocks appear which explain the high drag of these shapes. Indeed, for the point P_1 a large deformation appears at the rear of the airfoil causing a strong acceleration of the flow on this part and thus a transonic shock on the rear of the airfoil; see Figure 5.20 a). The point P_2 deforms the optimal shape near the leading edge and curves the airfoil in this area. The P_2 deformed shape looks approximately like the original NACA 0012 airfoil and thus leads to a transonic shock; see Figure 5.20 b). The point P_3 results from a negative curvature at the leading edge side and a bump at the trailing edge side. This deformation creates two transonic shocks as shown in Figure 5.20 c).

5.3.4.4 Performance of the proposed method

For this application, the setting of the SLSQP algorithm differs from Section 5.3.1.3 as the maximum number of iterations in the algorithm is lowered to 50 instead of 1000. The precision goal for the value of the objective function in the stopping criterion is set to 10^{-6} for the first optimization Eq. (5.1) and to 10^{-4} for the second optimization Eq. (5.2). Furthermore, the value of the maximum number of iterations p_{max} of the CE-AMIS algorithm is lowered to 5 instead of 10. The algorithm is run only three times. The displayed failing points P_1, P_2 and P_3 of the Section 5.3.4.3 are those found in the first run. The performance of the proposed algorithm is summarized in Table 5.5, with $N = 1000$ and $\kappa_0 = 50$. The coefficients of variation are the theoretical ones, assuming the samples are independent.

Table 5.5: Comparison of the proposed algorithm with the Monte Carlo method for g_4.

Failure probability of the airfoil shape							
Run 1		Run 2		Run 3		Monte Carlo	
5.30×10^{-4}		5.11×10^{-4}		4.87×10^{-4}		5.61×10^{-4}	
CV	9.8%	CV	5.9%	CV	6.2%	CV	12.9%
N_{samp}	13000	N_{samp}	13000	N_{samp}	13000	N_{samp}	106770
$N_{\mathrm{opt},1}$	4717	$N_{\mathrm{opt},1}$	3575	$N_{\mathrm{opt},1}$	1597		
$N_{\mathrm{opt},2}$	918	$N_{\mathrm{opt},2}$	617	$N_{\mathrm{opt},2}$	438		

One can note that the failure probability estimates found by the proposed algorithm in Run 1 and Run 2 are very close to the Monte Carlo reference, and have a smaller coefficient of variation. However, Run 3 gives an estimate with a smaller failure probability value than the MC reference. For the three runs, the proposed method found 3 failure regions. The pressure coefficient distributions resulting of the configurations of the 2D airfoil with the failing points $(P_k)_{k=1,2,3}$ are displayed in Figure 5.21, for each of the three runs.

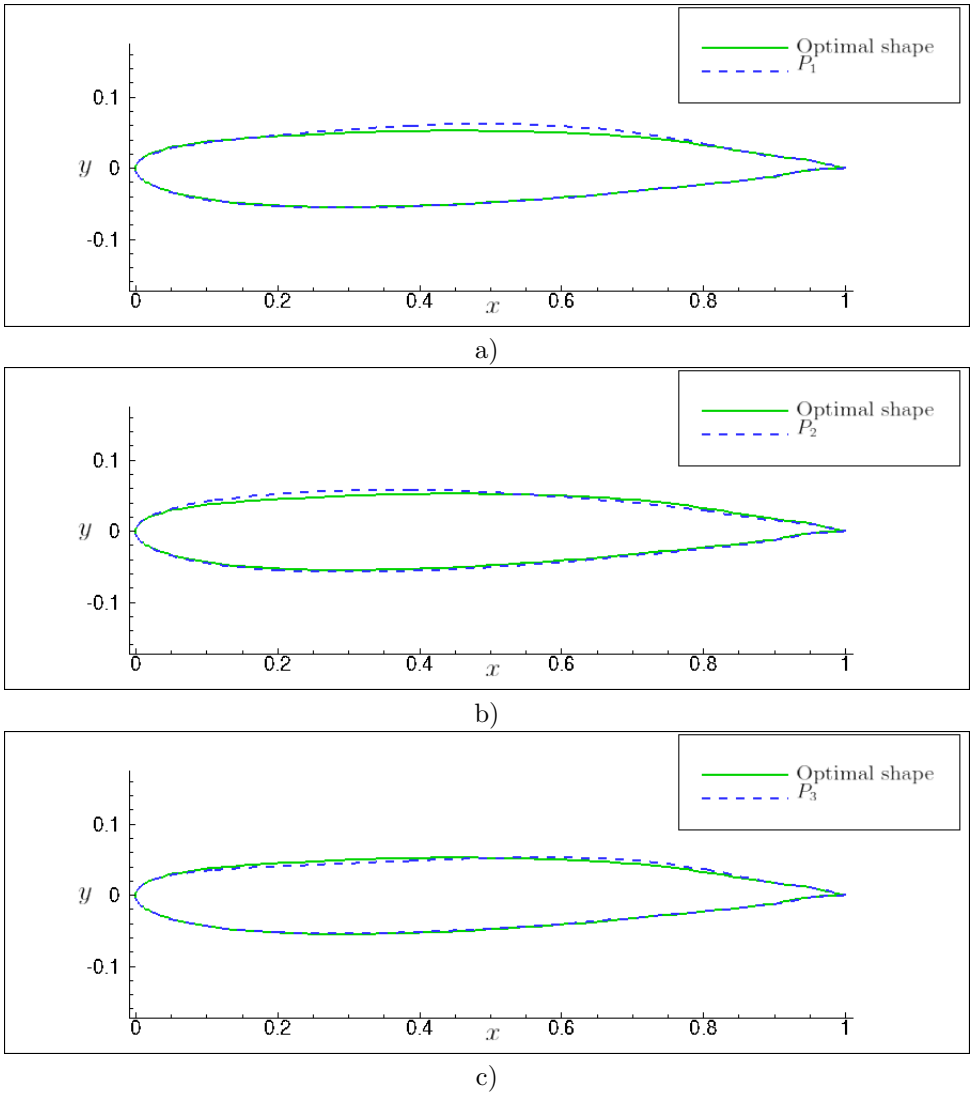
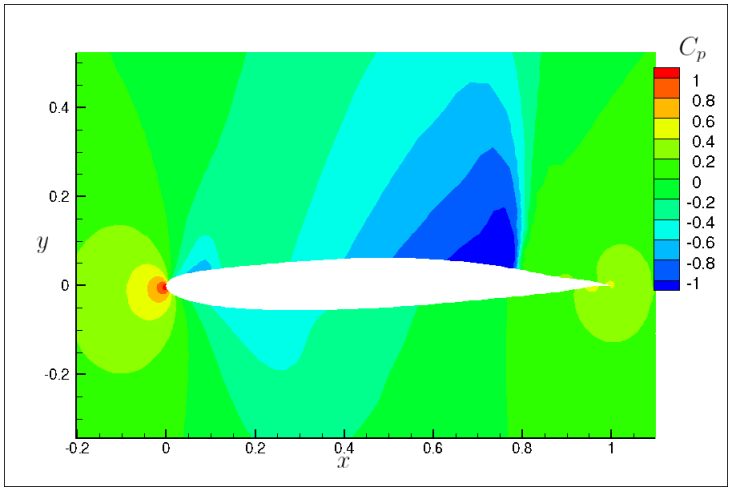


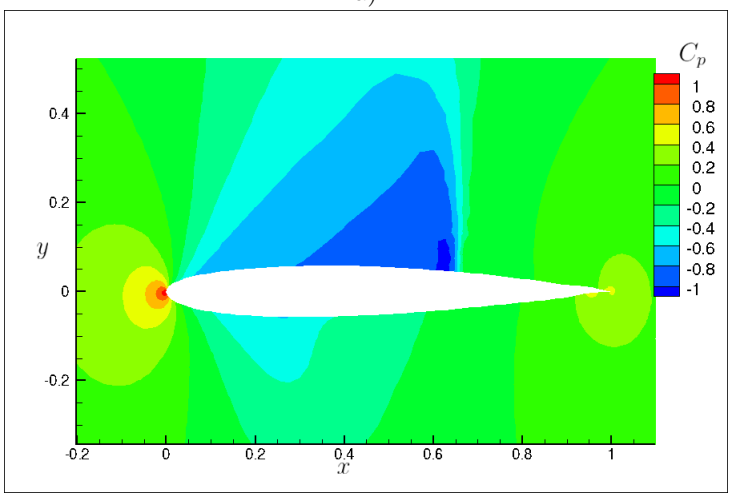
Figure 5.19: Shape comparison between the optimal airfoil and the critical point found by the proposed approach: a) P_1, b) P_2, c) P_3.

The first failing point P_1 found is exactly the same for the three runs. The CE-AMIS converges in 3 steps for this failure region, for all runs. The final converged value of κ equals $\kappa_{1,n_1} = 37$. The resulting failure probability estimates of this region are close to $\hat{P}_{f,1} = 5 \times 10^{-4}$. The second failing point P_2 is quite similar for Run 1 and Run 2 but there is a small deviation in Run 3. This results from the first optimization Eq. (5.1): in Run 3, P_2 is found at a higher norm than LB. This could be a consequence of the low convergence criterion of the SLSQP as the algorithm necessarily stops at 50 iterations. For the three runs, the CE-AMIS algorithm reaches the maximum value of iterations $p_{max} = 5$ without reaching the CV criterion. The final converged value of κ equals $\kappa_{2,n_2} = 30$ except for Run 3 where it is equal to $\kappa_{2,n_2} = 70$. The resulting failure probability estimates of this region are close to $\hat{P}_{f,2} = 2 \times 10^{-5}$ except for Run 3 where $\hat{P}_{f,2} = 3 \times 10^{-9}$. Finally, the last failing point P_3 is similar for all the runs. Once more the CE-AMIS algorithm reaches the maximum value of iterations $p_{max} = 5$ without reaching the CV criterion. The final converged value of κ equals $\kappa_{3,n_3} = 40$. The resulting failure probability estimates of this region are close to $\hat{P}_{f,1} = 1 \times 10^{-5}$. Therefore, the first failure region contributes the most to the global failure probability. The other two regions contribute quite equally, with, maybe, the failure region associated with P_2 a bit more relevant. The poor coordinates of P_2 in Run 3 explains why for this run, the global failure probability estimate is a bit lower.

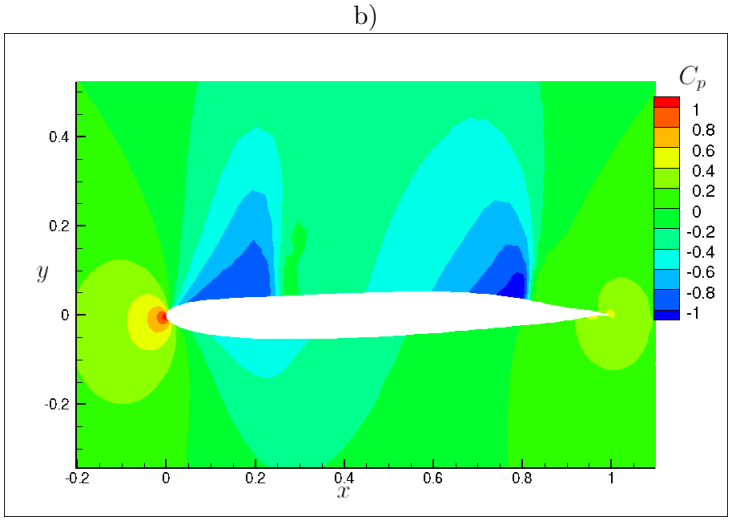
The numbers $N_{\text{opt},1}$ and $N_{\text{opt},2}$ vary a lot for each run. This is an indicator that the optimization



a)



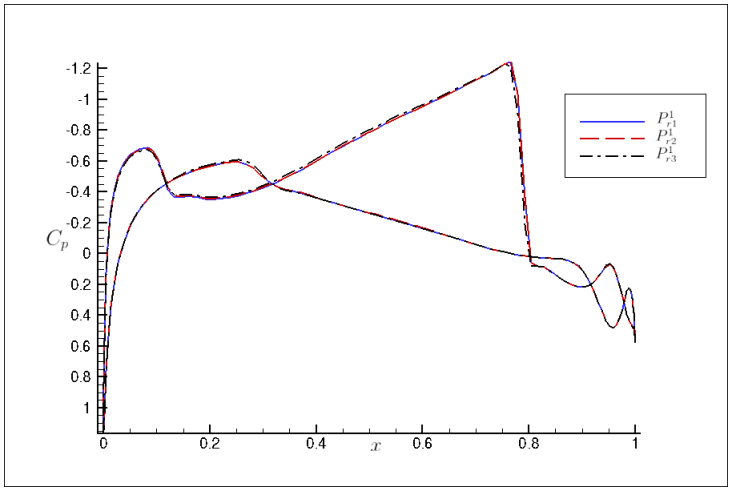
b)



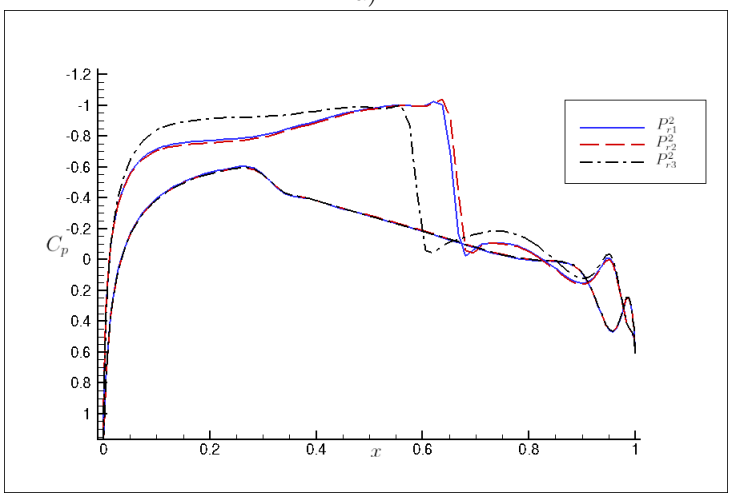
c)

Figure 5.20: Pressure coefficient fields at the critical point found by the proposed approach: a) P_1, b) P_2, c) P_3.

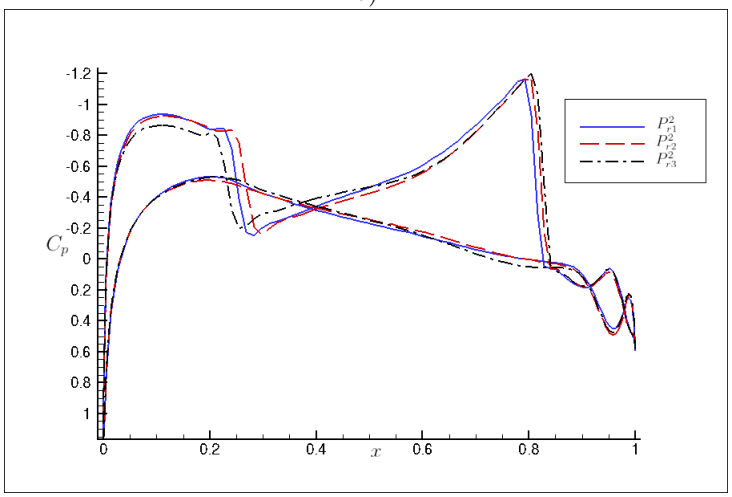
results seem to depend a lot on the random LHS employed to perform the random multi-start optimizations of the various identification problems. It is reminded here that the LHS performed is similar to the one presented in Section 5.3.1.3, but each observation is multiplied by the standard deviation



a)



b)



c)

Figure 5.21: Pressure coefficient distributions on both sides of the 2D airfoil at the critical points found by the proposed approach: a) P_1, b) P_2, c) P_3. The critical points found by Run 1 are represented in blue, those of Run 2 are represented in red and those of Run 3 in black.

5×10^{-4}. These numbers can be quite large, like in Run 1 where $N_{\text{opt},1}$ represents more than a third of the simulation budget N_{samp}

Compared to the Monte Carlo estimation, the simulation budget is divided by a factor close to 7, which is a very interesting gain for such a costly numerical application. This example also shows the value of finding the important failure region locations in the input space as this information can further help to understand the critical failure mechanisms and help the design.

Remarks on the simulation budget
As previously mentioned in Section 5.2.6.2, this application is an example where the structure of the simulation budget does not reflect the contribution of each failure region. Indeed, the first failure region is the most important, but the simulation budget for this failure region is only $3N = 3000$. Whereas for the other two failure regions, since the p_{max} criterion of the CE-AMIS algorithm is reached, they require each $5N = 5000$ evaluations of the costly lsf, for smaller contributions to the global failure probability. However, these failure regions are then not left unexplored, which is the strong asset of the proposed algorithm.

In Run 3, the failure region associated with P_2 is badly sampled. The coordinates of P_2 result in a large value of $r_{\text{opt},2}$, as P_2 is not found in the lower bound of the important ring. The radial auxiliary density is thus equal to $h_R = f_{R|R>r_{\text{opt}}}$, instead of $h_R = f_R$ when $\|P_2\|$ =LB. The resulting failure probability of this region is then underrated and the global failure probability as well. Consequently, the failure region associated with P_2 should not be discarded, even if its contribution is 10 times lower than the contribution of the failure region associated with P_1.

5.4 Conclusion

In this chapter we presented a new method to estimate the failure probability of a system in a high-dimensional standard elliptical space, for failure domain encompassing several failure regions. Here we discuss some final remarks and outlooks.

5.4.1 Necessity of standard elliptical inputs

On can wonder if the proposed algorithm is in fact only restricted to standard elliptical inputs, or if it could be applicable to other kinds of input space. Several points must be mentioned to answer this.

5.4.1.1 Definition of an important ring

First of all, without standard elliptical inputs, then an important ring might not be defined. The important ring of the elliptical inputs comes from the independence between the random variable R, which is the Euclidean norm of $\mathbf{X}$, and the random vector $\mathbf{T}$. With other inputs, such an independence might not exist. Suppose the optimization Eq. (5.1) of Section 5.2.2.2 is not constrained to an important ring and searches for failing points in the whole space $\mathbb{R}^d$. This search could then be endless since d is large. Indeed, the bulges built around the failing regions as well as the failure cones most likely do not cover the whole failure domain of $\mathbb{R}^d$. So, for large values of $\|\mathbf{X}\|$, there could always be a failing point in the failure domain, away for the bulges, who does not belong to any failure cone. Furthermore, the associated probability mass of such a failing point would be negligible as its norm is so large. Consequently, trying to sample in the vicinity of this failing point would be a waste of simulation budget. In fact, with the portfolio loss application, this phenomenon was encountered for wider important ring settings.

Therefore, the important ring is essential in the optimization problem for two reasons. First, it allows to focus on the most important part of the input space. Second, the search for the failure regions inside of it necessarily comes to an end at one point, as the important ring is a restricted space. Without elliptical inputs, once should focus on another criterion to restrict the space for the optimization. This restriction must not be too strong, as we are estimating rare events, thus the failure domain is supposed to be located in distant parts of the input space.

5.4.1.2 Selection of the radial auxiliary density

Next, without standard elliptical inputs, the choice of the radial auxiliary density h_R is not as obvious. The idea to condition a univariate law to being greater than r_{opt} is still valid but the choice of this law is not straightforward. One easy solution could be to use kernel density estimates (kde) (see Chapter 3

section 3.3.3.1), since they are very efficient in dimension one, restricted to $(r_{\text{opt}}, +\infty)$. A pre-sampling step should then be performed to build the kde.

5.4.1.3 Degeneration of the MIS weights

Finally, without elliptical inputs there are higher chances that the *curse of dimensionality* affects the MIS weights of the CE-AMIS algorithm. As underlined in Eq. (5.5) of the CE-AMIS algorithm, for standard elliptical inputs, there is a simplification in the MIS weights since the density f_R is evaluated both in the numerator and the denominator. Moreover, the density $f_{\mathbf{T}}$ is a constant for each $\left(\mathbf{T}_\ell^{(j)}\right)$. In other words, the expression of the MIS weights is quite simple in our method with standard elliptical inputs and would not be with other inputs. Consequently, the efficiency of the proposed method could be heavily affected by the dimension of the inputs.

5.4.2 Influence of IS auxiliary density in high-dimensional spaces

For the three applications tested in high-dimensional standard normal space, the failure regions always start before the important ring and spread across it. Thus the radial auxiliary density h_R is left unchanged : $h_R = f_R$ (the situation is referred to as Case 2). The accuracy of the proposed method suggests that in this particular space, the directional component of the IS density is the most influential. This becomes even more valid as the dimension increases and the variance of the random variable R decreases. However, it should be noted that with spherical failure domain in this particular space, then the proposed method would be quite inefficient. One solution could be to learn a better radial auxiliary density, either with kde as mentioned in the previous section, or with the Nakagami distribution employed in [Papaioannou et al., 2019]. The simplification in the MIS weights would not occur and a pre-sampling step should be necessary but the resulting IS density would perform better.

In contrast, in the application with standard multivariate Student inputs, the influence of h_R is the strongest and a study showed that if h_R is poorly constructed, the proposed method is not as efficient; cf Section 5.3.3.5. Consequently, in both random spaces, taking into account the stochastic representation of the elliptical inputs in the construction of the IS auxiliary density is relevant.

5.4.3 Proposed method without the gradient of the limit state function

In this manuscript, it is assumed that the gradient of the limit state function is available for the optimization problems of Section 5.2.2.2. When the gradient is not available, these optimizations are very difficult to perform in high-dimensional space and the proposed algorithm may be inefficient. However, many recent papers can be found in the literature which make it possible to find the design points for the FORM in high-dimensional standard normal space without employing the gradient [Zhong et al., 2020, Zhu et al., 2022]. Combining these new optimization schemes with the important ring constraint could be a promising gradient-free alternative.

In moderate dimension, the use of surrogates [Echard et al., 2011] could be appropriate: the limit state function g is first approximated with a meta model $\tilde{g}$ constructed from training points, then the gradient of this surrogate $\tilde{g}$ is available [Torii et al., 2017]. However, it should be noted that the accurate computation of the gradient of the limit state function is crucial in the proposed method as it may lead to poor coordinates of P_k, decreasing the efficiency of the algorithm, cf Section 5.3.3.5. In high dimension however, it is very difficult to build accurate meta models [Tabandeh et al., 2022], this solution is thus not an option.

Estimation de la sensibilité locale de la probabilité de défaillance par régression polynomiale hétéroscédastique

Nous présentons ici une nouvelle manière d'estimer la dérivée de la probabilité de défaillance du système, par rapport aux entrées déterministes. Il est supposé que la probabilité de défaillance a déjà été estimée par une méthode de simulation. Il s'agit alors de réutiliser le plus possible les évaluations de la fonction d'état limite nécessaires pour le calcul de la probabilité, afin d'augmenter le moins possible le budget de simulation. Le calcul de la probabilité ainsi que sa dérivée se font dans l'espace standard elliptique.

Expression de la dérivée dans un développent en série de Taylor

L'approche présentée s'inspire de l'approche Faible [Torii, 2020], puisqu'une approximation de la fonction indicatrice du domaine de défaillance du système est utilisée. Plus précisément, cette approximation est choisie comme un fonction de répartition continue Ξ_σ, définie avec un paramètre $\sigma > 0$ vérifiant la propriété suivante

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \mathbb{I}_{D_f(\mathbf{s})}(\mathbf{x}) = \mathbb{I}_{y \leq 0}(g(\mathbf{s}, \mathbf{x})) = \lim_{\sigma \to 0} \Xi_\sigma \left(-g(\mathbf{s}, \mathbf{x})\right),$$

où 0 est contenu dans l'intérieur du support de Ξ_σ. La fonction de probabilité $\overline{P_f(\mathbf{s}, \cdot)}$ est alors définie sur $\mathbb{R}^+ \backslash \{0\}$ ainsi

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \quad \overline{P_f(\mathbf{s}, \sigma)} = \int_{\mathbb{R}^d} \Xi_\sigma \left(-g(\mathbf{s}, \mathbf{x})\right) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \mathbb{E}_{f_{\mathbf{X}}} \left[\Xi_\sigma \left(-g(\mathbf{s}, \mathbf{X})\right)\right].$$

La limite de cette fonction de probabilité est égale à $P_f(\mathbf{s})$ lorsque σ tend vers 0. De même, la dérivée de cette fonction tend vers la quantité d'intérêt $\partial P_f(\mathbf{s}) / \partial s_\ell$, lorsque $\sigma \to 0$.

En travaillant sur l'expression de $\overline{P_f(\mathbf{s}, \sigma)}$ avec le changement de variable $H_{\mathbf{s}} = g(\mathbf{s}, \mathbf{X})$ dans la mesure image, il est possible de réécrire l'intégrale en faisant apparaître la fonction de répartition $F_{H_{\mathbf{s}}}$. En dérivant cette intégrale, une autre expression de $\partial \overline{P_f(\mathbf{s}, \sigma)} / \partial s_\ell$ s'obtient

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \quad \frac{\partial \overline{P_f(\mathbf{s}, \sigma)}}{\partial s_\ell} = \mathbb{E}_{\xi_\sigma} \left[\frac{\partial F_{H_{\mathbf{s}}}(-Z)}{\partial s_\ell}\right],$$

où Z est une variable aléatoire univariée de fonction de densité ξ_σ (la dérivée de Ξ_σ). Un développement en série de Taylor en 0 de la fonction dans l'espérance, notée T, permet d'aboutir à la décomposition suivante

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \quad \frac{\partial \overline{P_f(\mathbf{s}, \sigma)}}{\partial s_\ell} = \frac{\partial P_f(\mathbf{s})}{\partial s_\ell} + \frac{\mathbb{E}_{\xi_\sigma}[Z]}{1!} T'(0) + \frac{\mathbb{E}_{\xi_\sigma}\left[Z^2\right]}{2!} T''(0) + \dots + \frac{\mathbb{E}_{\xi_\sigma}\left[Z^n\right]}{n!} T^{(n)}(0) + \mathbb{E}_{\xi_\sigma}\left[R_n(Z)\right].$$

Ainsi, la quantité d'intérêt $\partial P_f(\mathbf{s})/\partial s_\ell$ se trouve être le coefficient d'ordre 0 d'une expression polynomiale des moments de Z. Selon la fonction de répartition Ξ_σ choisie, les moments de Z sont fonctions de σ. Par exemple, si Ξ_σ est la fonction de répartition d'une loi normale centrée de variance σ^2, alors les moments impairs de Z sont nuls. Le polynôme obtenu est ainsi un polynôme pair. L'identification des coefficients de ce polynôme, grâce à une régression polynomiale, permet alors d'estimer la quantité d'intérêt.

Afin de réaliser la régression polynomiale, m évaluations de $\partial\overline{P_f(\mathbf{s},\sigma)}/\partial s_\ell$ sont nécessaires, pour m valeurs de σ différentes. Il s'agit alors d'estimer l'intégrale suivante

$$\forall\sigma \in \mathbb{R}^+\backslash\{0\} \qquad \frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell} = \int_{\mathbb{R}} -\frac{\partial g(\mathbf{s},\mathbf{x})}{\partial s_\ell}\xi_\sigma(-g(\mathbf{s},\mathbf{x}))f_{\mathbf{x}}(\mathbf{x})\mathrm{d}\mathbf{x},$$

qui est l'expression directe de la dérivée, obtenue sans avoir effectué au préalable le changement de variable mentionné. Cette estimation pour un σ fixé est faisable avec des méthodes de simulations, c'est le principe de l'approche Faible [Torii, 2020]. Chaque évaluation de la fonction d'état limite nécessaire à l'estimation de la probabilité de défaillance est donc réutilisée. Il faut cependant aussi évaluer la dérivée de la fonction d'état limite par rapport à s_ℓ pour chaque observation, ce qui augmente le budget de simulation. Grâce à une approche de Monte-Carlo classique, l'estimateur obtenu est sans biais et sa variance théorique est une fonction strictement décroissante de σ, pour un budget de simulation constant. Ainsi, chaque estimateur $\partial\widehat{\overline{P_f(\mathbf{s},\sigma_l)}}/\partial s_\ell$, pour $l = 1,\ldots,m$ a un bruit différent. Plus σ est faible, plus le bruit est élevé. La régression polynomiale est alors hétéroscédastique.

Il est important de remarquer que la méthode proposée est théoriquement biaisée, puisque le modèle polynomial utilisé pour la régression sera forcément de degré fini, noté n. L'égalité dans l'équation du développement en série de Taylor est valable pour un nombre infini de termes. Puisqu'ici seuls les $n+1$ premiers termes seront considérés, le polynôme n'est alors qu'une approximation de $\partial\overline{P_f(\mathbf{s},\sigma)}/\partial s_\ell$. Plus la valeur de n est grande, plus le biais est faible, pour un σ fixé. De même, pour un n fixé, plus σ est faible, plus le biais est faible.

Régression polynomiale hétéroscédastique

Dans cette thèse, nous avons choisi de recourir aux méthodes linéaires des moindres carrés pour effectuer la régression polynomiale [Watson, 1967]. Il est néanmoins rappelé que les régressions polynomiales ne sont pas le sujet principal de ce chapitre, ni de cette thèse. Nous nous intéressons ici au calcul de la dérivée de la probabilité de défaillance. Ainsi, les ajustements détaillés dans la suite résultent de choix simples obtenus après plusieurs tests.

Pour ne pas augmenter le budget de simulation, chaque estimateur $\partial\widehat{\overline{P_f(\mathbf{s},\sigma_l)}}/\partial s_\ell$ est calculé avec les mêmes observations $\left(\mathbf{X}^{(j)}\right)_{j=1,\ldots,N}$ qui ont servi à l'estimation de la probabilité de défaillance. Cependant, une utilisation directe du même échantillon rend les estimateurs fortement corrélés entre eux. Le modèle de régression adapté est alors la méthode des moindres carrés généralisée (Generalized Least Squares GLS). De plus, puisque la matrice de covariance des estimateurs est inconnue et doit être aussi estimée, le terme "méthode GLS réalisable (Feasible GSL, FGSL)" est utilisé. Le cadre FGLS requiert l'inversion de la matrice de covariance. Ce calcul est particulièrement complexe, surtout dans notre contexte où l'estimation de la matrice de covariance par des méthodes de simulation est imparfaite et bruitée, puisque le budget de simulation est constant.

Afin de simplifier le cadre de régression, la technique de bootstrap [Horowitz, 2001] est ainsi utilisée. Cette technique permet, à partir de l'échantillon $\left(\mathbf{X}^{(j)}\right)_{j=1,\ldots,N}$ de générer d'autres échantillons indépendants de même propriété statistique. Cette technique n'augmente pas le budget de simulation et permet donc de rendre les estimateurs indépendants entre eux. La matrice de covariance estimée devient diagonale et est aisément inversible : la régression polynomiale s'effectue alors avec une méthode des moindres carrés pondérée (Weighted Least Squares WSL).

Avant de définir les paramètres de régression de la méthode WSL, les réponses de la fonction d'état limite $\left(Y^{(j)} = g\left(\mathbf{s},\mathbf{X}^{(j)}\right)\right)_{j=1,\ldots,N}$ sont adimensionnées afin de rendre le paramétrage aussi général que possible. L'intervalle de régression est alors défini grâce à un critère de coefficient de variation (CV) théorique. La régression ne concerne que les valeurs de σ_l telles que le CV théorique de l'estimateur

$\widehat{\partial P_f(\mathbf{s}, \sigma_l)}/\partial s_\ell$ ne dépasse pas une valeur seuil. Le degré du polynôme est quant à lui déterminé en faisant plusieurs expérimentations. En effet, grâce à la méthode de bootstrap, le budget de simulation est le même peu importe le degré du polynôme. Ainsi, plusieurs degrés peuvent être examinés, en commençant avec un degré faible. Dès lors que la valeur de l'estimateur de la dérivée ne varie plus, cela signifie que le biais peut être considéré comme négligeable et que la valeur correcte a été atteinte. De cette manière, un contrôle du biais est possible grâce au degré du polynôme.

L'application de cette approche avec une méthode de Monte-Carlo conduit à un intervalle de régression assez large, où la contrainte de CV joue seulement sur la borne inférieure de l'intervalle. Une illustration avec un exemple jouet permet de valider l'intérêt de l'approche proposée, cf Tables 6.1 et 6.2. Trois budgets de simulation différents sont expérimentés. Pour un budget de simulation similaire, les estimateurs de dérivées obtenus avec la méthode proposée ont toujours un biais empirique plus faible que ceux de l'approche Faible. Lorsque le budget de simulation est élevé, leur coefficient de variation est aussi plus faible que ceux de l'approche Faible.

En adaptant l'approche proposé aux méthodes d'échantillonnage préférentiel, la régression n'est pas tout à fait la même. En effet, pour des valeurs de σ élevées, la variance des estimateurs $\widehat{\partial P_f(\mathbf{s}, \sigma_l)}/\partial s_\ell$ augmente. La contrainte de CV joue alors sur les deux bornes de l'intervalle de régression. Pour l'exemple jouet étudié, l'intérêt de l'approche est une fois de plus démontré. Les estimateurs obtenus avec l'approche proposée ont un biais très faible, et un CV plus bas que ceux obtenus avec l'approche Faible, cf Table 6.3.

Il est important de noter que l'approche proposée permet aussi de calculer la variance théorique de la dérivée obtenue par régression. En effet, le cadre de régression WSL admet l'estimation de cette variance, avec une formule théorique. Néanmoins, pour l'exemple jouet étudié, cet estimateur de variance donne alors une valeur plus faible que la variance réelle observée.

Application numérique

La performance de cette nouvelle approche est mise à l'épreuve avec trois applications numériques. Pour toutes les applications les variables d'entrées sont normales. Les deux premières concernent des systèmes de dimension faible qui sont fréquemment étudiés dans la littérature et permettent donc de valider l'approche. La troisième application concerne un système de grande dimension, possédant deux régions de défaillance. Cette troisième application permet d'expérimenter l'approche proposée dans le cadre spécifique du manuscrit. Pour toutes les applications, l'approximation de la fonction indicatrice choisie est la fonction de répartition d'une loi normale centrée de variance σ^2, ainsi les polynômes sont pairs. Trois degrés de polynôme différents sont examinés : $n = 2$, $n = 4$ et $n = 6$.

La première application [Papaioannou et al., 2018,Torii, 2020] étudie le comportement d'une poutre en porte-à-faux soumise à un fléchissement biaxial. La dimension des variables aléatoires est égale à 4. Deux fonctions d'état limite sont alors définies. Elles dépendent respectivement de deux et trois paramètres déterministes de conception du système. Il s'agit alors de calculer les dérivées des probabilités de défaillance par rapport à ces deux ou trois paramètres de conception. Afin d'estimer la première probabilité de défaillance ainsi que ses dérivées, l'approche proposée est associée à un algorithme d'échantillonnage préférentiel adaptatif non paramétrique (NAIS [Zhang, 1996]). La valeur de la dérivée obtenue est constante pour les trois degrés de polynôme, cf Table 6.4. Ainsi une régression de degré égal à $n = 2$ suffit pour aboutir à un estimateur de dérivée de valeur correcte. Les résultats soulignent l'intérêt de la méthode par rapport à l'approche Faible: la valeur des estimateurs est correcte pour un CV plus faible. Pour la deuxième fonction d'état limite, l'approche proposée est associé à un algorithme d'échantillonnage préférentiel adaptatif paramétrique (iCE-SG [Papaioannou et al., 2019]). La valeur de la dérivée obtenue est constante dès lors que $n \geq 4$, cf Table 6.5. Ainsi une régression de degré égal à $n = 4$ est nécessaire pour aboutir à un estimateur de dérivée de valeur correcte. Les estimateurs obtenus avec l'approche proposée sont alors moins biaisés que ceux obtenus avec l'approche Faible. Cependant, leur variance est alors légèrement supérieure.

La deuxième application représente une ferme de toit soumise à une charge aléatoire [Song et al., 2009]. La dimension des variables aléatoires est égale à 6. Pour cet exemple, les variables aléatoires normales de départ ne sont pas standards. Il s'agit alors de calculer la probabilité de défaillance et ses dérivées par rapport aux paramètres de distribution des entrées de départ du système. Une transformation isoprobabiliste est alors nécessaire pour calculer la probabilité ainsi que sa dérivée

dans l'espace standard normal. Afin d'estimer la probabilité de défaillance ainsi que ses dérivées, l'approche proposée est de nouveau associée à l'algorithme NAIS. Douze dérivées sont estimées. Pour dix paramètres de distribution, la valeur de la dérivée obtenue est constante pour les trois degrés de polynôme, cf Tables 6.6 et 6.7. Ainsi une régression de degré égal à $n = 2$ suffit pour aboutir à un estimateur de dérivée de valeur correcte pour ces dix paramètres. L'approche proposée permet alors d'obtenir un estimateur de biais bas et de CV inférieur à celui obtenu avec l'approche Faible. Pour les deux autres paramètres de distributions, les valeurs des estimateurs obtenus avec l'approche proposée ne sont pas constantes et leur CV sont très larges, particulièrement pour $n = 6$. Lorsque $n = 6$, ces résultats sont la conséquence du phénomène du surapprentissage. La régression polynomiale ne parvient alors pas à lisser le bruit des estimateurs $\widehat{\partial P_f(\mathbf{s}, \sigma_l)}/\partial s_\ell$, cf Figure 6.9.

La dernière application s'intéresse au déplacement maximal d'un oscillateur de Duffing [Zuev, 2009]. La dimension des variables aléatoires est égale à 100. La fonction d'état limite dépend de deux paramètres de conception du système. Il s'agit alors de calculer la probabilité de défaillance et ses dérivées par rapport à ces deux paramètres de conception. Afin d'estimer la probabilité et ses dérivées, l'approche proposée est associée à l'algorithme d'échantillonnage préférentiel multiple adaptatif paramétrique présenté au Chapitre 5 du manuscrit. La valeur de la dérivée obtenue varie très peu pour les trois degrés de polynôme, cf Table 6.8. Ainsi une régression de degré égal à $n = 2$ suffit pour aboutir à un estimateur de dérivée de valeur correcte. Les estimateurs obtenus avec l'approche proposée sont alors bien moins biaisés que ceux obtenus avec l'approche Faible. Cependant, leur CV est alors plus large. Néanmoins, puisque le biais de l'estimateur est contrôlé grâce à l'étude sur le degré du polynôme, l'approche proposée peut tout de même être considérée comme une amélioration de l'approche Faible. Le CV plus faible des estimateurs obtenus avec l'approche Faible ne garantit pas la qualité de l'estimation vis-à-vis du biais.

Pour les trois applications, l'estimateur de la variance des dérivées était à chaque fois sous-évalué. De plus, la variation de cet estimateur était très élevée. Ainsi, comme précédemment mentionné, il semble que l'estimation de la variance, possible dans le cadre WSL, ne soit pas efficace en pratique.

Conclusion

Dans ce chapitre, une nouvelle manière d'estimer la dérivée de la probabilité de défaillance par rapport aux paramètres déterministes du système a été présentée. Cette dérivée s'obtient grâce à une régression polynomiale hétéroscédastique. Les paramètres déterministes peuvent représenter les paramètres de conception du système ou bien les paramètres de distribution du vecteur aléatoire de départ. Le calcul de la probabilité ainsi que sa dérivée se font dans l'espace standard elliptique, après une transformation isoprobabiliste. Cette nouvelle approche, qui s'inspire de l'approche Faible, est présentée comme une amélioration de cette dernière. Nous abordons ici quelques remarques finales et ainsi que des perspectives.

La principale innovation proposée par l'approche présentée consiste à exprimer la dérivée de la probabilité comme le terme d'ordre zéro d'un développement en série de Taylor. Ce terme constant peut alors être identifié grâce à une régression polynomiale. Puisque le budget de simulation ne doit pas être augmenté, le paramétrage de la régression a été fait de la manière la plus simple possible afin de respecter cette contrainte. En conséquence, il semble que la méthode pourrait être améliorée suite à une étude plus approfondie sur les régressions polynomiales hétéroscédastiques.

Plus particulièrement, une des perspectives les plus intéressantes consiste à améliorer la stabilité des estimateurs de dérivée obtenus. En effet, pour de nombreuses applications, la valeur des estimateurs étaient très proches de la valeur de référence, mais leur CV était un peu trop large par rapport à l'approche Faible. Puisque la méthode proposée est présentée comme une amélioration de l'approche Faible, arriver à baisser le niveau de CV des estimateurs en plus du biais serait favorable.

De même, l'estimateur de la variance des dérivées semble toujours être sous-évalué. Un estimateur de la variance efficace permet d'évaluer la qualité de l'estimation de la dérivée avec une unique simulation. Puisqu'un tel estimateur est disponible dans l'approche Faible, il semble aussi nécessaire d'améliorer cette estimation dans l'approche proposée.

Cette approche a été testée en grande dimension dans un espace standard normal avec un système possédant deux régions de défaillance. Bien que les résultats ne soient pas aussi précis que dans d'autres exemples, notamment à cause du CV des estimateurs qui est assez élevé, nous ne pensons pas que cela vienne de la grande dimension du système ou bien de la multitude des régions de défaillance. En

effet, l'approche proposée ici considère seulement l'espace unidimensionnel de la sortie de la fonction d'état limite $Y = g(\mathbf{s}, \mathbf{X})$. Ainsi, elle est indépendante de la dimension des entrées $\mathbf{X}$ ainsi que de la configuration du domaine de défaillance. De même, nous ne pensons pas que la distribution elliptique initiale des entrées aléatoires influence les résultats. Ainsi, une autre perspective intéressante serait d'appliquer la méthode proposée à d'autres systèmes de grande dimension, possédant plusieurs régions de défaillance, dont les entrées suivent une autre distribution elliptique que la loi normale.

Enfin, les différentes applications présentées dans ce chapitre ont souligné la possibilité de combiner l'approche proposée à plusieurs méthodes de simulation. En effet, seuls le vecteur des évaluation de la fonction d'état limite et celui de ses dérivées sont nécessaires. La densité ayant permis de générer l'échantillon n'est pas utilisée. De sorte, nous supposons que l'approche présentée peut être associée à de nombreuse méthodes de simulation, autre que la méthode de Monte-Carlo et l'échantillonnage préférentiel. Puisque la méthode proposée est ainsi valable pour des systèmes de faible et large dimensions, possédant plusieurs régions de défaillance et dont la distribution des variables aléatoires importe peu, le fait de pouvoir la combiner avec de nombreuses méthodes de simulation la rend particulièrement flexible à l'étude de différents systèmes.

Chapter 6

Sensitivity estimation through polynomial regression

Contents

6.1 Introduction . . . 148
6.2 New sensitivity expression through Taylor series expansion . . . 148
6.2.1 Baseline of the mathematical framework of the approach . . . 149
6.2.1.1 Approximation of the failure indicator function . . . 149
6.2.1.2 Change of variable . . . 149
6.2.1.3 Taylor series expansion . . . 150
6.2.2 Illustration with a one-dimensional toy example . . . 151
6.2.2.1 Reliability analysis of the toy example . . . 151
6.2.2.2 Choice of the indicator function approximation . . . 151
6.2.2.3 New expression of the Taylor series expansion . . . 152
6.2.3 Combining sampling methods and polynomial regression to derive the failure probability sensitivity . . . 152
6.2.3.1 Regression with crude Monte Carlo method . . . 154
6.2.3.2 Theoretical bias of the sensitivity estimate . . . 155
6.3 Heteroscedastic polynomial regression . . . 155
6.3.1 Linear least squares method in our specific context . . . 155
6.3.1.1 Generalized Least Squares regression framework . . . 156
6.3.1.2 Pre-processing of the covariance matrix . . . 156
6.3.1.3 Remark on the bias of the estimate . . . 157
6.3.2 Settings of the regression parameters . . . 157
6.3.2.1 Scaling of the limit state function . . . 157
6.3.2.2 Choice of the regression interval and the polynomial degree . . . 157
6.3.2.3 Illustration with the toy example . . . 158
6.3.3 Proposed approach with importance sampling methods . . . 162
6.3.3.1 Change of behavior of the variance function . . . 163
6.3.3.2 Illustration with the toy example . . . 163
6.4 Numerical investigation . . . 165
6.4.1 Cantilever beam . . . 166
6.4.1.1 Presentation of the application . . . 166
6.4.1.2 Sensitivity analysis for the first failure of the system . . . 166
6.4.1.3 Sensitivity analysis for the second failure of the system . . . 167
6.4.2 Roof truss . . . 168
6.4.2.1 Presentation of the application . . . 168
6.4.2.2 Sensitivity analysis of the system . . . 169
6.4.2.3 Focus on the derivatives with respect to δ_{Z_2} and δ_{Z_6} . . . 169

6.4.3 Duffing oscillator 171
6.4.3.1 Presentation of the application 171
6.4.3.2 Sensitivity analysis of the system 172
6.5 Conclusion 173
6.5.1 Discussion on the heteroscedastic polynomial regression 173
6.5.1.1 Settings of the polynomial regression 174
6.5.1.2 Theoretical variance of the regression estimate 174
6.5.2 High-dimensional elliptical space with multiple failure regions 174
6.5.3 Generalization of the proposed approach 174

6.1 Introduction

In this chapter we present an original approach to estimate the derivatives of P_f with respect to the deterministic inputs $\mathbf{s} \in \mathbb{R}^p$ of the system. The main highlights of the proposed method are as follows:

- the derivatives of the failure probability are derived in a context close to the Weak approach described in Chapter 4 Section 4.3.3.1, as an approximation of the failure indicator function is employed to compute the sensitivity;
- the resulting derivative is the coefficient of order zero of a Taylor series expansion, which can be recovered with a heteroscedastic polynomial regression;
- the dimension d of the random inputs as well as the number of failure regions in the failure domain have no influence on the method as the heteroscedastic polynomial regression is performed in the one-dimensional space of the limit state function output $g(\mathbf{s}, \mathbf{X})$;
- the final simulation budget is the same as in the Weak approach context, as the derivatives of the costly limit state function with respect to s_ℓ have to be evaluated, for $s_\ell = 1, \ldots, p$. It should be noted that the evaluation of these derivatives are mandatory as long as $\mathbf{s}$ are not distribution parameters only, cf Chapter 4 Section 4.3;
- even though the proposed method is theoretically biased, the resulting bias of the sensitivity estimates can be controlled by increasing the degree of the polynomial used for the regression;
- the proposed method can be seen as an improvement of the Weak approach, as it is expected to lead to sensitivity estimates less biased and more precise in terms of CV.

This approach is suited for both distribution parameters as well as design parameters and is performed in the standard elliptical space. Consequently, the standardized limit state function g depends upon the parameters $\mathbf{s}$ whereas the pdf of the inputs $f_{\mathbf{X}}$ is parameter-free, cf Chapter 4 Section 4.3. It is assumed that the estimation of the failure probability is already performed with a simulation method. The observations generated during the procedure are then reused in the proposed approach to increase as little as possible the simulation budget, as will be detailed.

This chapter is organized as follows. Sections 6.2 and 6.3 describe the proposed algorithm. In order to illustrate the new approach in these first sections, a one-dimensional toy example is detailed. Section 6.2 focuses on the mathematical background leading to the new expression of the sensitivities while Section 6.3 presents the heteroscedastic polynomial regression. Then Section 6.4 presents three numerical examples, with standard normal inputs. The results are summarized and some conclusions are drawn in Section 6.5.

6.2 New sensitivity expression through Taylor series expansion

The sensitivity analysis presented here is performed in the standard elliptical space. Therefore, no matter the nature of vector $\mathbf{s}$, its influence is expressed in the failure domain D_f of the system,

through the limit state function g. We recall here that the failure probability is defined with the following domain integral and expected value

$$P_f(\mathbf{s}) = \int_{D_f(\mathbf{s})} f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} \mathbb{I}_{D_f(\mathbf{s})}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \mathbb{E}_{f_{\mathbf{X}}}\left[\mathbb{I}_{D_f(\mathbf{s})}(\mathbf{X})\right],$$

where $D_f(\mathbf{s}) = \{\mathbf{x} \in \mathbb{R}^d | \, g(\mathbf{s}, \mathbf{x}) \leq 0\}$. The mathematical framework of the approach is first presented. Next a one-dimensional toy example serves as an example to illustrate the different mathematical quantities introduced. Finally the main idea of the approach is introduced, which is to use sampling methods to compute the derivative of the failure probability as a byproduct of a heteroscedastic polynomial regression.

6.2.1 Baseline of the mathematical framework of the approach

The baseline of the mathematical framework of the approach, detailed in this section, is summarized as follows. As in the Weak approach presented in Chapter 4 Section 4.3.3.1, the failure domain indicator function is first approximated with a smoother function. Then, a change of random variable in the image measure is introduced. The resulting integral is differentiable and its derivative with respect to s_ℓ with $\ell = 1, \ldots, p$ is expressed as an expected value. Next, the Taylor series expansion of the expected value is derived and the sensitivity with respect to s_ℓ is finally identified amongst the Taylor series coefficients.

6.2.1.1 Approximation of the failure indicator function

As the indicator function is not differentiable, several smoother functions have been used in the literature as surrogate [Torii, 2020, Papaioannou et al., 2013, Lacaze et al., 2015]; cf Chapter 4 Section 4.3.3. Here, we focus on approximations which are continuous cumulative distribution functions Ξ_σ, defined with a parameter $\sigma > 0$ which verify the following property

$$\forall \mathbf{x} \in \mathbb{R}^d \quad \mathbb{I}_{D_f(\mathbf{s})}(\mathbf{x}) = \mathbb{I}_{y \leq 0}(g(\mathbf{s}, \mathbf{x})) = \lim_{\sigma \to 0} \Xi_\sigma\left(-g(\mathbf{s}, \mathbf{x})\right), \tag{6.1}$$

where 0 is contained in the interior of the support of Ξ_σ. The failure probability function $\overline{P_f(\mathbf{s}, \cdot)}$ is then defined on $\mathbb{R}^+ \backslash \{0\}$ such as

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \quad \overline{P_f(\mathbf{s}, \sigma)} = \int_{\mathbb{R}^d} \Xi_\sigma\left(-g(\mathbf{s}, \mathbf{x})\right) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \mathbb{E}_{f_{\mathbf{X}}}\left[\Xi_\sigma\left(-g(\mathbf{s}, \mathbf{X})\right)\right], \tag{6.2}$$

and we further assume the cdf Ξ_σ to be regular enough to verify the following properties

$$P_f(\mathbf{s}) = \lim_{\sigma \to 0} \overline{P_f(\mathbf{s}, \sigma)} \quad \text{and} \quad \frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = \lim_{\sigma \to 0} \frac{\partial \overline{P_f(\mathbf{s}, \sigma)}}{\partial s_\ell}.$$

In the Weak approach framework presented in Chapter 4 Section 4.3.3.1, the sensitivity of P_f is obtained by differentiating Eq. (6.2) with respect to s_ℓ, $\ell = 1, \ldots, p$, for a fixed value of σ denoted $\widetilde{\sigma}$. Here, we introduce a modification in the approach with a change of random variable to derive another expression of the failure probability function Eq. (6.2).

6.2.1.2 Change of variable

It is assumed in the rest of this chapter that the random response of the system $g(\mathbf{s}, \mathbf{X})$ is an absolutely continuous univariate random variable. Considering the following change of random variable in the image measure: $H_{\mathbf{s}} = g(\mathbf{s}, \mathbf{X})$, the failure probability function Eq. (6.2) becomes

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \quad \overline{P_f(\mathbf{s}, \sigma)} = \mathbb{E}_{f_{H_{\mathbf{s}}}}\left[\Xi_\sigma\left(-H_{\mathbf{s}}\right)\right] = \int_{\mathbb{R}} \Xi_\sigma\left(-h\right) f_{H_{\mathbf{s}}}(h) \mathrm{d}h, \tag{6.3}$$

where $f_{H_{\mathbf{s}}}$ is the unknown density of the univariate random variable $H_{\mathbf{s}}$. It should be noted that the failure probability function is thus written as an integral defined on $\mathbb{R}$ rather than $\mathbb{R}^d$. Denoting ξ_σ the

pdf associated with the Ξ_σ distribution, the failure probability function Eq. (6.3) can then be rewritten in the following way

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \overline{P_f(\mathbf{s}, \sigma)} = \int_{\mathbb{R}} \Xi_\sigma(-h) f_{H_\mathbf{s}}(h) \mathrm{d}h = \int_{\mathbb{R}} \left(\int_{-\infty}^{-h} \xi_\sigma(z) \, \mathrm{d}z \right) f_{H_\mathbf{s}}(h) \mathrm{d}h \tag{6.4}$$

$$= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{I}_{\{z \leq -h\}} \xi_\sigma(z) \, \mathrm{d}z \right) f_{H_\mathbf{s}}(h) \mathrm{d}h. \tag{6.5}$$

Applying the theorem of Fubini to integral Eq. (6.5) results in

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \overline{P_f(\mathbf{s}, \sigma)} = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{I}_{\{h \leq -z\}} f_{H_\mathbf{s}}(h) \mathrm{d}h \right) \xi_\sigma(z) \, \mathrm{d}z = \int_{\mathbb{R}} F_{H_\mathbf{s}}(-z) \xi_\sigma(z) \, \mathrm{d}z,$$

where $F_{H_\mathbf{s}}$ is the unknown cdf of the univariate random variable $H_\mathbf{s}$. Employing this new expression of the failure probability function $\overline{P_f(\mathbf{s}, \cdot)}$ to compute the derivative with regard to s_ℓ with $\ell = 1, \ldots, p$ gives

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \frac{\partial \overline{P_f(\mathbf{s}, \sigma)}}{\partial s_\ell} = \int_{\mathbb{R}} \frac{\partial F_{H_\mathbf{s}}(-z)}{\partial s_\ell} \xi_\sigma(z) \mathrm{d}z = \mathbb{E}_{\xi_\sigma} \left[\frac{\partial F_{H_\mathbf{s}}(-Z)}{\partial s_\ell} \right], \tag{6.6}$$

where Z is a univariate random variable of pdf ξ_σ. Eq. (6.6) is a new expression of the failure probability derivative function in a modified Weak approach context, in which the cdf nature of the approximation function Ξ_σ is taken advantage of. Instead of trying to evaluate this expression of the derivative function at a specific value of σ, we use the Taylor series expansion to remove the dependence in σ, as detailed in the next section.

6.2.1.3 Taylor series expansion

Taylor series expansion is a powerful tool to help derive the expression of expected values [Ang and Tang, 2007, Khan, 2004, Biau and Mason, 2015]. In our specific context, we derive the Taylor series expansion in the neighborhood of 0 of the function T defined as $T(z) = \partial F_{H_\mathbf{s}}(-z)/\partial s_\ell$ for all $z \in \mathbb{R}$. Assuming $F_{H_\mathbf{s}}$ is $\mathcal{C}^{n+2}$ in 0, the Taylor series expansion leads to

$$T(Z) = T(0) + \frac{Z}{1!} T'(0) + \frac{Z^2}{2!} T''(0) + \ldots + \frac{Z^n}{n!} T^{(n)}(0) + R_n(Z), \tag{6.7}$$

where the remainder is expressed in the integral form with [Apostol, 1967]

$$R_n(Z) = \int_0^Z \frac{(Z-t)^n}{n!} T^{(n+1)}(t) \mathrm{d}t.$$

Applying the expectation to both sides of the equality Eq. (6.7) results in the following equation

$$\mathbb{E}_{\xi_\sigma}[T(Z)] = T(0) + \frac{\mathbb{E}_{\xi_\sigma}[Z]}{1!} T'(0) + \frac{\mathbb{E}_{\xi_\sigma}[Z^2]}{2!} T''(0) + \ldots + \frac{\mathbb{E}_{\xi_\sigma}[Z^n]}{n!} T^{(n)}(0) + \mathbb{E}_{\xi_\sigma}[R_n(Z)]. \tag{6.8}$$

The first term $T(0)$ on the right side of the equality is equal to the derivative of the function $F_{H_\mathbf{s}}$ with respect to s_ℓ, evaluated in $z = 0$. This term is equal to $\partial P_f(\mathbf{s})/\partial s_\ell$ as detailed in the Appendix D.1. Therefore, the following expression of $\partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$ is obtained

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \frac{\partial \overline{P_f(\mathbf{s}, \sigma)}}{\partial s_\ell} = \frac{\partial P_f(\mathbf{s})}{\partial s_\ell} + \frac{\mathbb{E}_{\xi_\sigma}[Z]}{1!} T'(0) + \frac{\mathbb{E}_{\xi_\sigma}[Z^2]}{2!} T''(0) + \ldots$$
$$+ \frac{\mathbb{E}_{\xi_\sigma}[Z^n]}{n!} T^{(n)}(0) + \mathbb{E}_{\xi_\sigma}[R_n(Z)]. \tag{6.9}$$

From Eq. (6.9) the derivative of P_f with respect to s_ℓ appears to be the constant term of a polynomial expression of the moments of Z, where Z is a random univariate variable of cdf Ξ_σ and pdf ξ_σ.

6.2.2 Illustration with a one-dimensional toy example

Let X be a univariate standard normal variable. The limit state function of the toy example is defined as $g : X \mapsto aX + b$ where $\mathbf{s} = [a, b] \in \mathbb{R}^* \times \mathbb{R}$ is the deterministic vector. It is noticeable that the response of the limit state function $Y = g([a, b], X)$ follows then a normal law of mean b and variance a^2, $Y \sim \mathcal{N}(b, a^2)$.

As the toy example is very simple, all the quantities mentioned in the equations above can be analytically computed. First the reliability analysis of the toy example is briefly presented. A particular choice of approximation function Ξ_σ is then detailed, leading to the expression of the Taylor series expansion.

6.2.2.1 Reliability analysis of the toy example

The failure probability of the toy example is given by

$$P_f(\mathbf{s}) = \mathbb{P}(aX + b < 0) = \mathbb{P}\left(X < -\frac{b}{a}\right) = \Psi\left(-\frac{b}{a}\right),$$

where Ψ is the cdf of a standard normal random variable. Denoting ϕ the pdf of a standard normal variable, the derivatives of P_f are equal to

$$\frac{\partial P_f(\mathbf{s})}{\partial a} = \frac{b}{a^2}\phi\left(\frac{b}{a}\right) \quad \text{and} \quad \frac{\partial P_f(\mathbf{s})}{\partial b} = -\frac{1}{a}\phi\left(\frac{b}{a}\right).$$

The random variable $H_\mathbf{s}$ follows a normal distribution of mean b and variance a^2, therefore $F_{H_\mathbf{s}}$ is $\mathcal{C}^\infty$.

6.2.2.2 Choice of the indicator function approximation

The selected approximation of the indicator function is the cdf of a centered normal variable of variance σ^2. Therefore $\forall y \in \mathbb{R}\,,\ \Xi_\sigma(y) = \Psi(y/\sigma)$ where Ψ is the cdf of a univariate standard normal variable. This particular choice of smooth approximation of the failure indicator function has been studied in [Lacaze et al., 2015] and employed in various methods [Papaioannou et al., 2013, Papaioannou et al., 2018, Torii, 2020].

With this particular approximation function, the failure probability function $\sigma \to \partial\overline{P_f(\mathbf{s},\sigma)}/\partial a$ and $\sigma \to \partial\overline{P_f(\mathbf{s},\sigma)}/\partial b$ can be analytically computed. From Eq. (6.2), one has

$$\forall \sigma \in \mathbb{R}^+\backslash\{0\} \qquad \overline{P_f(\mathbf{s},\sigma)} = \int_\mathbb{R} \Psi\left(\frac{-ax-b}{\sigma}\right)\phi(x)\mathrm{d}x,$$

since $f_X = \phi$, the univariate standard normal pdf. Consequently, the failure probability derivative functions are expressed as

$$\forall \sigma \in \mathbb{R}^+\backslash\{0\} \quad \frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial a} = -\int_\mathbb{R} \frac{x}{\sigma}\phi\left(\frac{-ax-b}{\sigma}\right)\phi(x)\mathrm{d}x = \frac{1}{\sqrt{a^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2+\sigma^2}\right)\right)\frac{ab}{a^2+\sigma^2}$$

$$\forall \sigma \in \mathbb{R}^+\backslash\{0\} \quad \frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial b} = -\int_\mathbb{R} \frac{1}{\sigma}\phi\left(\frac{-ax-b}{\sigma}\right)\phi(x)\mathrm{d}x = -\frac{1}{\sqrt{a^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2+\sigma^2}\right)\right),$$

where the detailed calculus can be found in the Appendix D.2. It should be noted that the convergence of the failure probability derivative function proven in [Papaioannou et al., 2013] is indeed verified with this example as one has

$$\lim_{\sigma\to 0}\frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial a} = \frac{b}{a^2}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2}\right)\right) = \frac{\partial P_f(\mathbf{s})}{\partial a}$$

$$\lim_{\sigma\to 0}\frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial b} = -\frac{1}{a}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2}\right)\right) = \frac{\partial P_f(\mathbf{s})}{\partial b}.$$

Furthermore, since the variable Z of the Taylor series expansion is a centered normal variable of variance σ^2, its moments are known and we recall here their expression

$$\forall k \in \mathbb{N} \qquad \mathbb{E}_{\xi_\sigma}\left[Z^{2k}\right] = \frac{(2k)!}{2^k k!}\sigma^{2k} \quad \text{and} \quad \mathbb{E}_{\xi_\sigma}\left[Z^{2k+1}\right] = 0.$$

Consequently, the expression of Eq. (6.9) can be further detailed and gives

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell} = \frac{\partial P_f(\mathbf{s})}{\partial s_\ell} + \frac{\sigma^2}{2} T''(0) + \frac{\sigma^4}{4 \times 2!} T^{(4)}(0) + \dots + \frac{\sigma^{2n}}{2^n n!} T^{(2n)}(0) + \mathbb{E}_{\xi_\sigma}\left[R_{2n}(Z)\right] \quad (6.10)$$

Therefore, the failure probability sensitivity is equal to the constant coefficient of an even polynomial in σ. It should be noted that the polynomial is even as the pdf ξ_σ is an even function. Therefore, if ones chooses Ξ_σ as the cdf of a uniform random variable in the interval $[-\sigma/2, \sigma/2]$ like in [Torii, 2020, Torii and Novotny, 2021], the resulting polynomial will also be even.

6.2.2.3 New expression of the Taylor series expansion

Since an analytical expression of the function $F_{H_\mathbf{s}}$ is available, the other coefficients of the polynomial which are derivatives of the function T multiplied by the moments of Z can be analytically derived as well. Computing the derivatives of order 2 and 4 of the function T (see Appendix D.3) gives the following analytical formula

$$T_a''(0) = \frac{b}{a^4}\phi\left(\frac{b}{a}\right)\left(-3 + \frac{b^2}{a^2}\right) \quad \text{and} \quad T_a^{(4)}(0) = \frac{b}{a^6}\phi\left(\frac{b}{a}\right)\left(15 - 10\frac{b^2}{a^2} + \frac{b^4}{a^4}\right),$$
$$T_b''(0) = \frac{1}{a^3}\phi\left(\frac{b}{a}\right)\left(1 - \frac{b^2}{a^2}\right) \quad \text{and} \quad T_b^{(4)}(0) = \frac{1}{a^5}\phi\left(\frac{b}{a}\right)\left(-3 + 6\frac{b^2}{a^2} - \frac{b^4}{a^4}\right).$$

Given these formulas, we introduce the polynomials P_2 and P_4 of order 2 and order 4 defined on $\mathbb{R}^+$ as

$$P_2(\sigma) = \frac{\partial P_f(\mathbf{s})}{\partial a} + \frac{\sigma^2}{2}\frac{b}{a^4}\phi\left(\frac{b}{a}\right)\left(-3 + \frac{b^2}{a^2}\right) \text{ and } P_4(\sigma) = P_2(\sigma) + \frac{\sigma^4}{4 \times 2!}\frac{b}{a^6}\phi\left(\frac{b}{a}\right)\left(15 - 10\frac{b^2}{a^2} + \frac{b^4}{a^4}\right),$$

for parameter a and

$$P_2(\sigma) = \frac{\partial P_f(\mathbf{s})}{\partial b} + \frac{\sigma^2}{2}\frac{1}{a^3}\phi\left(\frac{b}{a}\right)\left(1 - \frac{b^2}{a^2}\right) \text{ and } P_4(\sigma) = P_2(\sigma) + \frac{\sigma^4}{4 \times 2!}\frac{1}{a^5}\phi\left(\frac{b}{a}\right)\left(-3 + 6\frac{b^2}{a^2} - \frac{b^4}{a^4}\right),$$

for parameter b. An illustration of these polynomials is shown in Figure 6.1, with the derivative functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s},\sigma)}/\partial a$ and $\sigma \mapsto \partial \overline{P_f(\mathbf{s},\sigma)}/\partial b$ plotted as well.

From the curves shown in Figure 6.1, it seems polynomials P_2 and P_4 are quite similar and close to the failure probability derivative function for $\sigma \in [0, 0.5]$, however for higher values of σ, the plots begin to diverge, with a stronger divergence between P_2 and the failure probability derivative function than between P_4 and the failure probability derivative function. This result is not surprising as for low values of σ, the first coefficients of the Taylor series expansion matter more than for higher values of σ, where higher order coefficients are dominant.

Consequently, if one focuses on a small interval of low σ values, the failure probability derivative function can be approximated with a polynomial of order 2 or 4 for this particular toy example, whose constant coefficient is equal to the quantity of interest: the probability sensitivity. Therefore, if the coefficients of this polynomial are identified, the probability sensitivity is identified as well. This identification is the main idea of the new approach proposed.

6.2.3 Combining sampling methods and polynomial regression to derive the failure probability sensitivity

In this section and in the rest of this chapter, we denote P_n the polynomial of order n such as

$$\forall \sigma \in \mathbb{R}^+ \qquad P_n(\sigma) = \frac{\partial P_f(\mathbf{s})}{\partial s_\ell} + \frac{\mathbb{E}_{\xi_\sigma}[Z]}{1!}T'(0) + \frac{\mathbb{E}_{\xi_\sigma}\left[Z^2\right]}{2!}T''(0) + \dots + \frac{\mathbb{E}_{\xi_\sigma}\left[Z^n\right]}{n!}T^{(n)}(0).$$

Consequently one has

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell} - P_n(\sigma) = \mathbb{E}_{\xi_\sigma}\left[R_n(Z)\right] \quad \text{and for a small } \sigma \quad \frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell} \approx P_n(\sigma).$$

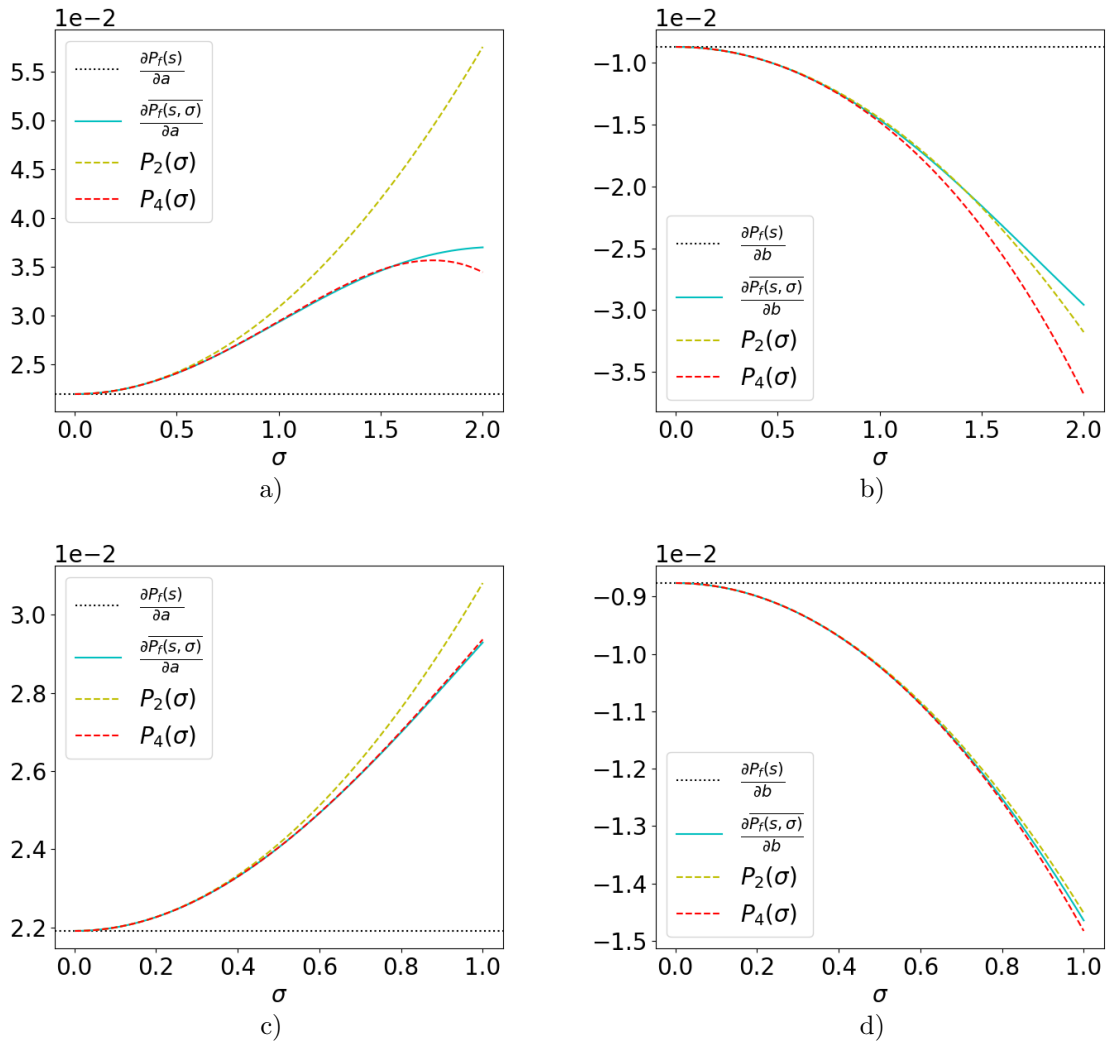


Figure 6.1: Illustration of the polynomials P_2 and P_4 for both of the derivatives of P_f. The real derivative value is in black dotted line and the derivative function $\sigma \mapsto \partial\overline{P_f(\mathbf{s},\sigma)}/\partial s_\ell$ is in blue. The parameters of the toy example are set as follows: $a = 2$ and $b = 5$. Figures a) and c) display the derivative with respect to a while Figures b) and d) display the derivative with respect to b. Figures a) and b) show the evolution of the different functions for a range of σ in $(0, 2)$ while figures c) and d) show the evolution for a range of σ in $(0, 1)$. The failure probability is equal to 6.21×10^{-3}, the derivative with respect to a is equal to 2.19×10^{-2} and the derivative with respect to b is equal to 8.76×10^{-3}.

The smaller the σ, the lower n is needed to reach an accurate equivalence between P_n and $\partial\overline{P_f(\mathbf{s},\cdot)}/\partial s_\ell$, as illustrated with the toy example. Furthermore, in order to evaluate the coefficients of the polynomial P_n, one must have at least $n+1$ evaluations of the function $\sigma \mapsto P_n(\sigma)$. The idea of the proposed approach is thus to find the coefficient of P_n of order zero, which is the probability sensitivity, by performing a polynomial regression with m evaluations of $\left(\partial\overline{P_f(\mathbf{s},\sigma_l)}/\partial s_\ell\right)_{l=1,\dots,m}$ with $m \geq n+1$, since the two functions are equivalent for small σ. Recalling that

$$\forall \sigma \in \mathbb{R}^+\backslash\{0\} \qquad \frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell} = \int_{\mathbb{R}^d} -\frac{\partial g(\mathbf{s},\mathbf{x})}{\partial s_\ell}\xi_\sigma(-g(\mathbf{s},\mathbf{x}))f_{\mathbf{x}}(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{6.11}$$

is a domain integral, it can be estimated with Monte Carlo methods as in the Weak approach framework presented in Chapter 4 Section 4.3.3. We briefly review here the properties of this estimate.

6.2.3.1 Regression with crude Monte Carlo method

The crude Monte Carlo estimate of Eq. (6.11) is given by the following formula

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \widehat{\frac{\partial P_f(\mathbf{s},\sigma)}{\partial s_\ell}}^{\mathrm{MC}} = -\frac{1}{N}\sum_{j=1}^{N} \frac{\partial g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)}{\partial s_\ell} \xi_\sigma\left(-g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)\right), \tag{6.12}$$

where the observations $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$ are iid from $f_{\mathbf{X}}$. These observations are reused from the estimation of the failure probability with an MC estimate. Therefore, the only additional simulation budget concerns the evaluation of the limit state function derivative with respect to s_ℓ for each observation. The MC estimate is unbiased and its variance is given by

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial P_f(\mathbf{s},\sigma)}{\partial s_\ell}}^{\mathrm{MC}}\right) = \frac{1}{N}\mathbb{V}\mathrm{ar}\left(\frac{\partial g\left(\mathbf{s}, \mathbf{X}\right)}{\partial s_\ell} \xi_\sigma\left(-g\left(\mathbf{s}, \mathbf{X}\right)\right)\right). \tag{6.13}$$

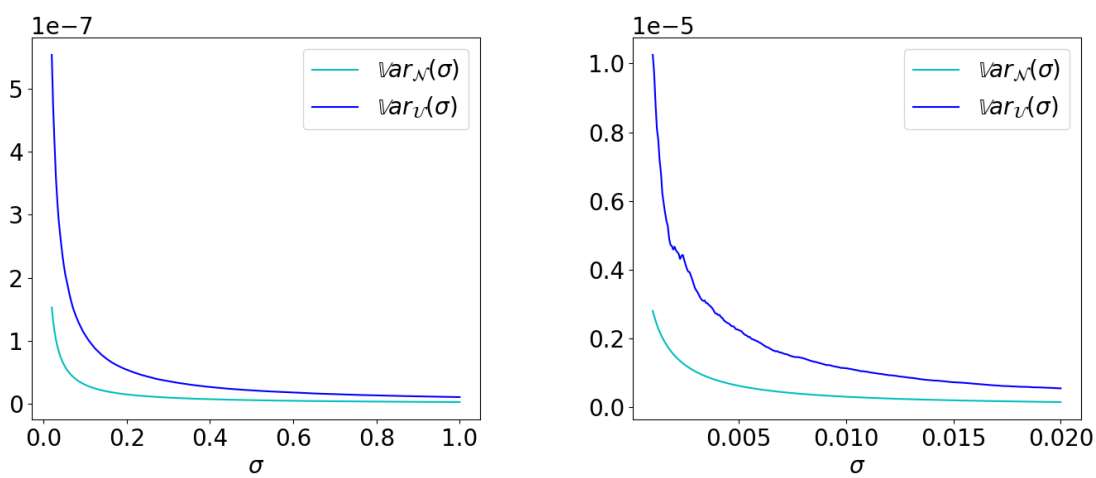
Depending on the pdf ξ_σ, the expression of this variance can be further detailed. For instance, if ξ_σ is the pdf of a centered normal variable of variance σ^2, then $\forall y \in \mathbb{R}\ \ \xi_\sigma(y) = (1/\sigma)\phi(y/\sigma)$ and the variance becomes

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial P_f(\mathbf{s},\sigma)}{\partial s_\ell}}^{\mathrm{MC}}\right) = \frac{1}{N\sigma^2}\mathbb{V}\mathrm{ar}\left(\frac{\partial g\left(\mathbf{s}, \mathbf{X}\right)}{\partial s_\ell} \phi\left(\frac{-g\left(\mathbf{s}, \mathbf{X}\right)}{\sigma}\right)\right).$$

If ξ_σ is the cdf of a uniform variable defined on $[-\sigma/2, \sigma/2]$, then $\forall y \in [-\sigma/2, \sigma/2]\ \ \xi_\sigma(y) = 1/\sigma$ and 0 elsewhere and the variance becomes

$$\forall \sigma \in \mathbb{R}^+ \backslash \{0\} \qquad \mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial P_f(\mathbf{s},\sigma)}{\partial s_\ell}}^{\mathrm{MC}}\right) = \frac{1}{N\sigma^2}\mathbb{V}\mathrm{ar}\left(\frac{\partial g\left(\mathbf{s}, \mathbf{X}\right)}{\partial s_\ell} \mathbb{I}_{\left[-\frac{\sigma}{2}, \frac{\sigma}{2}\right]}\left(-g\left(\mathbf{s}, \mathbf{X}\right)\right)\right).$$

For a fixed N, these variance functions are derived in Figure 6.2. From Figure 6.2, it is obvious that the



a) b)

Figure 6.2: Illustration of the variance functions for both of the mentioned indicator approximation functions Ξ_σ with the toy example ($a = 2, b = 5$) with respect to $s_\ell = a$ and with $N = 10^6$. The normal approximation is plotted in cyan while the uniform approximation is plotted in navy blue. Figure a) shows the evolution of the different functions for a range of σ in $(0.02, 1)$ while Figure b) shows the evolution for a range of σ in $(0, 0.02)$. The derivative with respect to b gives roughly the same variance function evolution.

variance Eq. (6.13) of the MC estimate is a decreasing function of σ, for both approximation functions. This conclusion is valid for all approximation functions of the indicator function. Indeed, as they verify the convergence property Eq. (6.1), $\Xi_\sigma(y)$ tends to the indicator function as $\sigma \to 0$. Consequently the domain integral Eq. (6.11) converges to a surface integral over the limit state surface $g(\mathbf{s}, \mathbf{x}) = 0$ as $\sigma \to 0$. Most of the observations of Eq. (6.12) are then located away from this surface and thus have a negligible weight in the mean sample value; cf Chapter 4 Section 4.3.3.2.

Therefore, for a fixed N, each estimate $\widehat{\partial \overline{P_f(\mathbf{s}, \sigma_l)}}/\partial s_\ell$, for $l = 1, \ldots, m$, has a different noise, and the smaller the σ_l the bigger the noise. In order to identify the coefficients of the polynomial P_n, a heteroscedastic polynomial regression must then be performed. This heteroscedastic polynomial regression is further detailed in the next section. The theoretical bias of the proposed approach is first briefly addressed here, as it is an important aspect of the proposed method.

6.2.3.2 Theoretical bias of the sensitivity estimate

Each MC estimate of Eq. (6.2) is unbiased. Nevertheless, as the polynomial P_n is of degree n and n is finite, the proposed approach is theoretically biased as there is no equality between P_n and $\partial \overline{P_f(\mathbf{s}, \cdot)}/\partial s_\ell$, no matter the sampling method. Indeed, for each σ_l, it is recalled that

$$\frac{\partial \overline{P_f(\mathbf{s}, \sigma_l)}}{\partial s_\ell} = P_n(\sigma_l) + \mathbb{E}_{\xi_{\sigma_l}}\left[R_n(Z)\right] = P_n(\sigma_l) + \mathbb{E}_{\xi_{\sigma_l}}\left[\int_0^Z \frac{(Z-t)^n}{n!} T^{(n+1)}(t)\mathrm{d}t\right],$$

where $T(z) = \partial F_{H_\mathbf{s}}(-z)/\partial s_\ell$ for all $z \in \mathbb{R}$. The higher the n, the smaller the theoretical bias, for a fixed σ_l. Furthermore, for a fixed n, since several values of σ_l are considered, the remainder $\mathbb{E}_{\xi_{\sigma_l}}\left[R_n(Z)\right]$ is then more negligible for the smaller values of σ_l than for the higher values. This evolving theoretical bias has to be taken into account when performing the heteroscedastic polynomial regression. The discussion on the bias of the proposed method continues on the next sections.

6.3 Heteroscedastic polynomial regression

The main parameters of the heteroscedastic polynomial regression are the selected degree of the polynomial n, the number m of estimates $\left(\widehat{\partial \overline{P_f(\mathbf{s}, \sigma_l)}}/\partial s_\ell\right)_{l=1,\ldots,m}$ with $m \geq n+1$ and the interval of regression $[\sigma_{\min}, \sigma_{\max}]$, given that n and the interval are dependent, as underlined in the toy example Section 6.2.2.3. However, the correlation between the different estimates $\left(\widehat{\partial \overline{P_f(\mathbf{s}, \sigma_l)}}/\partial s_\ell\right)_{l=1,\ldots,m}$ is first discussed, as it directly influences the regression framework.

In the rest of this chapter, the following notations will be used. We denote $\mathbf{V}$ the m-length vector of the estimates: $\mathbf{V} = \left(\widehat{\partial \overline{P_f(\mathbf{s}, \sigma_l)}}/\partial s_\ell\right)_{l=1,\ldots,m}$, $\mathbf{S}$ the Vandermonde matrix of σ of size $m \times (n+1)$ such as $\mathbf{S}_{l,i} = \sigma_l^i$, for $l = 1 \ldots, m$ and $i = 0, \ldots, n$ and $\boldsymbol{\alpha}$ the polynomial coefficients such as the regression is written

$$\mathbf{V} = \mathbf{S}\boldsymbol{\alpha} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\epsilon}$ is the m-length vector of random errors ϵ_l of expected value $\mathbb{E}_{f_\mathbf{X}}\left[\epsilon_l\right] = 0$, as the estimates V_l are unbiased. The errors variance $\mathbb{V}\mathrm{ar}\left(\epsilon_l\right)$ depends on σ_l cf Eq. (6.13). The σ-vector is denoted in the ascending order: $\sigma_1 > \sigma_2 > \ldots > \sigma_m$. This linear regression is addressed in our manuscript with linear least square methods [Watson, 1967, Lai et al., 1979]. However, it should be noted that there exist other methods that aim to identify the coefficients of a polynomial; for instance the polynomial interpolation framework [Werner, 1984]. As this identification is not the main subject of this thesis, those methods are not derived here.

6.3.1 Linear least squares method in our specific context

As previously mentioned, the observations $\left(\mathbf{X}^{(j)}\right)_{j=1,\ldots,N}$ needed for each estimate V_l are reused from the failure probability estimation procedure, in order to minimize the additional simulation budget. However, reusing the exact same sample for each V_l results in a highly correlated database

$(\sigma_l, V_l)_{l=1,\dots,m}$. The regression framework required is then the Generalized Least Squares (GLS) framework that we describe here in short.

6.3.1.1 Generalized Least Squares regression framework

The Generalized Least Squares (GLS) framework is the generalization of the Ordinary Least Squares (OLS) framework, to the cases where the variance of the random error ϵ_l is not constant (heteroscedasticity) and the estimates vector $\mathbf{V}$ is not uncorrelated. We denote $\Sigma_{\boldsymbol{\epsilon}}$ the covariance matrix of the random errors $\boldsymbol{\epsilon}$. The GLS estimate of the coefficient $\boldsymbol{\alpha}$ is then obtained with the following equations

$$\widehat{\boldsymbol{\alpha}} = \left(\mathbf{S}^\top \Sigma_{\boldsymbol{\epsilon}}^{-1} \mathbf{S}\right)^{-1} \left(\mathbf{S}^\top \Sigma_{\boldsymbol{\epsilon}}^{-1} \mathbf{V}\right), \tag{6.14}$$

and it is the best linear unbiased estimate [Taboga, 2021]. It is possible to derive the covariance matrix estimate of $\widehat{\boldsymbol{\alpha}}$ with the following formula

$$\widehat{\mathbb{V}\mathrm{ar}}\left(\widehat{\boldsymbol{\alpha}}\right) = \left(\mathbf{S}^\top \Sigma_{\boldsymbol{\epsilon}}^{-1} \mathbf{S}\right)^{-1}. \tag{6.15}$$

In our specific context, the covariance matrix $\Sigma_{\boldsymbol{\epsilon}}$ is not analytically known and has to be estimated with the sample $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$. Therefore, in the formulas Eqs. (6.14) and (6.15), replacing $\Sigma_{\boldsymbol{\epsilon}}$ with $\widehat{\Sigma_{\boldsymbol{\epsilon}}}$ gives the final framework suited for the heteroscedastic polynomial regression needed in our approach, called the Feasible GLS (FGLS).

From formulas Eqs. (6.14) and (6.15), one can notice that the inverse of the covariance matrix is required to compute both quantities. In our specific case, this covariance matrix is an estimation of the real covariance matrix. Therefore, each coefficient already comes with an estimation error. All those errors combined make it very difficult to accurately compute $\widehat{\Sigma_{\boldsymbol{\epsilon}}}^{-1}$. Consequently, although the FGLS framework is theoretically the best suited for the proposed method, it is practically inapplicable. For this reason, a regression framework simplification must be considered, which is detailed next.

6.3.1.2 Pre-processing of the covariance matrix

The matrix $\widehat{\Sigma_{\boldsymbol{\epsilon}}}$ is particularly hard to inverse as its non-diagonal coefficients are non-zero. Therefore, the simplification proposed here is to decrease the correlation of the vector of estimates $\mathbf{V}$. One solution could be to generate m samples of size N independently and then estimate each V_l with a different sample. The resulting vector $\mathbf{V}$ would be completely independent. However, as the goal is to increase as little as possible the simulation budget, this solution is not applicable in our context, especially if the number m is high.

The solution selected to decrease the correlation of the vector $\mathbf{V}$ is to employ bootstrap [Horowitz, 2001]. The logic of bootstrapping is to learn an empirical discrete cdf from a vector of iid observations in order to generate a new sample. The new sample shares the same property as the original sample but they are independent. Here, we apply bootstrap to the iid observations $\left(\mathbf{X}^{(j)}\right)_{j=1,\dots,N}$ already available from the failure probability estimation. The first estimate V_1 is computed with the original sample and for each σ_l with $l > 1$, bootstrap is employed. The resulting vector $\mathbf{V}$ is thus independent. Bootstrap does not require any additional call to the limit state function or its derivatives.

The resulting covariance matrix $\widehat{\Sigma_{\boldsymbol{\epsilon}}}$ is thus diagonal, and numerically invertible. The regression framework is then referred to as Weighted Least Squares (WSL). Denoting $\mathbf{W}$ the diagonal matrix such as $W_{l,l} = 1/\widehat{\mathbb{V}\mathrm{ar}}\left(V_l\right)$, where $\widehat{\mathbb{V}\mathrm{ar}}\left(V_l\right)$ is given by Eq. (6.13), the WSL estimate is then written

$$\widehat{\boldsymbol{\alpha}}_W = \left(\mathbf{S}^\top \mathbf{W} \mathbf{S}\right)^{-1} \left(\mathbf{S}^\top \mathbf{W} \mathbf{V}\right), \tag{6.16}$$

and it is the best linear unbiased estimate [Taboga, 2021]. It is possible to derive the covariance matrix estimate of $\widehat{\boldsymbol{\alpha}}_W$ with the following formula

$$\widehat{\mathbb{V}\mathrm{ar}}\left(\widehat{\boldsymbol{\alpha}}_W\right) = \left(\mathbf{S}^\top \mathbf{W} \mathbf{S}\right)^{-1}. \tag{6.17}$$

Therefore, with only one simulation run, an estimation of the variance of the sensitivity estimate is available. The sensitivity estimate is equal to the coefficient of order zero of the polynomial, which is the first component of the vector $\widehat{\boldsymbol{\alpha}}_W$ and its theoretical variance estimate is the component $\widehat{\mathbb{V}\mathrm{ar}}\left(\widehat{\boldsymbol{\alpha}}_W\right)_{1,1}$.

6.3.1.3 Remark on the bias of the estimate

Even if the estimate given by Eq. (6.16) is unbiased, it is reminded here that because of the remainder of the Taylor series expansion $\mathbb{E}_{\xi_\sigma}\left[R_n(Z)\right]$, there is still a bias in the proposed approach, as previously mentioned. This bias decreases as the values of $(\sigma_l)_{l=1,\dots,m}$ decrease and with an increase of the degree of the polynomial P_n, as illustrated in the toy example Section 6.2.2.3. Consequently, this bias can be controlled, as it highly depends on the setting of the regression parameters, which are the regression interval $[\sigma_{\min}, \sigma_{\max}]$ and the polynomial degree n of P_n. The setting of the regression parameters is further detailed in the next section as well as the control of the bias of the sensitivity estimate.

6.3.2 Settings of the regression parameters

The settings presented here result of several tests. However, as previously mentioned, polynomial regression is not the main subject of this chapter, nor this thesis. Consequently, the parameter are set in the simplest manner possible. Before presenting the settings, the scaling of the limit state function is first addressed, as it greatly influences the evolution of the failure probability derivative functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$.

6.3.2.1 Scaling of the limit state function

Depending on the limit state function, the order of magnitude of $\mathbf{x} \mapsto g(\mathbf{s}, \mathbf{x})$ and its derivatives with respect to $\mathbf{s}$ can significantly vary, especially in the vicinity of the failure surface, as underlined in [Lacaze et al., 2015] and [Papaioannou et al., 2018]. As a result, the behavior of the failure probability derivative functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$ can considerably vary as well. Here, general regression settings are presented, which aim to be applied to various limit state functions. Consequently, the following scaling is first performed.

Let $\left(Y^{(j)} = g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)\right)_{j=1,\dots,N}$ be the lsf values vector obtained from the estimation of the failure probability. The scaling proposed here is to divide the vector $\mathbf{Y}$ as well as the vector of derivatives $\left(\partial g\left(\mathbf{s}, \mathbf{X}^{(j)}\right)/\partial s_\ell\right)_{j=1,\dots,N}$ by the standard deviation δ obtained from the negative lsf values only of $\mathbf{Y}$

$$\delta^2 = \sum_{j=1}^{N} \bar{w}^{(j)} \left(Y^{(j)} - \mu\right)^2 \quad \text{with} \quad \mu = \sum_{j=1}^{N} \bar{w}^{(j)} Y^{(j)}, \tag{6.18}$$

where $\bar{w}^{(j)} = w^{(j)} / \sum_{i=1}^{N} w^{(i)}$ and $w^{(j)} = \mathbb{I}_{Y<0}\left(Y^{(j)}\right)$. Therefore, δ represents the order of magnitude of the lsf in the failure domain.

This division does not affect the value of the failure probability nor the failure probability sensitivity. Indeed provided $\delta > 0$, we have $P_f(\mathbf{s}) = \mathbb{P}\left(g(\mathbf{s}, \mathbf{X}) < 0\right) = \mathbb{P}\left(\frac{g(\mathbf{s}, \mathbf{X})}{\delta} < 0\right)$ and the derivatives do not change either. However, this division influences the value of the coefficients of order superior to zero in the Taylor series expansion of T. Since these coefficients are not of interest in the proposed approach, the consequences are negligible. Thanks to this scaling, the order of magnitude of the lsf is expected to have little influence on the polynomial regression settings presented in the next paragraphs.

6.3.2.2 Choice of the regression interval and the polynomial degree

Definition of the regression interval $[\sigma_{\min}, \sigma_{\max}]$
After several tests, it appeared that the selection of the interval is crucial in the regression process, as it greatly influences the quality of the sensitivity estimate. From the graphs shown in Figure 6.2, it is obvious that if the interval is too close to zero, then the estimates $(V_l)_{l=1,\dots,m}$ will be very noisy and their variance might be inaccurately estimated, as N is fixed. Consequently the WLS framework might lead to erroneous results. However, if the regression interval is too far from zero, a higher degree polynomial is needed to correctly approximate the Taylor series expansion as shown in Figure 6.1. The polynomial regression is then harder to achieve and leads to a probability sensitivity of higher variance. Therefore there is a trade-off between the accuracy of the estimates $(V_l)_{l=1,\dots,m}$ and $\left(\widehat{\mathbb{V}\mathrm{ar}}\left(V_l\right)\right)_{l=1,\dots,m}$ and the variance of the sensitivity estimate obtained with the polynomial regression. In the proposed approach, it was decided to define the interval bounds inside the interval $\sigma \in [0.01, 1]$. Thanks to the

scaling process mentioned above, restricting the search for the regression bounds in this interval has proven to be efficient for the various limit state functions tested.

Within this interval, the regression bounds are set depending on the theoretical CV estimate of the MC estimate V_1 of Eq. (6.11), whose theoretical variance is given by Eq. (6.13). This CV is thus a function of σ. Choosing a CV criterion to set the interval bounds is inspired by [Papaioannou et al., 2018] where the optimal $\tilde{\sigma}$ of the Weak approach framework is selected with a target CV technique; see Chapter 4 Section 4.3.3.3. Here, the values of $\sigma_{\min}$ and $\sigma_{\max}$ are thus respectively the lowest and the highest value of σ such as $\mathrm{CV}(\sigma) < \mathrm{CV}_{\text{target}}$ for all $\sigma \in [\sigma_{\min}, \sigma_{\max}]$. This CV criterion ensures the regression is performed with estimates $(V_l)_{l=1,\dots,m}$ and $\left(\widehat{\mathbb{V}\mathrm{ar}}\,(V_l)\right)_{l=1,\dots,m}$ that have reasonable noise.

For the choice of $\mathrm{CV}_{\text{target}}$, selecting an arbitrary value, e.g. $\mathrm{CV}_{\text{target}} = 10\%$ could lead to very narrow intervals or in the contrary the whole restricted interval $[0.01, 1]$. It depends on N as well as the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$. While performing the polynomial regression in the whole restricted interval $[0.01, 1]$ does not affect the quality of the regression (thanks to the scaling previously mentioned), if the interval is very small then the quality of the regression cannot be guaranteed. For this reason, the $\mathrm{CV}_{\text{target}}$ is set dependently on the minimum theoretical CV estimate of the V_1 estimate with

$$\mathrm{CV}_{\text{target}} = \min_{\sigma \in [0.01,1]} \mathrm{CV}_{V_1}(\sigma) + 5\%.$$

The additional 5% ensures a sufficiently large interval and is an arbitrary value. It has proven to be efficient after several tests. This threshold is independent of the quality of the estimation of P_f, as it has been underlined that P_f and its derivatives do not depend on the same quantities; cf Chapter 4, section 4.3.3.2. Furthermore, this threshold changes for each deterministic inputs s_ℓ and allows the regression interval to be specifically suited for the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$.

Once this bounds are computed, the values of the vector $(\sigma_l)_{l=1,\dots,m}$ are set uniformly in the interval $[\sigma_{\min}, \sigma_{\max}]$.

Selection of the polynomial degree n and number of estimations m

The choice of the polynomial degree n influences the global bias of the proposed approach. The higher the degree n is, the lower the bias, for a fixed regression interval. However, as previously mentioned, a high degree n induces a more intricate polynomial regression, as the polynomial P_n is then more complex, resulting in a sensitivity estimate with higher variance. Therefore, there is a trade-off between the value of the theoretical bias of the sensitivity estimate and its variance.

The choice of n affects the number m of estimations V_l required for the regression, as there are $n+1$ coefficients that have to be estimated with the regression. For simplification purposes, the number m of estimations of the failure probability derivatives $V_l = \left(\widehat{\partial \overline{P_f(\mathbf{s}, \sigma_l)}/\partial s_\ell}\right)_{l=1,\dots,m}$ is set to $m = n + 2$, as several tests showed that an increase in this number did not improve the quality of the sensitivity estimates.

It is reminded here that these m estimations are obtained with bootstrap, therefore they do not increase the simulation budget. Consequently, several tests can be performed to assess the minimal degree n needed for a correct sensitivity estimate, without affecting the simulation budget. If, with a higher degree polynomial, the resulting sensitivity estimate value no longer changes, we can assume the estimate has reached the correct value and the bias is controlled since it has become negligible. This control of the bias depends naturally on the level of accuracy wanted for the estimation of the sensitivity.

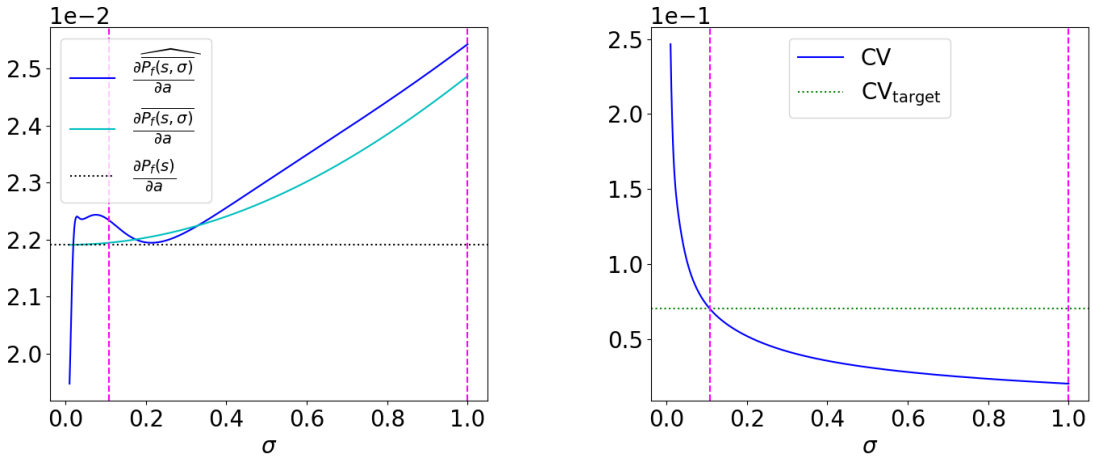
6.3.2.3 Illustration with the toy example

It is recalled here that the selected indicator approximation function for the toy example is the cdf of a centered normal variable of variance σ^2, therefore the resulting polynomial P_n is even: $n = 2k$, consequently $m = k + 2$ as $k + 1$ coefficients must be estimated. The toy example is tested for various values of sample sizes N. As the scaling mentioned in Section 6.3.2.1 has been performed, the graphs are not exactly similar to the ones shown in Figure 6.1.

Illustration with a large Monte Carlo sample

The settings of the different parameters are illustrated with the toy example, with $N = 100000$. The

limit state function is evaluated N times and its derivatives with respect to s_ℓ are evaluated also N times. The failure probability is equal to 6.21×10^{-3} with $a = 2$ and $b = 5$. The graphs of the theoretical CV estimate of the MC estimate are shown in Figure 6.3 with the value of $\mathrm{CV}_{\mathrm{target}}$ and of the regression interval $[\sigma_{\min}, \sigma_{\max}]$, for the derivative with respect to a.



a) b)

Figure 6.3: Illustration of the regression bounds plotted in dashed magenta lines corresponding to the values of σ such as the theoretical CV estimate of the MC estimate is below $\mathrm{CV}_{\mathrm{target}}$. Here the CV target is around 7.5% for the derivative with respect to a. Figure a) represents the evolution of the MC estimate according to σ in navy blue as well as the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ in clear blue. Figure b) shows the evolution of the theoretical CV estimate of the MC estimate and the value of $\mathrm{CV}_{\mathrm{target}}$.

From Figure 6.3 b), it appears that the CV criterion only shortly reduces the regression interval from $[0.01, 1]$, as the variance of the MC estimate is low, since the simulation budget is high. The MC estimate is quite close to the function $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ for large values of σ in Figure 6.3 a). However as σ tends to zero the functions diverge, which is expected because of the increased variance. Furthermore, we can observe for larger σ that the behavior of both functions is similar.

An illustration of the polynomial regression for $n = 2$ is shown in Figure 6.4 for a sample of size $N = 100000$. From these curves, it appears that the polynomial regression has succeeded in smoothing the noise of the MC estimates for small values of σ, resulting in a sensitivity estimate very close to the real value. The evolutions of P_2 and $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ are very close in the entire regression interval. Consequently, with this setting, it seems the MC estimates $(V_l)_{l=1,2,3}$ allow to find a polynomial of degree 2 quite similar to $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$. It is reminded here that the MC estimates V_1 is computed directly from the original sample, with $\sigma_1 = \sigma_{\min}$, while V_l is computed from a sample obtained with bootstrap for $l > 1$.

The estimates of the proposed method over 500 independent simulations, for three polynomial degrees $2k = 2$, $2k = 4$ and $2k = 6$ are given in Table 6.1. The last column is the mean value and CV obtained with the estimate V_1, for which $\sigma_1 = \sigma_{\min}$. Therefore, this last column displays the value of the sensitivity estimate obtained with the Weak approach, by selecting $\tilde{\sigma} = \sigma_{\min}$; cf Chapter 4 Section 4.3.3. The given CV is the empirical CV computed over the 500 independent simulations while the number in parentheses is the mean value of the theoretical CV estimate obtained with the theoretical variance estimate of the sensitivity, available in the WSL framework with Eq. (6.17).

From Table 6.1, the sensitivity estimates of both the proposed approach and the Weak approach are very close to the real value of the failure probability sensitivities. For the three degrees selected and the two derivatives, the value of the estimate is quite constant. Therefore, it seems that with this particular setting, a polynomial of degree 2 is sufficient to correctly approximate the coefficient of order zero of the Taylor series expansion of the functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$. The sensitivity estimate obtained with $2k = 2$ has the smallest variance, as expected. Furthermore, the estimates resulting from the polynomial of degrees $2k = 2$ and $2k = 4$ both have a lower CV than the one of the Weak

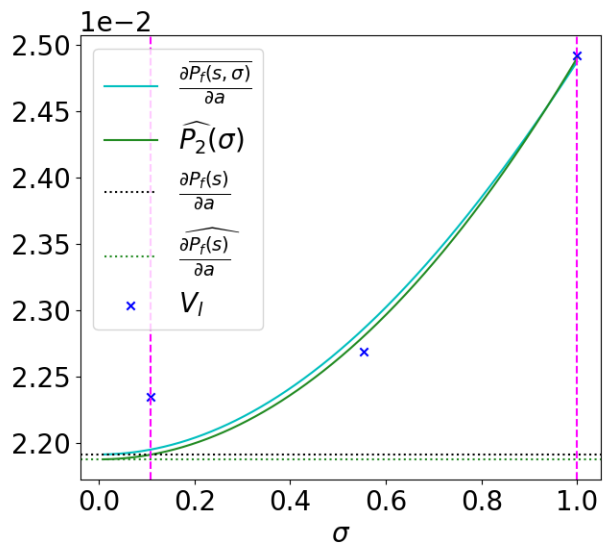


Figure 6.4: Illustration of the polynomial regression obtained from the estimates $(V_l)_{l=1,2,3}$ displayed with navy blue crosses. The resulting polynomial $\widehat{P_2}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line. The clear blue line is the function $\sigma \mapsto \partial \overline{P_f(\mathbf{s},\sigma)}/\partial a$. The regression bounds are plotted in dashed magenta lines.

Table 6.1: Comparison of the result of the polynomial regression with 3 different degrees, for $N = 100000$. The failure probability is equal to 6.22×10^{-3} with an empirical CV of 3.8%. The reference values of the sensitivities are $\partial P_f/\partial a = 2.19 \times 10^{-2}$ and $\partial P_f/\partial b = -8.76 \times 10^{-3}$.

Toy example								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial a}}$	2.19×10^{-2}		2.19×10^{-2}		2.19×10^{-2}		2.20×10^{-2}	
	CV	5.4% (3.9%)	CV	7.0% (5.1%)	CV	7.5% (6.2%)	CV	7.2%
$\widehat{\frac{\partial P_f}{\partial b}}$	-8.75×10^{-3}		-8.78×10^{-3}		-8.76×10^{-3}		-8.82×10^{-3}	
	CV	5.4% (3.9%)	CV	7.0% (5.1%)	CV	7.5% (6.2%)	CV	7.1%

approach, which validates the polynomial regression process: for no additional simulation budget, the proposed polynomial regressions lead to a more precise sensitivity estimate. For the derivative with respect to b, the value of the sensitivity estimates obtained with the proposed method is even slightly less biased. However, when $2k = 6$, the CV is then moderately larger than those of the Weak approach estimate.

The theoretical CV estimate of the proposed method is a little underrated, however it gives an accurate idea of the real CV of the sensitivity estimate. These CV estimates have small variations, as their empirical CV, which are not displayed here, are all below 9%. Therefore the accuracy of the proposed method can be correctly determined with only one simulation run.

Illustration with smaller Monte Carlo samples

The robustness of the proposed method is tested for smaller values of sample size: $N = 50000$ and $N = 10000$. For these sample sizes, noisier MC estimates are expected. Figure 6.5 illustrates the regression interval and the resulting polynomial regression for $n = 2$. Compared to the graphs in Figure 6.3 a) of the crude MC estimates, the Figure 6.5 a) and d) diverge more of $\sigma \mapsto \partial \overline{P_f(\mathbf{s},\sigma)}/\partial a$ in the entire interval, and especially for small values of σ, as expected. Once again, the resulting

polynomials allow to smooth the noise of the MC estimates, however, they are not as close as before to the function $\sigma \mapsto \partial\overline{P_f(\mathbf{s},\sigma)}/\partial a$. This can be explained by the quality of the polynomial regression procedure that has slightly decreased since the variance estimates are less accurate. It can also simply be a result of the poor quality of each V_l, particularly V_1.

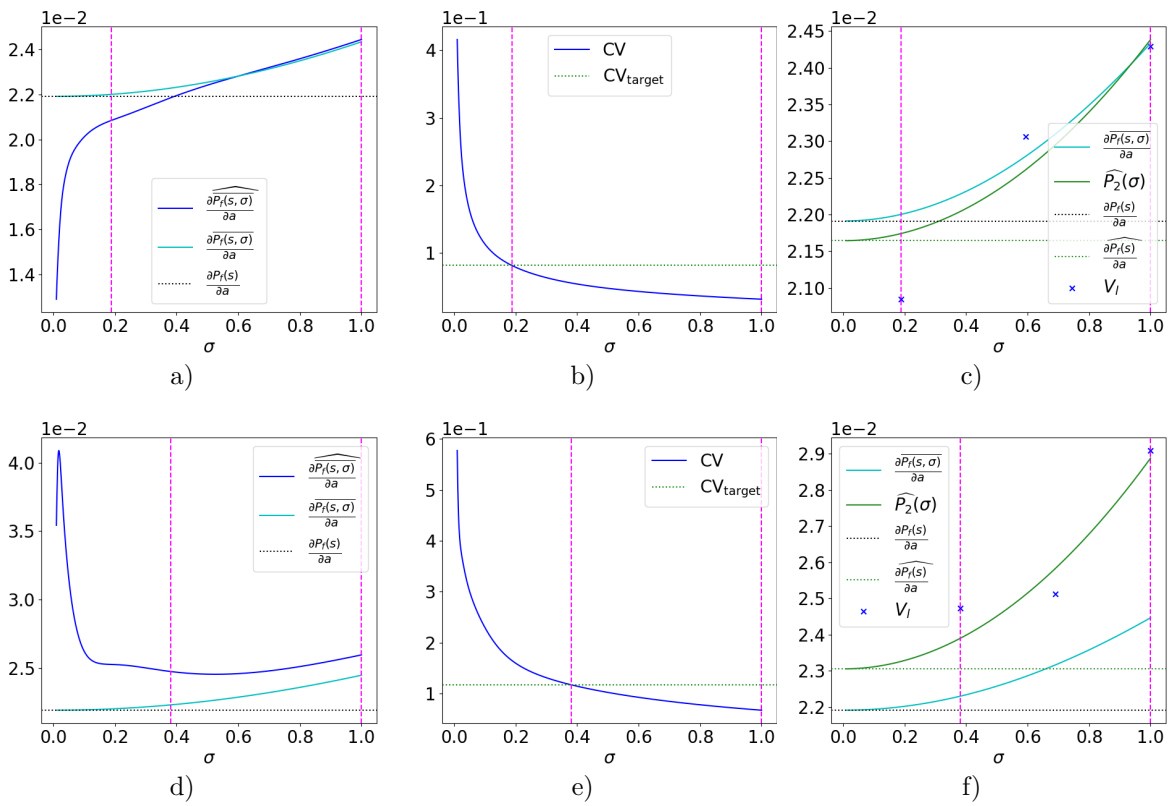


Figure 6.5: Illustration of the proposed method with the toy example for two values of sample size N. Figures a), b) and c) illustrate the method for $N = 50000$, while Figures d), e) and f) illustrate the method for $N = 10000$. Figures a) and d) represent the evolution of the MC estimate according to σ in navy blue as well as the evolution of $\sigma \mapsto \partial\overline{P_f(\mathbf{s},\sigma)}/\partial a$ in clear blue. Figures b) and e) show the evolution of the theoretical CV estimate of the MC estimate and the value of $\mathrm{CV}_{\mathrm{target}}$, close to 9% for $N = 50000$ and close to 12% for $N = 10000$. Figures c) and f) represent the estimates $(V_l)_{l=1,2,3}$ displayed with navy blue crosses obtained with bootstrap. The resulting polynomial $\widehat{P_2}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line.

Compared to the large MC sample, the size of the regression interval has reduced, as the lower bound $\sigma_{\min}$ gets larger with smaller N. Only the less noisy part of the interval is kept for the polynomial regression.

The results are presented in Table 6.2, where only the sensitivity estimates obtained from the polynomial degree $2k = 2$ are given. For both sample size values, the estimates obtained with polynomials of higher degree are not displayed; they have equal mean value but higher CV, for both parameters a and b, as in Table 6.1. Consequently, as in the previous study, a polynomial of degree 2 is still sufficient to correctly approximate the coefficient of order zero of the Taylor series expansion of the functions $\sigma \mapsto \partial\overline{P_f(\mathbf{s},\sigma)}/\partial s_\ell$ for $\ell = 1, 2$.

When $N = 50000$, the proposed method with $2k = 2$ leads to a sensitivity estimate more precise than the Weak approach and less biased for the derivative with respect to b. However, when $N = 10000$, the proposed approach leads to an estimate of higher CV. Nevertheless, the resulting estimates are then empirically less biased than the estimates obtained with Weak approach, for both derivatives. The derivative with respect to b estimated with the Weak approach is particularly biased. This comes from the setting of the value of $\sigma_{\min}$ which leads to a $\sigma_{\min}$ very large, thus $\tilde{\sigma} = \sigma_{\min}$ is very large as well and the Weak approach estimates are no longer accurate.

Table 6.2: Comparison of the result of the polynomial regression, for $N = 50000$ and $N = 10000$. The failure probability is equal to 6.19×10^{-3} with an empirical CV of 5.7% for $N = 50000$. The failure probability is equal to 6.23×10^{-3} with an empirical CV of 13.3% for $N = 10000$. The reference values of the sensitivities are $\partial P_f/\partial a = 2.19 \times 10^{-2}$ and $\partial P_f/\partial b = -8.76 \times 10^{-3}$.

Toy example								
	$N = 50000$				$N = 10000$			
	Regression $2k = 2$		Weak approach $\tilde{\sigma} = \sigma_{\min}$		Regression $2k = 2$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial a}}$	2.19×10^{-2}		2.21×10^{-2}		2.19×10^{-2}		2.24×10^{-2}	
	CV	7.1% (5.2%)	CV	7.8%	CV	14.1% (11.7%)	CV	11.3%
$\widehat{\frac{\partial P_f}{\partial b}}$	-8.74×10^{-3}		-8.85×10^{-3}		-8.73×10^{-3}		-9.12×10^{-3}	
	CV	7.2% (5.3%)	CV	7.7%	CV	14.1% (11.9%)	CV	10.8%

Consequently, the polynomial regression procedure is still relevant, even for smaller sample sizes. The resulting sensitivity estimates are empirically less biased than if the polynomial regressions were not performed. Furthermore, the bias can be controlled: as the value of the estimate does not change with higher polynomial degrees, then this value can be considered accurate. However, it appears that the proposed method requires a certain simulation budget to provide estimates with small CVs. As in the previous study, the theoretical CV estimate of the proposed approach gives a slightly underrated value. The empirical CVs of these theoretical CV estimates, not displayed in Table 6.2, are below 11% for $N = 50000$ but close to 18% for $N = 10000$ and $2k = 2$. This is discussed below.

Discussion about the theoretical CV

The theoretical CV estimate available from the WSL framework seems to undervalue the actual CV, independently of the sample size N. One reason for this could be the inverse covariance matrix W, which can be accurate enough to obtain a correct sensitivity estimate, but not precise enough for the estimation of the variance, even with a large simulation budget as $N = 100000$. As the evolution of the variance is only slightly decreasing in the regression interval, it is interesting to compare the results of Table 6.1 and Table 6.2, with the sensitivity estimates obtained from an ordinary least square (OLS) approach. The OLS framework is recalled in the Appendix E.1. In this framework, the variance is supposed to be equal to a constant η^2. Here this constant is taken as the maximum of the variance estimates of the $(V_l)_{l=1,\dots,m}$, thus $\eta^2 = \max_{l=1,\dots,m} \left(\widehat{\mathbb{V}\mathrm{ar}}(V_l)\right)$.

The sensitivity estimates with the OLS framework, using the same simulation samples as those of Table 6.1 and Table 6.2 are given in the Appendix E.2 in Table E.1 and Table E.2. Performing the polynomial regression in the OLS framework seems to lead to equivalent results for this toy example, as the sensitivity estimates are also empirically unbiased and with an equivalent CV. However, the theoretical CV estimates are then closer to the actual values, with slightly higher empirical CV. Consequently, although the OLS framework is less theoretically correct, as the variance of the MC estimates are not constant, it leads to a less biased variance estimation of the failure probability sensitivity for this particular example.

6.3.3 Proposed approach with importance sampling methods

The crude Monte Carlo method is practical to study the framework of the proposed method. However, this simulation method is inapplicable in the context of our manuscript, as the simulation budget required is too large and we assume the lsf to be numerically expensive. Therefore, the proposed method is adapted to the IS framework, with the following equations. The failure probability derivative

function $\sigma \mapsto \partial\overline{P_f(\mathbf{s},\sigma)}/\partial s_\ell$ is computed with the following estimate

$$\forall \sigma \in \mathbb{R}^+\backslash\{0\} \qquad \widehat{\frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell}}^{\mathrm{IS}} = -\frac{1}{N}\sum_{j=1}^{N}\frac{\partial g\left(\mathbf{s},\mathbf{X}^{(j)}\right)}{\partial s_\ell}\xi_\sigma\left(-g\left(\mathbf{s},\mathbf{X}^{(j)}\right)\right)\frac{f_{\mathbf{X}}\left(\mathbf{X}^{(j)}\right)}{h\left(\mathbf{X}^{(j)}\right)}, \tag{6.19}$$

where the observations $\mathbf{X}^{(j)}$ are iid from h, the IS auxiliary density. The IS estimate is unbiased, and its variance is written

$$\forall \sigma \in \mathbb{R}^+\backslash\{0\} \qquad \mathbb{V}\mathrm{ar}\left(\widehat{\frac{\partial\overline{P_f(\mathbf{s},\sigma)}}{\partial s_\ell}}^{\mathrm{IS}}\right) = \frac{1}{N}\mathbb{V}\mathrm{ar}_h\left(\frac{\partial g\left(\mathbf{s},\mathbf{X}\right)}{\partial s_\ell}\xi_\sigma\left(-g\left(\mathbf{s},\mathbf{X}\right)\right)\frac{f_{\mathbf{X}}\left(\mathbf{X}\right)}{h\left(\mathbf{X}\right)}\right). \tag{6.20}$$

The scaling process of Section 6.3.2.1 is also adapted as the weights $w^{(j)}$ of Eq. (6.18) are set to $w^{(j)} = \mathbb{I}_{Y<0}\left(Y^{(j)}\right) f_{\mathbf{X}}\left(\mathbf{X}^{(j)}\right)/h\left(\mathbf{X}^{(j)}\right)$ with the same notation as before. As the IS population is more concentrated around the limit state surface than a classic MC population, some aspects of the proposed method change, mainly the evolution of the variance function Eq. (6.20). These changes and their consequences are first detailed, and an illustration is provided with the toy example, as we combined the propose method with the NAIS algorithm described in Chapter 3 Section 3.3.3.1.

6.3.3.1 Change of behavior of the variance function

The population generated from h is assumed to be concentrated around the limit state surface $g(\mathbf{s},\mathbf{x}) = 0$, therefore the values of $g\left(\mathbf{s},\mathbf{X}^{(j)}\right)$ are expected to be close to zero and have a small magnitude. This phenomenon does not occur when the population is generated from $f_{\mathbf{X}}$. The consequence of this phenomenon is that the variance estimate of Eq. (6.20) is expected to have a smaller value than the MC variance estimate, when σ is close to zero. Indeed, more observations are close to the limit state surface. However, the variance estimate of Eq. (6.20) is also expected to be greater than the MC variance estimate for larger σ. In fact, for larger values of σ, the indicator approximation function Ξ_σ is then a lot smoother. This results in a density ξ_σ very wide, less concentrated around 0, with possibly heavier tails. Therefore, the integral $\partial\overline{P_f(\mathbf{s},\sigma)}/\partial s_\ell$ is defined over a larger domain. The population generated from h might not cover the whole domain, which results in a higher variance estimation, and also a larger empirical bias of the IS estimate Eq. (6.19).

The consequences of this phenomenon on the proposed approach concern the definition of the regression interval. Indeed, with crude Monte Carlo, as the variance estimate is a decreasing function of σ, we have $\sigma_{\max} = 1$, and the $\mathrm{CV}_{\mathrm{target}}$ affects only $\sigma_{\min}$. With IS, we expect $\mathrm{CV}_{\mathrm{target}}$ to affect both $\sigma_{\min}$ and $\sigma_{\max}$. The value of $\sigma_{\min}$ should be less shifted towards high values of σ than with crude MC. The value $\sigma_{\max}$ should no longer be equal to 1, reducing the regression interval on the higher values of σ. Furthermore, the noisiest estimates V_l will be located at both bounds of the regression interval, instead of only the lower bound. The global noise of the polynomial regression procedure is thus expected to be higher.

6.3.3.2 Illustration with the toy example

This phenomenon is illustrated with the toy example. The NAIS algorithm presented in Chapter 3 Section 3.3.3.1 is employed to construct an optimal IS density h, with $\rho = 20\%$. Let t be the number of steps required for the algorithm to converge, with h_t the last kernel density constructed, which serves for the failure probability estimation. Using the last population generated from h_t, we employ the proposed method to compute the probability sensitivity. Therefore, the simulation budget is equal to $t \times N$ evaluations of g to obtain the failure probability estimation and an additional N evaluations of the derivatives of g with respect to s_ℓ, for $\ell = 1, 2$ to obtain the failure probability sensitivity estimates.

Figure 6.6 illustrates the different graphs in the NAIS context, with $2k = 4$. Compared to the CV curves with crude MC, the regression interval is shifted towards the lower σ values, as both $\sigma_{\min}$ and $\sigma_{\max}$ decrease, as expected. The value of the threshold $\mathrm{CV}_{\mathrm{target}}$ allows to perform the polynomial regression where the IS estimates are the most precise. As there is a large noise at both bounds of the interval, the performed polynomial regression differs from the one performed with crude MC estimates. With a degree equal to 4, the polynomial regression succeeds in smoothing the noise of

the first IS estimate V_1, leading to a sensitivity estimate closer to the reference value. However the general behavior of the polynomial P_4 is quite distinct from the behavior of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$. The two functions are not as similar as in Figure 6.4 with a large MC, nevertheless, the resulting sensitivity estimate is still correct.

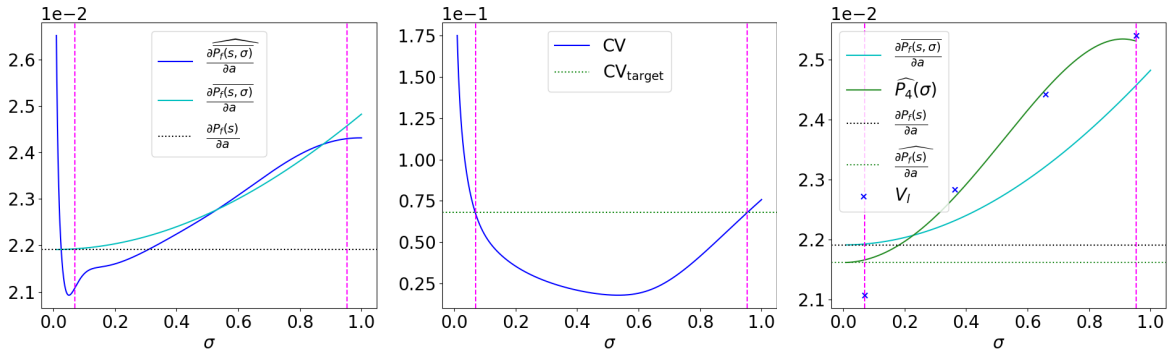


Figure 6.6: Illustration of the proposed method combined with NAIS with the toy example for $N = 2000$. Figure a) represents the evolution of the IS estimate according to σ in navy blue as well as the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ in clear blue. Figure b) shows the evolution of the theoretical CV estimate of the IS estimate and the value of the $\mathrm{CV}_{\mathrm{target}}$, close to 7%. Figure c) represents the estimates $(V_l)_{l=1,\dots,4}$ displayed with navy blue crosses obtained with bootstrap. The resulting polynomial $\widehat{P_4}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line.

The proposed method combined with NAIS is tested with $N = 2000$. The results are given in Table 6.3, for three different polynomial degrees, $2k = 2$, $2k = 4$ and $2k = 6$. The estimates are computed over 500 independent simulations, the given CV is the empirical CV while the number in parentheses is the mean value of the theoretical CV estimate obtained with the theoretical variance estimate of the sensitivity available in the WSL framework with Eq. (6.17). We compare the proposed method with the Weak approach as in the previous studies.

Table 6.3: Comparison of the result of the polynomial regression with NAIS, with 3 different degrees, for $N = 2000$. The failure probability is equal to 6.20×10^{-3} with an empirical CV of 4.8%. The reference values of the sensitivities are $\partial P_f/\partial a = 2.19 \times 10^{-2}$ and $\partial P_f/\partial b = -8.76 \times 10^{-3}$.

Toy example with NAIS								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial a}}$	2.23×10^{-2}		2.19×10^{-2}		2.19×10^{-2}		2.18×10^{-2}	
	CV	5.4% (4.0%)	CV	6.0% (4.8%)	CV	6.8% (5.7%)	CV	7.0%
$\widehat{\frac{\partial P_f}{\partial b}}$	-8.92×10^{-3}		-8.75×10^{-3}		-8.74×10^{-3}		-8.74×10^{-3}	
	CV	5.8% (4.1%)	CV	6.1% (4.9%)	CV	6.9% (5.8%)	CV	7.0%

From Table 6.3, it appears that for this example with the NAIS method, a polynomial of degree 2 is no longer sufficient to correctly estimate the coefficient of order zero of the Taylor series expansion of the functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$. Indeed, for both derivatives, the resulting sensitivity estimates are quite different from the ones obtained with a regression of degree $2k = 4$ and $2k = 6$. The difference is particularly obvious for the derivative with respect to b. This difference implies that the degree 2 is insufficient to obtain an accurate estimate with this particular setting. Since the sensitivity estimates obtained with $2k = 4$ and $2k = 6$ are similar, we can conclude that $2k = 4$ is the minimum degree needed for an accurate approximation with this particular setting. This increase in the minimum degree

required could be a consequence of the heavier noise at the lower bound of the regression interval.

For $2k \geq 4$, the sensitivity estimates obtained with the proposed method and the Weak approach are very close to the reference values. The CV of the sensitivity estimate obtained with the Weak approach is the highest. Consequently, the relevance of the polynomial regression procedure is once more underlined: for no additional simulation budget, the proposed method leads to a more precise sensitivity estimate. Compared to the results with crude MC, the estimates obtained with the Weak approach are more accurate, particularly with respect to b. This is due to the value of $\widetilde{\sigma} = \sigma_{\min}$, which is then smaller, as the IS population allows to have a better estimate V_1.

Discussion about the theoretical CV in the IS context

The theoretical CV estimate is still slightly underrated, but the difference between the theoretical CV estimate and the empirical CV is smaller than with crude MC; cf Table 6.1 and Table 6.2. However, the empirical CVs of these estimates, not displayed in Table 6.3, are larger than with crude MC. They are close to 40% for all degrees. Therefore the value of these theoretical CV estimates cannot be considered as very accurate.

As in the previous study, the sensitivity estimates obtained with an OLS approach with $\eta^2 = \max_{l=1,\dots,m} \left(\widehat{\mathbb{V}\mathrm{ar}}(V_l)\right)$ were computed as well, and are given in the Appendix E.3 in Table E.3. Performing the polynomial regression in the OLS framework leads to equivalent results in terms of empirical bias and empirical CV. Nevertheless, in contrast to the studies in the OLS framework with crude MC, the theoretical CV estimate is overrated and not particularly close to the actual value. And their resulting empirical CV is also close to 40%. Consequently, for the numerical investigation detailed next, the WSL framework is favored as it suits best the theoretical context.

6.4 Numerical investigation

The performance of the proposed approach is investigated with various numerical applications, taken from the failure probability sensitivity literature. All the examples involve standard normal inputs. It is reminded here that the inputs are supposed to follow a standard elliptical distribution in this thesis. The first two examples focus on applications in a rather low-dimensional space, with a single failure region: the cantilever beam [Papaioannou et al., 2013, Torii, 2020, Torii and Novotny, 2021, Papaioannou et al., 2018] and the roof truss [Torii and Novotny, 2021, Papaioannou et al., 2018, Proppe, 2021, Song et al., 2009]. As they are very common examples in the sensitivity literature, they are relevant to present the proposed approach. The last example is a Duffing oscillator [Zuev, 2009] in high-dimensional space with two failure regions and is more related to the specific context of our manuscript.

In the first and third examples the deterministic inputs $\mathbf{s}$ are design parameters while in the second example, they are the distribution parameters of the original inputs denoted $\mathbf{Z}$. It is recalled here that an isoprobabilistic transformation allows to transform the original inputs $\mathbf{Z}$ into the standard normal inputs $\mathbf{X}$; Chapter 2 Section 2.2.2.2. The limit state function in the original space is denoted $g_{\mathbf{Z}}$ while the transformed limit state function is denoted g.

The proposed approach is combined with IS algorithms for the numerical investigation. The IS method is employed to estimate the failure probability, as it results in a lower simulation budget than crude MC. The NAIS algorithm and the iCE-AIS algorithm presented in Chapter 3 are employed for the applications in small dimension. The proposed algorithm presented in Chapter 5 is employed for the high-dimensional application. For the NAIS algorithm and the iCE-AIS algorithm, the last generated sample of the adaptive procedure is used to compute the sensitivity with the new approach. For the algorithm presented in Chapter 5, since the MIS framework is used, all the generated samples of the adaptive procedure are used to compute the sensitivity with the new approach.

For all of these applications, the indicator approximation function Ξ_σ selected is the cdf of a centered normal random variable of variance σ^2. As previously detailed in Section 6.2.2.2, this particular choice results in even polynomials P_n, therefore $n = 2k$ and the number of V_l estimates is fixed to $m = k+2$. The proposed approach is studied for three different polynomial degrees $n = 2$, $n = 4$ and $n = 6$. It is compared to the Weak approach presented in Chapter 4 Section 4.3.3, when $\tilde{\sigma} = \sigma_{\min}$, the lower bound of the regression interval. For comparison purpose, 500 independent simulation runs are performed to calculate the statistics of the probability estimates and the other quantities of interest.

It should be noted that the mathematical statements needed in the theory of the proposed approach, i.e. that $F_{H_{\mathbf{s}}}$ is $\mathcal{C}^{n+2}$ in 0 and that all the different functions are differentiable with respect to s_ℓ, $\ell \in [1, \dots, p]$, are difficult to verify in practice, as we suppose the lsf to be a black-box function with no analytical expression. Therefore, these hypothesis are assumed to be validated in this section, even if it might not be the case, in order to assess the performance of the proposed method in *more realistic* scenarios.

6.4.1 Cantilever beam

6.4.1.1 Presentation of the application

The first example is a cantilever beam subject to biaxial bending as illustrated in Figure 6.7. This example is quite popular in sensitivity analysis [Papaioannou et al., 2013, Torii, 2020, Torii and Novotny, 2021, Papaioannou et al., 2018] and was first studied in [Yang and Gu, 2004] in a RBDO context. Two

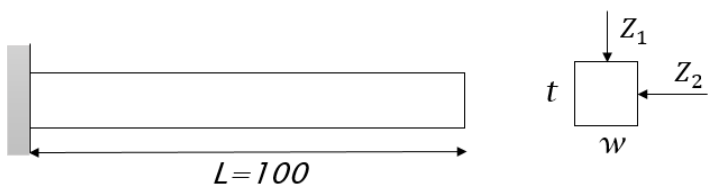


Figure 6.7: Illustration of a cantilever beam subject to biaxial bending.

different limit state functions are considered for this example, both defined in the original space $\mathbf{Z}$ of dimension 4. The first one $g^{(1)}_{\mathbf{Z}}$ represents yielding at the fixed end of the beam, with $\mathbf{s} = [w, t]$ respectively the width and the height of the cross-section beam

$$g^{(1)}_{\mathbf{Z}}(\mathbf{s}, \mathbf{Z}) = Z_3 - \left(\frac{600}{wt^2} Z_1 + \frac{600}{w^2 t} Z_2 \right).$$

The second limit state function $g^{(2)}_{\mathbf{Z}}$ restricts the maximum allowed displacement at the tip of the beam to the value d_0. Therefore, one has with $\mathbf{s} = [w, t, d_0]$

$$g^{(2)}_{\mathbf{Z}}(\mathbf{s}, \mathbf{Z}) = d_0 - \frac{4L^3}{Z_4 wt} Z_3 \sqrt{\left(\frac{Z_1}{t^2}\right)^2 + \left(\frac{Z_2}{w^2}\right)^2}$$

where $L = 100$m. The random variables Z_1 and Z_2 represent the loads, Z_3 is the yield strength of the beam and Z_4 is the Young's modulus. We assume the vector $\mathbf{Z}$ to be an independent normal vector. For comparison purpose, the distribution of each random variable is the same as in [Yang and Gu, 2004], therefore, denoting μ_{Z_i} the mean value and δ_{Z_i} the standard deviation, one has

$$\begin{aligned}
(\mu_{Z_1}, \delta_{Z_1}) &= (1000, 100) \\
(\mu_{Z_2}, \delta_{Z_2}) &= (500, 100) \\
(\mu_{Z_3}, \delta_{Z_3}) &= (40000, 2000) \\
(\mu_{Z_4}, \delta_{Z_4}) &= (29 \times 10^6, 1.45 \times 10^6).
\end{aligned}$$

Nevertheless, it should be noted that the normal distribution is strictly not an appropriate choice for modeling Z_3 and Z_4 as they represent physical variables of positive support. The isoprobabilistic transformation is quite simple in this case, as the inputs are already independent. Therefore $X_i = (Z_i - \mu_{Z_i})/\delta_{Z_i}$ for $i = 1, \dots, 4$; this transformation is linear.

The design parameters are fixed as $[w, t, d_0] = [2.4, 3.9, 2.5]$ which corresponds to the optimal reliability based design [Papaioannou et al., 2018].

6.4.1.2 Sensitivity analysis for the first failure of the system

As the first limit state function $g^{(1)}_{\mathbf{Z}}$ is linear and the transformation from $\mathbf{Z}$ to $\mathbf{X}$ is linear as well, the transformed lsf $g^{(1)}$ is also linear in the random variables $\mathbf{X}$. Consequently, the failure probability as

well as its sensitivities can be exactly determined with a FORM analysis; see Chapter 4 Section 4.3.1.1. This FORM analysis gives the following reference values [Papaioannou et al., 2018]: $P_f^{(1)} = 3.03 \times 10^{-3}$, $\partial P_f^{(1)}/\partial w = -5.76 \times 10^{-2}$ and $\partial P_f^{(1)}/\partial t = -3.53 \times 10^{-2}$.

For this first application, the NAIS algorithm is combined with the proposed method, with $\rho = 20\%$ and $N = 2000$. The mean simulation budget required for the probability estimation is near 8000 (4 iterations in the NAIS algorithm). The results are presented in Table 6.4, with the empirical CVs given, as well as the theoretical CV estimates of the proposed method in parentheses.

Table 6.4: Comparison of the result of the polynomial regression with 3 different degrees, for the first lsf of the cantilever beam with NAIS. The failure probability is equal to 2.97×10^{-3} with an empirical CV of 8.4%. The reference values of the sensitivities are $\partial P_f^{(1)}/\partial w = -5.76 \times 10^{-2}$ and $\partial P_f^{(1)}/\partial t = -3.53 \times 10^{-2}$.

Cantilever Beam 1								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial w}}$	-5.76×10^{-2}		-5.74×10^{-2}		-5.76×10^{-2}		-5.76×10^{-2}	
	CV	5.1% (3.4%)	CV	6.1% (4.6%)	CV	7.1% (5.7%)	CV	7.0%
$\widehat{\frac{\partial P_f}{\partial t}}$	-3.53×10^{-2}		-3.52×10^{-2}		-3.53×10^{-2}		-3.53×10^{-3}	
	CV	5.1% (3.4%)	CV	6.1% (4.6%)	CV	7.1% (5.7%)	CV	7.0%

From Table 6.4, the sensitivity estimates of both the proposed approach and the Weak approach are very close to the reference values of the failure probability sensitivities. For the three degrees selected and the two derivatives, the value of the estimate is quite constant. Therefore a polynomial of degree 2 is sufficient to correctly approximate the coefficient of order zero of the Taylor series expansion of the function $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$. The sensitivity estimate obtained with $2k = 2$ has a smaller CV than the one obtained with the Weak approach. Consequently, this application illustrates how the proposed method is an improvement of the Weak approach; it allows to have a more precise estimate, without any additional simulation budget. However, the theoretical CV estimates are still slightly underrated. They have a large variation (above 45%, not displayed in Table 6.4), therefore they cannot be considered accurately estimated.

6.4.1.3 Sensitivity analysis for the second failure of the system

The limit state function $g_{\mathbf{Z}}^{(2)}$ is nonlinear in all the random variables therefore FORM cannot be applied to compute accurately the failure probability and its sensitivities. The reference values used here are thus the one given in [Papaioannou et al., 2018], which results of a large line sampling simulation; see Chapter 4 Section 4.3.2.2. The LS simulation gives the following values: $P_f^{(2)} = 2.54 \times 10^{-4}$, $\partial P_f^{(2)}/\partial w = -8.84 \times 10^{-3}$, $\partial P_f^{(2)}/\partial t = -2.95 \times 10^{-3}$ and $\partial P_f^{(2)}/\partial d_0 = -3.27 \times 10^{-3}$.

For this second application, the proposed approach is combined with the iCE-AIS presented in Chapter 2 Section 3.3.3.4, with a single Gaussian density as auxiliary density, $N = 1000$ and a target CV parameter equal to $\delta = 2.0$. The mean simulation budget required for the probability estimation is near 4000 (4 iterations in the iCE-SG algorithm). The results are presented in Table 6.5, with the empirical CVs given, as well as the theoretical CV estimates of the proposed method in parentheses.

From Table 6.5, it is noticeable that the polynomial of degree $2k = 2$ leads to sensitivities whose values are slightly different than those obtained with higher degrees, for the three parameters. Therefore it appears that a polynomial of degree 2 is not sufficient to accurately estimate the coefficient of order zero of the Taylor series expansion of the functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$.

The estimate obtained with $2k = 2$ shares then the same properties than the Weak approach estimate. Their bias are quite similar along with their CVs. For $2k \geq 4$ the estimates obtained with the proposed approach have a smaller bias than the estimate obtained with the Weak approach, but they have a slightly higher CV. The theoretical CV estimates are still slightly underrated. They have

Table 6.5: Comparison of the result of the polynomial regression with 3 different degrees, for the second lsf of the cantilever beam with iCE-SG. The failure probability is equal to 2.51×10^{-3} with an empirical CV of 4.5%. The reference values of the sensitivities are $\partial P_f^{(2)}/\partial w = -8.84 \times 10^{-4}$, $\partial P_f^{(2)}/\partial t = -2.95 \times 10^{-3}$ and $\partial P_f^{(2)}/\partial d_0 = -3.27 \times 10^{-3}$.

Cantilever Beam 2								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial w}}$	-8.78×10^{-3}		-8.82×10^{-3}		-8.83×10^{-3}		-8.93×10^{-3}	
	CV	7.3% (5.0%)	CV	8.6% (6.8%)	CV	9.9% (8.8%)	CV	7.7%
$\widehat{\frac{\partial P_f}{\partial t}}$	-2.92×10^{-3}		-2.94×10^{-3}		-2.94×10^{-3}		-2.98×10^{-3}	
	CV	7.4% (5.0%)	CV	8.6% (6.8%)	CV	9.9% (8.8%)	CV	7.7%
$\widehat{\frac{\partial P_f}{\partial d_0}}$	-3.24×10^{-3}		-3.26×10^{-3}		-3.26×10^{-3}		-3.29×10^{-3}	
	CV	7.3% (5.0%)	CV	8.6% (6.8%)	CV	9.9% (8.8%)	CV	7.7%

a moderate variation for $2k \leq 4$, with an empirical CV close to 25% (not displayed in Table 6.5). For $2k = 6$ their variation is quite high (39%).

Consequently, for this application, the proposed approach brings improvement to the Weak approach framework as estimates with smaller bias can be obtained and the bias is globally controlled. For $2k \geq 4$, since the values of the estimates no longer vary, it can be assumed that the estimates have reached accurate values. Such analysis cannot be performed with the Weak approach, where the bias is not properly managed.

6.4.2 Roof truss

6.4.2.1 Presentation of the application

The second example is a roof truss subject to random loading as illustrated in Figure 6.8. This example is also very commonly used in sensitivity analysis [Torii and Novotny, 2021, Papaioannou et al., 2018, Proppe, 2021, Song et al., 2009] and we keep the same framework as presented in [Song et al., 2009]. The top boom and the compression bars are reinforced by concrete, the bottom boom and the tension bars are made of steel. The perpendicular deflection of the peak of the structure must

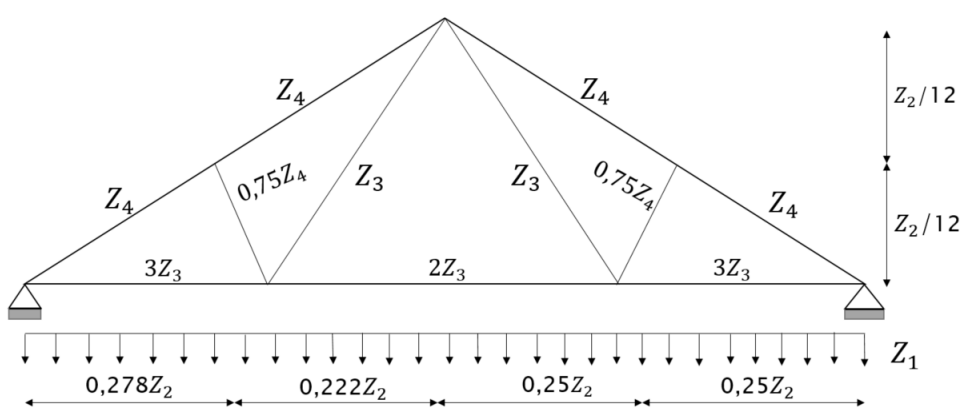


Figure 6.8: Illustration of a roof truss subject to random loading.

not exceed 3cm [Song et al., 2009]. Consequently the limit state function is defined in the original

space $\mathbf{Z}$ of dimension 6 with the following equation

$$g_{\mathbf{Z}}^{(3)}(\mathbf{s}, \mathbf{Z}) = 0.03 - \frac{Z_1 Z_2^2}{2}\left(\frac{3.81}{Z_4 Z_6} + \frac{1.13}{Z_3 Z_5}\right)$$

where $\mathbf{s} = [\mu_{Z_1}, \delta_{Z_1}, \ldots, \mu_{Z_6}, \delta_{Z_6}]$ with δ_{Z_i} denoting again the standard deviation of Z_i. The random variable Z_1 represents the uniformly distributed load applied on the roof truss and Z_2 is the roof span. Z_3 is the cross section area of the bottom boom and the tension bars made of steel, whose Young's modulus is Z_5. Z_4 is the cross section area of the top boom and compression bars reinforced in concrete, whose Young's modulus is Z_6. We assume the vector $\mathbf{Z}$ to be an independent normal vector. For comparison purpose, the distribution of each random variable is the same as in [Song et al., 2009]. Therefore we have

$$\begin{aligned}
(\mu_{Z_1}, \delta_{Z_1}) &= (20000, 1400)\\
(\mu_{Z_2}, \delta_{Z_2}) &= (12, 0.12)\\
(\mu_{Z_3}, \delta_{Z_3}) &= (9.82 \times 10^{-4}, 5.892 \times 10^{-5})\\
(\mu_{Z_4}, \delta_{Z_4}) &= (0.04, 0.0048)\\
(\mu_{Z_5}, \delta_{Z_5}) &= (1 \times 10^{11}, 6 \times 10^{9})\\
(\mu_{Z_6}, \delta_{Z_6}) &= (2 \times 10^{10}, 1.2 \times 10^{9})
\end{aligned}$$

As for the previous application, it should be noted that the normal distribution is strictly not an appropriate choice for modeling Z_2, Z_3, Z_4, Z_5 and Z_6 as they represent physical variables of positive support. Once more, the isoprobabilistic transformation is simple in this case, as the inputs are already independent and it results in $X_i = (Z_i - \mu_{Z_i})/\delta_{Z_i}$ for $i = 1, \ldots, 6$; this transformation is linear.

6.4.2.2 Sensitivity analysis of the system

The reference values of the failure probability and its sensitivities are taken the same as in [Papaioannou et al., 2018], obtained with the score function method combined with IS; see Chapter 4 Section 4.2.3.2. The resulting failure probability is equal to 9.38×10^{-3} and the sensitivities are equal to

$$\begin{aligned}
(\partial P_f/\partial \mu_{Z_1}, \partial P_f/\partial \delta_{Z_1}) &= \left(1.11 \times 10^{-5}, 1.59 \times 10^{-5}\right)\\
(\partial P_f/\partial \mu_{Z_2}, \partial P_f/\partial \delta_{Z_2}) &= \left(4.03 \times 10^{-2}, 1.80 \times 10^{-2}\right)\\
(\partial P_f/\partial \mu_{Z_3}, \partial P_f/\partial \delta_{Z_3}) &= \left(-1.86 \times 10^{2}, 2.05 \times 10^{2}\right)\\
(\partial P_f/\partial \mu_{Z_4}, \partial P_f/\delta_{Z_4}) &= (-2.14, 2.56)\\
(\partial P_f/\partial \mu_{Z_5}, \partial P_f/\delta_{Z_5}) &= \left(-1.83 \times 10^{-12}, 2.00 \times 10^{-12}\right)\\
(\partial P_f/\partial \mu_{Z_6}, \partial P_f/\delta_{Z_6}) &= \left(-3.77 \times 10^{-12}, 2.03 \times 10^{-12}\right)
\end{aligned}$$

For this application, the NAIS algorithm is combined with the proposed method, with $\rho = 20\%$ and $N = 2000$. The mean simulation budget required for the probability estimation is near 6000 (3 iterations in the NAIS algorithm). The results are presented in Table 6.6 and Table 6.7, with the empirical CVs given, as well as the theoretical CV estimates of the proposed method in parentheses.

From Table 6.6 and Table 6.7, the sensitivity estimates of both the proposed approach (omitting the estimates with respect to δ_{Z_2} and δ_{Z_6} in red) and the Weak approach are very close to the reference values. For the three degrees selected, the value of the estimate is quite constant, except for the derivatives with respect to δ_{Z_2} and δ_{Z_6}. Therefore a polynomial of degree 2 is sufficient to correctly approximate the coefficient of order zero of the Taylor series expansion of the functions $\sigma \mapsto \partial \overline{P_f}(\mathbf{s}, \sigma)/\partial s_\ell$ for $s_\ell \neq \delta_{Z_2}$ and $s_\ell \neq \delta_{Z_6}$. The sensitivity estimates obtained with $2k = 2$ have then a smaller CV than the one obtained with the Weak approach, for an equivalent bias. Consequently, for this application, the proposed method is an improvement of the Weak approach for 10 derivatives out of 12.

6.4.2.3 Focus on the derivatives with respect to δ_{Z_2} and δ_{Z_6}

The sensitivity estimates with respect to the distribution parameters δ_{Z_2} and δ_{Z_6} obtained with the proposed approach have a larger CV than those obtained with the Weak approach, no matter the

Table 6.6: Comparison of the result of the polynomial regression with 3 different degrees, for the roof truss with NAIS. The failure probability is equal to 9.28×10^{-3} with an empirical CV of 12.5%. The reference values of the sensitivities are $(\partial P_f/\partial \mu_{Z_1}, \partial P_f/\partial \delta_{Z_1}) = (1.11 \times 10^{-5}, 1.59 \times 10^{-5})$, $(\partial P_f/\partial \mu_{Z_2}, \partial P_f/\partial \delta_{Z_2}) = (4.03 \times 10^{-2}, 1.80 \times 10^{-2})$ and $(\partial P_f/\partial \mu_{Z_3}, \partial P_f/\partial \delta_{Z_3}) = (-1.86 \times 10^{2}, 2.05 \times 10^{2})$.

Roof Truss, Part 1								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial \mu_{Z_1}}}$	1.10×10^{-5}		1.10×10^{-5}		1.10×10^{-5}		1.11×10^{-5}	
	CV	7.4% (5.2%)	CV	9.0% (7.2%)	CV	14.4% (9.2%)	CV	8.2%
$\widehat{\frac{\partial P_f}{\partial \delta_{Z_1}}}$	1.57×10^{-5}		1.57×10^{-5}		1.58×10^{-5}		1.58×10^{-5}	
	CV	8.2% (5.8%)	CV	10.4% (8.2%)	CV	15.6% (11.2%)	CV	9.2%
$\widehat{\frac{\partial P_f}{\partial \mu_{Z_2}}}$	4.01×10^{-2}		4.03×10^{-2}		4.00×10^{-2}		4.05×10^{-2}	
	CV	7.4% (5.2%)	CV	9.0% (7.2%)	CV	16.0% (9.2%)	CV	8.3%
$\widehat{\frac{\partial P_f}{\partial \delta_{Z_2}}}$	1.79×10^{-2}		1.88×10^{-2}		2.04×10^{-2}		1.85×10^{-2}	
	CV	18.7% (15.1%)	CV	36.1% (26.0%)	CV	375% (12.8%)	CV	14.4%
$\widehat{\frac{\partial P_f}{\partial \mu_{Z_3}}}$	-1.85×10^{2}		-1.85×10^{2}		-1.84×10^{2}		-1.86×10^{2}	
	CV	7.4% (5.2%)	CV	9.1% (7.2%)	CV	14.8% (9.2%)	CV	8.3%
$\widehat{\frac{\partial P_f}{\partial \delta_{Z_3}}}$	2.02×10^{2}		2.03×10^{2}		2.05×10^{2}		2.03×10^{2}	
	CV	10.0% (7.3%)	CV	13.6% (9.1%)	CV	26.1% (14.0%)	CV	10.1%

polynomial degree. For $2k \leq 4$ the estimates have a small bias and a moderate CV. However, when $2k = 6$ the bias is higher and the CV considerably increases to reach values above 300%. It should also be noted that the theoretical CV estimates are meaningless, as they all have an empirical CV above 100% (not displayed here), for the three degrees.

These poor results can be explained by the poor quality of the estimates V_l and their variance estimates, but it is mostly due to the phenomenon of polynomial overfitting. As previously underlined, polynomial regressions of higher degrees are harder to perform and the resulting polynomial is more flexible than with smaller degrees. For these two distribution parameters, the noise of each V_l is quite high and the regression interval can be particularly narrow, as illustrated in Figure 6.9.

The resulting polynomials shown in Figure 6.9 c) and f) have much more fluctuations in the regression interval than the original IS estimates shown in Figure 6.9 a) and d). For these parameters, the polynomial regression does not smooth the noise of the IS estimate V_l. The polynomials are too flexible and they take excessively into account the noise of each IS estimate: it is the phenomenon of overfitting. This phenomenon does not occur for lower degrees, as the polynomial are then less flexible. Consequently, when $2k = 6$, the polynomial regression can lead to sensitivity estimate very biased, which explains the large CV.

It should be underlined that in other studies focusing on this roof truss application, these two distribution parameters always lead to sensitivity estimates with significantly higher CVs compared to the other distribution parameters [Proppe, 2021, Song et al., 2009, Papaioannou et al., 2018].

Table 6.7: Comparison of the result of the polynomial regression with 3 different degrees, for the roof truss with NAIS. The failure probability is equal to 9.28×10^{-3} with an empirical CV of 12.5%. The reference values of the sensitivities are $(\partial P_f/\partial \mu_{Z_4}, \partial P_f/\delta_{Z_4})=(-2.14, 2.56)$, $(\partial P_f/\partial \mu_{Z_5}, \partial P_f/\delta_{Z_5}) = (-1.83 \times 10^{-12}, 2.00 \times 10^{-12})$ and $(\partial P_f/\partial \mu_{Z_6}, \partial P_f/\delta_{Z_6})=(-3.77 \times 10^{-12}, 2.03 \times 10^{-12})$.

Roof Truss, Part 2								
	Regression $2k=2$		Regression $2k=4$		Regression $2k=6$		Weak approach $\tilde{\sigma}=\sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial \mu_{Z_4}}}$	-2.11		-2.11		-2.11		-2.13	
	CV	7.5% (5.4%)	CV	9.0% (7.4%)	CV	10.6% (9.6%)	CV	8.3%
$\widehat{\frac{\partial P_f}{\partial \delta_{Z_4}}}$	2.49		2.48		2.50		2.50	
	CV	10.7% (8.1%)	CV	13.3% (11.6%)	CV	19.1% (18.2%)	CV	10.8%
$\widehat{\frac{\partial P_f}{\partial \mu_{Z_5}}}$	-1.82×10^{-12}		-1.82×10^{-12}		-1.81×10^{-12}		-1.83×10^{-12}	
	CV	7.5% (5.2%)	CV	9.1% (7.2%)	CV	16.5% (9.6%)	CV	8.3%
$\widehat{\frac{\partial P_f}{\partial \delta_{Z_5}}}$	2.01×10^{-12}		2.02×10^{-12}		1.98×10^{-12}		2.02×10^{-12}	
	CV	10.2% (7.3%)	CV	13.2% (10.5%)	CV	48.5% (15.3%)	CV	10.4%
$\widehat{\frac{\partial P_f}{\partial \mu_{Z_6}}}$	-3.74×10^{-12}		-3.75×10^{-12}		-3.74×10^{-12}		-3.77×10^{-12}	
	CV	7.4% (5.3%)	CV	9.3% (7.3%)	CV	16.4% (9.2%)	CV	8.2%
$\widehat{\frac{\partial P_f}{\partial \delta_{Z_6}}}$	1.97×10^{-12}		2.00×10^{-12}		2.32×10^{-12}		2.02×10^{-12}	
	CV	15.2% (13.2%)	CV	45.1% (20.9%)	CV	320% (41.2%)	CV	13.4%

6.4.3 Duffing oscillator

6.4.3.1 Presentation of the application

This application is the same as the one presented in Chapter 5, Section 5.3.2, it is a nonlinear elastic system as described in [Zuev, 2009]. We recall here that the Duffing oscillator is modeled by the equation

$$m\ddot{z}(t) + c\dot{z}(t) + k\left\{z(t) + \gamma z(t)^3\right\} = f(t),$$

where $z(t)$, $\dot{z}(t)$, and $\ddot{z}(t)$ are the displacement, velocity, and acceleration of the oscillator at time t, respectively, and the constants are taken as in [Zuev, 2009], with $m = 1000$ kg, $c = 200\pi$ Ns/m, $k = 1000(2\pi)^2$ N/m and $\gamma = 1$ m^{-2}. The oscillator is subjected to the random loading $f(t, \mathbf{X})$, discretized in the frequency domain as in [Papaioannou et al., 2019]

$$f(t, \mathbf{X}) = -m\sigma \sum_{i=1}^{d/2} \left\{X_i \cos(\omega_i t) + X_{d/2+i} \sin(\omega_i t)\right\},$$

where $\omega_i = i\Delta\omega$ with $\Delta\omega = 30\pi/d$ and $\sigma = \sqrt{2S\Delta\omega}$, where $S = 0.005$ m^2/s^3 is the intensity of the loading. Here, $\mathbf{X}$ are independent standard Gaussian random variables of dimension 100. The lsf $g^{(4)}$ is the maximum displacement of the oscillator at $t = 2$ s defined as

$$g^{(4)}(\mathbf{s}, \mathbf{X}) = \min\{z_{\text{crit},1} - z(2s), z(2s) - z_{\text{crit},2}\},$$

where $z_{\text{crit},1} = 0.1$ and $z_{\text{crit},2} = -0.06$ and $\mathbf{s} = [m, k]$ which represent respectively the mass and the stiffness of the oscillator. The initial values are set to $z(0) = 0$ and $\dot{z}(0) = 1.5$. The limit state function along with its gradient are computed with an Euler method. The reference value is computed with a

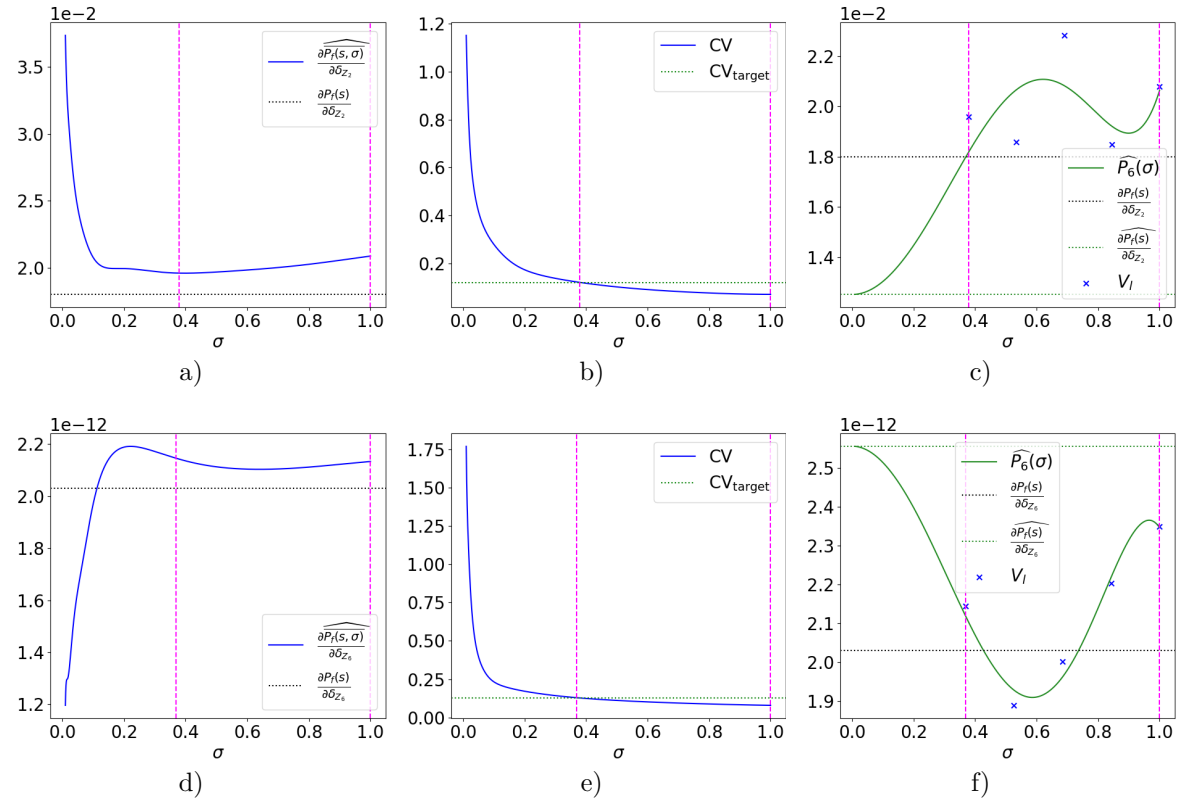


Figure 6.9: Illustration of the proposed method with $2k = 6$ for the roof truss application, with a NAIS algorithm of size $N = 2000$, with $\rho = 20$. Figures a), b) and c) illustrate the method with respect to the distribution parameter $s_\ell = \delta_{Z_2}$, while Figures d), e) and f) illustrate the method with respect to the distribution parameter $s_\ell = \delta_{Z_6}$. Figures a) and d) represent the evolution of the IS estimate according to σ in navy blue. Figures b) and e) show the evolution of the theoretical CV estimate of the IS estimate and the value of the $\text{CV}_{\text{target}}$. Figures c) and f) represent the estimates $(V_l)_{l=1,\dots,5}$ displayed with navy blue crosses obtained with bootstrap. The resulting polynomial $\widehat{P_6}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line.

Monte Carlo method and is equal to 4.28×10^{-4} with an empirical CV of 4.9% with a sample of size of 10^6. This application allows to link the two algorithms presented in this manuscript.

6.4.3.2 Sensitivity analysis of the system

The reference value of the failure probability sensitivities are computed with finite difference schemes, as there is no reference values available in the literature for this example; see Chapter 4 Section 4.3.4 for the finite difference schemes. The resulting sensitivities are equal to $\partial P_f/\partial m = -2.45 \times 10^{-5}$ and $\partial P_f/\partial k = 6.83 \times 10^{-7}$ with the Direct approach and the central formula.

For this application, the algorithm presented in Chapter 5 is combined with the proposed method with $N = 1000$ and $\kappa_0 = 50$. The final simulation budget N_{samp} is near 6000 (3 iterations in the CE-AMIS for each failure region). We suppose the dependence between the various samples generated during the MIS procedure of the failure probability estimation to be negligible. All of these samples are used in the proposed method to compute the sensitivity estimates. The derivatives of the lsf with respect to m and k are thus evaluated for all samples. Therefore, the simulation budget is greater than with other IS algorithms where only the last sample is used for the sensitivity process. The results are presented in Table 6.8 with the empirical CVs given as well as the theoretical CV estimates of the proposed method in parentheses.

All the sensitivity estimates displayed in Table 6.8 have a very large CV, which indicates that the sensitivity analysis of this particular system is especially difficult to derive. The sensitivity estimates

Table 6.8: Comparison of the result of the polynomial regression with 3 different degrees, for the Duffing oscillator. The failure probability is equal to 4.28×10^{-4} with an empirical CV of 4.4%. The reference values of the sensitivities are $\partial P_f/\partial m = -2.45 \times 10^{-5}$ and $\partial P_f/\partial k = 6.83 \times 10^{-7}$.

Duffing Oscillator								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial m}}$	-2.43×10^{-5}		-2.44×10^{-5}		-2.42×10^{-5}		-2.54×10^{-5}	
	CV	18.2% (16.1%)	CV	21.8% (28.5%)	CV	46.1% (145%)	CV	14.1%
$\widehat{\frac{\partial P_f}{\partial k}}$	6.76×10^{-7}		6.79×10^{-7}		6.72×10^{-7}		7.06×10^{-7}	
	CV	17.4% (15.2%)	CV	20.5% (25.9%)	CV	39.3% (217%)	CV	13.7%

obtained with the Weak approach are quite biased, and their CVs are the smallest.

The sensitivity estimates obtained with the proposed method are all closer to the reference values, for the three degrees selected, but they have a large CV. Since the regression estimates all have similar value, it seems that a polynomial of degree 2 is sufficient to estimate the coefficient of order zero of the Taylor series expansion of the functions $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$.

The empirical CVs of the theoretical CV estimates in parenthesis are equal to $(25.8\%, 71.8\%, 183\%)$ for the sensitivity estimates with respect to m. They are equal to $(24.8\%, 62.4\%, 152\%)$ for the sensitivity estimates with respect to k. Consequently, only for $2k = 2$ can they be considered as moderately accurate. For $2k \geq 4$ these estimates are meaningless and do not reflect the real variations of the sensitivity estimates, especially for $2k = 6$. Therefore, for this example, the theoretical variance is particularly poorly estimated. It should be noted that the polynomial regressions were also performed in the OLS framework for this particular example, and led to very poor variance estimates as well.

As in the application Section 6.4.1.3, the proposed method is still an improvement of the Weak approach as the bias of the estimate is smaller and can be controlled. Since the value of the estimates, for both parameters, is quite similar for the three degrees, one can assume that this value is correct. However, the variation of the estimates is very large, especially for greater polynomial degrees. On the other hand, the estimates obtained with the Weak approach have a smaller CV but their accuracy cannot be assessed. A smaller CV cannot guarantee the quality of the estimate with regard to the bias. Therefore, focusing on the bias, the proposed method allows to ameliorate the Weak approach.

It should be noted that compared to the other applications, as the estimates obtained with the Weak approach have a large CV, it seems all Weak approach inspired methods struggle to perform the sensitivity analysis of this particular system.

6.5 Conclusion

In this chapter a new method to compute the local sensitivity of the failure probability of a system was presented, based on a heteroscedastic polynomial regression. The derivatives with regard to deterministic inputs of the system are estimated. These deterministic inputs can either be design parameters or distribution parameters. The method is suited for standard elliptical inputs. This approach is inspired by the Weak approach framework and is presented as an improvement of the latter. Here we discuss some final remarks and outlooks.

6.5.1 Discussion on the heteroscedastic polynomial regression

The main innovation of the proposed approach is to express the sensitivity estimate as the constant coefficient of a Taylor series expansion, which can be recovered with a polynomial regression. However, as previously mentioned, we underline here that the main subject of this chapter is not the study of polynomial regressions, but the sensitivity analysis of the system with respect to deterministic inputs.

6.5.1.1 Settings of the polynomial regression

The proposed method uses basic settings in order to perform a heteroscedastic polynomial regression. This polynomial regression must keep the simulation budget constant, and therefore it cannot require new evaluations of the limit state function or its derivatives, compared to the Weak approach. These constraints led to the choices made to perform the regression as described. However, it is assumed that this method could be further improved with a more detailed study on heteroscedastic polynomial regressions, leading to better regression interval and other regression settings.

In particular, one of the main outlooks of the proposed method should be to improve the stability of the resulting sensitivity estimate. Indeed, in several applications, the mean value of the estimates is very close to the reference value but the CV is slightly too large for the method to be definitely an upgrade of the Weak approach; see 6.4.1.3 and 6.4.3 for instance.

6.5.1.2 Theoretical variance of the regression estimate

In all the applications presented, including the toy example, the theoretical variance of the regression estimate is biased in the heteroscedastic framework. The bias is not very large but it appears the theoretical variance is always underestimated. This could be a consequence of the small inaccuracy of the variance of each V_l which is evaluated with a MC or IS sample. The approximation of the variances is not taken into account in the regression and could explain the resulting bias. The results of the proposed method in the OLS framework, applied to the toy example with crude MC (in the Appendix E.2), seem to validate the fact that the variance estimates might be the reason of such bias.

Therefore, another interesting outlook of the proposed approach is to obtain a better estimation of the theoretical variance of the sensitivity estimate. Indeed, with the Weak approach, the theoretical variance is already available with formulas Eq. (6.13) for crude MC and Eq. (6.20) for IS. Consequently, since the proposed method is presented as an improvement of the Weak approach, the availability of an accurate estimation of the theoretical variance is of great interest.

6.5.2 High-dimensional elliptical space with multiple failure regions

This method was tested in high-dimensional normal space with a system with two failure regions, as it is the specific context of this manuscript. Although the results were not as precise as with other numerical applications in terms of CV, we do not think that it is linked to the high-dimensional inputs or the multiple failure regions of the failure domain. Indeed, the proposed method only focuses on the one-dimensional response of the system $Y = g(\mathbf{s}, \mathbf{X})$. It is therefore independent of the input space dimension and the shape of the failure domain. The results, less precise in terms of CV, could be a consequence of the limit state function being more intricate and more complex than with the other applications.

Furthermore, as the proposed approach focuses on the output of the limit state function, we assume that the initial elliptical distribution of the inputs does not matter. Consequently, it could be interesting to investigate the performance of the proposed algorithm in the context of high-dimensional systems with multiple failure regions, with other elliptical distributions as well.

6.5.3 Generalization of the proposed approach

Here, it is assumed that the failure probability has first been estimated with a simulation method. As underlined with the various examples throughout the chapter, the proposed approach can be combined with different simulation method. In fact, the proposed approach only requires the vector value of the limit state function outputs (with weights in the IS framework) and the vector value of its derivative with respect to s_ℓ, for a specific sample. The density used to generate such a sample is not needed.

Therefore, we assume that the proposed approach can be applied to many other simulation methods than MC and IS. As mentioned above, it is also assumed that the dimension of the system, the multiple failure regions, as well as the distribution of the inputs do not matter for the proposed approach to be efficient. Consequently, this approach can be adapted to a great number of systems.

Conclusion

La probabilité de défaillance du système et sa sensibilité locale sont des quantités d'intérêt importantes en fiabilité. Si la dimension du système est particulièrement grande, c'est-à-dire plus grande que 50, et que le domaine de défaillance des entrées possède plusieurs régions de défaillance, l'estimation de ces quantités est plus complexe. Dans cette thèse, deux approches originales sont proposées pour estimer dans l'espace standard elliptique la probabilité de défaillance d'une part et ses dérivées d'une autre part. Les deux approches emploient le cadre de l'échantillonnage préférentiel, qui consiste à introduire une densité auxiliaire dans le calcul des intégrales de la probabilité de défaillance et de ses dérivées.

L'algorithme d'estimation de probabilité de défaillance est détaillé au Chapitre 5 du manuscrit. Il a été le fruit d'une collaboration avec le Professeur Christian Genest, suite à un stage de 5 mois à l'Université McGill à Montréal sous sa supervision. Cet algorithme a été présenté à deux conférences, ECCOMAS 2022 et ESREL 2022 et a donné lieu à la publication suivante

- Chiron M., Genest C., Morio, J. and Dubreuil S. (2023). Failure probability estimation through high-dimensional elliptical distribution modeling with multiple importance sampling. *Reliability Engineering & System Safety*, Vol. 235, 109238.

Le principe de l'algorithme consiste à chercher progressivement dans l'espace standard elliptique les régions de défaillance. Pour chaque région de défaillance trouvée, une densité auxiliaire est construite et optimisée de manière adaptative. Cette densité tient compte de la représentation stochastique des entrées elliptiques puisqu'elle s'écrit comme le produit d'une densité radiale et d'une densité directionnelle. La probabilité de défaillance finale est estimée grâce à un estimateur d'échantillonnage préférentiel multiple, avec un mélange de toutes les densités auxiliaires précédemment construites. Plusieurs perspectives intéressantes concernant cet algorithme sont brièvement mentionnées ici. Tout d'abord, la densité auxiliaire radiale sélectionnée dans l'algorithme peut ne pas être optimale pour des domaines de défaillance à symétrie sphérique. Ainsi, une autre densité auxiliaire radiale pourrait être envisagée, comme une densité de noyau ou bien une loi de Nakagami comme présentée dans [Papaioannou et al., 2019]. Une étape d'échantillonnage supplémentaire serait alors nécessaire pour paramétrer ces densités. Ensuite, il pourrait être utile d'adapter l'algorithme proposé au cas où le gradient de la fonction d'état limite n'est pas disponible pour la recherche des régions de défaillance. En effet, dans ce cas, les optimisations nécessaires à cette recherche sont alors bien plus complexes à réaliser. Récemment, plusieurs algorithmes ont été présentés dans la littérature [Zhong et al., 2020, Zhu et al., 2022] permettant de localiser les points de conception d'un système dans un espace de grande dimension sans le gradient. Ces méthodes innovantes offrent donc de nouvelles possibilités à ce sujet. Enfin, l'approche globale proposée ici, consistant à chercher chaque région de défaillance au fur et à mesure et à optimiser une unique densité, pourrait être adaptée à d'autres algorithmes d'échantillonnage préférentiel. En effet, la logique du "diviser pour mieux régner" semble pertinente et pourrait aboutir à des algorithmes d'échantillonnage préférentiel pour régions de défaillance multiples plus efficaces.

L'algorithme d'estimation de la sensibilité locale de la probabilité de défaillance par rapport aux entrées déterministes du système est détaillé au Chapitre 6 du manuscrit. Il a été présenté au séminaire ETICS 2020 ainsi qu'à la conférence RESIM 2021 avec un poster intitulé *Rare event probability derivatives of a parametric complex system*. Un article est en cours de rédaction. Le principe de cette nouvelle approche consiste à exprimer la dérivée de la probabilité comme le terme d'ordre zéro d'un développement de Taylor. Une estimation de la dérivée est alors possible grâce à une régression polynomiale hétéroscédastique, dont la base de données est obtenue par des méthodes de simulation. L'estimateur est biaisé mais le biais est contrôlable grâce aux paramètres de la régression : le degré du polynôme choisi et l'intervalle de régression. La régression polynomiale détaillée dans le Chapitre 6 résulte de choix simples. Ainsi, nous supposons que la méthode pourrait être améliorée après une

étude plus approfondie sur les régressions polynomiales hétéroscédastiques dans le contexte spécifique présenté. En particulier, il semble que les estimateurs de dérivées obtenus aient une variance trop élevée par rapport à l'approche Faible. Ainsi, parvenir à stabiliser la valeur des estimateurs semble un axe de recherche important. Il en est de même pour l'estimateur de la variance théorique, dont le biais est de plus non négligeable. Dans le cadre de l'approche proposée, seul le terme constant du polynôme identifié est examiné, puisqu'il s'agit de la quantité d'intérêt. Cependant, une étude des autres coefficients du polynôme pourrait apporter une information supplémentaire quant à l'influence des paramètres déterministes sur la probabilité de défaillance. Enfin, l'approximation de la fonction indicatrice du domaine de défaillance est une pratique courante [Papaioannou et al., 2016]. L'analyse mathématique de l'expression de la probabilité de défaillance approchée, qui a été conduite ici, pourrait donner lieu à d'autres nouvelles expressions exploitables.

La combinaison des deux algorithmes a été examinée avec une application à la fin du Chapitre 6, dont les variables aléatoires standards normales sont de dimension 100. Il semble que pour cet exemple, l'algorithme de calcul de dérivée de probabilité ne soit pas aussi performant que pour d'autres applications, en termes de coefficient de variation. Nous ne pensons pas que cela soit dû à la grande dimension des entrées du système, ou à la multitude de régions de défaillance. Ainsi, il serait intéressant de tester la combinaison des deux algorithmes proposés avec d'autres exemples en grande dimension possédant plusieurs zones de défaillance, dont les variables aléatoires sont modélisées par d'autres distributions elliptiques standards que la loi normale.

Chapter 7

Conclusion

Summary of the main contributions

The failure probability and its local sensitivity are important quantities of interest when evaluating the reliability of the system. If the dimension of the system is particularly high, say superior to fifty, and the failure domain of the inputs is quite complex, encompassing several failure regions, the estimation of these quantities becomes more demanding. In this thesis, two original approaches to estimate respectively the failure probability and its derivatives are presented. Both approaches employ the important sampling framework, which consist in introducing an auxiliary density in the computation of both the failure probability integral and the sensitivity integral.

Estimation of the failure probability with elliptical modeling

This algorithm, specifically tailored for standard elliptical inputs, is described in detail in Chapter 5. The search for the failure regions is performed with an optimization inspired by the FORM/SORM framework. However, the constraint of searching for the failure region inside the important ring of the standard elliptical inputs is added to the optimization. For each failure region, the stochastic representation of the inputs is taken advantage of to propose an auxiliary density built as the product of a parametric radial density and a parametric directional density. The parametric radial density is a conditional density based on the original radial distribution of the elliptical inputs. The choice of this radial density makes it possible to simplify the importance sampling weights. The parametric directional density is a von Mises–Fisher density. The parameters of both densities are set thanks to the coordinates of the closest failing point of the failure region, which belongs as well to the important ring, along with a cross-entropy based optimization. Depending on the location of the failure region inside the importance ring, the parameter setting changes. Therefore two distinct situations were introduced.

This gradual search for the failure regions combined with adaptive importance sampling is quite original, particularly in high-dimensional spaces. The failure probability of each failure region can be computed along and allows a better understanding of the behavior of the inputs leading to the failure of the system. The final global failure probability is computed with a mixture of all the auxiliary densities previously constructed, as in the multiple importance sampling framework. All the samples generated and evaluated with the costly limit state function in the cross-entropy based optimization procedures are taken advantage of, as they are reused in the final failure probability estimation. It should be underlined that the proposed algorithm presents a general method suitable for other elliptical distributions than the standard normal distribution.

The algorithm is tested with four numerical applications in which the inputs follow a standard normal distribution or a multivariate Student distribution. The dimension of the inputs goes from 100 to 1000 in these examples. The first three applications are inspired by the failure probability literature while the last one is a realistic aerodynamic engineering application. These examples make it possible to experiment the robustness of the proposed algorithm for various shapes of failure domains, encompassing multiple regions. For all the numerical applications, the algorithm succeeds in finding all the failure regions of the failure domain, without any prior information. The last example underlines the performance of the proposed algorithm with a real numerically expensive black-box model.

This contribution is the result of a collaboration with Professor Christian Genest, following a five-month research internship at the University of McGill in Canada under his supervision. The algorithm was presented at two conferences, ECCOMAS 2022 and ESREL 2022 and resulted in the following journal publication

- Chiron M., Genest C., Morio, J. and Dubreuil S. (2023). Failure probability estimation through high-dimensional elliptical distribution modeling with multiple importance sampling. *Reliability Engineering & System Safety*, Vol. 235, 109238.

Estimation of the failure probability sensitivity with polynomial regression

This algorithm is described in detail in Chapter 6 and concerns any type of deterministic inputs; the sensitivity analysis is performed in the standard elliptical space. The proposed approach is inspired by the Weak approach, as an approximation of the failure domain indicator function is employed, namely a cumulative distribution function. The originality of the proposed approach is to take into account the nature of such an approximation, in order to write the approximated sensitivity integral with another expression, whose Taylor series expansion is of interest. Indeed, the sensitivity appears as the constant coefficient of a specific polynomial, and can be recovered with a polynomial regression. A toy example allows to illustrate the main idea of the approach, where the coefficients of the polynomial are analytically computed. As the Taylor series expansion has an infinite number of terms, the proposed approach is theoretically biased since the polynomial used for the regression has necessarily a finite degree. However, this bias can be controlled with the different settings of the polynomial regression such as the regression interval and the chosen polynomial degree.

Using simulation methods to build the polynomial regression database, the variance of each estimate of the polynomial evaluation is not constant, therefore the heteroscedastic framework of the linear least squares method is necessary, and bootstrap is employed to decrease the correlation. This polynomial regression does not require any additional evaluation of the limit state function or its gradient. The method is first experiment with crude Monte Carlo estimates and is then adapted to the importance sampling framework. With this adaptation, it is then possible to combined both algorithms developed during the thesis to estimate the failure probability and its sensitivity with respect to deterministic inputs.

The proposed approach is tested with three numerical applications, in which the inputs follow a standard normal distribution, their dimension goes from 3 to 100. The first two examples are taken from the failure probability sensitivity literature in small dimension and allow to validate the proposed method, as it is quite original. The last example concerns one of the numerical applications used to test the first algorithm described in Chapter 5, and allows to link the two algorithms presented in this manuscript. The importance sampling algorithms employed for the different numerical applications differ, which underlines the adaptability of the proposed method to compute the failure probability sensitivities, no matter the initial failure probability estimation procedure selected. For most of the numerical applications, the proposed algorithm is an improvement of the Weak approach as it results in sensitivity estimates with smaller bias and smaller CV without any additional simulation budget. Moreover, for all the applications, the bias of the estimate can be effectively controlled with the degree of the polynomial used for the regression.

This contribution was presented at the ETICS 2020 seminar and the RESIM 2021 conference with posters under the title *Rare event probability derivatives of a parametric complex system.* A paper is soon to be sent to a scientific journal, in order to be published.

Perspectives

Several aspects of the two contributions could be enhanced and are presented here. Because of time constraints, these outlooks could not be further studied during the thesis but they provide interesting subjects for future work.

Outlooks concerning the first algorithm on failure probability estimation

In several numerical applications, the radial distribution selected as auxiliary density in the construction of the importance sampling density of a failure region is left unchanged. This situation occurs

when the failure region starts before the beginning of the importance ring and spreads across it. In this situation, if the failure domain happens to have a spherical symmetry, the proposed algorithm could then be quite inefficient. Therefore, another parametric radial density could be introduced in the proposed algorithm. This other density would lead to more intricate multiple importance sampling weights, since there would be no longer a simplification. However the resulting algorithm might be more flexible to different shapes of failure domain.

For instance, kernel densities could be considered, as they are particularly efficient in one-dimensional space. Another solution could be the Nakagami distribution employed in [Papaioannou et al., 2019]. It should be noted that a recent study [Leng et al., 2022] focused exclusively on finding an importance sampling auxiliary density for the radial component of the standard normal inputs, leaving the directional density as the original uniform distribution over the unit hypersphere. Consequently, this subject is already acknowledged by the scientific community and such studies could inspire improvements for the proposed algorithm.

On another subject, it is assumed in this manuscript that the gradient of the limit state function with respect to the random variables is available. When this gradient is not available, the optimizations to search for the failure regions are very difficult to perform in high-dimensional space and the proposed algorithm may be inefficient or inaccurate. However, many recent papers can be found in the literature which make it possible to find the design points for the FORM in high-dimensional standard normal space without employing the gradient [Zhong et al., 2020, Zhu et al., 2022]. These new optimization schemes combined with the important ring constraint could be a promising gradient-free alternative, in the proposed algorithm procedure.

On a more general note, it could be interesting to adapt the search for the failure regions presented in the proposed algorithm to other failure probability estimation methods, as well as the *divide and conquer* logic implemented here. Indeed, if the failure domain encompasses several failure regions, most importance sampling algorithms in the literature rely on an auxiliary density built as a mixture of densities optimized as a whole [Wang and Song, 2016, Papaioannou et al., 2019, Kurtz and Song, 2013, Zhang et al., 2022, Geyer et al., 2019]. The number K of densities in the mixture is then difficult to set *a priori* and the optimization concerns a very large number of parameters, since it is proportional to K. The method proposed here finds this number of failure region gradually, building the mixture of densities at the end of the procedure. Furthermore, each optimization concerns then a smaller number of parameter, as only one density is optimized. Therefore, adapting this logic to other importance sampling methods could lead to improvements of the latter.

Outlooks concerning the second algorithm on failure probability sensitivity estimation

The performance of the proposed algorithm described in Chapter 6 depends on the quality of the polynomial regression procedure. The different settings of the proposed regression are the consequences of straightforward simple choices. Consequently it is assumed that the proposed approach could be further improved with a more detailed study on heteroscedastic polynomial regressions, leading to a better regression interval and other regression settings.

In particular, one of the main outlooks of the proposed method should be to improve the stability of the resulting sensitivity estimate. Indeed, in several applications, the mean value of the estimate is very close to the reference but the coefficient of variation is slightly too large for the method to be definitely considered as an improvement of the Weak approach.

Furthermore, in all the different applications displayed, including the toy example, the theoretical variance estimate of the derivative is biased in the heteroscedastic framework. The bias is not very large but it appears the theoretical variance is always underestimated. The variation of this theoretical variance estimate is also quite large. Therefore, another interesting outlook of the proposed method is to be able to have a better estimation of the theoretical variance of the sensitivity estimate. Since a correct theoretical variance estimate is available in the Weak approach framework, the accessibility of an accurate estimation of the theoretical variance in the proposed method is of great interest.

On another subject, in the proposed approach, only the constant term of the polynomial has been studied and examined as it represents the failure probability sensitivity which is the quantity of interest. However, it could be interesting to study the other coefficients of the polynomial obtained with the regression. These coefficients might bring additional information concerning the influence of

each parameter s_ℓ on the failure probability.

On a more general note, the approximation of the failure domain indicator function selected in this approach is used in other methods, such as in the Sequential Importance Sampling (SIS) framework [Papaioannou et al., 2016, Papaioannou et al., 2018] for instance. However, the nature of this approximation, which is a cumulative distribution function, is not taken advantage of, as in the proposed approach. Therefore, it could be appealing to examine if the mathematical study proposed here of the approximated failure probability integral could lead to other innovative expressions.

Outlooks concerning the combination of both algorithms

The combination of both algorithms was tested in the last numerical application of Chapter 6, with a high-dimensional system whose failure domain encompasses two failure regions. This numerical application lead to failure probability sensitivity estimates of inferior quality regarding their coefficient of variation, compared to the other applications presented in Chapter 6, with small-dimensional systems.

However, a direct link between the performance of the first algorithm and these results could not be particularly underlined, as the failure probability is accurately estimated. Furthermore, the high-dimensional space as well as the multiple regions of the failure domain should not play an important role in the quality of the resulting sensitivity estimates as the polynomial regression is performed in the one-dimensional space of the output of the limit state function.

Consequently, further studies should focus on understanding the slight decrease of the sensitivity estimate quality for this particular numerical application. The combination of the algorithms developed during the thesis should also be tested on other numerical applications in high-dimensional spaces. For example, the performance of the combination of both algorithms could be addressed with the application described in Chapter 5, modeling an airfoil in inviscid transonic flow, which is closer to realistic engineering applications. Since the first algorithm has already proven to be relevant to correctly estimate the failure probability of the system for this application, the estimation of the failure probability sensitivity with respect to the design parameters as well as the distribution parameters could be of interest to examine the performance of the second algorithm. The combination of both algorithms should also be tested in other high-dimensional elliptical spaces than the standard normal space.

Appendix A

MCMC algorithms

A.1 Original Metropolis algorithm

Let $\boldsymbol{\theta}^{(1)}$ be a current sample (one of the *seeds*) and let $p^*(\boldsymbol{\zeta}|\boldsymbol{\theta})$, called the proposal pdf, be a d-dimensional pdf for $\boldsymbol{\zeta}$ centered at $\boldsymbol{\theta}$ with the symmetry property $p^*(\boldsymbol{\zeta}|\boldsymbol{\theta}) = p^*(\boldsymbol{\theta}|\boldsymbol{\zeta})$. The proposal pdf depends on a spread parameter w. Generate a sequence of observations $(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots)$ starting from the given sample $\boldsymbol{\theta}^{(1)}$ by computing $\boldsymbol{\theta}^{(k+1)}$ from $\boldsymbol{\theta}^{(k)}$ as follows

1. Generation of a candidate state $\bar{\boldsymbol{\theta}}$. Simulate $\boldsymbol{\zeta}$ according to $p^*\left(\cdot|\boldsymbol{\theta}^{(k)}\right)$ and compute the ratio $r_t = f_{\mathbf{X}}(\boldsymbol{\zeta})/f_{\mathbf{X}}\left(\boldsymbol{\theta}^{(k)}\right)$. Set $\bar{\boldsymbol{\theta}} = \boldsymbol{\zeta}$ with probability $\min(1, r_t)$ or set $\bar{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(k)}$ otherwise.

2. Acceptation/Rejection of $\bar{\boldsymbol{\theta}}$. If $\bar{\boldsymbol{\theta}} \in F_i$ accept it as the next step thus $\boldsymbol{\theta}^{(k+1)} = \bar{\boldsymbol{\theta}}$. Otherwise reject it and take the current sample as the next sample, thus $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)}$.

A.2 Modified Metropolis algorithm

Let $\boldsymbol{\theta}^{(1)}$ be a current sample (one of the *seeds*) and for every $j = 1, \ldots, d$ let $p_j^*(\zeta_j|\theta_j)$, called the proposal pdf, be a univariate pdf for ζ_j centered at θ_j with the symmetry property $p_j^*(\zeta_j|\theta_j) = p_j^*(\theta_j|\zeta_j)$. Each proposal pdf depends on a spread parameter w_j. Generate a sequence of observations $(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots)$ starting from the given sample $\boldsymbol{\theta}^{(1)}$ by computing $\boldsymbol{\theta}^{(k+1)}$ from $\boldsymbol{\theta}^{(k)}$ as follows

1. Generation of a candidate state $\bar{\boldsymbol{\theta}}$. For each component $j = 1, \ldots, d$, simulate ζ_j according to $p_j^*\left(\cdot|\theta_j^{(k)}\right)$ and compute the ratio $r_{t,i} = f_{X_j}(\zeta_j)/f_{X_j}\left(\theta_j^{(k)}\right)$. Set $\bar{\theta}_j = \zeta_j$ with probability $\min(1, r_{t,i})$ or set $\bar{\theta}_j = \theta_j^{(k)}$ otherwise.

The second step of the algorithm is the same as the step 2. of the original Metropolis algorithm. For a large d, the probability that the next state in the Markov Chain will be equal to the current state is thus very small; it would mean that all the d components ζ_j are rejected candidate state components, which is highly improbable.

A.3 MCMC algorithm developed by Katafygiotis et al.

Let $\boldsymbol{\theta}^{(1)}$ be a current sample (one of the *seeds*). Generate a sequence of observations $(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots)$ starting from the given sample $\boldsymbol{\theta}^{(1)}$ by computing $\boldsymbol{\theta}^{(k+1)}$ from $\boldsymbol{\theta}^{(k)}$ as follows. Using the stochastic representation of the current state $\boldsymbol{\theta}^{(k)} = r\mathbf{t}$

1. Generation of the first candidate state. Generate a candidate state with same radius as the current state but a different direction $\check{\boldsymbol{\theta}} = r\bar{\mathbf{t}}$.

2. Generation of the second candidate state. Starting from $\check{\boldsymbol{\theta}}$, generate another candidate state with the same direction but a changed radius $\bar{\boldsymbol{\theta}} = \bar{r}\bar{\boldsymbol{t}}$.

This final candidate $\bar{\boldsymbol{\theta}}$ is accepted with a probability of one as it lays in the failure domain thus $\boldsymbol{\theta}^{(k+1)} = \bar{\boldsymbol{\theta}}$, see [Katafygiotis and Cheung, 2007] for the full description of the algorithm

A.4 Statistical properties of the Metropolis algorithms

Here we derive some properties of the Markov chains obtained with both the original and the modified Metropolis algorithms.

Both Metropolis algorithms fail to generate iid observations. Indeed, the algorithms introduce dependence between the observations as the next state $\boldsymbol{\theta}^{(k+1)}$ is located on the vicinity of the current state $\boldsymbol{\theta}^{(k)}$. The dependence is controlled with the spread parameter of the proposal (univariate or multivariate) pdfs. A small spread allows a high acceptance rate at step 2, as there are higher chances that the next state is also in the intermediate failure domain, but it induces a high level of dependence between the observations. A large spread may reduce the acceptance rate of step 2 and thus may increase the number of repeated Markov chain observations which slows down the convergence of the SS algorithm and induces even higher levels of dependence in the chain. A trade-off between the acceptance rate and the dependence must then be studied in order to set the spread parameter (see [Bourinet, 2018] for practical values of the spread).

Because of the dependence, the convergence rate of the intermediate failure probability estimate is reduced compared to the case where the samples would be totally independent. However, this dependence amongst the sample $\left(\mathbf{X}_{i+1}^{(j)}\right)_{j=1,\ldots,N}$ does not induce bias in the estimation of the intermediate failure probability estimate $\mathbb{P}(F_{i+1}|F_i)$ [Au and Beck, 2001].

Appendix B

MSE for the IS estimate in the score function method

Using the same arguments as in [Torii and Novotny, 2021] we want to derive an upper bound for $\mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}}$ written

$$\mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}} = \frac{1}{N}\left(\mathbb{E}_h\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\, Q_h(\mathbf{Z})^2\right] - \left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial \theta_\ell}\right)^2\right),$$

where $Q_h(\cdot) = (f_{\mathbf{Z}}/h) \times (\partial \ln(f_{\mathbf{Z}}(\cdot;\boldsymbol{\theta}))/\partial\theta_\ell)$. We notice that

$$\mathbb{E}_h\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\, Q_h(\mathbf{Z})^2\right] = \mathbb{E}_h\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}Q(\mathbf{Z})\right)^2\right],$$

where $Q(\cdot) = \partial \ln\left(f_{\mathbf{Z}}(\cdot;\boldsymbol{\theta})\right)/\partial\theta_\ell$. Using the Cauchy-Schwarz inequality, we have

$$\left|\mathbb{E}_h\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}Q(\mathbf{Z})\right)^2\right]\right| \leq \sqrt{\mathbb{E}_h\left[\left(\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\right)^2 \frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}\right]\mathbb{E}_h\left[\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}\right)^3 Q(\mathbf{Z})^4\right]},$$

since $\mathbb{I}_{D_{f_{\mathbf{Z}}}}$ and $\left(\frac{f_{\mathbf{Z}}}{h}Q\right)^2$ are non-negative we can write

$$\begin{aligned}\mathbb{E}_h\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}Q(\mathbf{Z})\right)^2\right] &\leq \sqrt{\mathbb{E}_h\left[\mathbb{I}_{D_{f_{\mathbf{Z}}}}(\mathbf{Z})\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}\right]\mathbb{E}_h\left[\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}\right)^3 Q(\mathbf{Z})^4\right]} \\ &\leq \sqrt{P_f\,\mathbb{E}_h\left[\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}\right)^3 Q(\mathbf{Z})^4\right]}.\end{aligned}$$

Therefore, we have the following upper bound for $\mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}}$

$$\mathrm{E}_{\mathrm{mse}}^{\mathrm{IS}} \leq \frac{1}{N}\left(\sqrt{P_f}\sqrt{\mathbb{E}_h\left[\left(\frac{f_{\mathbf{Z}}(\mathbf{Z})}{h(\mathbf{Z})}\right)^3 Q(\mathbf{Z})^4\right]} - \left(\frac{\partial P_f(\boldsymbol{\theta})}{\partial\theta_\ell}\right)^2\right).$$

Appendix C

Discussion on the optimizations in the important ring

In the case of high-dimensional standard normal inputs, the representative failing points could be searched directly in the hypersphere of radius $\sqrt{d}$ rather than LB in Eq. (5.2). The first optimization of Eq. (5.1) could be skipped to save lsf evaluations. Indeed, since the design points in this space most likely do not belong to the important ring [Katafygiotis and Zuev, 2008], and since the hypersphere of radius $\sqrt{d}$ maximizes the density f_R of the radial component of $\mathbf{X}$, this is where one should look for important directions, especially for nonlinear limit state functions. Nevertheless, we explain here why both optimizations were kept in the proposed method.

As the dimension increases, the difference between LB and $\sqrt{d}$ decreases as the variance of R converges to 1/2; cf Chapter 2 Section 2.2.2.1, therefore this difference is negligible in very high dimensions. Furthermore, once the first optimization of Eq. (5.1) is performed, the failing point P^* is located in both D_f and the important ring, as it is located in the lower bound of the important ring. Consequently, performing the second optimization in the hypersphere of radius LB makes sense: P^* is the starting point of this optimization and the result $\tilde{P}^*$ necessarily belongs to both D_f and the important ring as well, since such failing points exist. If the optimization of Eq. (5.2) is directly performed in the hypersphere of radius $\sqrt{d}$, there is theoretically no proof that the optimization will converge to a point in D_f. There could be particular settings of the failure domain D_f, where no failing point is located in the hypersphere of radius $\sqrt{d}$. Moreover, in smaller dimensions, $d < 50$, the important ring encompasses the whole space and the search for important failing directions in the hypersphere of radius $\sqrt{d}$ would not give any result (in the 4-branch system of dimension 2 for instance, there is no failing point before the hypersphere of radius 3). Finally, the importance of the hypersphere of radius $\sqrt{d}$ is valid for standard normal inputs but not for the other standard elliptical distributions; see Chapter 2 Section 2.2.2.1.

As we try to propose a general method fit for all standard elliptical distributions, that performs particularly well in high-dimensional spaces without being irrelevant in small and moderate dimensions, the optimization of Eq. (5.2) is thus performed on the lower bound of the important ring, if the first optimization of Eq. (5.1) gave a failing point of norm LB. For nonlinear limit state functions, decreasing the width of the important interval will most likely give better important directions. However, if the important ring is too narrow, some failure regions could be missed, that are not necessarily negligible, especially if the failure probability is very low.

Appendix D

Calculus for the new sensitivity estimate

D.1 Expression of the derivative in the Taylor series expansion

We derive here the expression of $T(0)$, the derivative of the function $z \to F_{H_\mathbf{s}}(-z)$ with respect to s_ℓ, evaluated in 0. Recall that $\forall z \in \mathbb{R}$, $F_{H_\mathbf{s}}(z) = \mathbb{P}(H_\mathbf{s} \leq z)$, therefore

$$\forall z \in \mathbb{R} \qquad F_{H_\mathbf{s}}(-z) = \int_{-\infty}^{-z} f_{H_\mathbf{s}}(h)\mathrm{d}h,$$

where $f_{H_\mathbf{s}}$ is the unknown density of the univariate random variable $H_\mathbf{s}$. Taking the derivatives of the above equation with respect to s_ℓ gives

$$\forall z \in \mathbb{R} \qquad T(z) = \frac{\partial F_{H_\mathbf{s}}(-z)}{\partial s_\ell} = \int_{-\infty}^{-z} \frac{\partial f_{H_\mathbf{s}}(h)}{\partial s_\ell}\mathrm{d}h \qquad \text{thus} \qquad T(0) = \int_{-\infty}^{0} \frac{\partial f_{H_\mathbf{s}}(h)}{\partial s_\ell}\mathrm{d}h.$$

Now recall that $P_f(\mathbf{s}) = \mathbb{P}(g(\mathbf{s}, \mathbf{X}) \leq 0) = \mathbb{P}(H_\mathbf{s} \leq 0)$, thus

$$P_f(\mathbf{s}) = \int_{-\infty}^{0} f_{H_\mathbf{s}}(h)\mathrm{d}h.$$

Taking the derivatives of the above equation with respect to s_ℓ gives

$$\frac{\partial P_f(\mathbf{s})}{\partial s_\ell} = \int_{-\infty}^{0} \frac{\partial f_{H_\mathbf{s}}(h)}{\partial s_\ell}\mathrm{d}h.$$

Consequently, one has the following equality

$$T(0) = \frac{\partial P_f(\mathbf{s})}{\partial s_\ell}.$$

D.2 Expression of the failure probability derivative functions

We derive here the expression of the failure probability derivative functions $\sigma \to \partial\overline{P_f(\mathbf{s}, \sigma)}/\partial a$ and $\sigma \to \partial\overline{P_f(\mathbf{s}, \sigma)}/\partial b$ defined on $\mathbb{R}^+\backslash\{0\}$. The derivative with respect to b is first computed, with ϕ the pdf of a univariate standard normal variable

$$\begin{aligned}
\frac{\partial\overline{P_f(\mathbf{s}, \sigma)}}{\partial b} &= \int_{\mathbb{R}} -\frac{1}{\sigma}\phi\left(-\frac{ax+b}{\sigma}\right)\phi(x)\mathrm{d}x = -\frac{1}{\sigma}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(ax+b)^2}{2\sigma^2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x \\
&= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{a^2x^2}{\sigma^2} + 2\frac{abx}{\sigma^2} + \frac{b^2}{\sigma^2} + x^2\right)\right)\mathrm{d}x \\
&= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(x^2\left(\frac{a^2}{\sigma^2}+1\right) + 2\frac{abx}{\sigma^2} + \frac{b^2}{\sigma^2}\right)\right)\mathrm{d}x \\
&= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(x^2U^2 + 2xUV + V^2 - V^2 + \frac{b^2}{\sigma^2}\right)\right)\mathrm{d}x,
\end{aligned}$$

where $U = \sqrt{\frac{a^2}{\sigma^2} + 1}$ and $V = \frac{ab}{U\sigma^2} = \frac{ab}{\sigma\sqrt{a^2+\sigma^2}}$. Therefore

$$\frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial b} = -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right)\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(Ux+V\right)^2\right)\mathrm{d}x.$$

With the change of variable $y = Ux$, which is bijective on $\mathbb{R}$ as U is strictly positive, one gets

$$\begin{aligned}\frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial b} &= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right)\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y+V\right)^2\right)\frac{\mathrm{d}y}{U}\\ &= -\frac{1}{U\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right),\end{aligned}$$

where the last equality comes from the fact that the integral represents a density of a normal random variable of variance 1 and mean $-V$ integrated over the whole domain of definition, thus it is equal to 1. Finally one has

$$\forall\sigma \in \mathbb{R}^+\backslash\{0\} \qquad \frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial b} = -\frac{1}{\sqrt{a^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2+\sigma^2}\right)\right).$$

The derivative with respect to a is obtained as follows, with the same transformation than above

$$\begin{aligned}\frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial a} &= \int_{\mathbb{R}} -\frac{x}{\sigma}\phi\left(-\frac{ax+b}{\sigma}\right)\phi(x)\mathrm{d}x\\ &= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right)\int_{\mathbb{R}}\frac{x}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(Ux+V\right)^2\right)\mathrm{d}x.\end{aligned}$$

Therefore, with the same change of variable $y = Ux$, one has

$$\begin{aligned}\frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial a} &= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right)\int_{\mathbb{R}}\frac{y}{U\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y+V\right)^2\right)\frac{\mathrm{d}x}{U}\\ &= -\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right)\frac{\mathbb{E}[\mathcal{N}(-V,1)]}{U^2}\\ &= \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{\sigma^2} - V^2\right)\right)\frac{V}{U^2}\\ &= \frac{1}{\sqrt{a^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2+\sigma^2}\right)\right)\frac{V}{U} = \frac{1}{\sqrt{a^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2+\sigma^2}\right)\right)\frac{ab}{a^2+\sigma^2}.\end{aligned}$$

Consequently, the derivative is

$$\forall\sigma \in \mathbb{R}^+\backslash\{0\} \qquad \frac{\partial \overline{P_f(\mathbf{s},\sigma)}}{\partial a} = \frac{1}{\sqrt{a^2+\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{b^2}{a^2+\sigma^2}\right)\right)\frac{ab}{a^2+\sigma^2}.$$

D.3 Expression of the other Taylor coefficients of the toy example

We derive here the computation of the Taylor coefficients of order 2 and 4 for parameters a and b through the derivatives of the T function. It is recalled that $\forall z \in \mathbb{R}$, $F_{H_\mathbf{s}}(-z) = \Psi\left(\frac{-z-b}{a}\right)$ where Ψ is the standard normal cdf. The functions T_a and T_b are fist defined as

$$\begin{aligned}T_a(z) &= \frac{\partial F_{H_\mathbf{s}}(-z)}{\partial a} = -\frac{-z-b}{a^2}\phi\left(\frac{-z-b}{a}\right) = \frac{z+b}{a^2}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(z+b)^2}{2a^2}\right)\\ T_b(z) &= \frac{\partial F_{H_\mathbf{s}}(-z)}{\partial b} = -\frac{1}{a}\phi\left(\frac{-z-b}{a}\right) = -\frac{1}{a}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(z+b)^2}{2a^2}\right),\end{aligned}$$

where ϕ is the standard normal pdf. Therefore, one has

$$\begin{aligned}
T_a'(z) &= \frac{1}{a^2}\phi\left(\frac{z+b}{a}\right)\left(1-\frac{(z+b)^2}{a^2}\right)\\
T_a''(z) &= \frac{1}{a^2}\phi\left(\frac{z+b}{a}\right)\left(-2\frac{(z-b)}{a^2}-\frac{(z+b)}{a^2}\left(1-\frac{(z+b)^2}{a^2}\right)\right)\\
T_a''(z) &= \frac{1}{a^2}\phi\left(\frac{z+b}{a}\right)\left(-3\frac{(z+b)}{a^2}+\frac{(z+b)^3}{a^4}\right).
\end{aligned}$$

The evaluation in 0 results in

$$T_a''(0) = \frac{1}{a^2}\phi\left(\frac{b}{a}\right)\left(-3\frac{b}{a^2}+\frac{b^3}{a^4}\right) = \frac{b}{a^4}\phi\left(\frac{b}{a}\right)\left(-3+\frac{b^2}{a^2}\right).$$

Now for the coefficient of order 4,

$$\begin{aligned}
T_a^{(3)}(z) &= \frac{1}{a^2}\phi\left(\frac{z+b}{a}\right)\left(-\frac{3}{a^2}+3\frac{(z+b)^2}{a^4}-\frac{(z+b)}{a^2}\left(-3\frac{(z+b)}{a^2}+\frac{(z+b)^3}{a^4}\right)\right)\\
T_a^{(3)}(z) &= \frac{1}{a^4}\phi\left(\frac{z+b}{a}\right)\left(-3+6\frac{(z+b)^2}{a^2}-\frac{(z+b)^4}{a^4}\right)\\
T_a^{(4)}(z) &= \frac{1}{a^4}\phi\left(\frac{z+b}{a}\right)\left(12\frac{(z+b)}{a^2}-4\frac{(z+b)^3}{a^4}-\frac{(z+b)}{a^2}\left(-3+6\frac{(z+b)^2}{a^2}-\frac{(z+b)^4}{a^4}\right)\right)\\
T_a^{(4)}(z) &= \frac{1}{a^6}\phi\left(\frac{z+b}{a}\right)\left(15(z+b)-10\frac{(z+b)^3}{a^2}+\frac{(z+b)^5}{a^4}\right).
\end{aligned}$$

The evaluation in 0 results in

$$T_a^{(4)}(0) = \frac{b}{a^6}\phi\left(\frac{b}{a}\right)\left(15-10\frac{b^2}{a^2}+\frac{b^4}{a^4}\right).$$

For the coefficient of the derivative with respect to b, one has

$$\begin{aligned}
T_b'(z) &= \frac{z+b}{a^3}\phi\left(\frac{z+b}{a}\right)\\
T_b''(z) &= \frac{1}{a^3}\phi\left(\frac{z+b}{a}\right)\left(1-\frac{(z+b)^2}{a^2}\right).
\end{aligned}$$

Therefore the evaluation in 0 gives

$$T_b''(0) = \frac{1}{a^3}\phi\left(\frac{b}{a}\right)\left(1-\frac{b^2}{a^2}\right)$$

and for the coefficient of order 4

$$\begin{aligned}
T_b^{(3)}(z) &= \frac{1}{a^3}\phi\left(\frac{z+b}{a}\right)\left(-2\frac{(z+b)}{a^2}-\frac{(z+b)}{a^2}\left(1-\frac{(z+b)^2}{a^2}\right)\right)\\
T_b^{(3)}(z) &= \frac{1}{a^3}\phi\left(\frac{z+b}{a}\right)\left(-3\frac{(z+b)}{a^2}+\frac{(z+b)^3}{a^4}\right) = \frac{1}{a^5}\phi\left(\frac{z+b}{a}\right)\left(-3(z+b)+\frac{(z+b)^3}{a^2}\right)\\
T_b^{(4)}(z) &= \frac{1}{a^5}\phi\left(\frac{z+b}{a}\right)\left(-3+3\frac{(z+b)^2}{a^2}-\frac{(z+b)}{a^2}\left(-3(z+b)+\frac{(z+b)^3}{a^2}\right)\right)\\
T_b^{(4)}(z) &= \frac{1}{a^5}\phi\left(\frac{z+b}{a}\right)\left(-3+6\frac{(z+b)^2}{a^2}-\frac{(z+b)^4}{a^4}\right).
\end{aligned}$$

Finally

$$T_b^{(4)}(0) = \frac{1}{a^5}\phi\left(\frac{b}{a}\right)\left(-3+6\frac{b^2}{a^2}-\frac{b^4}{a^4}\right).$$

Appendix E

Polynomial regressions in the OLS framework

E.1 Introduction of the Ordinary Least Squares (OLS) method

We briefly derive here the OLS framework. In the OLS framework, $\mathbb{E}_{f_{\mathbf{X}}}[\boldsymbol{\epsilon}] = \mathbf{0}$ and the covariance matrix $\Sigma_{\boldsymbol{\epsilon}}$ is diagonal with constant components: $\Sigma_{\boldsymbol{\epsilon}} = \eta^2 \mathbf{I}_m$ where η^2 is the constant variance of the error ϵ_l, for $l = 1, \ldots, m$ and $\mathbf{I}_m$ is the identity matrix of dimension m. The OLS estimate is written

$$\widehat{\boldsymbol{\alpha}}_O = \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \left(\mathbf{S}^\top \mathbf{V}\right),$$

and it is the best linear unbiased estimate according to the Gauss–Markov theorem, as it has the lowest variance within the class of linear unbiased estimate [Theil, 1971]. The covariance matrix estimate of $\widehat{\boldsymbol{\alpha}}_O$ is given by the following formula

$$\hat{\mathbb{V}}(\hat{\beta}_{OLS}) = \eta^2 \left(\mathbf{S}^\top \mathbf{S}\right)^{-1}.$$

The OLS framework is suited for homoscedastic uncorrelated data.

E.2 Toy example with various MC in the OLS framework

The results of Tables E.1 and E.2 are obtained with the same simulation samples used for Table 6.1 and Table 6.2, but the sensitivity estimates are computed in the OLS framework.

Table E.1: Comparison of the result of the polynomial regression with 3 different degrees, for $N = 100000$. The failure probability is equal to 6.22×10^{-3} with an empirical CV of 3.8%. The reference values of the sensitivities are $\partial P_f / \partial a = 2.19 \times 10^{-2}$ and $\partial P_f / \partial b = -8.76 \times 10^{-3}$.

Toy example with OLS								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial a}}$	2.20×10^{-2}		2.20×10^{-2}		2.20×10^{-2}		2.20×10^{-2}	
	CV	6.0% (5.9%)	CV	6.9% (6.6%)	CV	7.4% (7.3%)	CV	7.2%
$\widehat{\frac{\partial P_f}{\partial b}}$	-8.78×10^{-3}		-8.79×10^{-3}		-8.79×10^{-3}		-8.82×10^{-3}	
	CV	5.9% (5.9%)	CV	6.9% (6.5%)	CV	7.3% (7.2%)	CV	7.1%

The theoretical CV estimates have an empirical CV below 14% for $N = 100000$, below 15% for $N = 50000$ and close to 20% for $N = 10000$ and $2k = 2$.

Table E.2: Comparison of the result of the polynomial regression, for $N = 50000$ and $N = 10000$. The failure probability is equal to 6.19×10^{-3} with an empirical CV of 5.7% for $N = 50000$. The failure probability is equal to 6.23×10^{-3} with an empirical CV of 13.3% for $N = 10000$. The reference values of the sensitivities are $\partial P_f / \partial a = 2.19 \times 10^{-2}$ and $\partial P_f / \partial b = -8.76 \times 10^{-3}$.

Toy example with OLS								
	$N = 50000$				$N = 10000$			
	Regression $2k = 2$		Weak approach $\tilde{\sigma} = \sigma_{\min}$		Regression $2k = 2$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial a}}$	2.20×10^{-2}		2.21×10^{-2}		2.20×10^{-2}		2.24×10^{-2}	
	CV	7.2% (6.9%)	CV	7.8%	CV	13.3% (12.9%)	CV	11.3%
$\widehat{\frac{\partial P_f}{\partial b}}$	-8.77×10^{-3}		-8.85×10^{-3}		-8.76×10^{-3}		-9.12×10^{-3}	
	CV	7.2% (6.9%)	CV	7.7%	CV	13.3% (13.1%)	CV	10.8%

E.3 Toy example with NAIS in the OLS framework

The results in Table E.3 are obtained with the same simulation samples used for Table 6.3, but the sensitivity estimates are computed in the OLS framework.

Table E.3: Comparison of the result of the polynomial regression with NAIS, for $N = 2000$. The failure probability is equal to 6.20×10^{-3} with an empirical CV of 4.8%. The reference values of the sensitivities are $\partial P_f / \partial a = 2.19 \times 10^{-2}$ and $\partial P_f / \partial b = -8.76 \times 10^{-3}$.

Toy example with NAIS and OLS								
	Regression $2k = 2$		Regression $2k = 4$		Regression $2k = 6$		Weak approach $\tilde{\sigma} = \sigma_{\min}$	
$\widehat{\frac{\partial P_f}{\partial a}}$	2.20×10^{-2}		2.19×10^{-2}		2.18×10^{-2}		2.18×10^{-2}	
	CV	5.5% (6.6%)	CV	6.2% (7.3%)	CV	6.7% (7.9%)	CV	7.0%
$\widehat{\frac{\partial P_f}{\partial b}}$	-8.82×10^{-3}		-8.74×10^{-3}		-8.73×10^{-3}		-8.74×10^{-3}	
	CV	5.7% (6.7%)	CV	6.3% (7.5%)	CV	6.8% (8.1%)	CV	7.0%

List of Figures

2.1 The von Mises–Fisher distribution on the unit circle $\mathcal{T}^2$ with $\boldsymbol{\nu} = [0.71, 0.71]$ and a) $\kappa = 10$ b) $\kappa = 75$. The blue line represents the mean direction $\boldsymbol{\nu}$. The red dots represent a sample generated with the chosen κ. 26

2.2 The density function of the radial component R when a) $d = 250$ b) $d = 500$. When $d = 250$, $\mathbb{E}[R] = 15.7956$, while $\sqrt{250} = 15.8114$, and $\mathbb{V}\mathrm{ar}(R) = 0.4995$ 28

2.3 The density function of the radial component $R \sim \sqrt{d \times \mathcal{F}(d, \nu)}$ of a Student distribution for different values of d and ν and of a Cauchy distribution for different values of d. 29

2.4 Illustration on a two-dimensional example of the map of the input space, split in the failure domain D_f and the safety domain D_s by the limit state surface in blue. The dotted lines represent isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g. 30

2.5 Illustration of the location of the two failure regions on an example in dimension 2 a) where the failure regions are disjoint b) where the failure regions are joint. The blue line is the limit state surface. The dotted lines represent isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g. 31

3.1 Illustration on a two-dimensional example of the design point of the failure domain D_f. The dark blue line is the limit state surface. The dotted lines represent the isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g. 41

3.2 Illustration on a two-dimensional example of the FORM and SORM approximations of the limit state surface. The dark blue line is the limit state surface. The dotted lines represent the isovalues of the joint pdf of the inputs. The color map represents the value of the limit state function g. 43

3.3 Illustration of the search for the design points on an example in dimension 2 with four failure regions a) the original limit state function, b) the first design point found P_1^*, c) the second design point found P_2^*, d) the third design point found P_3^*, e) the fourth design point found P_4^*, f) the final modified limit state function; as the optimization finds a spurious design point, the procedure stops. The dark blue line is the limit state surface. The color map represents the value of the limit state function $\tilde{g}$. The red cross represents the starting point $\mathbf{x}_0$ of the optimization. The dotted lines represent the isovalues of the joint pdf of the inputs. 44

3.4 Illustration of the main steps of the SS algorithm with $p_0 = 0.1$ a) the first threshold $s_0 = 1.4$ is computed from observations generated from $f_{\mathbf{X}}$, b) the second threshold $s_1 = 0.5$ is computed from observations resulting from the MCMC algorithm, c) the third threshold s_2 computed from observations resulting from the MCMC algorithm is smaller than 0 thus it is set back to 0, the loop stops and the failure probability is estimated with the observations in the failure domain. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dashed white lines represent the intermediate limit state surfaces. The dotted lines represent the isovalues of the joint pdf of the inputs $f_{\mathbf{X}}$. The magenta dots represent the observations located in the intermediate failure domains at each step. 48

3.5 Illustration of the main steps of the NAIS algorithm with $\rho = 0.1$ a) the initial threshold $s_0 = 1.4$ is computed from observations generated from $f_{\mathbf{X}}$, b) the next ρ-quantile computed from observations generated from $\hat{h}_1$ is already smaller than 0 thus the loop stops as $s_1 = 0$ and the failure probability is estimated with the observations in the failure domain. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dashed white lines represent the intermediate limit state surfaces. The dotted lines represent the isovalues of the densities $f_{\mathbf{X}}$ and $\hat{h}_1$. The magenta dots represent the observations of non-zero weight at each step. 53

3.6 Illustration of the main steps of the CE-AIS algorithm with GM, with $K = 4$ and $\rho = 0.1$ a) the initial threshold $s_0 = 1.4$ is computed from observations generated from $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_0) = f_{\mathbf{X}}$, b) the second threshold $s_1 = 0.5$ computed from observations generated from the mixture of 4 Gaussian densities $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_1)$ c) the third ρ-quantile is already smaller than 0 thus the loop stops as $s_2 = 0$ and the failure probability is estimated with the observations in the failure domain generated from the mixture of 4 Gaussian densities $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_2)$. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dashed white lines represent the intermediate limit state surfaces. The dotted lines represent the isovalues of the densities $f_{\mathbf{X}}$, $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_1)$ and $h_{\mathrm{GM}}(\cdot\,;\boldsymbol{\Theta}_2)$. The magenta dots represent the observations of non-zero weight at each step. 57

3.7 Illustration of the main steps of the iCE-vMFNM algorithm, with $K = 4$ and $\delta = 4$ a) the initial CV: $\mathrm{CV}_0 = 28.9$ is computed from observations generated from $h_{\mathrm{vMFNM}}(\cdot\,;\boldsymbol{\Theta}_0) = f_{\mathbf{X}}$ b) the second CV computed from the observations generated from the mixture density $h_{\mathrm{vMFNM}}(\cdot\,;\boldsymbol{\Theta}_1)$ is already smaller than δ thus the loop stops c) the failure probability is estimated with the observations in the failure domain generated from $h_{\mathrm{vMFNM}}(\cdot\,;\boldsymbol{\Theta}_1)$. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dotted lines represent the isovalues of the densities $f_{\mathbf{X}}$ and $h_{\mathrm{vMFNM}}(\cdot\,;\boldsymbol{\Theta}_1)$. The magenta dots represent the observations generated at each step in a) and b) and only the observations lying in the failure domain in c). 62

3.8 Illustration on a two-dimensional example of the DS main idea. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The colored lines represent different directions $\mathbf{T}^{(j)}$, the blue dots represent the intersection between the directions and the limit state surface. The associated roots r_j are the distance between the blue dots and the origin for each $\mathbf{T}^{(j)}$. 64

3.9 Illustration on a two-dimensional example of the RBIS main idea a) the observations are generated from $f_{R|R\geq\beta}$ where $\beta = \|P^*\|$ b) the failure probability is estimated with the observations in the failure domain generated from $f_{R|R\geq\beta}$. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dotted lines represent the isovalues of the inputs density $f_{\mathbf{X}}$. The magenta dots represent observations generated from $f_{R|R\geq\beta}$ in a) and only those belonging to the failure domain in b). 66

3.10 Illustration on a two-dimensional example of the LS main idea. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The black arrows are the axes of the rotated space $\mathbf{V}$. The dotted lines represent the isovalues of the standard normal inputs density $f_{\mathbf{X}}$. 67

3.11 Illustration on a two-dimensional example of the IS auxiliary mixture built with design points a) the observations are generated from h of Eq. (3.27), b) the failure probability is estimated with the observations in the failure domain generated from h. The dark blue line is the limit state surface. The color map represents the value of the limit state function g. The dotted lines represent the isovalues of the standard normal mixture density h. The magenta dots represent observations generated from h in a) and only those in the failure domain in b). 69

4.1 Approximation of the indicator function $\mathbb{I}_{y>0}$ with the function $y \to \Psi(y/\sigma)$ for three different values of σ, where Ψ is the cdf of a standard normal univariate variable. 87

4.2 Illustration of the relation between $\widetilde{E}^{\mathrm{MC}}_{\mathrm{mse}}$ and σ in log-log scale, for a fixed sample size N, with the approximations derived in [Torii, 2020] in a one-dimensional case. 89

5.1 Scheme of the proposed method to find and model every failure region. 103

5.2 Illustration of the location of the failure region and the important ring in a) Case 1 where the closest failing point of the failure region is inside the important ring, and in b) Case 2 where the closest failing point of the failure region is outside the important ring. The red circles delineate the lower bound and the upper bound of the important ring. In both examples, there is only one failure region in the failure domain, located in the upper right corner delimited in dark blue. The color legend represents the value of the limit state function. The representation is in dimension two, however the aim of the algorithm is high-dimensional spaces. 105

5.3 Illustration of the failing point representative of the failure region P^* in a) Case 1 and $\tilde{P}^*$ in b) Case 2. The blue line represents the Euclidean distance between the representative point and the origin. The red circles delineate the lower bound and upper bound of the important ring. In both examples, there is only one failure region in the failure domain, located in the upper right corner delimited in dark blue. The color legend represents the value of the limit state function. 106

5.4 Illustration of the bulge built around the representative point P_1 of the failure region and the resulting modified failure domain $\tilde{D}_f$ in a) Case 1 and b) Case 2. This illustration should be compared to Figure 5.3 for a better understanding. The red circles represent the lower bound and upper bound of the important ring. In both examples, there is only one failure region in the failure domain, located in the upper right corner delimited in dark blue. The color legend represents the value of the modified limit state function $\tilde{g}$. 112

5.5 Limit state function for the 4-branch system. This system has four failure regions located at the four corners of the input space. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface. 115

5.6 Sampling of the first failure region of the 4-branch system. One has $P_1 = (-2.12, -2.12)$, $r_{\mathrm{opt},1} = 2.99$ and $\boldsymbol{\nu}_1 = (-0.71, -0.71)$. The population M_1 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{1,0} = 20$ and $\kappa_{1,1} = 33.5$ and is represented with pink dots. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface. 116

5.7 Sampling of a) the second and b) third failure regions of the 4-branch system. One has $P_2 = (2.12, 2.12)$, $r_{\mathrm{opt},2} = 2.99$, and $\boldsymbol{\nu}_2 = (0.71, 0.71)$. Population M_2 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{2,0} = 20$ and $\kappa_{2,1} = 27.6$. Moreover, $P_3 = (2.74, -2.47)$, $r_{\mathrm{opt},3} = 3.49$, and $\boldsymbol{\nu}_3 = (0.71, -0.71)$. Population M_3 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{3,0} = 20$ and $\kappa_{3,1} = 18.1$. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface. 117

5.8 The modified limit state function $\tilde{g}$ after the additions of Bulge_1, Bulge_2, and Bulge_3. The color legend represents the value of the modified limit state function $\tilde{g}$. The dark blue line is the modified limit state surface. 117

5.9 a) Sampling of the fourth failure region of the 4-branch system. One has $P_4 = (-2.74, 2.47)$, $r_{\mathrm{opt},4} = 3.49$, and $\boldsymbol{\nu}_4 = (-0.71, 0.71)$. Population M_4 is of size $2N$ as only two iterations were needed in the algorithm, with $\kappa_{4,0} = 20$ and $\kappa_{4,1} = 16.4$. The color legend represents the value of the limit state function g. The dark blue line is the limit state surface. b) The modified limit state function after the addition of Bulge_4 to the previous modified limit state function. The color legend represents the value of the modified limit state function $\tilde{g}$. The dark blue line is the modified limit state surface. 118

5.10 Illustration of the important interval of the $\chi(d)$ distribution for the three dimensions tested a) $d = 200$, b) $d = 300$ c) $d = 400$. The dashed line represents the value of $\sqrt{d}$, the red dotted lines are the two important interval bounds. 119

5.11 a) Evolution of the failure probability estimate with respect to κ_0, for dimensions 200, 300, and 400. b) Evolution of the number of lsf evaluations N_{samp} with respect to κ_0. c) Evolution of the number K of failure regions found with respect to κ_0. d) Evolution of variance of the failure probability estimate with respect to κ_0. 121

5.12 Illustration of the direction of the failing points P_1 and P_2 representative of the two failure regions) in dimension 100 b) in dimension 200 c) in dimension 300. 125

5.13 Illustration of the important intervals of the $\chi(100)$ distribution for the two different settings. The dashed line represents the value $\sqrt{100}$, the red dotted lines are the two important interval bounds of the first setting, the purple dotted lines are the two important interval bounds of the second setting. 127

5.14 Illustration of the important interval of the $\sqrt{250 \times \mathcal{F}(250, 4)}$ distribution. The red dotted lines are the two important interval bounds. 129

5.15 Illustration of the direction of the failing points P_1, in the space of $\mathbf{X}$ in green and the space of $\mathbf{Z}$ in brown. 130

5.16 Density of $h_R = f_{R|R>r_{\text{opt},1}}$ with $r_{\text{opt},1} = 21.19$ plotted with an histogram of the norm of the points in D_f obtained with a MC sampling. 131

5.17 a) Close view on the initial computational mesh of the NACA 0012 airfoil. b) Pressure coefficient field showing the transonic shock on the NACA 0012 airfoil. 132

5.18 a) Comparison between the optimal airfoil solution and the baseline NACA 0012 airfoil. b) Pressure coefficient field around the optimal solution. 133

5.19 Shape comparison between the optimal airfoil and the critical point found by the proposed approach: a) P_1, b) P_2, c) P_3. 135

5.20 Pressure coefficient fields at the critical point found by the proposed approach: a) P_1, b) P_2, c) P_3. 136

5.21 Pressure coefficient distributions on both sides of the 2D airfoil at the critical points found by the proposed approach: a) P_1, b) P_2, c) P_3. The critical points found by Run 1 are represented in blue, those of Run 2 are represented in red and those of Run 3 in black. 137

6.1 Illustration of the polynomials P_2 and P_4 for both of the derivatives of P_f. The real derivative value is in black dotted line and the derivative function $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial s_\ell$ is in blue. The parameters of the toy example are set as follows: $a = 2$ and $b = 5$. Figures a) and c) display the derivative with respect to a while Figures b) and d) display the derivative with respect to b. Figures a) and b) show the evolution of the different functions for a range of σ in $(0, 2)$ while figures c) and d) show the evolution for a range of σ in $(0, 1)$. The failure probability is equal to 6.21×10^{-3}, the derivative with respect to a is equal to 2.19×10^{-2} and the derivative with respect to b is equal to 8.76×10^{-3}. 153

6.2 Illustration of the variance functions for both of the mentioned indicator approximation functions Ξ_σ with the toy example $(a = 2, b = 5)$ with respect to $s_\ell = a$ and with $N = 10^6$. The normal approximation is plotted in cyan while the uniform approximation is plotted in navy blue. Figure a) shows the evolution of the different functions for a range of σ in $(0.02, 1)$ while Figure b) shows the evolution for a range of σ in $(0, 0.02)$. The derivative with respect to b gives roughly the same variance function evolution. . . 154

6.3 Illustration of the regression bounds plotted in dashed magenta lines corresponding to the values of σ such as the theoretical CV estimate of the MC estimate is below $\text{CV}_{\text{target}}$. Here the CV target is around 7.5% for the derivative with respect to a. Figure a) represents the evolution of the MC estimate according to σ in navy blue as well as the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ in clear blue. Figure b) shows the evolution of the theoretical CV estimate of the MC estimate and the value of $\text{CV}_{\text{target}}$. 159

6.4 Illustration of the polynomial regression obtained from the estimates $(V_l)_{l=1,2,3}$ displayed with navy blue crosses. The resulting polynomial $\widehat{P_2}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line. The clear blue line is the function $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$. The regression bounds are plotted in dashed magenta lines. 160

6.5 Illustration of the proposed method with the toy example for two values of sample size N. Figures a), b) and c) illustrate the method for $N = 50000$, while Figures d), e) and f) illustrate the method for $N = 10000$. Figures a) and d) represent the evolution of the MC estimate according to σ in navy blue as well as the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ in clear blue. Figures b) and e) show the evolution of the theoretical CV estimate of the MC estimate and the value of $\mathrm{CV}_{\mathrm{target}}$, close to 9% for $N = 50000$ and close to 12% for $N = 10000$. Figures c) and f) represent the estimates $(V_l)_{l=1,2,3}$ displayed with navy blue crosses obtained with bootstrap. The resulting polynomial $\widehat{P_2}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line. 161
6.6 Illustration of the proposed method combined with NAIS with the toy example for $N = 2000$. Figure a) represents the evolution of the IS estimate according to σ in navy blue as well as the evolution of $\sigma \mapsto \partial \overline{P_f(\mathbf{s}, \sigma)}/\partial a$ in clear blue. Figure b) shows the evolution of the theoretical CV estimate of the IS estimate and the value of the $\mathrm{CV}_{\mathrm{target}}$, close to 7%. Figure c) represents the estimates $(V_l)_{l=1,\ldots,4}$ displayed with navy blue crosses obtained with bootstrap. The resulting polynomial $\widehat{P_4}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line. . 164
6.7 Illustration of a cantilever beam subject to biaxial bending. 166
6.8 Illustration of a roof truss subject to random loading. 168
6.9 Illustration of the proposed method with $2k = 6$ for the roof truss application, with a NAIS algorithm of size $N = 2000$, with $\rho = 20$. Figures a), b) and c) illustrate the method with respect to the distribution parameter $s_\ell = \delta_{Z_2}$, while Figures d), e) and f) illustrate the method with respect to the distribution parameter $s_\ell = \delta_{Z_6}$. Figures a) and d) represent the evolution of the IS estimate according to σ in navy blue. Figures b) and e) show the evolution of the theoretical CV estimate of the IS estimate and the value of the $\mathrm{CV}_{\mathrm{target}}$. Figures c) and f) represent the estimates $(V_l)_{l=1,\ldots,5}$ displayed with navy blue crosses obtained with bootstrap. The resulting polynomial $\widehat{P_6}$ is displayed in green along with the failure probability sensitivity value plotted in a dotted green line. . 172

List of Tables

2.1 Distributions of R^2 and density generators of three elliptical laws: the normal distribution, the Student distribution with ν degree of freedom and the Cauchy distribution. . . 23

2.2 Cumulative distribution function of R, generator function and survival function associated to the independence copula, an example of the links between ψ, F_R and $H_{\mathbf{X}}$ for the Archimedean copulas. 24

2.3 Example of important interval definitions for the standard normal vector, according to the dimension d and the integer n such as $\mathbb{P}(R \in \mathrm{II}) \geq 1 - 10^{-n}$. 28

5.1 Comparison of the proposed algorithm with the iCE-vMFNM algorithm and with a SS algorithm for g_1. The reference probability is equal to 1.15×10^{-6} for the three dimensions considered. The symbol [$**$] indicates that fewer than 500 simulations converged. . 122

5.2 Comparison of the proposed algorithm with the iCE-vMFNM algorithm and with a SS algorithm for g_1, with $\beta = 3.5$. The reference probability is equal to 9.30×10^{-6} for the $d = 200$. 123

5.3 Comparison of the proposed algorithm with the iCE-vMFNM algorithm and with a SS algorithm for g_2. The reference probability is equal to 4.28×10^{-4} for the three dimensions considered. 126

5.4 Comparison of the proposed algorithm with the VM algorithm of [Chan and Kroese, 2012] and with a classical Monte Carlo for $\bar{g}_3$. The reference probability is equal to 8.12×10^{-3}. 130

5.5 Comparison of the proposed algorithm with the Monte Carlo method for g_4. 134

6.1 Comparison of the result of the polynomial regression with 3 different degrees, for $N = 100000$. The failure probability is equal to 6.22×10^{-3} with an empirical CV of 3.8%. The reference values of the sensitivities are $\partial P_f / \partial a = 2.19 \times 10^{-2}$ and $\partial P_f / \partial b = -8.76 \times 10^{-3}$. 160

6.2 Comparison of the result of the polynomial regression, for $N = 50000$ and $N = 10000$. The failure probability is equal to 6.19×10^{-3} with an empirical CV of 5.7% for $N = 50000$. The failure probability is equal to 6.23×10^{-3} with an empirical CV of 13.3% for $N = 10000$. The reference values of the sensitivities are $\partial P_f / \partial a = 2.19 \times 10^{-2}$ and $\partial P_f / \partial b = -8.76 \times 10^{-3}$. 162

6.3 Comparison of the result of the polynomial regression with NAIS, with 3 different degrees, for $N = 2000$. The failure probability is equal to 6.20×10^{-3} with an empirical CV of 4.8%. The reference values of the sensitivities are $\partial P_f / \partial a = 2.19 \times 10^{-2}$ and $\partial P_f / \partial b = -8.76 \times 10^{-3}$. 164

6.4 Comparison of the result of the polynomial regression with 3 different degrees, for the first lsf of the cantilever beam with NAIS. The failure probability is equal to 2.97×10^{-3} with an empirical CV of 8.4%. The reference values of the sensitivities are $\partial P_f^{(1)} / \partial w = -5.76 \times 10^{-2}$ and $\partial P_f^{(1)} / \partial t = -3.53 \times 10^{-2}$. 167

6.5 Comparison of the result of the polynomial regression with 3 different degrees, for the second lsf of the cantilever beam with iCE-SG. The failure probability is equal to 2.51×10^{-3} with an empirical CV of 4.5%. The reference values of the sensitivities are $\partial P_f^{(2)} / \partial w = -8.84 \times 10^{-4}$, $\partial P_f^{(2)} / \partial t = -2.95 \times 10^{-3}$ and $\partial P_f^{(2)} / \partial d_0 = -3.27 \times 10^{-3}$. 168

6.6 Comparison of the result of the polynomial regression with 3 different degrees, for the roof truss with NAIS. The failure probability is equal to 9.28×10^{-3} with an empirical CV of 12.5%. The reference values of the sensitivities are $(\partial P_f/\partial \mu_{Z_1}, \partial P_f/\partial \delta_{Z_1}) = (1.11 \times 10^{-5}, 1.59 \times 10^{-5})$, $(\partial P_f/\partial \mu_{Z_2}, \partial P_f/\partial \delta_{Z_2}) = (4.03 \times 10^{-2}, 1.80 \times 10^{-2})$ and $(\partial P_f/\partial \mu_{Z_3}, \partial P_f/\partial \delta_{Z_3}) = (-1.86 \times 10^{2}, 2.05 \times 10^{2})$. 170

6.7 Comparison of the result of the polynomial regression with 3 different degrees, for the roof truss with NAIS. The failure probability is equal to 9.28×10^{-3} with an empirical CV of 12.5%. The reference values of the sensitivities are $(\partial P_f/\partial \mu_{Z_4}, \partial P_f/\delta_{Z_4}) = (-2.14, 2.56)$, $(\partial P_f/\partial \mu_{Z_5}, \partial P_f/\delta_{Z_5}) = (-1.83 \times 10^{-12}, 2.00 \times 10^{-12})$ and $(\partial P_f/\partial \mu_{Z_6}, \partial P_f/\delta_{Z_6}) = (-3.77 \times 10^{-12}, 2.03 \times 10^{-12})$. 171

6.8 Comparison of the result of the polynomial regression with 3 different degrees, for the Duffing oscillator. The failure probability is equal to 4.28×10^{-4} with an empirical CV of 4.4%. The reference values of the sensitivities are $\partial P_f/\partial m = -2.45 \times 10^{-5}$ and $\partial P_f/\partial k = 6.83 \times 10^{-7}$. 173

E.1 Comparison of the result of the polynomial regression with 3 different degrees, for $N = 100000$. The failure probability is equal to 6.22×10^{-3} with an empirical CV of 3.8%. The reference values of the sensitivities are $\partial P_f/\partial a = 2.19 \times 10^{-2}$ and $\partial P_f/\partial b = -8.76 \times 10^{-3}$. 191

E.2 Comparison of the result of the polynomial regression, for $N = 50000$ and $N = 10000$. The failure probability is equal to 6.19×10^{-3} with an empirical CV of 5.7% for $N = 50000$. The failure probability is equal to 6.23×10^{-3} with an empirical CV of 13.3% for $N = 10000$. The reference values of the sensitivities are $\partial P_f/\partial a = 2.19 \times 10^{-2}$ and $\partial P_f/\partial b = -8.76 \times 10^{-3}$. 192

E.3 Comparison of the result of the polynomial regression with NAIS, for $N = 2000$. The failure probability is equal to 6.20×10^{-3} with an empirical CV of 4.8%. The reference values of the sensitivities are $\partial P_f/\partial a = 2.19 \times 10^{-2}$ and $\partial P_f/\partial b = -8.76 \times 10^{-3}$. 192

List of Algorithms

1 Subset simulation . . . 47
2 NAIS . . . 52
3 CE-AIS . . . 55
4 CE-AMIS . . . 59
5 iCE-AIS . . . 61
6 Identification Problem . . . 107
7 Sampling in all failure regions of the ordered set L . . . 113
8 Search and sampling of all failure regions of the failure domain . . . 113

Bibliography

[Ang and Tang, 2007] Ang, A. H.-S. and Tang, W. H. (2007). *Probability concepts in engineering planning and design: Emphasis on application to civil and environmental engineering.* Wiley.

[Apostol, 1967] Apostol, T. (1967). Calculus.

[Apostolakis, 1990] Apostolakis, G. (1990). The concept of probability in safety assessments of technological systems. *Science*, 250(4986):1359–1364.

[Arnold and Lynch, 1982] Arnold, S. F. and Lynch, J. (1982). On Ali's characterization of the spherical normal distribution. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(1):49–51.

[Arthur and Vassilvitskii, 2006] Arthur, D. and Vassilvitskii, S. (2006). k-means++: The advantages of careful seeding. Technical report, Stanford.

[Atkinson, 1991] Atkinson, K. (1991). *An introduction to numerical analysis.* John Wiley & Sons.

[Au and Beck, 2003] Au, S.-K. and Beck, J. (2003). Important sampling in high dimensions. *Structural safety*, 25(2):139–163.

[Au and Beck, 2001] Au, S.-K. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277.

[Augustin et al., 2014] Augustin, T., Coolen, F. P., De Cooman, G., and Troffaes, M. C. (2014). *Introduction to imprecise probabilities.* John Wiley & Sons.

[Banerjee et al., 2005] Banerjee, A., Dhillon, I. S., Ghosh, J., Sra, S., and Ridgeway, G. (2005). Clustering on the unit hypersphere using von Mises–Fisher distributions. *Journal of Machine Learning Research*, 6(9).

[Bassamboo et al., 2008] Bassamboo, A., Juneja, S., and Zeevi, A. (2008). Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation. *Operations Research*, 56(3):593–606.

[Beaurepaire et al., 2013] Beaurepaire, P., Jensen, H., Schuëller, G., and Valdebenito, M. (2013). Reliability-based optimization using bridge importance sampling. *Probabilistic Engineering Mechanics*, 34:48–57.

[Biau and Mason, 2015] Biau, G. and Mason, D. M. (2015). High-dimensional p-norms. *Mathematical Statistics and Limit Theorems: Festschrift in Honour of Paul Deheuvels*, pages 21–40.

[Bjerager, 1988] Bjerager, P. (1988). Probability integration by directional simulation. *Journal of Engineering Mechanics*, 114(8):1285–1302.

[Bourinet, 2018] Bourinet, J.-M. (2018). *Reliability analysis and optimal design under uncertainty-Focus on adaptive surrogate-based approaches.* HDR. Université Clermont Auvergne.

[Breitung, 2019] Breitung, K. (2019). The geometry of limit state function graphs and subset simulation: Counterexamples. *Reliability Engineering & System Safety*, 182:98–106.

[Breitung et al., 1984] Breitung, K. et al. (1984). Asymptotic approximations for multinormal integrals. *Journal of Engineering Mechanics*, 110(3):357–366.

[Breitung, 2006] Breitung, K. W. (2006). *Asymptotic approximations for probability integrals*. Springer.

[Bucklew, 2004] Bucklew, J. A. (2004). *Introduction to rare event simulation*, volume 5. Springer.

[Canal, 2005] Canal, L. (2005). A normal approximation for the chi-square distribution. *Computational statistics & data analysis*, 48(4):803–808.

[Cérou et al., 2012] Cérou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Statistics and computing*, 22(3):795–808.

[Cérou et al., 2019] Cérou, F., Guyader, A., and Rousset, M. (2019). Adaptive multilevel splitting: Historical perspective and recent results. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(4):043108.

[Chabridon, 2018] Chabridon, V. (2018). *Reliability-oriented sensitivity analysis under probabilistic model uncertainty–Application to aerospace systems*. PhD thesis, Université Clermont Auvergne.

[Chan and Kroese, 2012] Chan, J. C. and Kroese, D. P. (2012). Improved cross-entropy method for estimation. *Statistics and Computing*, 22(5):1031–1040.

[Cheng et al., 2023] Cheng, K., Papaioannou, I., Lu, Z., Zhang, X., and Wang, Y. (2023). Rare event estimation with sequential directional importance sampling. *Structural Safety*, 100:102291.

[Cornuet et al., 2012] Cornuet, J.-M., Marin, J.-M., Mira, A., and Robert, C. P. (2012). Adaptive multiple importance sampling. *Scandinavian Journal of Statistics*, 39(4):798–812.

[de Angelis et al., 2015] de Angelis, M., Patelli, E., and Beer, M. (2015). Advanced line sampling for efficient robust reliability analysis. *Structural safety*, 52:170–182.

[Dempster et al., 1977] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

[Der Kiureghian and Dakessian, 1998] Der Kiureghian, A. and Dakessian, T. (1998). Multiple design points in first and second-order reliability. *Structural Safety*, 20(1):37–49.

[Ditlevsen and Madsen, 1996] Ditlevsen, O. and Madsen, H. O. (1996). *Structural reliability methods*, volume 178. Wiley New York.

[Ditlevsen et al., 1990] Ditlevsen, O., Melchers, R. E., and Gluver, H. (1990). General multi-dimensional probability integration by directional simulation. *Computers & Structures*, 36(2):355–368.

[Ditlevsen et al., 1986] Ditlevsen, O., Olesen, R., and Mohr, G. (1986). Solution of a class of load combination problems by directional simulation. *Structural Safety*, 4(2):95–109.

[Dong et al., 2014] Dong, Y.-g., Lu, H.-t., and Li, L.-l. (2014). Reliability sensitivity analysis based on multi-hyperplane combination method. *Defence Technology*, 10(4):354–359.

[Douc et al., 2007] Douc, R., Guillin, A., Marin, J.-M., and Robert, C. P. (2007). Convergence of adaptive mixtures of importance sampling schemes. *The Annals of Statistics*, 35(1):420–448.

[Dubourg and Sudret, 2014] Dubourg, V. and Sudret, B. (2014). Meta-model-based importance sampling for reliability sensitivity analysis. *Structural Safety*, 49:27–36.

[Dubourg et al., 2011] Dubourg, V., Sudret, B., and Bourinet, J.-M. (2011). Reliability-based design optimization using Kriging surrogates and subset simulation. *Structural and Multidisciplinary Optimization*, 44:673–690.

[Echard et al., 2011] Echard, B., Gayton, N., and Lemaire, M. (2011). AK-MCS: An active learning reliability method combining Kriging and Monte Carlo simulation. *Structural Safety*, 33(2):145–154.

[Economon et al., 2016] Economon, T. D., Palacios, F., Copeland, S. R., Lukaczyk, T. W., and Alonso, J. J. (2016). SU2: An open-source suite for multiphysics simulation and design. *AIAA Journal*, 54(3):828–846.

[El-Laham et al., 2019] El-Laham, Y., Martino, L., Elvira, V., and Bugallo, M. F. (2019). Efficient adaptive multiple importance sampling. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5. IEEE.

[El Masri et al., 2021] El Masri, M., Morio, J., and Simatos, F. (2021). Improvement of the cross-entropy method in high dimension for failure probability estimation through a one-dimensional projection without gradient estimation. *Reliability Engineering & System Safety*, 216:107991.

[Ester et al., 1996] Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231.

[Fang et al., 2002] Fang, H.-B., Fang, K.-T., and Kotz, S. (2002). The meta-elliptical distributions with given marginals. *Journal of multivariate analysis*, 82(1):1–16.

[Fang et al., 2018] Fang, K.-T., Kotz, S., and Ng, K. W. (2018). *Symmetric Multivariate and Related Distributions*. Chapman & Hall/CRC.

[Feng, 1990] Feng, Y. (1990). The computation of failure probability for nonlinear safety margin equations. *Reliability Engineering & System Safety*, 27(3):323–331.

[Ferson and Ginzburg, 1996] Ferson, S. and Ginzburg, L. R. (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2-3):133–144.

[Fujita and Rackwitz, 1988] Fujita, M. and Rackwitz, R. (1988). Updating first-and second-order reliability estimates by importance sampling. *Doboku Gakkai Ronbunshu*, 1988(392):53–59.

[Galambos, 1978] Galambos, J. (1978). The asymptotic theory of extreme order statistics. Technical report.

[Gebru et al., 2016] Gebru, I. D., Alameda-Pineda, X., Forbes, F., and Horaud, R. (2016). EM algorithms for weighted-data clustering with application to audio-visual scene analysis. *IEEE transactions on pattern analysis and machine intelligence*, 38(12):2402–2415.

[Genest et al., 2007] Genest, C., Favre, A.-C., Béliveau, J., and Jacques, C. (2007). Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resources Research*, 43(9).

[Genest and Nešlehová, 2012] Genest, C. and Nešlehová, J. (2012). Copulas and copula models. *Encyclopedia of Environmetrics*, 2:541–553.

[Geyer et al., 2019] Geyer, S., Papaioannou, I., and Straub, D. (2019). Cross entropy-based importance sampling using Gaussian densities revisited. *Structural Safety*, 76:15–27.

[Glad et al., 2007] Glad, I., Hjort, N., and Ushakov, N. (2007). Mean-squared error of kernel estimators for finite values of the sample size. *Journal of Mathematical Sciences*, 146(4):5977–5983.

[Grimaldi and Serinaldi, 2006] Grimaldi, S. and Serinaldi, F. (2006). Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*, 29(8):1155–1167.

[Grooteman, 2008] Grooteman, F. (2008). Adaptive radial-based importance sampling method for structural reliability. *Structural Safety*, 30(6):533–542.

[Guyader, 2011] Guyader, A. (2011). *Contributions à l'estimation non paramétrique et à la simulation d'événements rares*. HDR. Université Rennes 2.

[Harbitz, 1986] Harbitz, A. (1986). An efficient sampling method for probability of failure calculation. *Structural Safety*, 3(2):109–115.

[Hasofer and Lind, 1974] Hasofer, A. M. and Lind, N. C. (1974). Exact and invariant second-moment code format. *Journal of the Engineering Mechanics division*, 100(1):111–121.

[Hilton and Feigen, 1960] Hilton, H. H. and Feigen, M. (1960). Minimum weight analysis based on structural reliability. *Journal of the Aerospace Sciences*, 27(9):641–652.

[Hohenbichler, 1984] Hohenbichler, M. (1984). An asymptotic formula for the probability of intersections. *Berichte zur Zuverlassigkeitstheorie der Bauwerke*, (69):21–48.

[Hohenbichler and Rackwitz, 1988] Hohenbichler, M. and Rackwitz, R. (1988). Improvement of second-order reliability estimates by importance sampling. *Journal of Engineering Mechanics*, 114(12):2195–2199.

[Horowitz, 2001] Horowitz, J. L. (2001). The bootstrap. In *Handbook of econometrics*, volume 5, pages 3159–3228. Elsevier.

[Iooss and Saltelli, 2017] Iooss, B. and Saltelli, A. (2017). Introduction to sensitivity analysis. In *Handbook of uncertainty quantification*, pages 1103–1122. Springer.

[Jensen et al., 2009] Jensen, H. A., Valdebenito, M. A., Schuëller, G. I., and Kusanovic, D. S. (2009). Reliability-based optimization of stochastic systems using line search. *Computer methods in applied mechanics and engineering*, 198(49-52):3915–3924.

[Jia and Taflanidis, 2014] Jia, G. and Taflanidis, A. A. (2014). Sample-based evaluation of global probabilistic sensitivity measures. *Computers & Structures*, 144:103–118.

[Joe, 1996] Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. *Lecture notes-monograph series*, pages 120–141.

[Joe, 1997] Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC press.

[Katafygiotis and Cheung, 2007] Katafygiotis, L. S. and Cheung, S. (2007). Application of spherical subset simulation method and auxiliary domain method on a benchmark reliability study. *Structural Safety*, 29(3):194–207.

[Katafygiotis and Zuev, 2008] Katafygiotis, L. S. and Zuev, K. M. (2008). Geometric insight into the challenges of solving high-dimensional reliability problems. *Probabilistic Engineering Mechanics*, 23(2-3):208–218.

[Kenway et al., 2019] Kenway, G. K., Mader, C. A., He, P., and Martins, J. R. (2019). Effective adjoint approaches for computational fluid dynamics. *Progress in Aerospace Sciences*, 110:100542.

[Khan, 2004] Khan, R. A. (2004). Approximation for the expectation of a function of the sample mean. *Statistics*, 38(2):117–122.

[Klawonn et al., 2012] Klawonn, F., Höppner, F., and Jayaram, B. (2012). What are clusters in high dimensions and are they difficult to find? In *International Workshop on Clustering High-Dimensional Data*, pages 14–33. Springer.

[Kounias, 1968] Kounias, E. G. (1968). Bounds for the probability of a union, with applications. *The Annals of Mathematical Statistics*, 39(6):2154–2158.

[Koutsourelakis et al., 2004] Koutsourelakis, P.-S., Pradlwarter, H. J., and Schueller, G. I. (2004). Reliability of structures in high dimensions, part I: algorithms and applications. *Probabilistic Engineering Mechanics*, 19(4):409–417.

[Kroese et al., 2013] Kroese, D. P., Rubinstein, R. Y., and Glynn, P. W. (2013). The cross-entropy method for estimation. In *Handbook of statistics*, volume 31, pages 19–34. Elsevier.

[Kullback and Leibler, 1951] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

[Kurtz and Song, 2013] Kurtz, N. and Song, J. (2013). Cross-entropy-based adaptive importance sampling using Gaussian mixture. *Structural Safety*, 42:35–44.

[Lacaze et al., 2015] Lacaze, S., Brevault, L., Missoum, S., and Balesdent, M. (2015). Probability of failure sensitivity with respect to decision variables. *Structural and Multidisciplinary Optimization*, 52(2):375–381.

[Lai et al., 1979] Lai, T. L., Robbins, H., and Wei, C. Z. (1979). Strong consistency of least squares estimates in multiple regression II. *Journal of multivariate analysis*, 9(3):343–361.

[Lapuyade-Lahorgue, 2016] Lapuyade-Lahorgue, J. (2016). Lois elliptiques.

[Lebrun and Dutfoy, 2009] Lebrun, R. and Dutfoy, A. (2009). A generalization of the Nataf transformation to distributions with elliptical copula. *Probabilistic Engineering Mechanics*, 24(2):172–178.

[Lemaire, 2013] Lemaire, M. (2013). *Structural reliability*. John Wiley & Sons.

[Leng et al., 2022] Leng, Y., Lu, Z.-H., Cai, C.-H., Li, C.-Q., and Zhao, Y.-G. (2022). Ring simulation: A novel simple and efficient simulation method for structural reliability analysis. *Structural Safety*, 96:102182.

[Li et al., 2019] Li, L., Papaioannou, I., and Straub, D. (2019). Global reliability sensitivity estimation based on failure samples. *Structural Safety*, 81:101871.

[Lin et al., 2020] Lin, X., Jiang, Y., Peng, S., Chen, H., Tang, J., and Li, W. (2020). An efficient Nataf transformation based probabilistic power flow for high-dimensional correlated uncertainty sources in operation. *International Journal of Electrical Power & Energy Systems*, 116:105543.

[Ling, 1965] Ling, C.-H. (1965). Representation of associative functions. *Publicationes Mathematicae Debrecen*, 12:189–212.

[Lu et al., 2008] Lu, Z., Song, S., Yue, Z., and Wang, J. (2008). Reliability sensitivity method by line sampling. *Structural Safety*, 30(6):517–532.

[MacQueen, 1967] MacQueen, J. (1967). Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability*, pages 281–297.

[Madsen et al., 2006] Madsen, H. O., Krenk, S., and Lind, N. C. (2006). *Methods of structural safety*. Courier Corporation.

[Marin et al., 2012] Marin, J.-M., Pudlo, P., and Sedki, M. (2012). Consistency of the adaptive multiple importance sampling. *arXiv preprint arXiv:1211.2548*.

[Martinez et al., 2018] Martinez, E. N., Cutululis, N., and Sørensen, P. (2018). High dimensional dependence in power systems: A review. *Renewable and Sustainable Energy Reviews*, 94:197–213.

[McNeil and Nešlehová, 2009] McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *The Annals of Statistics*, 37(5B):3059–3097.

[Melchers and Ahammed, 2004] Melchers, R. and Ahammed, M. (2004). A fast approximate method for parameter sensitivity estimation in Monte Carlo structural reliability. *Computers & Structures*, 82(1):55–61.

[Metropolis et al., 1953] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.

[Morio, 2011] Morio, J. (2011). Influence of input PDF parameters of a model on a failure probability estimation. *Simulation Modelling Practice and Theory*, 19(10):2244–2255.

[Morio et al., 2014] Morio, J., Balesdent, M., Jacquemart, D., and Vergé, C. (2014). A survey of rare event simulation methods for static input–output models. *Simulation Modelling Practice and Theory*, 49:287–304.

[Moustapha and Sudret, 2019] Moustapha, M. and Sudret, B. (2019). Surrogate-assisted reliability-based design optimization: a survey and a unified modular framework. *Structural and Multidisciplinary Optimization*, 60:2157–2176.

[Owen, 2013] Owen, A. B. (2013). Monte Carlo Theory, Methods and Examples.

[Papaioannou et al., 2013] Papaioannou, I., Breitung, K., and Straub, D. (2013). Reliability sensitivity analysis with Monte Carlo methods. In *ICOSSAR 2013*.

[Papaioannou et al., 2018] Papaioannou, I., Breitung, K., and Straub, D. (2018). Reliability sensitivity estimation with sequential importance sampling. *Structural Safety*, 75:24–34.

[Papaioannou et al., 2019] Papaioannou, I., Geyer, S., and Straub, D. (2019). Improved cross entropy-based importance sampling with a flexible mixture model. *Reliability Engineering & System Safety*, 191:106564.

[Papaioannou et al., 2016] Papaioannou, I., Papadimitriou, C., and Straub, D. (2016). Sequential importance sampling for structural reliability analysis. *Structural safety*, 62:66–75.

[Papaioannou and Straub, 2021a] Papaioannou, I. and Straub, D. (2021a). Combination line sampling for structural reliability analysis. *Structural Safety*, 88:102025.

[Papaioannou and Straub, 2021b] Papaioannou, I. and Straub, D. (2021b). Variance-based reliability sensitivity analysis and the FORM α-factors. *Reliability Engineering & System Safety*, 210:107496.

[Paté-Cornell, 1996] Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety*, 54(2-3):95–111.

[Perrin and Defaux, 2019] Perrin, G. and Defaux, G. (2019). Efficient evaluation of reliability-oriented sensitivity indices. *Journal of Scientific Computing*, 79(3):1433–1455.

[Pradlwarter et al., 2005] Pradlwarter, H., Pellissetti, M., Schenk, C., Schueller, G., Kreis, A., Fransen, S., Calvi, A., and Klein, M. (2005). Realistic and efficient reliability estimation for aerospace structures. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16):1597–1617.

[Pradlwarter et al., 2007] Pradlwarter, H., Schueller, G. I., Koutsourelakis, P.-S., and Charmpis, D. C. (2007). Application of line sampling simulation method to reliability benchmark problems. *Structural safety*, 29(3):208–221.

[Proppe, 2021] Proppe, C. (2021). Local reliability based sensitivity analysis with the moving particles method. *Reliability Engineering & System Safety*, 207:107269.

[Qiu et al., 2008] Qiu, Z., Yang, D., and Elishakoff, I. (2008). Probabilistic interval reliability of structural systems. *International Journal of Solids and Structures*, 45(10):2850–2860.

[Quarteroni et al., 2010] Quarteroni, A., Sacco, R., and Saleri, F. (2010). *Numerical mathematics*, volume 37. Springer Science & Business Media.

[Robert and Casella, 1999] Robert, C. P. and Casella, G. (1999). *Monte Carlo statistical methods*, volume 2. Springer.

[Rosenblatt, 1952] Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3):470–472.

[Royset and Polak, 2004] Royset, J. and Polak, E. (2004). Reliability-based optimal design using sample average approximations. *Probabilistic Engineering Mechanics*, 19(4):331–343.

[Rubinstein, 1986] Rubinstein, R. Y. (1986). The score function approach for sensitivity analysis of computer simulation models. *Mathematics and Computers in Simulation*, 28(5):351–379.

[Rubinstein, 1997] Rubinstein, R. Y. (1997). Optimization of computer simulation models with rare events. *European Journal of Operational Research*, 99(1):89–112.

[Rubinstein, 1999] Rubinstein, R. Y. (1999). The cross-entropy method for combinatorial and continuous optimization. *Methodology and computing in applied probability*, 1(2):127–190.

[Rubinstein and Glynn, 2009] Rubinstein, R. Y. and Glynn, P. W. (2009). How to deal with the curse of dimensionality of likelihood ratios in Monte Carlo simulation. *Stochastic Models*, 25(4):547–568.

[Rubinstein and Kroese, 2004] Rubinstein, R. Y. and Kroese, D. P. (2004). *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning*, volume 133. Springer.

[Rubinstein and Kroese, 2016] Rubinstein, R. Y. and Kroese, D. P. (2016). *Simulation and the Monte Carlo method*. John Wiley & Sons.

[Rubinstein and Shapiro, 1990] Rubinstein, R. Y. and Shapiro, A. (1990). Optimization of static simulation models by the score function method. *Mathematics and Computers in Simulation*, 32(4):373–392.

[Sahin, 2019] Sahin, F. E. (2019). Open-source optimization algorithms for optical design. *Optik*, 178:1016–1022.

[Saltelli et al., 2008] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.

[Saltelli et al., 2004] Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., et al. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*, volume 1. Wiley Online Library.

[Sancetta and Satchell, 2004] Sancetta, A. and Satchell, S. (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric theory*, 20(3):535–562.

[Sarazin, 2021] Sarazin, G. (2021). *Analyse de sensibilité fiabiliste en présence d'incertitudes épistémiques introduites par les données d'apprentissage*. PhD thesis, Toulouse, ISAE.

[Sbert and Havran, 2017] Sbert, M. and Havran, V. (2017). Adaptive multiple importance sampling for general functions. *The Visual Computer*, 33(6):845–855.

[Sbert et al., 2018] Sbert, M., Havran, V., and Szirmay-Kalos, L. (2018). Multiple importance sampling revisited: breaking the bounds. *EURASIP Journal on Advances in Signal Processing*, 2018(1):1–15.

[Schueller et al., 2004] Schueller, G. I., Pradlwarter, H. J., and Koutsourelakis, P.-S. (2004). A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics*, 19(4):463–474.

[Schuëller and Stix, 1987] Schuëller, G. I. and Stix, R. (1987). A critical appraisal of methods to determine failure probabilities. *Structural Safety*, 4(4):293–309.

[Schweizer and Sklar, 1983] Schweizer, B. and Sklar, A. (1983). North-Holland Series in Probability and Applied Mathematics. In *Probabilistic Metric Spaces*. North-Holland Publishing Co., New York.

[Shayanfar et al., 2018] Shayanfar, M. A., Barkhordari, M. A., Barkhori, M., and Barkhori, M. (2018). An adaptive directional importance sampling method for structural reliability analysis. *Structural Safety*, 70:14–20.

[Shekhar and Xiong, 2007] Shekhar, S. and Xiong, H. (2007). *Encyclopedia of GIS*. Springer Science & Business Media.

[Shields et al., 2015] Shields, M. D., Teferra, K., Hapij, A., and Daddazio, R. P. (2015). Refined stratified sampling for efficient Monte Carlo based uncertainty quantification. *Reliability Engineering & System Safety*, 142:310–325.

[Shufang et al., 2009] Shufang, S., Zhenzhou, L., Weiwei, Z., and Zhengyin, Y. (2009). Reliability and sensitivity analysis of transonic flutter using improved line sampling technique. *Chinese Journal of Aeronautics*, 22(5):513–519.

[Silverman, 2018] Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

[Sklar, 1959] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris*, 8:229–231.

[Sobol', 1990] Sobol', I. M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, 2(1):112–118.

[Sobol, 2018] Sobol, I. M. (2018). *A primer for the Monte Carlo method*. CRC press.

[Song et al., 2009] Song, S., Lu, Z., and Qiao, H. (2009). Subset simulation for structural reliability sensitivity analysis. *Reliability Engineering & System Safety*, 94(2):658–665.

[Song et al., 2011] Song, S., Lu, Z., and Song, Z. (2011). Reliability sensitivity analysis involving correlated random variables by directional sampling. In *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pages 845–850. IEEE.

[Sues and Cesare, 2005] Sues, R. H. and Cesare, M. A. (2005). System reliability and sensitivity factors via the MPPSS method. *Probabilistic Engineering Mechanics*, 20(2):148–157.

[Tabandeh et al., 2022] Tabandeh, A., Jia, G., and Gardoni, P. (2022). A review and assessment of importance sampling methods for reliability analysis. *Structural Safety*, 97:102216.

[Taboga, 2021] Taboga, M. (2021). Generalized least squares. *Lectures on probability theory and mathematical statistics*.

[Thedy and Liao, 2021] Thedy, J. and Liao, K.-W. (2021). Multisphere-based importance sampling for structural reliability. *Structural Safety*, 91:102099.

[Theil, 1971] Theil, H. (1971). Best linear unbiased estimation and prediction. *Principles of Econometrics*, pages 119–124.

[Torii, 2020] Torii, A. J. (2020). On sampling-based schemes for probability of failure sensitivity analysis. *Probabilistic Engineering Mechanics*, 62:103099.

[Torii et al., 2017] Torii, A. J., Lopez, R. H., and Miguel, L. F. F. (2017). Probability of failure sensitivity analysis using polynomial expansion. *Probabilistic Engineering Mechanics*, 48:76–84.

[Torii and Novotny, 2021] Torii, A. J. and Novotny, A. A. (2021). *A priori* error estimates for local reliability-based sensitivity analysis with Monte Carlo simulation. *Reliability Engineering & System Safety*, 213:107749.

[Tvedt, 1989] Tvedt, L. (1989). Second order reliability by an exact integral. In *Reliability and Optimization of Structural Systems' 88*, pages 377–384. Springer.

[Uribe et al., 2021] Uribe, F., Papaioannou, I., Marzouk, Y. M., and Straub, D. (2021). Cross-entropy-based importance sampling with failure-informed dimension reduction for rare event simulation. *SIAM/ASA Journal on Uncertainty Quantification*, 9(2):818–847.

[Valdebenito et al., 2010] Valdebenito, M., Pradlwarter, H., and Schuëller, G. I. (2010). The role of the design point for calculating failure probabilities in view of dimensionality and structural nonlinearities. *Structural Safety*, 32(2):101–111.

[Valdebenito et al., 2018] Valdebenito, M. A., Jensen, H. A., Hernández, H., and Mehrez, L. (2018). Sensitivity estimation of failure probability applying line sampling. *Reliability Engineering & System Safety*, 171:99–111.

[Wang et al., 2013] Wang, P., Lu, Z., and Tang, Z. (2013). A derivative based sensitivity measure of failure probability in the presence of epistemic and aleatory uncertainties. *Computers & Mathematics with Applications*, 65(1):89–101.

[Wang and Song, 2016] Wang, Z. and Song, J. (2016). Cross-entropy-based adaptive importance sampling using von Mises–Fisher mixture for high dimensional reliability analysis. *Structural Safety*, 59:42–52.

[Wang and Song, 2018] Wang, Z. and Song, J. (2018). Hyper-spherical extrapolation method (HEM) for general high dimensional reliability problems. *Structural Safety*, 72:65–73.

[Watson, 1967] Watson, G. S. (1967). Linear least squares regression. *The Annals of Mathematical Statistics*, pages 1679–1699.

[Werner, 1984] Werner, W. (1984). Polynomial interpolation: Lagrange versus Newton. *Mathematics of computation*, 43(167):205–217.

[Williamson, 1956] Williamson, R. E. (1956). *Multiply monotone functions and their Laplace transforms*, volume 23. Duke Mathematical Journal.

[Wu, 1994a] Wu, Y.-T. (1994a). Adaptive importance sampling (AIS)-based system reliability sensitivity analysis method. In *Probabilistic Structural Mechanics: Advances in Structural Reliability Methods: IUTAM Symposium, San Antonio, Texas, USA June 7–10, 1993*, pages 550–564. Springer.

[Wu, 1994b] Wu, Y.-T. (1994b). Computational methods for efficient structural reliability and reliability sensitivity analysis. *AIAA journal*, 32(8):1717–1723.

[Yang and Gu, 2004] Yang, R. and Gu, L. (2004). Experience with approximate reliability-based optimization methods. *Structural and multidisciplinary optimization*, 26:152–159.

[Yoo et al., 2014] Yoo, D., Lee, I., and Cho, H. (2014). Probabilistic sensitivity analysis for novel second-order reliability method (SORM) using generalized chi-squared distribution. *Structural and Multidisciplinary Optimization*, 50:787–797.

[Yu et al., 2018] Yu, Y., Lyu, Z., Xu, Z., and Martins, J. R. (2018). On the influence of optimization algorithm and initial design on wing aerodynamic shape optimization. *Aerospace Science and Technology*, 75:183–199.

[Yun et al., 2020] Yun, W., Lu, Z., He, P., Dai, Y., and Feng, K. (2020). Adaptive subdomain sampling and its adaptive Kriging–based method for reliability and reliability sensitivity analyses. *Structural and Multidisciplinary Optimization*, 61:1107–1121.

[Zhang, 1996] Zhang, P. (1996). Nonparametric importance sampling. *Journal of the American Statistical Association*, 91(435):1245–1253.

[Zhang et al., 2022] Zhang, X., Lu, Z., and Cheng, K. (2022). Cross-entropy-based directional importance sampling with von Mises–Fisher mixture model for reliability analysis. *Reliability Engineering & System Safety*, 220:108306.

[Zhong et al., 2020] Zhong, C., Wang, M., Dang, C., Ke, W., and Guo, S. (2020). First-order reliability method based on Harris Hawks optimization for high-dimensional reliability analysis. *Structural and Multidisciplinary Optimization*, 62(4):1951–1968.

[Zhu et al., 2022] Zhu, S.-P., Keshtegar, B., Seghier, M. E. A. B., Zio, E., and Taylan, O. (2022). Hybrid and enhanced PSO: Novel first order reliability method-based hybrid intelligent approaches. *Computer Methods in Applied Mechanics and Engineering*, 393:114730.

[Zuev, 2009] Zuev, K. (2009). *Advanced stochastic simulation methods for solving high-dimensional reliability problems*. Hong Kong University of Science and Technology (Hong Kong).

[Zuev et al., 2012] Zuev, K. M., Beck, J. L., Au, S.-K., and Katafygiotis, L. S. (2012). Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions. *Computers & structures*, 92:283–296.

Résumé

Dans de nombreuses disciplines scientifiques, un système complexe est souvent modélisé avec une fonction $\mathcal{M}$, supposée simuler le comportement du système. La sortie de cette fonction est la réponse observée tandis que les entrées représentent les différentes variables qui influencent le comportement du système. Le système étant soumis à des aléas, certaines de ses entrées sont des variables aléatoires. Ainsi, la sortie de la fonction est également aléatoire. La défaillance du système correspond à une réponse extrême de la fonction, dont les conséquences sont critiques. Dans un contexte où la fonction $\mathcal{M}$ agit comme une boîte noire, le calcul de la probabilité de défaillance et de sa sensibilité locale est crucial pour l'analyse de fiabilité du système. L'estimation de la probabilité de défaillance et de sa sensibilité locale est particulièrement difficile pour des systèmes de grande dimension, dont le domaine de défaillance des entrées possède plusieurs régions de défaillance. Les entrées aléatoires sont modélisées ici avec des distributions elliptiques standards, comme la loi normale standard par exemple. Le comportement de telles distributions en grande dimension doit être pris en compte dans le processus d'estimation. Dans la littérature, de nombreux algorithmes d'échantillonnage préférentiel (EP) ont été spécifiquement adaptés à l'estimation de la probabilité de défaillance de systèmes possédant plusieurs régions de défaillance dans l'espace standard normal de grande dimension. Le but de l'EP est de construire une densité auxiliaire qui génère plus d'observations dans le domaine de défaillance que la densité de probabilité initiale des entrées aléatoires. Afin de prendre en compte la multiplicité des régions de défaillance, la densité auxiliaire est alors construite comme un mélange de densités. Cependant, le nombre de densités dans le mélange, correspondant au nombre de régions de défaillance, doit être fixé *a priori* et est difficile à paramétrer dans un contexte de boîte noire. De plus, dans l'espace standard elliptique, la sensibilité locale de la probabilité par rapport à une entrée déterministe s'écrit nécessairement comme une intégrale surfacique, dont l'estimation est laborieuse. En conséquence, ces sensibilités sont souvent approchées par un estimateur biaisé plus facile à obtenir, mais dont le biais est difficilement évaluable. Le but de cette thèse est alors d'améliorer l'estimation de la probabilité de défaillance et de sa sensibilité locale dans un tel contexte. L'algorithme proposé pour estimer la probabilité de défaillance repose sur une recherche progressive des régions de défaillance dans l'espace standard elliptique. Pour chaque région de défaillance identifiée, une densité auxiliaire est construite et optimisée de manière adaptative. Cette densité prend en compte le comportement des entrées en grande dimension puisque sa construction bénéficie de leur représentation stochastique. La probabilité de défaillance finale est estimée avec un échantillonnage préférentiel multiple, en utilisant toutes les densités auxiliaires précédemment construites. L'approche proposée pour estimer la sensibilité de la probabilité par rapport aux entrées déterministes repose sur une régression polynomiale hétéroscédastique. En effet, la sensibilité locale est identifiée comme le terme d'ordre zéro d'un développement de Taylor. En utilisant les méthodes de simulation pour construire la base de données de la régression, un nouvel estimateur de la sensibilité locale est ainsi obtenu. Cet estimateur est biaisé mais son biais est contrôlable grâce aux paramètres de régression : le degré du polynôme ainsi que l'intervalle de régression. Les deux algorithmes sont mis en pratique indépendamment avec différentes applications pour démontrer leur robustesse. Une étude de la fiabilité d'un oscillateur de Duffing permet de tester la combinaison des deux algorithmes pour un système de grande dimension possédant plusieurs régions de défaillance.

Mots-clés: analyse de fiabilité, estimation de probabilité de défaillance, sensibilité locale de probabilité de défaillance, échantillonnage préférentiel, régression polynomiale hétéroscédastique, loi standard elliptique, régions de défaillance multiples, système de grande dimension

Abstract

In several scientific disciplines, a complex system is typically modeled with a function $\mathcal{M}$, expected to simulate the behavior of the system. The output of this function is the observed response while the inputs represent the different variables that have an influence over the behavior of the system. As the system is subject to uncertainty, some of the inputs are random variables. Hence, the output of the function is random as well. The failure of the system corresponds to an extreme response of the function which can have severe consequences. In a context where the function $\mathcal{M}$ is seen as a black-box function, computing the probability of failure of the system, as well as its local sensitivity is crucial for reliability analysis. The estimation of the failure probability and its local sensitivity is particularly difficult for high-dimensional systems, whose input failure domain encompasses possibly several failure regions. Assuming the random inputs can be modeled with standard elliptical variables, the behavior of such distributions in high-dimensional spaces must be taken into account in the estimation process. In the literature, several Importance Sampling (IS) algorithms have been specifically tailored to estimate the failure probability of systems with several failure regions in the high-dimensional standard normal space, a commonly used elliptical space. The goal of IS is to build an auxiliary density which generates more observations in the failure domain than the original probability density function of the random inputs. In order to take into account the multiple failure regions, the auxiliary density is then built as a mixture of densities. However, the number of densities in the mixture, corresponding to the number of failure regions, has to be set *a priori* and is arduous to derive in a black-box context. Furthermore, in the standard elliptical space, the sensitivity of the failure probability with respect to a deterministic input is necessarily written as a surface integral, which is burdensome to estimate. Consequently, the local sensitivity is usually approximated with a biased estimate easier to obtain, whose bias is then difficult to examine. The goal of this thesis is to improve the estimation of both the failure probability and its local sensitivity in such a context. The algorithm proposed to estimate the failure probability relies on a gradual search for the failure regions in the standard elliptical space. For each identified failure region, an auxiliary density is built and adaptively optimized. This density takes into account the behavior of the random inputs in the high-dimensional space as their stochastic representation is taken advantage of. The final failure probability is then estimated with a multiple important sampling scheme, using every auxiliary density previously built. The approach proposed to estimate the failure probability sensitivity with respect to the deterministic inputs relies on a heteroscedastic polynomial regression. Indeed, the local sensitivity is identified as the term of order zero of a Taylor series expansion. Using simulation methods to build the regression database, a new sensitivity estimate is obtained. This estimate is biased, but its bias can be controlled thanks to the parameters of the regression: the polynomial degree and the regression interval. Both algorithms are applied to various applications to test their robustness independently. A study on the reliability of a Duffing oscillator allows to test the combination of both algorithms, for a high-dimensional system with several failure regions.

Keywords: reliability analysis, failure probability estimation, failure probability sensitivity, importance sampling, standard elliptical laws, heteroscedastic polynomial regression, multiple failure regions, high-dimensional system