Adaptive surrogate models for reliability analysis and reliability-based design optimization

Vincent Dubourg

N° d'ordre : D.U. : 2184
EDSPIC : 540

Université Blaise Pascal - Clermont II

École Doctorale
Sciences pour l'Ingénieur de Clermont-Ferrand

Thèse

présentée par

Vincent DUBOURG
Ingénieur IFMA

le **5 décembre 2011**

en vue d'obtenir le grade de

Docteur d'Université
(Spécialité : Génie Mécanique)

Méta-modèles adaptatifs pour l'analyse de fiabilité et l'optimisation sous contrainte fiabiliste

soutenue publiquement devant un jury composé de MM.

Pr. Pierre-Alain BOUCARD	LMT-Cachan, Cachan	Examinateur
Dr. Jean-Marc BOURINET	IFMA/LaMI, Clermont-Ferrand	Co-encadrant
Pr. Alaa CHATEAUNEUF	Polytech', Clermont-Ferrand	Examinateur
Pr. Josselin GARNIER	Université Paris VII, Paris	Examinateur
Dr. Bertrand IOOSS	EdF R&D, Chatou	Rapporteur
Dr. Rodolphe LE RICHE	ENSM-SE, Saint-Étienne	Rapporteur
Pr. John Dalsgaard SØRENSEN	Université d'Aalborg, Danemark	Président du jury
Dr. Bruno SUDRET	Phimeca Engineering S.A., Paris	Directeur de thèse

Laboratoire de Mécanique et Ingénieries (LaMI)
Université Blaise Pascal et Institut Français de Mécanique Avancée

Adaptive surrogate models for reliability analysis and reliability-based design optimization

A thesis submitted by **Vincent DUBOURG**
in partial fulfillment of the requirements for
the degree of doctor of philosophy

BLAISE PASCAL UNIVERSITY – CLERMONT II

Clermont-Ferrand, France

defended publicly on December 5, 2011 in front of a jury made up of:

Pr. Pierre-Alain BOUCARD	LMT-Cachan, Cachan	Examiner
Dr. Jean-Marc BOURINET	IFMA/LaMI, Clermont-Ferrand	Co-supervisor
Pr. Alaa CHATEAUNEUF	Polytech', Clermont-Ferrand	Examiner
Pr. Josselin GARNIER	Université Paris VII, Paris	Examiner
Dr. Bertrand IOOSS	EdF R&D, Chatou	Reviewer
Dr. Rodolphe LE RICHE	ENSM-SE, Saint-Étienne	Reviewer
Pr. John Dalsgaard SØRENSEN	Aalborg University, Denmark	President of the jury
Dr. Bruno SUDRET	Phimeca Engineering S.A., Paris	Supervisor

This manuscript was typeset with LaTeX-2ε (TeX Live v2009-7) using *Charter Bitstream with Math Design* as the body font. The source files were edited with Kile 2.1beta4 under Ubuntu Linux 10.04 LTS codename *Lucid Lynx*. Graphical illustrations were produced with Matplotlib (running Python 2.6), Ti*k*Z, Inkscape and Matlab ®.

BibTeX entry:

```
@PHDTHESIS{Dubourg2011,
  author = {Dubourg, V.},
  title = {Adaptive surrogate models for reliability analysis
           and reliability-based design optimization},
  school = {Universit\'e Blaise Pascal - Clermont II},
  year = {2011},
}
```

Earliest implementations of the algorithms detailed in this manuscript were done in the FERUM toolbox using the DACE Kriging toolbox for Matlab ®, but most illustrations were then produced with Python 2.6 using Numpy, Scipy, scikit learn, PyMC and OpenTURNS.

Abstract

This thesis is a contribution to the resolution of the reliability-based design optimization problem. This probabilistic design approach is aimed at considering the uncertainty attached to the system of interest in order to provide optimal and safe solutions. The safety level is quantified in the form of a probability of failure. Then, the optimization problem consists in ensuring that this failure probability remains less than a threshold specified by the stakeholders. The resolution of this problem requires a high number of calls to the limit-state design function underlying the reliability analysis. Hence it becomes cumbersome when the limit-state function involves an expensive-to-evaluate numerical model (*e.g.* a finite element model). In this context, this manuscript proposes a surrogate-based strategy where the limit-state function is progressively replaced by a Kriging meta-model. A special interest has been given to quantifying, reducing and eventually eliminating the error introduced by the use of this meta-model instead of the original model. The proposed methodology is applied to the design of geometrically imperfect shells prone to buckling.

Keywords: adaptive surrogate modelling • Kriging • Gaussian processes for regression and probabilistic classification • reliability analysis • rare event probabilities • importance sampling • reliability-based design optimization • probabilistic buckling • geometrically imperfect shells

Résumé

Cette thèse est une contribution à la résolution du problème d'optimisation sous contrainte de fiabilité. Cette méthode de dimensionnement probabiliste vise à prendre en compte les incertitudes inhérentes au système à concevoir, en vue de proposer des solutions optimales et sûres. Le niveau de sûreté est quantifié par une probabilité de défaillance. Le problème d'optimisation consiste alors à s'assurer que cette probabilité reste inférieure à un seuil fixé par les donneurs d'ordres. La résolution de ce problème nécessite un grand nombre d'appels à la fonction d'état-limite caractérisant le problème de fiabilité sous-jacent. Ainsi, cette méthodologie devient complexe à appliquer dès lors que le dimensionnement s'appuie sur un modèle numérique coûteux à évaluer (*e.g.* un modèle aux éléments finis). Dans ce contexte, ce manuscrit propose une stratégie basée sur la substitution adaptative de la fonction d'état-limite par un méta-modèle par Krigeage. On s'est particulièrement employé à quantifier, réduire et finalement éliminer l'erreur commise par l'utilisation de ce méta-modèle en lieu et place du modèle original. La méthodologie proposée est appliquée au dimensionnement des coques géométriquement imparfaites soumises au flambement.

Mots-clés: méta-modélisation adaptative • Krigeage • régression et classification probabiliste par processus Gaussiens • analyse de fiabilité • probabilités d'évènements rares • échantillonnage préférentiel • optimisation sous contrainte de fiabilité • flambage probabiliste • coques géométriquement imparfaites

Acknowledgements

Although it officially started in September 2008, this research work was ignited in September 2006. At that time, I was an undergraduate student at the French institute for advanced mechanics (IFMA) and I had to choose for a five-month long internship abroad. I met "Monsieur BOURINET" (at this point) on this occasion and he offered me the opportunity to visit Terje HAUKAAS at the University of British Columbia in Vancouver (Canada). Actually, this opportunity was conditioned by a first test project under his direction. Hence, he first introduced me with what would come to be my five-year research topic, namely: *reliability-based design optimization*. He certainly enjoyed the experience as much as I did since I eventually did visit Terje at UBC.

Back in France, as a fresh diplomed engineer, I asked Jean-Marc (from this point on) if he could set up a Ph. D. thesis on the same topic. Thanks to his enthusiasm, he managed to convince Maurice PENDOLA (co-founder and CEO of Phimeca Engineering SA) to support the project financially with the help of the French national agency for research and technology (ANRT) and the French shipbuilder DCNS in the person of Mikaël CAZUGUEL (engineer in the "Submarine Engineering" division). For this success, for your continued guidance since then and your friendly attitude to me, I sincerely thank you Jean-Marc.

This Ph. D. thesis would not have been possible without another experienced advisor in probabilistic engineering mechanics. Fortunately, during the same period Maurice PENDOLA decided to hire Bruno SUDRET on purpose to manage the research and development activity at Phimeca. Despite Bruno had not heard of my work before we met, he quickly caught up with the related literature to give me his relevant advice during these three years. No doubt he is the man responsible for most of the sixteen pages of references at the end of this manuscript! Indeed, beyond his technical assistance, he involved me into many conferences and seminars so that I could benefit from "live bibliography". For your tireless incentives at publishing my work and for your continued patience reviewing and improving our papers and this thesis, I sincerely thank you Bruno. Now I can say, I could not have imagined a better supervisor.

In addition to the supervision of Jean-Marc and Bruno I would like to acknowledge the equally valuable contribution of François DEHEEGER (research engineer at Phimeca). Besides the so many things you taught me about scientific programming in general (from pure Python/LaTeX scripting to the whole world around it), I wanted to thank you for being so available for discussion during these three years. I hope our collaboration will go on.

I would like to thank Bertrand IOOSS and Rodolphe LE RICHE who patiently reviewed this thesis, and who have been yet another source of inspiration. Thank you Bertrand for organizing the Mascot'NUM meetings. May the research group live for long (in this form or any other) and inspire many other Ph. D. students! Thank you Rodolphe for your interest in my work and for the discussions we had about it. I would also like to thank Pierre-Alain BOUCARD, Alaa CHATEAUNEUF, Josselin GARNIER and John SØRENSEN who accepted to examine my work. Thank you all for your comments and for coming to Clermont-Ferrand, on a monday and so early in the morning!

During these three years I split my time almost equally between two institutions: the mechanical engineering laboratory at IFMA (Clermont-Ferrand) and Phimeca (Clermont-Ferrand and Paris). It gave me the occasion to meet many friendly people who helped me making it to the end! Thank you Claudine for your time teaching me how to use EVE as you were writing your thesis. Thank you Gilles for your relevant expertise in computational mechanics (did you check that mesh?). Thank you Yann, Géraud and Ben for being my conference mates. Thank you Cécile, Agnès, Guillaume, Nicolas, Manu, Pierre, Alex, Julien, Mathieu, Pascal, Valérie, Thierry, ...

Let me also thank Alex GRAMFORT and Gaël VAROQUAUX who reviewed my contribution to the scikit-learn project: "machine learning in Python". I want to point out the great work the dev team is accomplishing for promoting the use of statistical learning techniques by non-experts.

Last but not least, I thank my parents, Jacques and Dominique, for their never-failing emotional support along this great adventure.

Vincent Dubourg, December 8, 2011

Résumé étendu

Contexte

La simulation numérique est omniprésente dans l'ingénierie moderne. Son champ d'application s'étend des sciences physiques à l'ingénierie financière. Il s'agit d'imiter le comportement d'un système physique ou abstrait à l'aide de modèles mathématiques. La fidélité de ces modèles a été améliorée de façon significative au cours des dernières décennies. Néanmoins, du fait de leur manque de représentativité, on observe encore des écarts entre les comportements simulés et réels.

Par définition, ces modèles sont des idéalisations de la réalité et ils négligent par conséquent certains aspects du phénomène étudié. Une contribution majeure aux écarts observés est liée à l'isolement du système étudié de son environnement. En effet, lorsque l'on étudie un système, son environnement immédiat est réduit à un ensemble fini de configurations idéalisées (ou nominales). Par exemple, en mécanique des structures, l'état initial d'une structure, sa géométrie, les propriétés du matériau la constituant ainsi que les chargements sont réduits à un ensemble fini de configurations qui peut manquer d'exhaustivité. Cette réduction du monde réel à un modèle synthétique s'avère pourtant nécessaire à la bonne résolution du problème (en l'état actuel de nos connaissances).

Dans ce contexte, un cadre probabiliste semble plus approprié puisqu'il permet de spécifier *un ensemble infini de configurations probables* sous la forme d'une seule distribution de probabilité. Cependant, une telle approche soulève des questions théoriques notamment sur la construction de modèles probabilistes appropriés et sur les mesures de performance associées, ces dernières étant la base de tout processus de prise de décision.

Heureusement, au cours des deux dernières décennies, un travail conséquent a été mené en *ingénierie probabiliste* dans le but de formaliser ces problèmes et de former les ingénieurs à leur résolution. En effet, la littérature scientifique fournit désormais une pléiade de techniques pour la *quantification des incertitudes*, et les élèves-ingénieurs sont sensibilisés à la théorie des probabilités ainsi qu'à ses applications potentielles aux problèmes industriels.

Formulation du problème

Cette thèse s'intéresse plus particulièrement au problème du dimensionnement et de l'optimisation des structures en présence d'incertitudes tel que présenté dans le livre de Tsompanakis et al. (2008). Souvent réduit au terme de *conception robuste* dans l'industrie, ce dernier problème est en fait plus général et peut se présenter sous plusieurs formes. L'*optimisation robuste* cherche à minimiser l'influence des conditions environnementales incertaines sur la performance du système à concevoir, tandis que l'*optimisation sous contrainte de fiabilité* vise à concevoir pour la sûreté par rapport à des évènements extrêmes (rares). Ce manuscrit se concentre sur la résolution de ce dernier problème basé sur la fiabilité (*reliability-based design optimization*, RBDO).

On suppose que l'on dispose d'un modèle analytique ou numérique noté $\mathscr{M}$ permettant de prédire une ou plusieurs grandeurs physiques caractérisant l'état du système par le biais de fonctions dites *de performance* notées $\mathfrak{g}_l$, $l = 1, \dots, n_p$. Ces fonctions sont construites de telle manière que *la défaillance du système* suivant le l-ième scénario soit caractérisée par les valeurs négatives de la l-ième fonction de performance $\mathfrak{g}_l$. On a par ailleurs construit un modèle probabiliste du système et des conditions environnementales dans lequel il évolue, et on suppose que ce dernier se présente sous la forme d'un vecteur aléatoire $\boldsymbol{X}$ de densité $f_{\boldsymbol{X}}$. De plus, on suppose que le vecteur aléatoire $\boldsymbol{X}$ est paramétré par des valeurs caractéristiques maîtrisables regroupées dans le vecteur des *variables de conception* noté $\boldsymbol{d}$.

Dans ce contexte, le degré de sûreté du système à concevoir est communément mesuré par ses *probabilités de défaillance* définies comme suit :

$$p_{f\,l}(\boldsymbol{d}) = \mathbb{P}\left[\mathfrak{g}_l\left(\boldsymbol{X}, \mathscr{M}(\boldsymbol{X})\right) \leq 0 \,|\, \boldsymbol{d}\right] = \int_{\mathbb{F}_l} f_{\boldsymbol{X}}\left(\boldsymbol{x} \,|\, \boldsymbol{d}\right) \mathrm{d}\boldsymbol{x}, \quad l = 1, \dots, n_p, \tag{1}$$

où $\mathbb{F}_l = \left\{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}_l\left(\boldsymbol{x}, \mathscr{M}(\boldsymbol{x})\right) \leq 0\right\}$ est le l-ième *domaine de défaillance*.

Ainsi, le dimensionnement optimal sous contrainte de fiabilité est celui qui minimise une fonction objectif notée $\mathfrak{c}$ (*e.g.* un coût de construction) et tel que ses probabilités de défaillance n'excèdent pas les seuils prescrits par les donneurs d'ordres et notés $p^0_{f\,l}$, $l = 1, \dots, n_p$. Mathématiquement, ce problème prend la forme suivante :

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d} \in \mathbb{D}} \mathfrak{c}(\boldsymbol{d}) \;:\; \begin{cases} \mathfrak{f}_i\left(\boldsymbol{d}\right) \leq 0, \; i = 1, \dots, n_c \\ p_{f\,l}(\boldsymbol{d}) \leq p^0_{f\,l}, \; l = 1, \dots, n_p \end{cases}, \tag{2}$$

où apparaissent le domaine de conception admissible $\mathbb{D}$ et un ensemble de contraintes déterministes auxiliaires $\mathfrak{f}_i, i = 1, \dots, n_c$ visant à restreindre la recherche de l'optimum recherché à des dimensionnements réalisables aux yeux des experts. Il est à noter que l'on s'est limité ici au cas où les variables de conception $\boldsymbol{d}$ sont des paramètres de la distribution de probabilité du vecteur $\boldsymbol{X}$.

Verrous scientifiques et objectifs

Un obstacle majeur à la résolution de l'Eq. (2) réside dans le nombre important de sollicitations du modèle $\mathscr{M}$ pour l'analyse de fiabilité d'un seul choix de conception (*i.e.* pour calculer l'intégrale de l'Eq. (1) pour une valeur donnée de $\boldsymbol{d}$). Aussi, dans un contexte industriel où le modèle $\mathscr{M}$ est coûteux à évaluer (*e.g.* un modèle aux éléments finis non linéaire), on souhaiterait limiter le nombre total d'appels à ce modèle à quelques milliers.

Ainsi, pour se plier à cette contrainte, de nombreuses techniques de résolution proposées dans la littérature sont basées sur la substitution des fonctions de performance par des *méta-modèles* moins coûteux à évaluer. A titre d'exemple, la méthode FORM, largement utilisée dans ce domaine, consiste à linéariser la surface d'état-limite $\mathscr{S} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) = 0\}$ au voisinage d'un point particulier baptisé point de défaillance le plus probable. Cependant, l'hypothèse d'unicité de ce dernier point ou même la linéarité de l'état-limite au voisinage de ce point peut s'avérer illicite, conduisant ainsi à une mauvaise estimation de la fiabilité. De plus, il est impossible de quantifier (même grossièrement) l'erreur commise en ayant substitué un hyperplan à l'état-limite réel.

Dans cette thèse, on s'est particulièrement attaché à quantifier, réduire *et finalement* éliminer *l'erreur introduite par l'utilisation d'un méta-modèle en lieu et place du modèle réel pour la résolution du problème d'optimisation sous contrainte de fiabilité.*

Méthodologie

Apprentissage statistique

Dans le but de quantifier l'erreur de substitution, on propose de placer l'exercice de construction du méta-modèle dans le cadre de la *théorie de l'apprentissage statistique* plutôt que celle des développements limités traditionnellement utilisée dans la littérature RBDO. La théorie de l'apprentissage statistique est présentée dans les ouvrages de références (voir *e.g.* Vapnik, 1995; Santner et al., 2003; Rasmussen et Williams, 2006) et l'intérêt de son application à la fiabilité des structures a été démontré par Hurtado (2004b).

De façon générale, cette théorie vise à construire un *émulateur* noté $\widetilde{\mathscr{M}}$ pour le *simulateur* $\mathscr{M}$ à partir d'une base de données $\mathscr{D} = \left\{(\boldsymbol{x}^{(i)}, y_i),\, i = 1, \ldots, m\right\}$ composée de m couples d'entrées $\boldsymbol{x} \in \mathbb{X} \subseteq \mathbb{R}^n$ et sorties $y \in \mathbb{Y} \subseteq \mathbb{R}$ telles que $y_i = \mathscr{M}(\boldsymbol{x}^{(i)})$, $i = 1, \ldots, m$. Les techniques de construction de ces émulateurs dépendent principalement de la nature de la *variable de supervision* y. Si elle prend des valeurs discrètes dans un ensemble fini dénombrable, on parle de *classification* et si elle prend des valeurs continues dans un ensemble indénombrable, on parle de *régression*. Dans le cadre de la fiabilité, on peut chercher à approximer la surface d'état-limite $\mathscr{S} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) = 0\}$ comme la frontière entre les points sûrs et les points défaillants, ce qui revient à résoudre un exercice de *classification*

binaire. Alternativement, on peut voir cet état-limite comme un contour particulier de la fonction de performance $\mathfrak{g}$ et ainsi résoudre un exercice de régression.

Dans cette thèse, on a choisi de maximiser l'information disponible en abordant le problème de méta-modélisation sous l'angle de la régression. Plus spécifiquement, la théorie de la *prédiction par processus gaussiens* (également connue sous le nom de *Krigeage*) a été retenue pour ses bonnes propriétés. Cette technique appartient aux méthodes de prédiction bayésiennes telles que présentées dans le livre de Santner et al. (2003). Il s'agit de reconstituer la fonction $\mathscr{M}$ comme si elle était une réalisation d'un processus gaussien dont on aurait collecté des observations sur un ensemble fini de points $\mathscr{X} = \left\{\boldsymbol{x}^{(i)}, i = 1, \ldots, m\right\}$. Finalement, on montre que sous un certain nombre d'hypothèses raisonnables, on peut prédire les valeurs prises par le modèle réel sous la forme d'une variable aléatoire gaussienne :

$$\widehat{Y}(\boldsymbol{x}) = \left[\mathscr{M}(\boldsymbol{x}) \,\middle|\, y_i = \mathscr{M}\left(\boldsymbol{x}^{(i)}\right), i = 1, \ldots, m\right] \sim \mathscr{N}\left(\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma^2_{\widehat{Y}}(\boldsymbol{x})\right), \tag{3}$$

dont la moyenne $\mu_{\widehat{Y}}$ et la variance (de prédiction) $\sigma^2_{\widehat{Y}}$ sont connues analytiquement. Outre son caractère *interpolant* et *consistant*, l'intérêt principal de cette prédiction est bien sa nature probabiliste qui permet de calculer des intervalles de confiance sur les prédictions fournies en général et sur la surface d'état-limite en particulier.

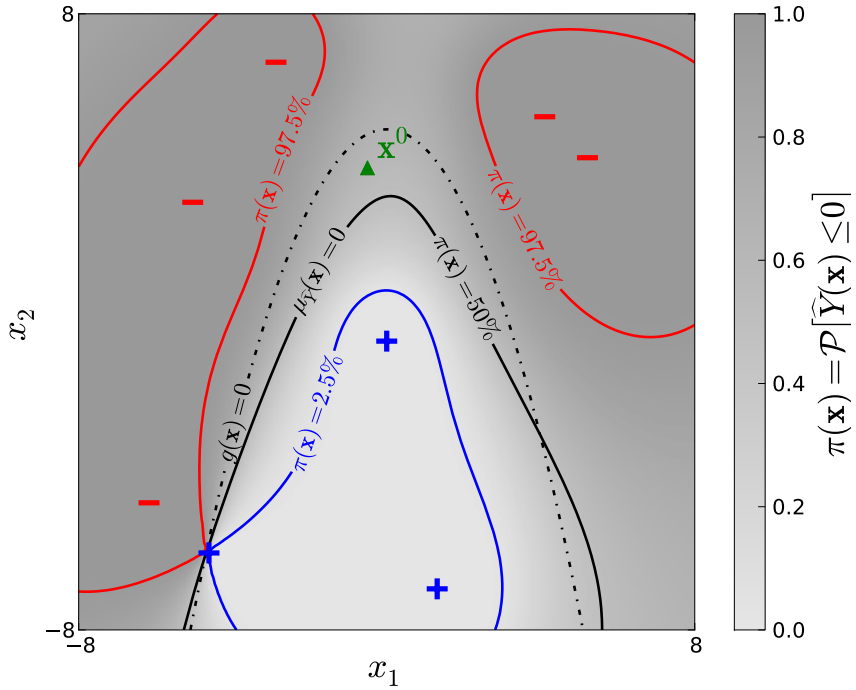


Figure 1 – Représentation 2D d'une prédiction par Krigeage d'un état-limite parabolique (ligne discontinue noire). La prédiction moyenne est représentée par la ligne continue noire tandis que la borne inférieure (resp. supérieure) de la marge d'incertitude à 95% est représentée par la ligne continue bleue (resp. rouge). Les points utilisés pour construire cette prédiction sont représentés par les signes "plus" bleus et "moins" rouges suivant le signe de la variable de supervision y associée. La fonction de classification probabiliste π vaut environ 60% au point $\boldsymbol{x}^0$ repéré par le triangle vert.

Par exemple sur la Figure 1, on a représenté la prédiction moyenne de la surface d'état-limite ainsi que les bornes inférieure et supérieure de l'intervalle de confiance à 95% sur cette prédiction. Dans ce manuscrit, on appelle cet intervalle la *marge d'incertitude* et elle est simplement définie comme suit :

$$\mathbb{M}_{95\%} = \left\{ \boldsymbol{x} \in \mathbb{X} : -1.96\,\sigma_{\widehat{Y}}(\boldsymbol{x}) \leq \mu_{\widehat{Y}}(\boldsymbol{x}) \leq 1.96\,\sigma_{\widehat{Y}}(\boldsymbol{x}) \right\}. \tag{4}$$

On peut aussi calculer la probabilité que le modèle réel prenne une valeur négative en un point $\boldsymbol{x}$ donné au moyen de la *fonction de classification probabiliste* suivante :

$$\pi(\boldsymbol{x}) = \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq 0\right] = \Phi\left(\frac{0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \tag{5}$$

et dont on a représenté les valeurs en nuances de gris sur la Figure 1.

Il est essentiel de distinguer la mesure de probabilité $\mathscr{P}$ dédiée à l'incertitude de méta-modélisation, de la mesure de probabilité $\mathbb{P}$ associée au vecteur aléatoire $\boldsymbol{X}$ dont il n'est pas encore question ici.

▷ *Cette partie est développée dans le Chapitre 1.*

Plans d'expériences adaptatifs

Les techniques d'enrichissement adaptatif de plans d'expériences visent à réduire l'erreur relative entre le méta-modèle et le modèle réel dans une région d'intérêt. Par exemple, en optimisation globale, il s'agira de s'assurer que la fonction est précise au voisinage de tous les minima locaux de la fonction substituée et au voisinage du minimum global recherché en particulier. C'est d'ailleurs dans ce but que Mockus (1994) et Jones et al. (1998) ont mis au point une technique d'enrichissement adaptatif de plans d'expériences pour localiser le minimum global d'une fonction coûteuse à évaluer. Cette technique utilise le caractère probabiliste de la prédiction par processus gaussien pour mesurer l'intérêt potentiel à explorer certaines régions de l'espace $\mathbb{X}$.

Depuis lors, l'idée générale a été étendue au problème d'approximation de contours pour l'optimisation sous contrainte d'inégalité ou pour l'analyse de fiabilité (voir *e.g.* Oakley, 2004; Picheny et al., 2010a). D'une façon générale, il s'agit de définir un critère qui mesure la proximité d'un point à l'état-limite en tenant compte de l'incertitude de méta-modélisation. Dans cette thèse, on propose d'utiliser la probabilité d'appartenir à la marge d'incertitude $\mathbb{M}_{95\%}$ comme mesure de proximité à l'état-limite $\mathscr{S} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) = 0\}$. Cette dernière est analytique grâce au caractère gaussien de la prédiction et se calcule comme suit :

$$\begin{aligned} \mathrm{MP}(\boldsymbol{x}) &= \mathscr{P}\left[-1.96\,\sigma_{\widehat{Y}}(\boldsymbol{x}) \leq \widehat{Y}(x) \leq 1.96\,\sigma_{\widehat{Y}}(\boldsymbol{x})\right] \\ &= \Phi\left(\frac{1.96\,\sigma_{\widehat{Y}}(\boldsymbol{x}) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \Phi\left(\frac{-1.96\,\sigma_{\widehat{Y}}(\boldsymbol{x}) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right). \end{aligned} \tag{6}$$

Fort de cette mesure de proximité, la plupart des stratégies d'enrichissement proposées dans la littérature consiste à trouver *le point* le plus proche de l'état-limite et qui permettra ainsi de réduire l'étendue de la marge d'incertitude sur la prédiction de ce dernier. Dans cette thèse, on soutient que ce sous-problème d'optimisation est relativement mal posé car les critères d'enrichissement $\mathscr{C}$ (dont MP n'est qu'un exemple) sont souvent hautement multi-modaux, ce qui signifie qu'il n'existe pas un seul meilleur point mais plutôt un ensemble de points plus ou moins intéressants et d'intérêt éventuellement égal au sens du critère $\mathscr{C}$ (surtout pour les critères d'approximation de contours).

Partant de ce constat, on propose d'interpréter les critères d'enrichissement (souvent positifs) comme des densités de probabilité pour les candidats potentiels à l'amélioration de la prédiction. Pour ce faire, il est nécessaire de s'assurer que leur intégrale sur $\mathbb{X}$ est finie. C'est pourquoi on propose de pondérer $\mathscr{C}$ par une densité w, de sorte que l'on peut définir le vecteur aléatoire *"candidat à l'enrichissement"* :

$$\boldsymbol{C} \sim f_{\boldsymbol{C}}(\boldsymbol{c}) \propto \mathscr{C}(\boldsymbol{c})\, w(\boldsymbol{c}). \tag{7}$$

On utilise ensuite des techniques d'échantillonnage par chaînes de Markov (Neal, 2003; Robert et Casella, 2004) pour simuler une population de points candidats à partir de la densité $f_{\boldsymbol{C}}$. Comme attendu, cette population se concentre principalement au voisinage de l'état-limite identifié (*i.e.* dans la marge d'incertitude $\mathbb{M}_{95\%}$).

Cependant, on ne peut pas évaluer tous ces points du fait du budget de calcul limité. C'est pour cette raison que l'on a recours au *K-means clustering* comme technique de réduction statistique d'échantillons. En effet, cette technique permet de localiser les K modes de la densité $f_{\boldsymbol{C}}$ (K étant donné). Ces K nouveaux points sont évalués pour venir enrichir la base de données $\mathscr{D}$ et réajuster la prédiction par Krigeage $\widehat{Y}$. La Figure 2 illustre l'ensemble de la stratégie proposée pas-à-pas. Cette procédure est répétée tant qu'un critère de précision n'est pas atteint. Ce critère dépend de l'usage de la prédiction et est évoqué dans la section suivante.

▷ *Cette partie est développée dans le Chapitre 2.*

Analyses de fiabilité basées sur l'apprentissage statistique

Dans le cadre d'une analyse de fiabilité, on vient classiquement évaluer la probabilité de défaillance sur la prédiction moyenne du Krigeage par simulations. C'est ce que l'on appelle *une analyse de fiabilité par substitution*. Dans le but de réduire encore le temps de calcul total, on propose d'utiliser une technique de réduction de variance robuste connue sous le nom de *subset sampling* (Au et Beck, 2001) en lieu et place des simulations de Monte Carlo classiques qui s'avèrent trop coûteuses lorsque la probabilité à estimer est faible (*i.e.* inférieure à 10^{-6}).

Cependant, la probabilité de défaillance évaluée sur la prédiction moyenne par Krigeage peut être biaisée par rapport à la probabilité de défaillance que l'on aurait pu estimer sur le modèle réel. Ce *biais* s'avère difficile à estimer en pratique.

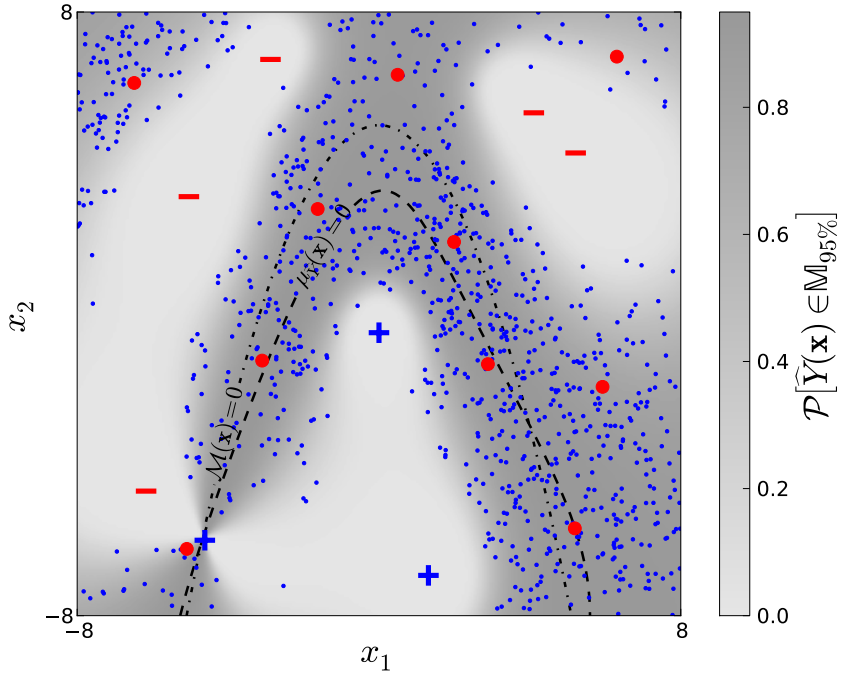


FIGURE 2 – Etapes clés de la procédure d'approximation adaptative de contour. (a) On construit une première prédiction par Krigeage à partir d'un plan d'expériences initial, ce qui permet également de définir le critère d'enrichissement $\mathscr{C}$ (ici MP défini à l'Eq. (6)) représenté en nuances de gris. (b) On échantillonne une population de taille $N = 1\,000$ suivant le critère $\mathscr{C}$ pondéré par la densité de la distribution uniforme sur le cube $\mathbb{X} = [-8;\ 8]^2$ par *slice sampling* (petits points bleus). (c) On réduit la population de taille N aux $K = 10$ centres de ses clusters par *K-means clustering* (gros points rouges). Le modèle $\mathscr{M}$ est ensuite évalué en ces points pour enrichir la base de données $\mathscr{D}$ et la prédiction $\widehat{Y}$.

Analyse de fiabilité par substitution

En première approximation, on propose d'appréhender l'ordre de grandeur de ce biais à l'aide de la variance de prédiction $\sigma^2_{\widehat{Y}}$. Pour ce faire, on définit les trois approximations suivantes du domaine de défaillance :

$$\widehat{\mathbb{F}}^{\,i}_{95\%} \equiv \left\{\boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) \leq i\, 1.96\, \sigma_{\widehat{Y}}(\boldsymbol{x})\right\}, \quad i = -1,\ 0, +1 \tag{8}$$

illustrées sur la Figure 3. Etant donné que l'écart-type de Krigeage est toujours positif, la proposition suivante est valide :

$$\mathbb{F}^{-1}_{95\%} \subseteq \mathbb{F}^{\,0}_{95\%} \subseteq \mathbb{F}^{+1}_{95\%} \quad \Rightarrow \quad p^{-1}_{f\,95\%} \leq p^{\,0}_{f\,95\%} \leq p^{+1}_{f\,95\%}, \tag{9}$$

où :

$$p^{\,i}_{f\,95\%} = \mathbb{P}\left[\boldsymbol{X} \in \mathbb{F}^{\,i}_{95\%}\right], \quad i = -1,\ 0, +1. \tag{10}$$

Il est néanmoins important de signaler ici qu'il n'existe aucune preuve que la probabilité de défaillance satisfasse cet encadrement de par les hypothèses faites jusqu'à présent. Pourtant, l'étendue de l'intervalle $[p^{-1}_{f\,95\%};\ p^{+1}_{f\,95\%}]$ s'avère être une bonne mesure pratique de la

précision de l'estimation de la probabilité de défaillance sur la prédiction par Krigeage. Il est donc proposé d'enrichir le méta-modèle tant que l'étendue logarithmique de l'intervalle $[p_{f\,95\%}^{-1};\ p_{f\,95\%}^{+1}]$ est plus grande qu'un seuil de tolérance fixé. La condition d'arrêt pour la procédure d'enrichissement itérative détaillée ci-avant s'écrit donc :

$$\varepsilon_{p_f} = \log_{10}\left(\frac{p_{f\,95\%}^{+1}}{p_{f\,95\%}^{-1}}\right) \leq \varepsilon_{p_f\,0}. \tag{11}$$

En pratique, on visera des étendues logarithmiques $\varepsilon_{p_f\,0}$ inférieures à 1, ce qui signifie que l'estimation de la probabilité de défaillance est au moins dans l'ordre de grandeur de la probabilité réelle.

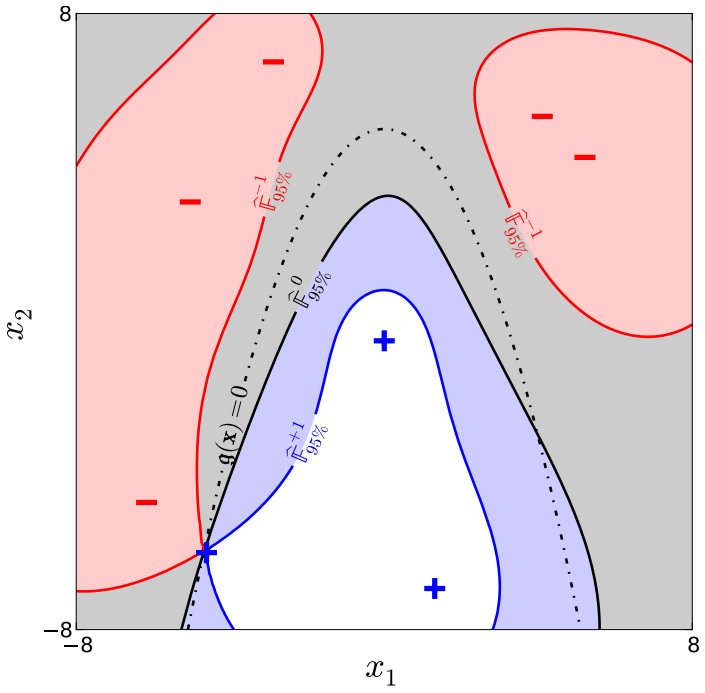


Figure 3 – Domaines de défaillance approchés pour l'approximation de l'erreur sur la probabilité de défaillance.

Quantification du biais par échantillonnage préférentiel

Afin de quantifier (et donc éliminer) le biais sur la probabilité de défaillance, on propose une *méthode de fiabilité hybride* qui utilise une prédiction par Krigeage pour approcher une estimation optimale de la probabilité de défaillance par échantillonnage préférentiel.

L'idée de l'échantillonnage préférentiel, pour l'analyse de fiabilité, est de favoriser l'échantillonnage dans le domaine de défaillance de sorte que même si la probabilité de défaillance est faible elle pourra être estimée précisément avec un nombre réduit de simulations. Pour ce faire, on vient réécrire la probabilité de défaillance comme suit :

$$p_f = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\,\frac{f_X(\boldsymbol{x})}{h(\boldsymbol{x})}\,h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \tag{12}$$

où $\mathbb{1}_{\mathbb{F}}$ est la fonction indicatrice de défaillance et h est la densité dite *instrumentale* qui va être utilisée pour estimer la probabilité de défaillance par simulations à la place de la densité originale $f_{\boldsymbol{X}}$. La densité instrumentale optimale pour réduire la variance d'estimation à zéro est donnée par Rubinstein et Kroese (2008) et s'écrit comme suit :

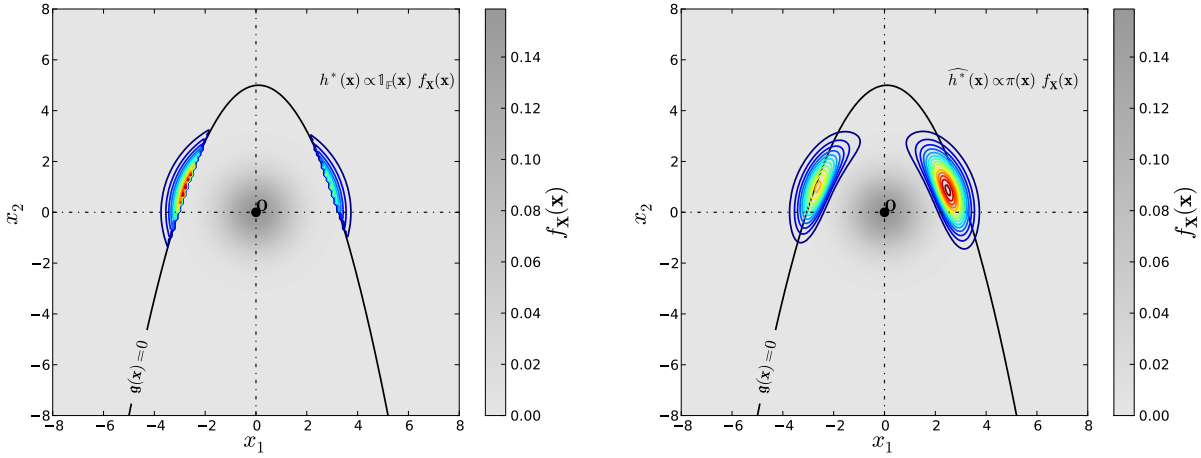
$$h^*(\boldsymbol{x}) = \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}} = \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}{p_f}. \tag{13}$$

Cependant, elle est inutilisable en pratique car elle requiert la connaissance *a priori* de la probabilité de défaillance p_f (au dénominateur de l'Eq. (13)) que l'on cherche précisément à estimer.

C'est pourquoi on propose ici d'approcher la densité instrumentale optimale par une autre densité *quasi-optimale* (mais utilisable) définie comme suit :

$$\widehat{h^*}(\boldsymbol{x}) = \frac{\pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\mathbb{X}} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}}. \tag{14}$$

Cette densité est construite à partir de la fonction de classification probabiliste π d'une prédiction par Krigeage de la fonction de performance $\mathfrak{g}$ (voir Eq. (5)). Les deux densités sont comparées sur la Figure 4. La densité optimale revient à *conditionner* l'échantillonnage par l'appartenance des individus simulés au domaine de défaillance. L'approximation, elle, ne fait que *favoriser* cette appartenance.



(a) Densité instrumentale optimale h^* (b) Approximation proposée $\widehat{h^*}$

FIGURE 4 – Comparaison entre la densité instrumentale optimale (inutilisable) et son approximation par Krigeage.

Finalement, en réinjectant la densité instrumentale proposée dans l'Eq. (12), on montre que la probabilité de défaillance se réécrit comme suit :

$$p_f = p_{f\,\varepsilon}\, \alpha_{\text{corr}}, \tag{15}$$

où $p_{f\,\varepsilon}$ est la probabilité de défaillance augmentée estimée sur la prédiction par Krigeage uniquement, et α_{corr} est le terme correctif qui vient quantifier le biais recherché. Le terme

correctif nécessite d'effectuer des simulations à la fois sur la prédiction par Krigeage et sur le modèle réel. Néanmoins, s'il est proche de 1 (absence de biais), il est très peu coûteux à estimer. On peut ainsi construire un estimateur non biaisé de la probabilité de défaillance basé sur une prédiction par Krigeage. On propose également une autre métrique que celle formulée à l'Eq. (11) pour stopper l'enrichissement adaptatif lorsque α_{corr} est proche de 1 et ainsi s'assurer de la quasi-optimalité de $\widehat{h}^*$.

▷ *Cette partie est développée dans le Chapitre 3.*

Optimisation sous contrainte de fiabilité utilisant Krigeage adaptatif et *subset sampling*

Pour résoudre le problème d'optimisation sous contrainte de fiabilité formulé initialement à l'Eq. (2), on se propose d'utiliser un algorithme issu de la programmation non-linéaire (Polak, 1997). Plus spécifiquement, on a choisi d'utiliser l'algorithme de Polak-He qui utilise une formulation de type min-max (*i.e.* qui consiste à minimiser la contrainte la moins respectée). Cet algorithme procède itérativement en deux étapes en améliorant une solution initiale $\boldsymbol{d}^{(0)}$. Il s'agit de déterminer *(i)* la meilleure direction d'optimisation au voisinage de la solution courante, et *(ii)* le meilleur pas à effectuer suivant cette direction. La direction est déterminée en résolvant un sous-problème d'optimisation quadratique approché à partir des gradients des fonctions objectif et contraintes. Le pas optimal est obtenu par une recherche linéique approchée au moyen de la règle de Goldstein-Armijo.

Analyse de sensibilités fiabilistes par simulations

L'utilisation de l'algorithme de Polak-He requiert le calcul des gradients des fonctions objectif et contraintes et *a fortiori* ceux des contraintes de fiabilité. Le calcul des gradients des probabilités de défaillance a donc fait l'objet de recherches bibliographiques. On a finalement retenu une approche initialement proposée par Rubinstein (1986) (voir aussi Wu, 1994; Rubinstein et Kroese, 2008) s'appuyant sur le concept de l'échantillonnage préférentiel. Cette approche est avantageuse car elle permet de calculer les dérivées partielles des probabilités à partir des résultats d'une analyse de fiabilité sans nécessiter de nouvelles simulations. De plus, l'approche peut également être utilisée avec des techniques de simulations par réduction de variance tel que l'échantillonnage préférentiel ou le *subset sampling* (Song et al., 2009).

Méta-modèles adaptatifs dans l'espace augmenté

Afin de réduire le coût de calcul induit par les analyses de fiabilité répétées au sein de la boucle d'optimisation, on remplace les fonctions de performance réelles par leurs prédictions par Krigeage. Ces prédictions sont construites dans un espace global qui permet de les réutiliser d'une itération d'optimisation à l'autre. Cet espace baptisé *"espace augmenté"* par Taflanidis (2007) est obtenu en considérant que l'incertitude sur les choix de conception

ne fait qu'augmenter l'étendue de la distribution des variables aléatoires sans changer la dimension du vecteur $\boldsymbol{x}$. En effet, en rappelant que les variables $\boldsymbol{d}$ définissent la densité conjointe du vecteur aléatoire $\boldsymbol{X}$, le *vecteur aléatoire augmenté* $\boldsymbol{V}$ a pour densité de probabilité :

$$f_V(\boldsymbol{v}) = \int_{\mathbb{D}} f_X(\boldsymbol{x} \mid \boldsymbol{d}) \, \pi(\boldsymbol{d}) \, \mathrm{d}\boldsymbol{d}, \tag{16}$$

où π est la densité de probabilité uniforme sur le domaine de conception admissible $\mathbb{D}$. Ensuite, il a fallu s'assurer que les prédictions par Krigeage des états-limites étaient suffisamment précises sur une région de confiance suffisamment large de la densité augmentée f_V.

▷ *Cette partie est développée dans le Chapitre 4.*

Résultats

Validation sur des exemples académiques

Toutes les stratégies proposées ont fait l'objet d'applications numériques sur des exemples simples issus de la littérature fiabiliste en mécanique de sorte que les calculs peuvent être reconduits par le lecteur. Ces applications ont permis de démontrer les propriétés théoriques énoncées ci-avant. On s'est notamment assuré que :

- la stratégie d'enrichissement adaptative par échantillonnage et *clustering* permet de réduire l'erreur de substitution du méta-modèle dans une analyse de fiabilité ;
- l'étendue du pseudo-intervalle de confiance sur la probabilité de défaillance estimée sur une prédiction par Krigeage permet une quantification suffisamment objective pour déclencher ou stopper l'enrichissement adaptatif ;
- l'estimateur par échantillonnage préférentiel permet de quantifier et éliminer le biais sur la probabilité de défaillance et permet également d'étendre le champ d'application du Krigeage à des problèmes de fiabilité de dimension raisonnablement grande (*i.e.* jusqu'à quelques dizaines de variables aléatoires) ;
- les sensibilités fiabilistes par simulation et les méta-modèles construits dans l'espace augmenté ont permis de résoudre le problème d'optimisation sous contrainte de fiabilité en quelques centaines d'appels au(x) modèle(s) mécanique(s).

Application au dimensionnement au flambement des coques imparfaites

Fort de ces applications académiques, la démarche de conception optimale sous contrainte de fiabilité a été appliquée au dimensionnement de coques imparfaites sensibles au phénomène de flambement. On s'est intéressé à la fiabilité d'un toit cylindrique dont les propriétés

sont modélisées par des champs aléatoires, et au dimensionnement fiabiliste d'une maille cylindrique raidie représentative d'une coque de sous-marin. La substitution d'un méta-modèle au modèle numérique de prédiction de la charge critique de flambement s'avère ici primordiale. Elle aura finalement permis de trouver des solutions de dimensionnement optimales associées à un haut niveau de fiabilité (*i.e.* $p_f \leq 10^{-9}$) tel que prescrit par les donneurs d'ordres dans ce domaine, et ce en un nombre raisonnable d'appels au modèle mécanique.

▷ *Les applications sont présentées dans les Chapitres 5 et 6.*

Conclusions et perspectives

Finalement, on a pu résoudre des problèmes d'optimisation sous contrainte de fiabilité faisant appel à des modèles numériques coûteux en un coût de calcul global raisonnable (*i.e.* moins de 1 000 appels au(x) modèle(s)). On s'est de plus attaché à maîtriser les erreurs induites par les hypothèses effectuées et notamment par la substitution d'un méta-modèle au modèle réel.

Ce travail ouvre de multiples perspectives. D'abord, il s'agit d'appliquer la méthode d'optimisation sous contrainte de fiabilité à de nouveaux cas de dimensionnement. Si la dimension du vecteur aléatoire $\boldsymbol{X}$ devient trop importante pour simplement substituer les prédictions par Krigeage aux fonctions de performance, il faudrait envisager l'utilisation de la technique d'échantillonnage préférentiel quasi-optimal au sein de la boucle d'optimisation. On pourra aussi chercher à optimiser d'autres valeurs caractéristiques que des valeurs moyennes. Par exemple, en tolérancement probabiliste, on chercherait plutôt à régler des coefficients de capabilité (proportionnels aux écarts-types des variables) de façon à garantir un taux de conformité cible (*i.e.* une probabilité) tout en minimisant les coûts de fabrication.

On pourrait également envisager d'appliquer la stratégie adaptative d'enrichissement de plans d'expériences par échantillonnage et *clustering* à des problèmes d'optimisation sous contraintes déterministes en se basant sur les critères d'enrichissement proposés par Mockus (1994) ou Jones et al. (1998).

Contents

Introduction

"As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality." A. Einstein

Context

Simulation is the sound basis of modern engineering. Its application field ranges from physical sciences to financial engineering. It basically consists in imitating the behaviour of a physical or abstract system using *mathematical models*. The fidelity of the latter models has been significantly improved during the past few decades. However, due to their lingering lack of comprehensiveness, discrepancies are still observed between simulation and reality.

By definition, models are idealizations of reality and thus neglect some aspects of the phenomenon of interest. One major contribution to the discrepancies between simulation and reality comes from the isolation of the system from its environment. Indeed, when studying a system, the environment in which it evolves is often reduced to a set of idealized (say *nominal*) configurations. For instance, in structural mechanics, the boundary conditions, the initial state of the system, its geometry, the properties of its constitutive material and the loads are reduced to a finite set of configurations (*e.g.* load cases) which may lack exhaustiveness. The reason for proceeding this way is mostly a matter of model scale and lack of knowledge. Indeed, for the sake of tractability, a model cannot account for the real world as a whole, and it has to summarize the immediate environment surrounding the system under study to a synthetic mathematical model.

In this context, the probabilistic framework seems more appropriate as it enables the consideration of *an infinite set of probable configurations* in the form of a single probability distribution. However, the probabilistic specification of a system also raises fundamental issues about *(i)* the construction of appropriate *probabilistic models* and *(ii)* the *metrics* to be used for assessing its performance, which is the basis of decision-making.

Fortunately, in the last two decades, a lot of work has been done in *probabilistic engineering* in order to formalize these problems and educate engineers to apprehend them. Indeed, the scientific literature now provides a large set of techniques for *uncertainty quantification*, and engineering students[1] are now made aware of probability theory and its potential applications to industrial problems.

[1] As far as the author is concerned.

Problem statement

This thesis is concerned with the problem of *structural design optimization in the presence of uncertainty* such as introduced in the book by Tsompanakis et al. (2008). It is more often referred to as *robust engineering* in the industry although this term refers to a specific formulation of the latter more general problem. Indeed, there exists two main formulations for the optimization of uncertain systems. *Robust design optimization* (RDO) primarily seeks to minimize the influence of the uncertain environmental conditions on the performance of a system. In contrast, the goal of *reliability-based design optimization* (RBDO) is to design for safety with respect to extreme (rare) events.

This manuscript focuses on the reliability-based formulation. Reliability is typically measured by a *failure probability* and the optimization consists in ensuring that the latter remains lower than a given threshold. The choice of this threshold depends on the system to be designed. For instance, in the field of part tolerancing a commonly admitted unit for assessing the non-conformity rates is the number of rejected *parts per million* (ppm). In other fields directly impacting human lives (such as the nuclear industry), the selection of an acceptable probability of failure is much more serious and beyond the technical scope of this manuscript.

The present work is primarily aimed at setting up a methodology for solving the reliability-based design optimization problem within a reasonable computational effort. Indeed, the estimation of low (admissible) failure probabilities through Monte Carlo simulation requires a large number of model evaluations which is often incompatible with the available computational resources. Nesting the estimation of such probabilities into a design optimization process makes the RBDO problem intractable when fine numerical models are used for assessing the performance of the system.

This remark motivated most of the work done in the past few years on RBDO, although most approaches are limited by the use of restrictive assumptions such as the linearity of the frontier (the limit-state surface) between safe and failed designs. The approach that is investigated in this manuscript relies on the use of the *statistical learning theory* as recently proposed by Hurtado (2004b). Statistical learning techniques aim at constructing emulators (or meta-models) of the original physical models that are much faster to evaluate. Nonetheless, the use of these emulators as surrogates for the original expensive-to-evaluate physical models raises a new issue about the bias introduced in estimating the failure probability. Indeed, despite this bias can be reduced by the use of meta-model refinement techniques, it remains hard to quantify.

Objectives and outline of the thesis

The reliability-based design optimization strategy developed in the sequel is aimed at satisfying the following three objectives:

(i) the computational procedure must be *parsimonious* with respect to the total number

of simulations of the physical models (less than a few thousands would be appreciated), for the sake of efficiency;

(ii) the use of emulators should be introduced in such a way that the potential bias on the final quantities of interest (namely the design and its failure probability) can be quantified, for the sake of safety;

(iii) the overall strategy is expected to remain applicable for admissible failure probabilities ranging from 10^{-2} down to 10^{-12}, for the sake of versatility with respect to industrial concerns.

The first objective is addressed in Chapters 1 and 2, while the other two underlie the reliability-based design methodologies developed in Chapters 3 and 4.

Chapter 1 intends to make the reader aware of a few elements of *statistical learning* theory. This chapter is mostly devoted to the presentation of the Gaussian process prediction (or Kriging) methodology such as presented in the related literature. A specific emphasis is put on the interrelation of this supervised learning technique with the more conventional least-squares regression methodology. This chapter does not feature any original contribution, although the author's interpretation of the Gaussian process predictor matters for the understanding of the methodologies developed in the subsequent chapters.

Chapter 2 is concerned with the selection of the experiments from which the Gaussian process predictors are built. The purpose is to select the minimal design of experiments for ensuring a good relative accuracy of the predictors with respect to the original models. The recent literature features a large amount of papers dealing with this problem so that a large part of Chapter 2 is devoted to a review of the so-called *adaptive design of experiments* techniques. Building on this review a novel refinement strategy is proposed. It concentrates on *the use* of the so-called *refinement criteria* (as a sampling density) rather than on the development of new ones as typically addressed in the literature.

Chapter 3 is devoted to the presentation of structural reliability methods. The purpose of these methods is to estimate the probability of failure of a structure given a probabilistic model of the system and its environment, and a failure scenario. State-of-the-art reliability methods are first being reviewed from an original point of view. Indeed, the stress is put on their interrelation with the so-called *importance sampling* estimation technique. Since the most robust of the latter techniques (namely subset sampling) still requires a few tens of thousands of calls to the physical model, surrogate-based approaches are then investigated for the sake of efficiency. As expected, these approaches reveal a lot more computationally efficient but they might introduce a significant bias in the failure probability estimates. Two heuristic error measures are thus proposed based on the variance of prediction of the Gaussian process predictors. Eventually, a novel hybrid approach named *meta-model-based importance sampling* is proposed in an attempt to get rid of the latter bias.

In **Chapter 4**, the reliability-based design optimization problem is explicitly formulated. This formulation closely follows the one used in the literature and it is proposed to solve it by means of a gradient-based optimization algorithm. After having introduced a few elements of inequality constrained optimization, this chapter focuses on the computation

of the gradient of the failure probability with respect to the design variables. The design variables being considered here as parameters (means) of the probabilistic model, this computation is performed by means of the so-called *score function approach*. Thanks to this importance-sampling-like trick, the computation reveals a simple post-processing of a sampling-based reliability analysis. In order to circumvent the computational burden induced by the nested reliability analyses, the physical models involved in the performance probabilistic constraints are then replaced by adaptive Kriging surrogates. These surrogates are built and refined in a so-called *augmented reliability space* that makes them reusable from one optimization iteration to the other. The overall approach was named *meta-model-based RBDO*.

Chapter 5 validates the three main contributions of this thesis on a chosen set of academic examples. Besides the meta-model-based importance sampling and RBDO strategies, the performance of the proposed sampling-based approach to adaptive design of experiments is also investigated. At last, **Chapter 6** applies the overall reliability-based design philosophy to the design of imperfect shells prone to buckling. This last application involves the use of an expensive-to-evaluate nonlinear finite element model.

Note that this manuscript is intended to be read in its chronological order as each chapter features fundamental concepts that are constantly being referred to in the subsequent ones. A deep understanding of Chapter 1 is not necessary though, even if it is important to remember the fundamental results about Gaussian process (or Kriging) predictors.

1

Gaussian process meta-modelling

Contents

1.1 Introduction

1.1.1 Statistical learning

In this thesis, *meta-modelling* refers to the mathematical discipline that has interest in emulating the statistical relationship between some input $\boldsymbol{x} \in \mathbb{X} \subseteq \mathbb{R}^n$ and some output $\boldsymbol{y} \in \mathbb{Y} \subseteq \mathbb{R}^d$ of some given mapping $\mathscr{M}$ from a set of *observations* $\mathscr{D} = \left\{(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}), i = 1, \dots, m\right\}$ and a limited set of *prior assumptions*. This is more often known as *supervised learning* in the *statistical learning* related literature (Rasmussen and Williams, 2006) as opposed to the *unsupervised learning* discipline which has interest in building models for a multivariate input only, *i.e.* without any supervision variable y. Throughout this manuscript, the supervised learning problem will be restricted to univariate (scalar) outputs meaning that $d = 1$. Consequently, the observations in the dataset are gathered in a vector $\boldsymbol{y} = (y_i, i = 1, \dots, m)^\intercal$. From now on, a meta-model (or emulator) will be denoted by $\widetilde{\mathscr{M}}$.

The motivations for the construction of such meta-models mostly depend on the field of application. In physics-oriented fields (which are in the scope of this manuscript), the emulator can be used as a *surrogate* of an existing but expensive-to-evaluate high fidelity *physical model* such as a finite element model. In this case, the emulator offers the interest of being both much faster to evaluate than the original model, and usually costless: indeed they are based on usual mathematical functions and do not require licensed software for their evaluation. In other fields that lack models to explain some observed phenomenon of interest (*e.g.* in medical imaging, pattern/speech recognition, Internet search engines, *etc.*), emulators may constitute the sole *empirical model* available yet.

The supervised learning discipline is usually split into two other sub-disciplines known as *regression* and *classification*. The distinction is based on the nature of the supervision variable y. In a regression problem, y spans a *continuous* input space $\mathbb{X} \subseteq \mathbb{R}^n$ as opposed to the classification problems where it takes its values in a *discrete* set of *labels* $\left\{L^{(\ell)}, \ell = 1, \dots, n_{\text{labels}}\right\}$. Note that in the context of reliability estimation, the classification problem is restricted to the so-called binary case for which the number of labels is $n_{\text{labels}} = 2$ since a design is either *failed* or *safe*.

1.1.2 A short state-of-the-art

There exists a wide variety of emulators in the supervised learning literature, amongst which are: general linear models, support vector machines and Gaussian process predictors. The work presented in this manuscript makes use of Gaussian process predictors so that this chapter is mostly devoted to the presentation of this regression technique. However the two other techniques are shortly reviewed and it is then argued why it has been decided to work with Gaussian process predictors in the sequel.

1.1.2.1 General linear models

The common assumption to *general linear models* is that they are defined as a linear combination of a finite set of p preselected functions $\boldsymbol{f} = \{f_i, i = 1, \ldots, p\}$:

$$\widetilde{\mathscr{M}}(\boldsymbol{x}) = \sum_{i=1}^{p} \beta_i f_i(\boldsymbol{x}) = \boldsymbol{\beta}^\mathsf{T} \boldsymbol{f}(\boldsymbol{x}), \tag{1.1}$$

where $\boldsymbol{\beta} = \big(\beta_i, i = 1, \ldots, p\big)^\mathsf{T} \in \mathbb{R}^p$ is a vector of weight coefficients to be determined from the dataset $\mathscr{D}$.

Note that such models sometimes feature a constant bias β_0 as an additional term, although this term can be incorporated in the functional set $\boldsymbol{f}$ by adding the constant function equal to one for any $\boldsymbol{x} \in \mathbb{X}$. The adjective *general* means here that the proposed model remains linear in the transformed input $\boldsymbol{f}(\boldsymbol{x})$ for any functional set $\boldsymbol{f}$. Commonly used functional sets mostly include polynomial, Fourier and wavelet series. In Figure 1.1, a polynomial basis is used to emulate the model $\mathscr{M}(x) = x\ \sin(x)$.

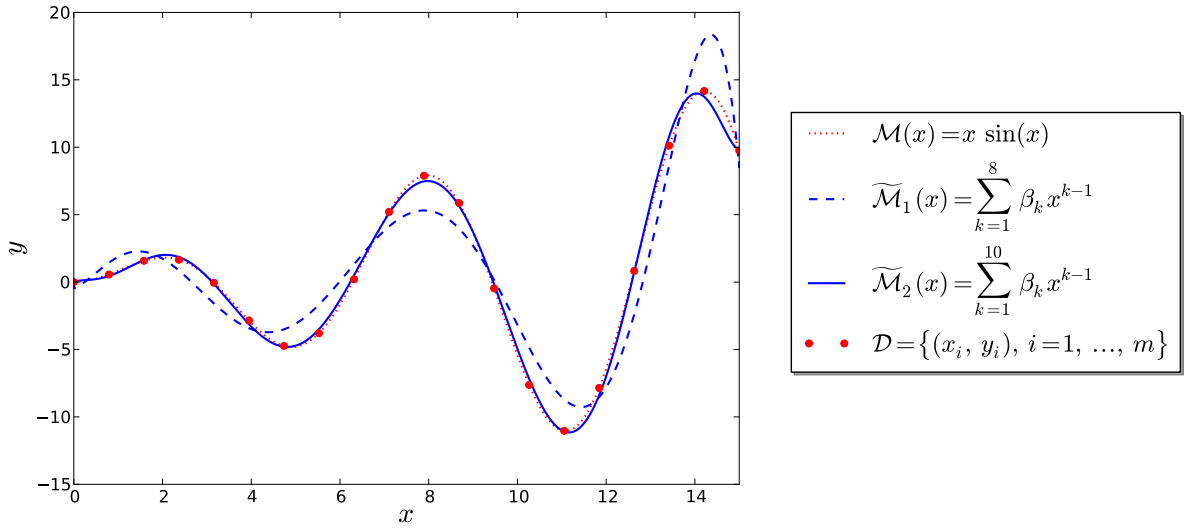


Figure 1.1: Two general linear models emulating the one-dimensional model $\mathscr{M}(x) = x\ \sin(x)$. They both use the full polynomial basis $\boldsymbol{f} = \left\{x^{k-1}, k = 1, \ldots, p\right\}$ where the order p is set equal to 8 for $\widetilde{\mathscr{M}_1}$ and 10 for $\widetilde{\mathscr{M}_2}$. The coefficients are computed by means of the ordinary least squares technique.

There exists many techniques to determine the unknown weight coefficients $\boldsymbol{\beta}$, the most famous being the least-squares regression technique which will be reviewed in more details in Section 1.2. However this technique requires that the size m of the dataset $\mathscr{D}$ is greater than the size p of the functional set $\boldsymbol{f}$. Moreover this size p conditions the capacity of the linear model in Eq. (1.1) to capture the nonlinearity in the original model $\mathscr{M}$. Starting from this premise, Efron et al. (2004) propose the least angle regression (LAR) procedure. This procedure enables a selection of the most significant functions in $\boldsymbol{f}$ to emulate the original

experiment $\mathscr{M}$ while avoiding the overfitting phenomenon which occurs as p approaches m. The final general linear model that is fitted by this approach involves a reduced number of non-zero regression coefficients $\{\beta_j \neq 0, j = 1, \ldots, p_{\text{non-zero}}\}$ with $p_{\text{non-zero}} \ll p$, and the corresponding functional set $\{f_j, j = 1, \ldots, p_{\text{non-zero}}\}$ is said to be *sparse*.

Despite such models proved efficient on a large set of examples, including uncertainty quantification examples, it is argued here that they are not well suited to the classification problem underlying a reliability analysis. This is the reason why it has been decided to move to kernel-based models such as support vector machines or Gaussian process predictors because they allow a better local refinement for approximating the limit-state surface which is defined as a specific contour at level $t \in \mathbb{Y}$:

$$\mathscr{S}_t = \{\boldsymbol{x} \in \mathbb{X} : \mathscr{M}(\boldsymbol{x}) = t\}. \tag{1.2}$$

1.1.2.2 Support vector machines

Support vector machines is certainly one of the most popular supervised learning techniques in the last two decades literature thanks to its computational advantages. It is used in a large field of applications ranging from Internet search engines to medical imaging, and it was recently brought up to the structural reliability community by Hurtado (2004b). The technique was initially designed to solve binary classification problems but it was then extended to multi-class and regression problems, although it does not feature the same efficiency for regression than for classification.

Consider a set of labeled observations $\mathscr{D} = \left\{(\boldsymbol{x}^{(i)}, L_i), i = 1, \ldots, m\right\}$ where $L_i = \pm 1$. Support vector binary classifiers assume that the two classes are *linearly separable* in some transformed input space $\mathbb{H}(K)$ named the *feature space*. This feature space is obtained through a transformation $\boldsymbol{\varphi}$ such that the inner product in that space can be computed by means of its (given) associated kernel function K in the former input space $\mathbb{X}$:

$$\langle \boldsymbol{\varphi}(\boldsymbol{x}), \boldsymbol{\varphi}(\boldsymbol{x}') \rangle_{\mathbb{H}(K)} \equiv K(\boldsymbol{x}, \boldsymbol{x}'), \quad (\boldsymbol{x}, \boldsymbol{x}') \in \mathbb{X}^2. \tag{1.3}$$

This relation is often referred to as the *kernel trick* due to the article by Aizerman et al. (1964). Admissible kernel functions satisfy Mercer's condition, meaning that they must be positive definite alike the forthcoming autocovariance functions for Gaussian processes[1] (Section 1.3). One widely used kernel function is the squared exponential kernel which is defined as:

$$K(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-\gamma \left(\boldsymbol{x}' - \boldsymbol{x}\right)^2\right), \quad (\boldsymbol{x}, \boldsymbol{x}') \in \mathbb{X}^2, \tag{1.4}$$

where $\gamma \in \mathbb{R}^{+*}$ is its so-called hyper-parameter. Formally, in functional analysis, $\mathbb{H}(K)$ is the reproducing kernel Hilbert space generated by the continuous real-valued kernel K.

The transformation $\boldsymbol{\varphi}$ features two essential properties for the classification problem at hand. The first one is that its explicit knowledge is not required to compute distances in the feature space $\mathbb{H}(K)$ thanks to Eq. (1.3). The other one is that the implicit mapping $\boldsymbol{\varphi}$ ensures *dimensionality explosion* (Hurtado, 2004b), meaning that the number of components in the

[1] As opposed to autocovariance functions Mercer's kernel may be non-symmetric though.

feature space is greater than the one in the original input space and theoretically infinite for some kernel functions (namely *strictly* positive definite kernel functions) like the squared exponential kernel. This *dimensionality explosion* is used to facilitate the linear classification in the feature space.

For the binary classification problem of interest, the linear boundary between the two classes is the hyperplane (in the feature space) of equation:

$$\begin{aligned} S(\boldsymbol{x}) &= \langle \varphi(\boldsymbol{w}), \varphi(\boldsymbol{x}) \rangle_{\mathbb{H}(K)} + b \\ &= K(\boldsymbol{w}, \boldsymbol{x}) + b. \end{aligned} \tag{1.5}$$

Hence, solving the classification problem again consists in finding the weight coefficients $\boldsymbol{w}$ and the intercept b. Another key concept of support vector machines is the *maximum margin* concept illustrated in Figure 1.2. The margin refers here to the distance (in the feature space) between the two classes.

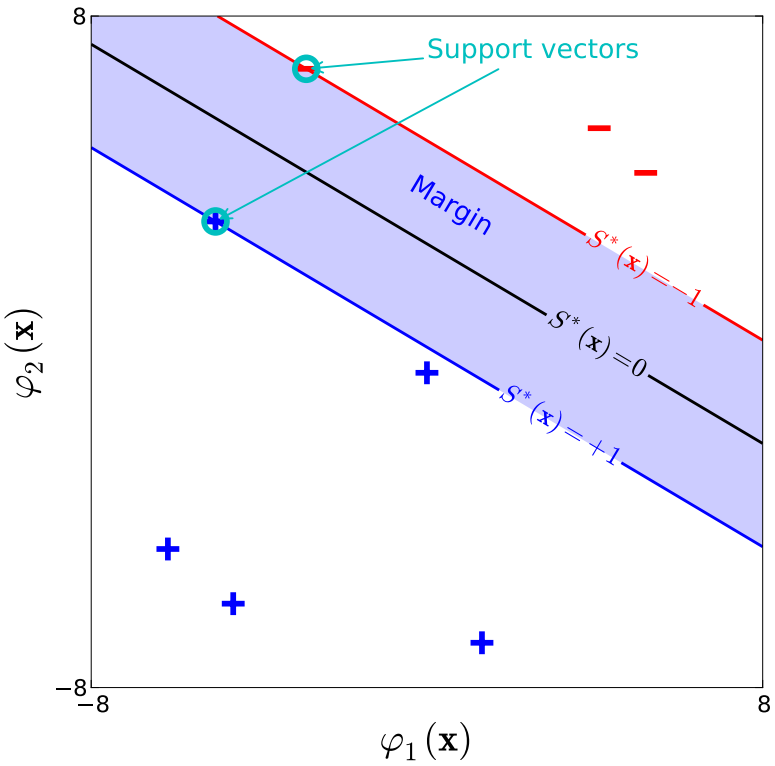


Figure 1.2: The optimal separating hyperplane in the feature space maximizes the margin (*i.e.* the distance) between the two classes.

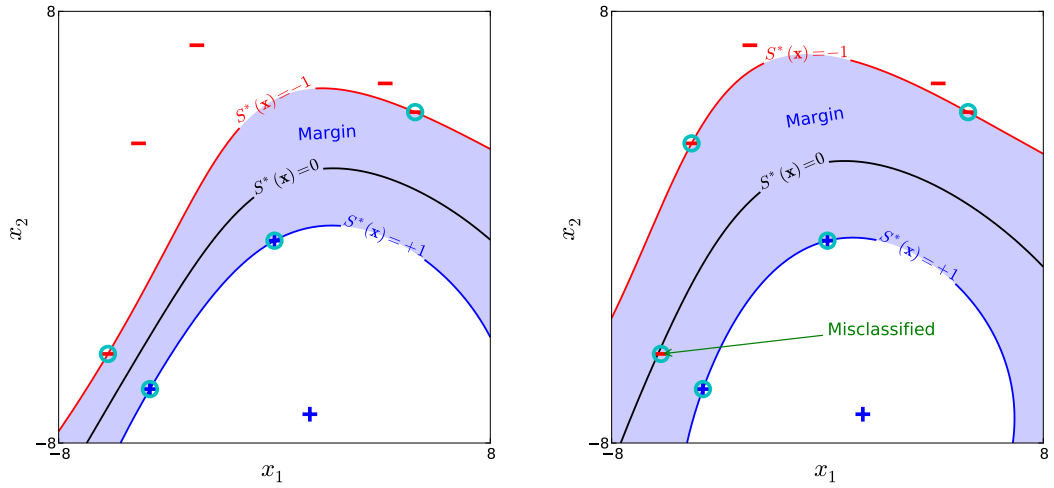
This distance is shown (see *e.g.* Gunn, 1998, pp. 6–7) to be proportional to the inverse of the norm (in the feature space) of the weight coefficients $K(\boldsymbol{w}, \boldsymbol{w})^{-1/2}$. Minimizing the half-quadratic norm $K(\boldsymbol{w}, \boldsymbol{w})/2$ is thus equivalent to maximizing the margin but it considerably simplifies the optimization. In addition to the maximum margin, one adds

right classification constraints so that the points in the dataset are accurately classified. All together, this lead to the establishment of the following constrained optimization problem:

$$\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} \frac{K(\boldsymbol{w},\boldsymbol{w})}{2} \quad \text{s.t.} \quad L_i\left(K(\boldsymbol{w},\boldsymbol{x}^{(i)})+b\right) \geq 1, \quad i=1,\ldots,m, \tag{1.6}$$

which is a linearly constrained quadratic optimization problem. As a matter of fact, it admits a unique solution $\boldsymbol{w}^*$ which comes analytically by introducing its associated Lagrangian and solving its first-order optimality conditions (see *e.g.* the technical report by Gunn, 1998).

In addition to their computational efficiency, one final interesting feature of support vector classifiers is their *sparsity* meaning that the final optimal separating hyperplane S^* is defined as a sum over a reduced dataset whose elements are referred to as the *support vectors*. Such support vectors are circled in cyan in Figures 1.2 and 1.3.



(a) Hard margin (b) Soft margin

Figure 1.3: The soft margin classifier was obtained with a penalty coefficient $C = 100$ whereas the "hard margin" classifier has indeed a soft margin except the penalty coefficient $C = 10,000$ is more severe. The hard margin does not contain any support vector whereas the soft one does. As a matter of fact soft margin classifiers do not guarantee the right classification of such points.

However, in the case of extremely nonlinear boundaries for which the dimensionality explosion provided by the kernel trick does not capture a sufficient amount of nonlinearity, one resorts to the so-called *soft margin classifiers*. It consists in authorizing a classification error by means of *slack variables* $\{\xi_i, i=1,\ldots,m\}$ and the optimization problem is recast as follows:

$$\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} \frac{K(\boldsymbol{w},\boldsymbol{w})}{2} + C\sum_{i=1}^{m}\xi_i \quad \text{s.t.} \quad L_i(K(\boldsymbol{w},\boldsymbol{x}^{(i)})+b) \geq 1-\xi_i, \quad i=1,\ldots,m, \tag{1.7}$$

where C is a given constant whose purpose is to penalize classification errors. The resolution of this problem is somewhat complicated by the m additional slack variables although

it can still be efficiently solved by means of the usual *sequential quadratic programming* algorithm (SQP). Nevertheless, it is worth mentioning that the sparsity property of the final separating hyperplane suffers from these additional soft constraints. Most practical applications of support vector classifiers use the latter more comprehensive formulation. Hard and soft margin classifiers are opposed in Figure 1.3.

Support vector regressors, whose purpose is to approximate real-valued models $\mathscr{M}$, use a similar formulation with a different (tunable) penalty term though (see *e.g.* Schölkopf and Smola, 2003; Smola and Schölkopf, 2006).

1.2 Generalized least-squares linear regression

The least-squares linear regression problem is certainly the most popular method for regression analysis due to its conceptual simplicity: it consists in minimizing the distance (in the classical $\mathscr{L}_2$ sense) between a predictive model and experimental measurements. Thanks to its computational efficiency, it arises as a subproblem in many disciplines. For instance:

- it is mainly used for model inference from *in situ* measurements;
- it has been widely used in the field of *uncertainty quantification* to fit *surrogate* models such as polynomial response surfaces (*e.g.* in Bucher and Bourgund, 1990) or polynomial chaos expansions (*e.g.* in Berveiller et al., 2006; Sudret, 2007; Blatman and Sudret, 2008b, 2010a).

In this section the problem is first reviewed from a frequentist viewpoint, and then from a Bayesian viewpoint as an introduction to the forthcoming Section 1.4 where it is used as the first stage of a two-stage prior model as in the book by Santner et al. (2003).

1.2.1 Elements of probability theory

As the vector of observed output $\boldsymbol{y} = (y_i, i = 1, \ldots, m)^\mathsf{T}$ will be considered uncertain, a *probabilistic framework* is now introduced.

1.2.1.1 Random variables

Let $(\Omega, \mathscr{F}, \mathscr{P})$ denote a *probability space* where Ω is the *event space* equipped with its σ-*algebra* $\mathscr{F}$ and a *probability measure* $\mathscr{P}$.

Throughout this manuscript, the probability measure $\mathscr{P}$ is dedicated to the epistemic uncertainty *associated with the construction of the meta-model $\widetilde{\mathscr{M}}$ whereas the probability measure $\mathbb{P}$ will later be devoted to the* aleatoric uncertainty *affecting the input vector $\boldsymbol{x}$ in a reliability analysis context.*

A real-valued random variable Y is formally defined as an application that maps the probability space $(\Omega, \mathscr{F}, \mathscr{P})$ in the *output space* $\mathbb{Y} \subseteq \mathbb{R}$:

$$Y : \left| \begin{array}{rcl} (\Omega, \mathscr{F}, \mathscr{P}) & \mapsto & \mathbb{Y} \\ \omega & \to & y \end{array} \right. , \tag{1.8}$$

where $y = Y(\omega)$ denotes a *realization*. Y may either be *discrete* or *continuous* depending on the nature of its output space $\mathbb{Y}$. A random variable is completely defined by its *cumulative distribution function* (CDF):

$$F_Y(y) = \mathscr{P}\left[Y \leq y\right]. \tag{1.9}$$

Continuous random variables additionally admit a *probability density function* (PDF) defined as follows:

$$f_Y(y) = \lim_{\substack{h \to 0 \\ h > 0}} \frac{\mathscr{P}\left[y \leq Y \leq y + h\right]}{h}, \tag{1.10}$$

hence:

$$f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y}. \tag{1.11}$$

The *mathematical expectation* will be denoted by $\mathbb{E}$, so that the mean value of the random variable Y is defined as:

$$\mu_Y = \mathbb{E}[Y] = \int_{\mathbb{Y}} y\, f_Y(y)\, \mathrm{d}y. \tag{1.12}$$

It is sometimes also denoted by $\mathbb{E}_Y$ to emphasize that the expectation is taken with respect to the random variable Y only. The moments (resp. centred moments) of given order $n > 1$ are defined as:

$$\mathbb{E}[Y^n] = \int_{\mathbb{Y}} y^n f_Y(y)\, \mathrm{d}y, \tag{1.13}$$

$$\mathbb{E}\left[(Y - \mu_Y)^n\right] = \int_{\mathbb{Y}} (y - \mu)^n f_Y(y)\, \mathrm{d}y, \tag{1.14}$$

provided these integrals exist. In particular, the centred second-order moment denoted by $\sigma_Y^2 = \mathrm{Var}[Y]$ is known as the *variance* and its square root σ_Y is the *standard deviation*. Provided the mean value is non-zero, the *coefficient of variation* is defined as the ratio between the standard deviation and the absolute mean value $\delta_Y \equiv \sigma_Y / |\mu_Y|$.

The covariance between two random variates X and Y is defined as follows:

$$\mathrm{Cov}[X, Y] = \mathbb{E}\left[(X - \mu_X)(Y - \mu_Y)\right], \tag{1.15}$$

and the corresponding correlation coefficient is defined as:

$$\rho_{XY} = \frac{\mathrm{Cov}[X, Y]}{\sigma_X \sigma_Y}. \tag{1.16}$$

The *vectorial space* of *second-order* real-valued random variables is denoted by $\mathscr{L}_2(\Omega, \mathscr{F}, \mathscr{P}; \mathbb{R})$ and a real-valued random variable Y belongs to this space if and only if it admits at least a finite second-order moment ($\mathbb{E}\left[Y^2\right] < \infty$). The expectation operator defines an *inner product* on this space:

$$\langle X, Y \rangle = \mathbb{E}\left[X\,Y\right], \quad (X, Y) \in \mathscr{L}_2(\Omega, \mathscr{F}, \mathscr{P}; \mathbb{R})^2, \tag{1.17}$$

so that the vectorial space $\mathscr{L}_2(\Omega, \mathscr{F}, \mathscr{P}; \mathbb{R})$ dotted with this inner product is a *Hilbert space*. In particular, two random variables X and Y are said *orthogonal* if and only if $\mathbb{E}\left[X\,Y\right] = 0$.

Example 1.2.1. *A Gaussian variable Y is a second-order random variate that will be denoted by:*

$$Y \sim \mathscr{N}_1(\mu, \sigma^2), \tag{1.18}$$

where μ is its mean value and σ^2 is its variance. The so-called standard Gaussian distribution is obtained for $\mu = 0$ and $\sigma = 1$. Standard Gaussian random variables are denoted by Ξ in the sequel. The standard Gaussian PDF is defined as follows:

$$\varphi(\xi) = \frac{1}{\sqrt{2\,\pi}} \exp\left(-\frac{\xi^2}{2}\right), \tag{1.19}$$

and its CDF is given by:

$$\Phi(\xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\,\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x, \tag{1.20}$$

and can only be estimated numerically using the so-called complementary error function*:*

$$\Phi(\xi) = \frac{1}{2}\,\mathrm{erfc}\left(-\frac{\xi}{\sqrt{2}}\right). \tag{1.21}$$

Hence it can be proven with a simple change of variable that the PDF and CDF of any given Gaussian variable $Y \sim \mathscr{N}_1(\mu, \sigma^2)$ are equal to:

$$f_Y(y) = \frac{1}{\sigma}\,\varphi\left(\frac{y-\mu}{\sigma}\right) \quad \textit{and} \quad F_Y(y) = \Phi\left(\frac{y-\mu}{\sigma}\right). \tag{1.22}$$

1.2.1.2 Random vectors

A real-valued random vector $\boldsymbol{Y}$ is formally defined as an application that maps the probability space $(\Omega, \mathscr{F}, \mathscr{P})$ in a d-dimensional *output space* $\mathbb{Y} \subseteq \mathbb{R}^d$:

$$\boldsymbol{Y} : \left| \begin{array}{rcl} (\Omega, \mathscr{F}, \mathscr{P}) & \mapsto & \mathbb{Y} \\ \omega & \to & \boldsymbol{y} \end{array} \right., \tag{1.23}$$

where $\boldsymbol{y} = \boldsymbol{Y}(\omega)$ denotes a *realization*. Any of its d components $\{Y_1, \ldots, Y_d\}$ are themselves random variables (the *margin variables*) such as previously defined. Continuous random vectors are either defined by their joint PDF $f_Y(\boldsymbol{y})$ or their joint CDF $F_Y(\boldsymbol{y})$.

Its expected value $\boldsymbol{\mu}_{\boldsymbol{Y}}$ is the vector containing the expectation of each component:

$$\boldsymbol{\mu}_{\boldsymbol{Y}} = \left(\mathbb{E}\left[Y_i\right], i = 1, \ldots, d\right)^{\mathsf{T}} \tag{1.24}$$

and its (symmetric positive-definite) covariance and correlation matrices respectively denoted by $\mathbf{C}$ and $\mathbf{R}$ have their terms defined as follows:

$$C_{ij} = \mathrm{Cov}\left[Y_i, Y_j\right], \quad i, j = 1, \ldots, d \tag{1.25}$$

$$R_{ij} = \frac{\mathrm{Cov}\left[Y_i, Y_j\right]}{\sigma_{Y_i}\sigma_{Y_j}}, \quad i, j = 1, \ldots, d. \tag{1.26}$$

Given a partitioned random vector $\boldsymbol{Y} = (\boldsymbol{Y}_1^{\mathsf{T}}, \boldsymbol{Y}_2^{\mathsf{T}})^{\mathsf{T}}$, the *marginal distribution* of $\boldsymbol{Y}_1$ is given by the following integral:

$$f_{\boldsymbol{Y}_1}(\boldsymbol{y}_1) = \int_{\mathbb{Y}_2} f_{\boldsymbol{Y}}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}_2, \tag{1.27}$$

while the *conditional distribution* of $\boldsymbol{Y}_1$ given $\boldsymbol{Y}_2$ is given by the following ratio:

$$f_{\boldsymbol{Y}_1|\boldsymbol{Y}_2}(\boldsymbol{y}_1) = \frac{f_{\boldsymbol{Y}}(\boldsymbol{y})}{f_{\boldsymbol{Y}_2}(\boldsymbol{y}_2)}, \tag{1.28}$$

in application of Bayes' theorem.

Example 1.2.2. *A Gaussian random vector $\boldsymbol{Y}$ is a second-order random vector that will be denoted by:*

$$\boldsymbol{Y} \sim \mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{1.29}$$

where μ is the vector of means and $\boldsymbol{\Sigma}$ is its covariance matrix. The properties of this fundamental distribution are reviewed in more details in Appendix A and used in the sequel.

1.2.2 The least-squares linear regression model

The least-squares linear regression model plays a fundamental role in supervised learning theory. It is defined as follows:

$$Y_i = \sum_{j=1}^{p} \beta_j f_j(\boldsymbol{x}^{(i)}) + Z_i, \quad i = 1, \ldots, m \tag{1.30}$$

where $\boldsymbol{Y} = (Y_i, i = 1, \ldots, m)^{\mathsf{T}}$ is the vector of observations, $\boldsymbol{\beta} = (\beta_j, j = 1, \ldots, p)^{\mathsf{T}}$ is the vector of weights, $\boldsymbol{f} = \{f_j, j = 1, \ldots, p\}$ is a collection of regression functions (or *regressors*) and $\boldsymbol{Z} = (Z_i, i = 1, \ldots, m)^{\mathsf{T}}$ is a Gaussian random vector. The number p of regression functions is assumed less than or equal to the number m of observations so that the problem is not under-determined (*i.e.* it does not lack equations to explain the unknowns). The second-order moments of the Gaussian vector $\boldsymbol{Z}$ are:

$$\mathbb{E}\left[\boldsymbol{Z}\right] = \mathbf{0} \quad \text{and} \quad \mathrm{Cov}\left[\boldsymbol{Z}, \boldsymbol{Z}\right] = \mathbb{E}\left[\boldsymbol{Z}\,\boldsymbol{Z}^{\mathsf{T}}\right] = \sigma^2\,\mathbf{R}, \tag{1.31}$$

where the variance σ^2 is to be determined. In this section the correlation matrix $\mathbf{R}$ is assumed to be given depending on the context, and it is worth mentioning that it is reduced to the identity matrix in most practical applications thus leading to the so-called *ordinary least-squares* (OLS) problem. In the sequel the equations are derived for the more general case of any symmetric positive definite correlation matrix $\mathbf{R}$ which is known as the *generalized least-squares* (GLS) problem. This formulation will be more relevant in Section 1.4.

Due to the general linear regression model in Eq. (1.30), the distribution of the observations is the multivariate normal distribution:

$$\boldsymbol{Y} \sim \mathcal{N}_m\left(\mathbf{F}\,\boldsymbol{\beta},\, \sigma^2\,\mathbf{R}\right) \tag{1.32}$$

where the terms in the regression matrix $\mathbf{F}$ are defined as follows:

$$F_{i\,j} = f_j(\boldsymbol{x}^{(i)}), \quad i = 1, \ldots, m,\ j = 1, \ldots, p. \tag{1.33}$$

Note that if $\boldsymbol{x} = x$ is one-dimensional and $\boldsymbol{f} = \left\{1, x, x^2, \ldots, x^p\right\}$ is the complete polynomial basis up to order p, the corresponding matrix $\mathbf{F}$ is the Vandermonde matrix:

$$\mathbf{F} = \begin{bmatrix} 1 & x^{(1)} & x^{(1)2} & \ldots & x^{(1)p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x^{(m)} & x^{(m)2} & \ldots & x^{(m)p} \end{bmatrix}. \tag{1.34}$$

The regression problem consists in finding the optimal set of parameters $(\widehat{\boldsymbol{\beta}}, \widehat{\sigma^2}) \in \mathbb{R}^p \times \mathbb{R}^{+*}$ according to some metric to be defined.

1.2.3 The frequentist viewpoint

The frequentist approach to solve the regression problem in the least-squares sense consists in *maximizing the likelihood* of both the weight vector $\boldsymbol{\beta}$ and the variance σ^2 given a realization $\boldsymbol{y}$ of the vector of (correlated) observations $\boldsymbol{Y}$.

The likelihood is defined using the multivariate normal probability density function:

$$\mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2\right) = \frac{1}{((2\,\pi\,\sigma^2)^m\ [\det \mathbf{R}])^{1/2}} \exp\left[-\frac{1}{2\,\sigma^2}\,(\boldsymbol{y} - \mathbf{F}\,\boldsymbol{\beta})^{\mathsf{T}}\,\mathbf{R}^{-1}\,(\boldsymbol{y} - \mathbf{F}\,\boldsymbol{\beta})\right]. \tag{1.35}$$

Since maximizing this quantity is equivalent to minimizing its opposite natural logarithm, the *maximum likelihood estimation* problem may be cast as follows:

$$(\widehat{\boldsymbol{\beta}}, \widehat{\sigma^2}) \equiv \arg \min_{(\boldsymbol{\beta}, \sigma^2) \in \mathbb{R}^p \times \mathbb{R}^{+*}} \; -\log \mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2\right). \tag{1.36}$$

The first-order optimality conditions for this unconstrained optimization problem read:

$$\begin{cases} \nabla_{\boldsymbol{\beta}} \log \mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2\right) &= \mathbf{0} \\ \dfrac{\partial \log \mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2\right)}{\partial \sigma^2} &= 0 \end{cases}, \tag{1.37}$$

and may be further expanded as follows:

$$\left\{\begin{array}{rcll} \dfrac{1}{\sigma^2}(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\boldsymbol{y}-\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F}\boldsymbol{\beta}) & = & \mathbf{0}, & \forall\sigma^2 \\ -\dfrac{m}{2}\dfrac{1}{\sigma^2}+\dfrac{1}{2\sigma^4}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta})^\mathsf{T}\mathbf{R}^{-1}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta}) & = & 0 & \end{array}\right. . \tag{1.38}$$

These two equations eventually lead to the so-called *generalized least-squares estimate* for the vector of weights:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\boldsymbol{y}, \tag{1.39}$$

and the variance estimate is:

$$\widehat{\sigma^2} = \frac{1}{m}(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})^\mathsf{T}\mathbf{R}^{-1}(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}}). \tag{1.40}$$

Note that the generalized least-squares solution reduces to the *ordinary least-squares* (OLS) solution when the observations are *uncorrelated*:

$$\mathbf{R}=\mathbf{I} \quad \Rightarrow \quad \widehat{\boldsymbol{\beta}}_{\text{OLS}} = (\mathbf{F}^\mathsf{T}\mathbf{F})^{-1}\mathbf{F}^\mathsf{T}\boldsymbol{y}. \tag{1.41}$$

1.2.4 The Bayesian viewpoint

1.2.4.1 Non-informative prior distributions

Generally speaking, the Bayesian approach allows the user to specify a prior distribution for the sought parameters (*e.g.* here $[\boldsymbol{\beta}, \sigma^2]$) and to derive a full posterior distribution conditioned on both the prior and the observations. This posterior distribution models the residual epistemic uncertainty due to the sparsity of the observations $\boldsymbol{y}$. In order to derive results in agreement with the frequentist approach, it is proposed to consider here a so-called *non-informative prior* distribution so that the posterior distribution depends solely on the observations.

As its name suggests, a non-informative prior distribution does not provide any prior information on the probability density function of the unknown parameters. This degree of information can be measured in rigorous mathematical terms thanks to the Fischer's information matrix. Provided a set of Q observations $\left\{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(Q)}\right\}$ of the inferred random variable $\boldsymbol{Y}$, this matrix is defined as follows:

$$\mathscr{I}(\boldsymbol{\theta}) = \mathbb{E}_X\left[\boldsymbol{\nabla}_{\boldsymbol{\theta}}\,\boldsymbol{\nabla}_{\boldsymbol{\theta}}^\mathsf{T}\log \mathsf{L}\left(\left\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(Q)}\right\}|\boldsymbol{\theta}\right)\right], \tag{1.42}$$

where L denotes the likelihood, and $\boldsymbol{\theta}$ is the unknown vector of parameters. This lead Jeffreys (1946) to show that the non-informative prior distribution for the parameter of any given parametric distribution satisfies:

$$p(\boldsymbol{\theta}) \propto \sqrt{[\det \mathscr{I}(\boldsymbol{\theta})]}. \tag{1.43}$$

In particular, the Jeffreys' non-informative prior for the location parameters of the multivariate normal distribution in Eq. (1.32) is shown to be proportional to:

$$p(\boldsymbol{\beta}) \propto 1, \qquad \forall \boldsymbol{\beta} \in \mathbb{R}^p. \tag{1.44}$$

Indeed, this prior simply means that each component of the vector $\boldsymbol{\beta}$ has uniform density on the whole real line *a priori*. It is also said to be *improper* because the integral of the constant 1 over $\mathbb{R}^p$ is infinite, hence the proportionality ($\propto$) in Eq. (1.44). Improper distributions may cause difficulties in a frequentist approach. However, in a Bayesian context an improper prior distribution may lead to a proper posterior and this will be the case in the sequel.

The Jeffreys' non-informative prior distribution for the variance of a Gaussian distribution is defined as follows:

$$p(\sigma^2) \propto \frac{1}{\sigma^2}, \qquad \forall \sigma^2 \in \mathbb{R}^{+*}. \tag{1.45}$$

To give some more insight about this prior on σ^2, consider instead the prior distribution of its natural logarithm $\log \sigma^2$. Indeed, if one chooses a uniform prior over the real line for $\log \sigma^2$, then it results in the so-called *logarithmic prior* given in Eq. (1.45) for σ^2 by means of a simple change of variable:

$$p(\log \sigma^2) \propto 1 \quad \Rightarrow \quad p(\sigma^2) \propto \left| \frac{\mathrm{d} \log \sigma^2}{\mathrm{d} \sigma^2} \right| = \frac{1}{\sigma^2}, \qquad \forall \sigma^2 \in \mathbb{R}^{+*}. \tag{1.46}$$

Even though this prior distribution is again improper (because the integral of its PDF over $\mathbb{R}^{+*}$ is not finite), it may result in a proper posterior distribution and it will be the case in the sequel.

The proposed joint probability density function of the non-informative prior finally reads:

$$p(\boldsymbol{\beta}, \sigma^2) = p(\boldsymbol{\beta})\, p(\sigma^2) \propto \frac{1}{\sigma^2}. \tag{1.47}$$

1.2.4.2 Generalized least-squares posterior distribution

According to Bayes' theorem, the posterior distribution of the parameters given the observations is defined as follows:

$$p(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}) \propto \mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2\right)\, p(\boldsymbol{\beta}, \sigma^2). \tag{1.48}$$

This can be further elicited by replacing the likelihood in Eq. (1.35) and the prior distribution by their respective expressions:

$$p(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}) \propto \frac{1}{((2\,\pi\,\sigma^2)^m\, [\det \mathbf{R}])^{1/2}} \exp\left[-\frac{1}{2\,\sigma^2} Q\right] \times \frac{1}{\sigma^2}, \tag{1.49}$$

where $Q \equiv (\boldsymbol{y} - \mathbf{F}\boldsymbol{\beta})^{\mathsf{T}}\, \mathbf{R}^{-1} (\boldsymbol{y} - \mathbf{F}\boldsymbol{\beta})$ has been introduced for the sake of brevity. This quantity can be further simplified by introducing the generalized least-squares solution $\widehat{\boldsymbol{\beta}}$

in Eq. (1.39) in its expression:

$$\begin{aligned} Q &\equiv (\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta})^{\mathsf{T}}\mathbf{R}^{-1}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta}) \\ &= \left[(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})-\mathbf{F}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\right]^{\mathsf{T}}\mathbf{R}^{-1}\left[(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})-\mathbf{F}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\right] \\ &= (\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})\mathbf{R}^{-1}(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})+(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})^{\mathsf{T}}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}) \\ &\quad -2\underbrace{(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})}_{Q_3}. \end{aligned} \tag{1.50}$$

At this point it can be seen that *(i)* the first term in this expression is proportional to the generalized least-squares estimate of the variance $\widehat{\sigma^2}$ in Eq. (1.40) *(ii)* the second term is a quadratic form centred on the generalized least-squares estimate of the regression weights $\widehat{\boldsymbol{\beta}}$ while *(iii)* the third term is proven to be zero by replacing $\widehat{\boldsymbol{\beta}}$ by its expression:

$$\begin{aligned} Q_3 &\equiv (\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}) \\ &= (\boldsymbol{y}-\mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y})^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}) \\ &= (\boldsymbol{y}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}-\boldsymbol{y}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}))(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}) \\ &= (\boldsymbol{y}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}-\boldsymbol{y}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}\mathbf{I})(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})=0. \qquad \text{(Q.E.D.)} \end{aligned}$$

Then,

$$Q = m\widehat{\sigma^2}+(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})^{\mathsf{T}}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}). \tag{1.51}$$

Substituting this new simplified expression for Q in the expression of the posterior distribution in Eq. (1.49), grouping the terms in $\boldsymbol{\beta}$ and σ^2 and neglecting the constant terms that do not depend on $\boldsymbol{\beta}$ nor σ^2 leads to:

$$p(\boldsymbol{\beta},\sigma^2\,|\,\boldsymbol{y}) \propto \frac{1}{\sigma^{m+2}}\exp\left[\frac{-1}{2\sigma^2}\left(m\widehat{\sigma^2}+(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})^{\mathsf{T}}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\right)\right]. \tag{1.52}$$

which is known as the Normal-Inverse-Gamma distribution.

From this joint posterior distribution, it appears immediately that the posterior distribution of the regression weights conditional on the observations $\boldsymbol{y}$ and the variance σ^2 is the multivariate Gaussian distribution centred on the generalized least-squares estimate:

$$[\boldsymbol{\beta}\,|\,\boldsymbol{y},\sigma^2] \sim \mathcal{N}_p\left(\widehat{\boldsymbol{\beta}},\sigma^2(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\right), \tag{1.53}$$

The computation of the marginal posterior distributions of $\boldsymbol{\beta}$ and σ^2 requires further integration of the joint posterior in Eq. (1.52):

$$p(\boldsymbol{\beta}\,|\,\boldsymbol{y}) = \int_{\mathbb{R}^{+*}} p(\boldsymbol{\beta},\sigma^2\,|\,\boldsymbol{y})\,\mathrm{d}\sigma^2, \tag{1.54}$$

$$p(\sigma^2\,|\,\boldsymbol{y}) = \int_{\mathbb{R}^p} p(\boldsymbol{\beta},\sigma^2\,|\,\boldsymbol{y})\,\mathrm{d}\boldsymbol{\beta}. \tag{1.55}$$

Zellner (1971, p. 67) carried out these integrations analytically and identified the posterior distributions of:

- the regression weights conditional on the observations $[\boldsymbol{\beta} \mid \boldsymbol{y}]$ as a *shifted multivariate Student distribution* centred on the generalized least-squares estimate $\widehat{\boldsymbol{\beta}}$;
- the variance conditional on the observations $[\sigma^2 \mid \boldsymbol{y}]$ as an *Inverse-Gamma distribution* (meaning the inverse of σ^2 has Gamma distribution) centred on the *unbiased* generalized least-squares estimate $\frac{m}{m-p}\widehat{\sigma^2}$.

It is important to note that the posterior distribution for $\boldsymbol{\beta}$ involves a non-necessarily diagonal covariance matrix, meaning that the regression weights are not necessarily independent although the non-informative prior assumed they were.

Finally, it is worth mentioning that analogous computations can be carried out with more informative priors in a Bayesian updating context. For instance, the availability of such informative priors may occur in an adaptive refinement procedure after a first generalized least-squares fit on an initial dataset. In this case, a meaningful and natural second stage prior when new data becomes available could be the posterior distribution (from a non informative prior) of the previous stage. Surprisingly, it can be shown (*e.g.* in Zellner, 1971, p. 70) that the updating of $\boldsymbol{\beta}$ and σ^2 from a joint Normal-Inverse-Gamma prior distribution leads again to a joint Normal-Inverse-Gamma distribution, thus meaning that the family of the joint posterior distributions is invariant by multiple updates while its scattering is expected to decrease.

1.3 Gaussian processes

1.3.1 Definition

Using the same formalism as in Section 1.2.1, a stochastic (or random) process is an application that maps the probability space $(\Omega, \mathscr{F}, \mathscr{P})$ times an index space $\mathbb{X}$ in an output space $\mathbb{Y}$:

$$Y(\boldsymbol{x}) : \left| \begin{array}{rcl} \mathbb{X} \times (\Omega, \mathscr{F}, \mathscr{P}) & \mapsto & \mathbb{Y} \\ (\boldsymbol{x}, \omega) & \to & y(\boldsymbol{x}) \end{array} \right. , \tag{1.56}$$

where $y(\boldsymbol{x}) \equiv Y(\boldsymbol{x}, \omega)$ denotes a *realization* (a *sample path*). In particular, $Y(\boldsymbol{x}, \omega_0)$ for some given $\omega_0 \in \Omega$ is a function of $\boldsymbol{x} \in \mathbb{X}$, and $Y(\boldsymbol{x}^{(0)}, \omega)$ for some given $\boldsymbol{x}^{(0)} \in \mathbb{X}$ is a random variable over the probability space $(\Omega, \mathscr{F}, \mathscr{P})$.

Less formally, if a random vector $\boldsymbol{Y}$ can be seen as a random variable indexed by a finite subset of natural integer indexes $[\![1; m]\!]$ then a random process $Y(\boldsymbol{x})$ is nothing but an infinite-dimensional random vector because it is indexed by a continuous parameter $\boldsymbol{x} \in \mathbb{X}$. Some other authors like Rasmussen and Williams (2006) view stochastic processes as a mean to describe probability distributions over functions, so that $\mathscr{M}$ can be considered as one particular sample of that distribution.

In particular, Rasmussen and Williams (2006) define a Gaussian process as an infinite collection of random variables, any finite number of which having a multivariate Gaussian distribution such as defined in Appendix A. This relationship between Gaussian processes

and jointly Gaussian random vectors is the sound basis of Gaussian process predictors such as introduced in Section 1.4.

A Gaussian process is a second-order stochastic process and it is completely defined by its first- and second-order statistics, namely:

- its mean function μ:

$$\mu(\boldsymbol{x}) \equiv \mathbb{E}\left[Y(\boldsymbol{x})\right], \quad \boldsymbol{x} \in \mathbb{X}, \tag{1.57}$$

- and its autocovariance function C:

$$C\left(\boldsymbol{x}, \boldsymbol{x}'\right) \equiv \mathbb{E}\left[\left(Y(\boldsymbol{x}) - \mu(\boldsymbol{x})\right)\left(Y(\boldsymbol{x}') - \mu(\boldsymbol{x}')\right)\right], \quad \left(\boldsymbol{x}, \boldsymbol{x}'\right) \in \mathbb{X} \times \mathbb{X}. \tag{1.58}$$

Consequently, a Gaussian process $Y(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{X}$, is denoted as follows:

$$Y(\boldsymbol{x}) \sim \mathcal{GP}\left(\mu(\boldsymbol{x}), C\left(\boldsymbol{x}, \boldsymbol{x}'\right)\right), \qquad \left(\boldsymbol{x}, \boldsymbol{x}'\right) \in \mathbb{X} \times \mathbb{X}. \tag{1.59}$$

As for the covariance matrix of random vectors (see Appendix A), the autocovariance function must be symmetric and (non-strictly) positive definite. This means that *admissible* autocovariance functions satisfy the following two relationships:

$$C\left(\boldsymbol{x}, \boldsymbol{x}'\right) = C\left(\boldsymbol{x}', \boldsymbol{x}\right), \quad \forall \left(\boldsymbol{x}, \boldsymbol{x}'\right) \in \mathbb{X} \times \mathbb{X}, \tag{1.60}$$

and

$$\sum_{i=1}^{m}\sum_{j=1}^{m} w_i\, w_j\, C\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right) \geq 0, \tag{1.61}$$

for any subset $\mathcal{X} = \left\{\boldsymbol{x}^{(i)}, i = 1, \ldots, m\right\}$ of $\mathbb{X}$ and any vector $\boldsymbol{w} = \left(w_i, i = 1, \ldots, m\right)^{\mathsf{T}} \in \mathbb{R}^m$.

The latter requirement can be understood by looking at the quantity on the left hand-side of Eq. (1.61) as the variance of any given linear combination of a second-order stochastic process Y observed on any finite subset $\mathcal{X}$ of $\mathbb{X}$, indeed:

$$\mathrm{Var}\left[\sum_{i=1}^{m} w_i\, Y\left(\boldsymbol{x}^{(i)}\right)\right] = \sum_{i=1}^{m}\sum_{j=1}^{m} w_i\, w_j\, C\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right), \tag{1.62}$$

and it is thus naturally required to be positive.

1.3.2 Simplifying assumptions

1.3.2.1 Stationarity

In a strict sense, stationarity refers to the property of a stochastic process which states that it is *invariant by translation*. As a consequence:

- its mean function reduces to a constant:

$$\mu(\boldsymbol{x}) = \mu_0, \quad \boldsymbol{x} \in \mathbb{X}, \tag{1.63}$$

- its autocovariance function rewrites:

$$C\left(\boldsymbol{x}, \boldsymbol{x}'\right) = \sigma^2 R\left(\boldsymbol{x} - \boldsymbol{x}'\right), \quad \left(\boldsymbol{x}, \boldsymbol{x}'\right) \in \mathbb{X} \times \mathbb{X}, \tag{1.64}$$

  where σ^2 is the constant variance and R is the so-called autocorrelation function which solely depends on the shift $\boldsymbol{x} - \boldsymbol{x}'$.

With slight abuse of language, one may sometimes refer to stationarity *w.r.t.* the mean, the variance and/or the autocorrelation.

1.3.2.2 Componentwise anisotropy and composed correlation functions

A stochastic process indexed on an input space $\mathbb{X}$ of dimension n is said to be isotropic if it is *invariant by rotation*. As a consequence, an isotropic stochastic process is stationary and, in addition its autocorrelation function rewrites:

$$R\left(\boldsymbol{x} - \boldsymbol{x}'\right) = R\left(\|\boldsymbol{x} - \boldsymbol{x}'\|_2\right), \quad (\boldsymbol{x}, \boldsymbol{x}') \in \mathbb{X} \times \mathbb{X} \tag{1.65}$$

where $\| \bullet \|_2$ denotes the usual $\mathscr{L}_2$ norm in $\mathbb{R}^n$.

A stochastic process is said to be *componentwise anisotropic* if its autocorrelation is defined as the tensor product of one-dimensional autocorrelation functions:

$$R\left(\boldsymbol{x} - \boldsymbol{x}'\right) = \prod_{i=1}^{n} R_i\left(x_i - x_i'\right), \quad \left(\boldsymbol{x}, \boldsymbol{x}'\right) \in \mathbb{X} \times \mathbb{X}. \tag{1.66}$$

In the present context of computer experiments, the isotropy assumption does not hold because of the various physical meaning of the margin variables composing the input vector $\boldsymbol{x}$. For instance, in structural mechanics, one component of $\boldsymbol{x}$ may represent a dimensionless Poisson's ratio of some steel alloy whose order of magnitude is 0.3 whereas another may represent its Young's modulus expressed in MPa whose order of magnitude is 10^5.

In order to circumvent this problem, a common practice, which is systematically employed throughout this manuscript, consists in normalizing (*whitening*) the input data with respect to the componentwise empirical means $\{m_{X_i},\, i = 1, \ldots, n\}$ and standard deviations $\{s_{X_i},\, i = 1, \ldots, n\}$ estimated from the dataset $\mathscr{X} = \{\boldsymbol{x}^{(i)}, i = 1, \ldots, m\}$. The mapped input vectors read:

$$\left\{ \overline{x_i} = \frac{x_i - m_{X_i}}{s_{X_i}},\, i = 1, \ldots, n \right\} \in \overline{\mathbb{X}} \tag{1.67}$$

say, where $\overline{\mathbb{X}}$ is the *centred* and *normalized* input space.

Nonetheless, accounting for componentwise anisotropy in the autocorrelation function may reveal interesting in a computer experiment context where a function $\mathscr{M}$ may not vary equally in all directions even in the transformed input space $\overline{\mathbb{X}}$. Indeed, if some components in $\boldsymbol{x}$ have small influence on the output y, it would be interesting to have a large-range autocorrelation function that tends to generalize the model along that component whereas short-range autocorrelations would be required in the other more fluctuating directions. The distinction between short- and long-range autocorrelations is illustrated in Section 1.3.4.

1.3.3 Representation techniques

The representation of stochastic processes consists in building a mathematical expression which enables the explicit computation of sample paths. There exists a wide variety of representation techniques for stochastic processes, although most of them do not apply when the processes are either non-Gaussian or non-stationary. Sudret and Der Kiureghian (2002) provide an interesting review of such techniques together with an extended set of examples and applications to structural reliability problems.

It is important to note here that the Karhunen-Loève representation (Loève, 1977; Ghanem and Spanos, 2003) theoretically holds for both non-Gaussian and non-stationary random fields. The so-called Karhunen-Loève expansion is defined as follows:

$$Y(\boldsymbol{x}) = \mu(\boldsymbol{x}) + \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \Xi_k\, \varphi_k(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{X}, \tag{1.68}$$

where the infinite countable set of eigenvalues and eigenfunctions $\{(\lambda_k, \varphi_k), k \in \mathbb{N}^{+*}\}$ are solution of the following equation:

$$\int_{\mathbb{X}} C(\boldsymbol{x}, \boldsymbol{x}')\, \varphi_k(\boldsymbol{x}')\, \mathrm{d}\boldsymbol{x}' = \lambda_k\, \varphi_k(\boldsymbol{x}), \quad k \in \mathbb{N}^{+*}, \boldsymbol{x} \in \mathbb{X}, \tag{1.69}$$

and $\{\Xi_k, k \in \mathbb{N}^{+*}\}$ is an infinite collection of zero-mean, unit-variance uncorrelated random variables.

In practice the Karhunen-Loève expansion is truncated to the M terms associated with the largest eigenvalues resulting in the following approximating expansion:

$$\widetilde{Y}(\boldsymbol{x}) = \mu(\boldsymbol{x}) + \sum_{k=1}^{M} \sqrt{\lambda_k}\, \Xi_k\, \varphi_k(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{X}, \tag{1.70}$$

where $\{(\lambda_k, \varphi_k), k \in \mathbb{N}^{+*}\}$ are ordered so that $\lambda_1 > \ldots > \lambda_M > \lambda_{M+1} > \ldots > 0$.

Practical implementation of the Karhunen-Loève expansion is not easy though because it requires *(i)* the resolution of a complex Fredholm integral equation of the second kind (Eq. (1.69)), and *(ii)* the determination of the distribution of the underlying random variables. The interested reader is referred to Sudret and Der Kiureghian (2002); Phoon et al. (2002); Ghanem and Spanos (2003) for a numerical alternative to the resolution of the integral equation in Eq. (1.69), and to Ghanem and Doostan (2006); Desceliers et al. (2007); Guilleminot et al. (2008); Dubourg et al. (2011f) for the identification of the distribution of the underlying random variables from experimental measurements.

The most basic representation for Gaussian processes illustrates the above-alluded relationship between Gaussian processes and jointly Gaussian vectors. Given a subset $\mathscr{X} = \{\boldsymbol{x}^{(i)}, i = 1, \ldots, m\}$ of $\mathbb{X}$, one may evaluate the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of the associated finite collection of random variables $\{Y_i, i = 1, \ldots, m\}$ whose terms read:

$$\mu_i = \mu\left(\boldsymbol{x}^{(i)}\right), \quad i = 1, \ldots, m \tag{1.71}$$

and

$$\Sigma_{ij} = C\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right), \quad i, j = 1, \ldots, m. \tag{1.72}$$

The corresponding vector of random observations $\boldsymbol{Y} = \left(Y_i, i = 1, \ldots, m\right)^{\mathsf{T}}$ is Gaussian by definition:

$$\boldsymbol{Y} \sim \mathcal{N}_m\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right), \tag{1.73}$$

so that one may use the usual sampling algorithms for Gaussian random vectors (such as the ones detailed in Section A.5 of Appendix A) in order to draw sample paths from the Gaussian process $Y(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{X}$.

1.3.4 Examples of stationary autocorrelation functions

All the stochastic processes illustrated in this section are strictly stationary with zero-mean and unit-variance (*i.e.* $\mu(\boldsymbol{x}) = 0$ and $\sigma^2(\boldsymbol{x}) = 1$). Only the autocorrelation function is being varied in order to illustrate its impact on the properties of the corresponding sample paths (functions of $\boldsymbol{x}$). These sample paths were obtained by means of the previously introduced most basic representation (see Section 1.3.3).

1.3.4.1 The nugget autocorrelation function

The nugget autocorrelation function is defined as follows:

$$R(\boldsymbol{x} - \boldsymbol{x}') = \delta(\boldsymbol{x} - \boldsymbol{x}'), \tag{1.74}$$

where δ is the Dirac function which is equal to one if $\boldsymbol{x} = \boldsymbol{x}'$ and zero otherwise. It is parameter-free.

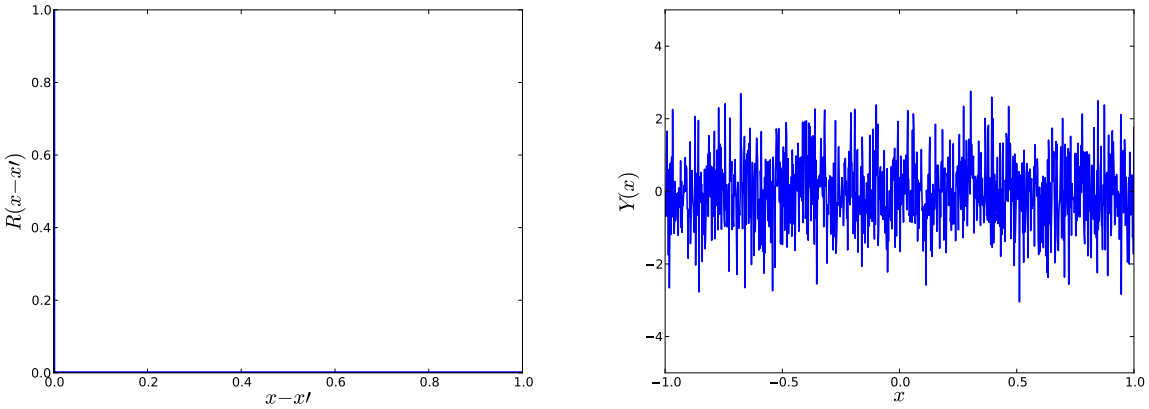


Figure 1.4: The nugget autocorrelation function and a sample path of the corresponding zero-mean unit-variance Gaussian process.

It models the absence of any correlation (coherence in the sample paths) meaning that all the observations of the associated stationary Gaussian process are *independent and identically distributed*. Such a Gaussian process is known as a *white noise* and its sample paths are discontinuous (see Figure 1.4). In computer experiments, it is sometimes used as an additive autocorrelation to account for noise in the dataset. It will also allow us to demonstrate the interrelation between the previously introduced ordinary least-squares regression procedure and the forthcoming Gaussian process regression methodology.

1.3.4.2 The linear autocorrelation function

The linear autocorrelation function is defined as follows (see Figure 1.5):

$$R(\boldsymbol{x}-\boldsymbol{x}') = \prod_{i=1}^{n} \max\left(0,\, 1 - \frac{\left|x_i - x_i'\right|}{\ell_i}\right), \tag{1.75}$$

where $\{\ell_i > 0,\, i = 1, \ldots, n\}$ are the so-called *scale parameters*.

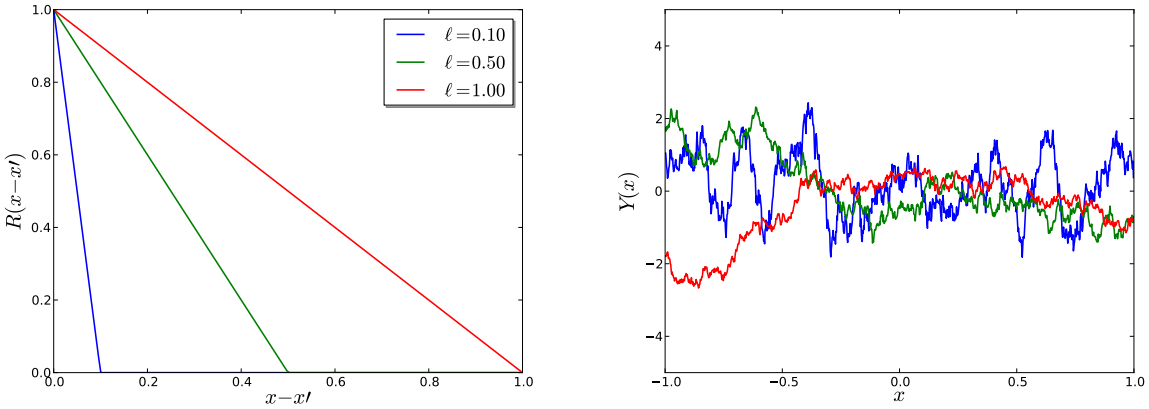


Figure 1.5: The linear autocorrelation function and some sample paths of the corresponding zero-mean unit-variance Gaussian process depending on its scale parameter.

1.3.4.3 The exponential autocorrelation function

The anisotropic exponential autocorrelation function is defined as follows (see Figure 1.6):

$$R(\boldsymbol{x}-\boldsymbol{x}', \boldsymbol{\ell}) = \exp\left(-\sum_{i=1}^{n} \frac{\left|x_i - x_i'\right|}{\ell_i}\right), \tag{1.76}$$

where $\{\ell_i > 0,\, i = 1, \ldots, n\}$ are the so-called *scale parameters*.

This autocorrelation function corresponds to a specific Gaussian process known as the Ornstein-Uhlenbeck process. Its sample paths are $\mathscr{C}^0$ (continuous but non-differentiable).

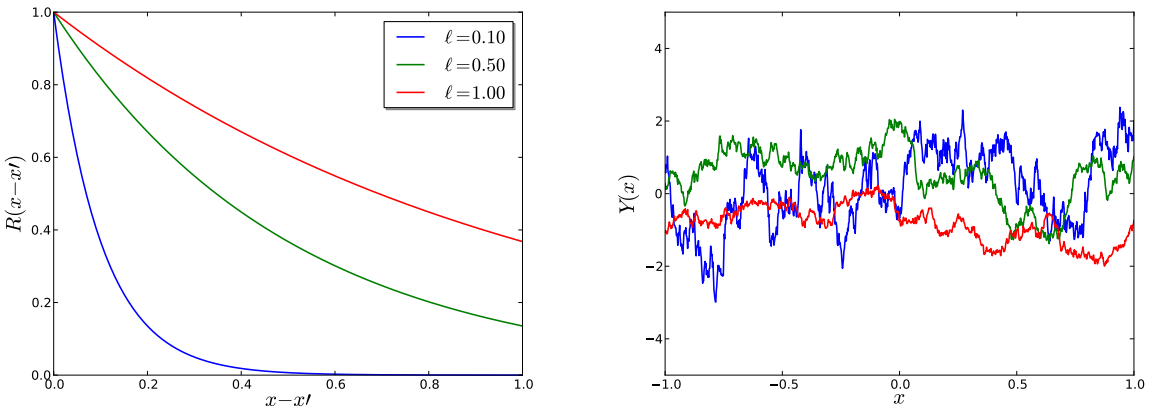


Figure 1.6: The exponential autocorrelation function and some sample paths of the corresponding zero-mean unit-variance Gaussian process depending on its scale parameter.

1.3.4.4 The squared exponential autocorrelation function

The anisotropic squared exponential autocorrelation function is defined as follows (see Figure 1.7):

$$R(\boldsymbol{x}-\boldsymbol{x}',\boldsymbol{\ell})=\exp\left(-\sum_{i=1}^{n}\left(\frac{x_i-x_i'}{\ell_i}\right)^2\right), \tag{1.77}$$

where $\{\ell_i>0,\, i=1,\ldots,n\}$ are the so-called *scale parameters*.

This autocorrelation provides an infinite degree of differentiability for the associated sample paths.

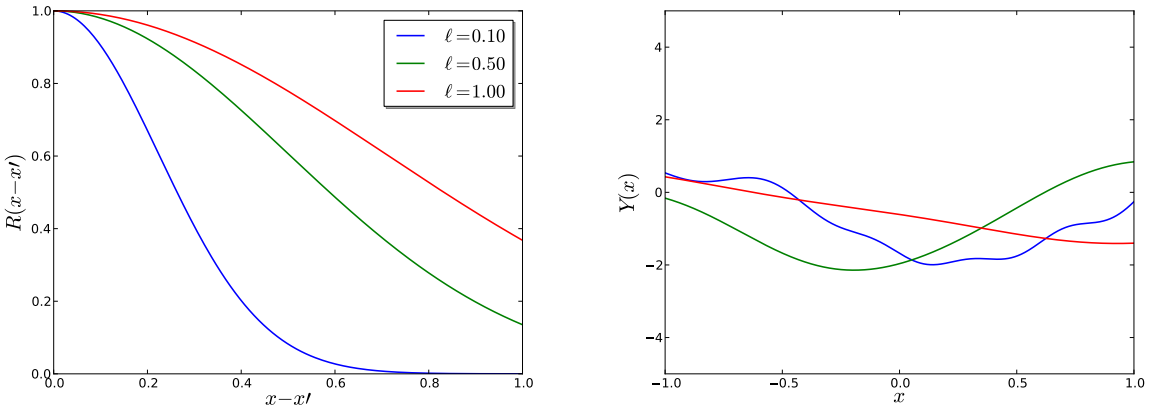


Figure 1.7: The squared exponential autocorrelation function and some sample paths of the corresponding zero-mean unit-variance Gaussian process depending on its scale parameter.

1.3.4.5 The Matérn autocorrelation function

The Matérn autocorrelation function is defined as follows:

$$R(\boldsymbol{x}-\boldsymbol{x}', \boldsymbol{\ell}, \nu) = \prod_{i=1}^{n} \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(2\sqrt{\nu} \frac{\left|x_i - x_i'\right|}{\ell_i} \right)^{\nu} \mathcal{K}_{\nu} \left(2\sqrt{\nu} \frac{\left|x_i - x_i'\right|}{\ell_i} \right), \qquad (1.78)$$

where $\{\ell_i > 0, i = 1, \ldots, n\}$ are the so-called *scale parameters*, $\nu \geq 1/2$ is the so-called *shape* parameter, Γ is the Euler Gamma function, and $\mathcal{K}_{\nu}$ is the modified Bessel function of the second kind (also known as the Bessel function of the third kind).

One interesting feature of this advanced autocorrelation function is that the sample paths from the corresponding Gaussian process are $\lceil \nu - 1 \rceil$ times differentiable (with $\lceil \bullet \rceil$ denoting the ceiling function). If $\nu = 1/2$, the Matérn autocorrelation function coincides with the exponential autocorrelation function which generates $\mathscr{C}^0$ sample paths (continuous but non-differentiable). On the contrary, as ν tends to infinity the Matérn autocorrelation function tends toward the squared exponential autocorrelation function which has $\mathscr{C}^{\infty}$ (infinitely differentiable) sample paths (see Figure 1.8).

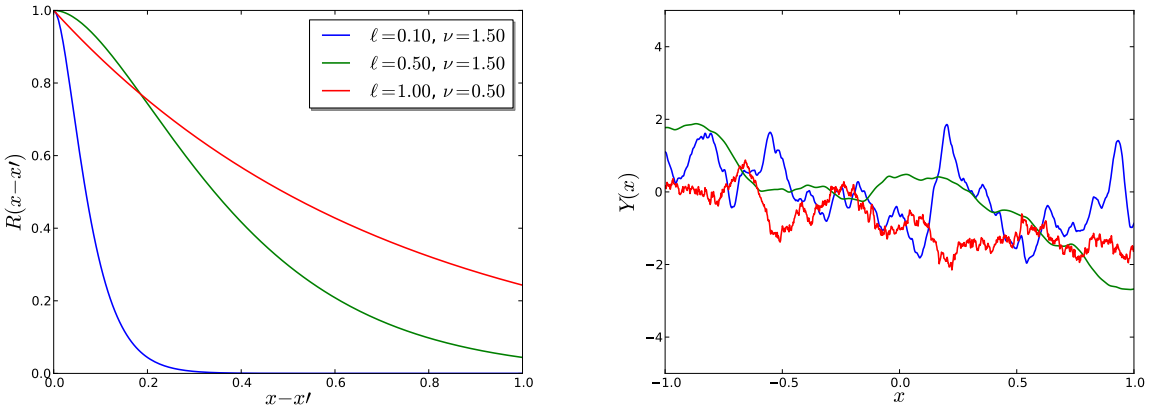


Figure 1.8: The Matérn autocorrelation function and some sample paths of the corresponding zero-mean unit-variance Gaussian process depending on its scale and shape parameters.

1.4 Gaussian process predictors

1.4.1 Prediction basics

This section is concerned with the so-called Bayesian prediction methodology such as defined in the book by Santner et al. (2003). It essentially consists in assuming that the observations gathered in the vector $\boldsymbol{y} \equiv \left(\mathcal{M}(\boldsymbol{x}^{(1)}), \ldots, \mathcal{M}(\boldsymbol{x}^{(m)})\right)^{\mathsf{T}}$ together with the unobserved quantity of interest $y_0 \equiv \mathcal{M}(\boldsymbol{x}^{(0)})$ is a realization of a random vector distributed

according to a joint parametric distribution $F \in \mathscr{F}$:

$$\left\{ \begin{array}{c} Y_0 \\ \boldsymbol{Y} \end{array} \right\} \sim F \in \mathscr{F}. \tag{1.79}$$

The Bayesian prediction methodology aims at deriving a (random) predictor $\widehat{Y}_0$ for the unobserved quantity y_0 by exploiting this statistical dependency.

Despite this general framework allows F to belong to a large class of distributions $\mathscr{F}$, it will be restricted to the multivariate Gaussian distribution in the sequel due to its convenient properties – see Appendix A for some of them.

Predictors $\widehat{Y}_0$ may have a general functional form *w.r.t.* the observations although most practical applications are reduced to *linear predictors*. A *linear predictor* $\widehat{Y}_0$ is defined as a linear combination of the observations $\boldsymbol{Y} = \left(Y_i, i = 1, \ldots, m\right)^\mathsf{T}$:

$$\widehat{Y}_0 = \sum_{i=1}^{m} a_i \, Y_i = \boldsymbol{a}^\mathsf{T} \boldsymbol{Y}, \tag{1.80}$$

where $\boldsymbol{a}$ is a weight vector of $\mathbb{R}^m$.

In addition to that linear property, it is important to derive predictors that *almost equal* the quantity of interest *everywhere* on the support of F. This property is expressed mathematically by requiring $\widehat{Y}_0$ to be *unbiased*. A predictor $\widehat{Y}_0$ is *unbiased* with respect to F if and only if:

$$\mathbb{E}_F \left[\widehat{Y}_0 - Y_0 \right] = 0. \tag{1.81}$$

Finally, a metric is required to compare the quality of competing predictors. In this prediction context, it is proposed to resort to the well-known *mean squared prediction error* (MSPE) which is defined with respect to F as follows:

$$\text{MSPE}(\widehat{Y}_0, F) = \mathbb{E}_F \left[(\widehat{Y}_0 - Y_0)^2 \right]. \tag{1.82}$$

The best MSPE predictor may now be defined as the predictor that minimizes the latter quantity amongst all other predictors.

The *fundamental theorem of prediction* (Santner et al., 2003, Theorem 3.2.1, p. 52) establishes a relationship between the conditional distribution of Y_0 given the observations $\boldsymbol{Y}$ and such best MSPE predictors. This link will be used in the subsequent sections to show that both frequentist and Bayesian viewpoints lead somehow to the same predictor.

Theorem 1.4.1. ***The fundamental theorem of prediction***
Suppose that $(Y_0, \boldsymbol{Y})$ *has a joint distribution* F *for which the conditional mean of* Y_0 *given* $\boldsymbol{Y}$ *exists. Then,*

$$\widehat{Y}_0 = \mathbb{E}\left[Y_0 \mid \boldsymbol{Y} \right] \tag{1.83}$$

is the best MSPE predictor of Y_0.

The proof consists in showing that the mean squared prediction error of any other predictor $Y_0^\star$ is greater than that of the best MSPE predictor $\widehat{Y}_0$. See the book by Santner et al. (2003, p. 52) for the complete proof and its interpretation.

1.4.2 The two-stage Gaussian process prior model

The two-stage Gaussian process prior model assumes that the functional relationship $\mathscr{M}$ between the input $\boldsymbol{x}$ and the scalar output y is a sample path from a Gaussian process Y to be characterized. It means that the general linear model in Eq. (1.30) still holds for both the unobserved value $Y_0 \equiv Y(\boldsymbol{x}^{(0)})$ and the observations $\boldsymbol{Y}$, except the additive Gaussian noise Z is now assumed to depend on $\boldsymbol{x}$. Indeed, the so-called two-stage Gaussian process prior model reads:

$$Y_i = \sum_{j=1}^{p} \beta_j f_j(\boldsymbol{x}^{(i)}) + Z(\boldsymbol{x}^{(i)}), \quad i = 0, \ldots, m, \tag{1.84}$$

where:

(i) the first stage consists in a linear combination on a given functional basis $\boldsymbol{f} = \left\{f_j, j = 1, \ldots, p\right\}$ with $p \leq m$ as for the generalized least-squares problem;

(ii) the second stage consists in the Gaussian process Z with zero mean:

$$\mathbb{E}\left[Z(\boldsymbol{x})\right] = 0, \quad \forall \boldsymbol{x} \in \mathbb{X}, \tag{1.85}$$

and stationary autocovariance:

$$\mathbb{E}\left[Z(\boldsymbol{x}), Z(\boldsymbol{x}')\right] = \sigma^2 R\left(\boldsymbol{x} - \boldsymbol{x}', \boldsymbol{\theta}\right), \quad \forall (\boldsymbol{x}, \boldsymbol{x}') \in \mathbb{X} \times \mathbb{X}. \tag{1.86}$$

In the latter expression σ^2 denotes the (constant) variance of the Gaussian process while R is its stationary autocorrelation which depends only on the difference $\boldsymbol{x} - \boldsymbol{x}'$ and its so-called *hyperparameters* grouped in $\boldsymbol{\theta}$. Examples of such autocorrelation functions and some of their properties have been detailed in Section 1.3.

Thanks to the Gaussian assumption in Eq. (1.84), the vector gathering the prediction Y_0 and the observations $\boldsymbol{Y}$ is normally distributed:

$$\left\{ \begin{array}{c} Y_0 \\ \boldsymbol{Y} \end{array} \right\} \sim \mathscr{N}_{1+m}\left(\left\{ \begin{array}{c} \boldsymbol{f}_0^{\mathsf{T}} \boldsymbol{\beta} \\ \mathbf{F}\boldsymbol{\beta} \end{array} \right\}, \sigma^2 \left[\begin{array}{cc} 1 & \boldsymbol{r}_0^{\mathsf{T}} \\ \boldsymbol{r}_0 & \mathbf{R} \end{array} \right] \right), \tag{1.87}$$

where:

$\boldsymbol{f}_0$ is the vector of regressors evaluated at $\boldsymbol{x}^{(0)}$;

$\mathbf{F}$ is the regression matrix as in Eq. (1.33);

$\boldsymbol{r}_0$ is the vector of cross-correlations between the point $\boldsymbol{x}^{(0)}$ where the prediction is to be performed and each one of the observations whose terms read:

$$r_{0\,i} = R(\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(i)}, \boldsymbol{\theta}), \quad i = 1, \ldots, m; \tag{1.88}$$

$\mathbf{R}$ is the correlation matrix of the observations whose terms read:

$$R_{i\,j} = R(\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(j)}, \boldsymbol{\theta}), \quad i, j = 1, \ldots, m. \tag{1.89}$$

1.4.3 The best linear unbiased predictor

In this section, it is assumed that the whole autocovariance $\sigma^2 R(\bullet, \boldsymbol{\theta})$ is known, either because it was chosen and/or estimated from empirical data as detailed later in Section 1.4.5.

The following theorem gives the expression for the best linear unbiased predictor of the unobserved quantity of interest Y_0. It is also known as the *universal Kriging* predictor in the geostatistics related literature. The name "Kriging" is due to Matheron (1962) who named it after Krige (1951), a South-African mining engineer who is widely recognized as the first person to have proposed this prediction methodology.

Theorem 1.4.2. ***The best linear unbiased (or universal Kriging) predictor***
The best linear unbiased predictor of the unobserved quantity of interest $y_0 \equiv \mathscr{M}(\boldsymbol{x}^{(0)})$ under the two-stage prior model in Eq. (1.84) is the Gaussian random variate $\widehat{Y}_0$ with mean:

$$\mu_{\widehat{Y}_0} = \mathbb{E}\left[\widehat{Y}_0\right] = \mathbb{E}\left[Y_0\right] = \boldsymbol{f}_0^{\mathsf{T}}\,\widehat{\boldsymbol{\beta}} + \boldsymbol{r}_0^{\mathsf{T}}\mathbf{R}^{-1}\left(\boldsymbol{y} - \mathbf{F}\,\widehat{\boldsymbol{\beta}}\right), \tag{1.90}$$

and minimal variance:

$$\sigma^2_{\widehat{Y}_0} = \mathbb{E}\left[(\widehat{Y}_0 - Y_0)^2\right] = \sigma^2\left(1 - \boldsymbol{r}_0^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 + \boldsymbol{u}_0^{\mathsf{T}}\left(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\boldsymbol{u}_0\right), \tag{1.91}$$

where:

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y} \tag{1.92}$$

is the generalized least-squares estimate of the underlying regression problem (the first stage), and:

$$\boldsymbol{u}_0 = \mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 - \boldsymbol{f}_0. \tag{1.93}$$

Proof 1.4.1. *By definition the* so-called *best linear unbiased predictor $\widehat{Y}_0$ of $Y_0 \equiv Y(\boldsymbol{x}^{(0)})$ has the following properties. It is:*

- *linear, meaning that given a vector of weights $\boldsymbol{a}_0 \equiv \boldsymbol{a}(\boldsymbol{x}^{(0)}) \in \mathbb{R}^m$:*

$$\widehat{Y}_0 = \sum_{i=1}^{m} a_{0\,i}\,Y_i = \boldsymbol{a}_0^{\mathsf{T}}\,\boldsymbol{Y}, \tag{1.94}$$

- *unbiased:*

$$\mathbb{E}\left[\widehat{Y}_0 - Y_0\right] = 0, \tag{1.95}$$

- *the best (in a mean-square sense) amongst all linear unbiased predictors:*

$$\widehat{Y}_0 = \arg\min_{Y_0^\star \text{ linear unbiased}} \mathbb{E}\left[\left(Y_0^\star - Y_0\right)^2\right]. \tag{1.96}$$

Note that the expectations in the two latter equations are taken with respect to the assumed joint Gaussian distribution of $\left(Y_0, \boldsymbol{Y}\right)^{\mathsf{T}}$.

The problem thus consists in finding the optimal weight vector $\boldsymbol{a}_0^*$ that satisfies these three definitions and which consequently reads:

$$\boldsymbol{a}_0^* \equiv \arg\min_{\boldsymbol{a}_0 \in \mathbb{R}^m} \mathbb{E}\left[\left(\boldsymbol{a}_0^{\mathsf{T}}\,\boldsymbol{Y} - Y_0\right)^2\right] \quad s.t. \quad \mathbb{E}\left[\boldsymbol{a}_0^{\mathsf{T}}\,\boldsymbol{Y} - Y_0\right] = 0. \tag{1.97}$$

The random deviation between Y_0 and its linear prediction can be further elicited by replacing $\boldsymbol{Y}$ and Y_0 by their assumed expressions from the two-stage Gaussian process prior in Eq. (1.84):

$$\begin{aligned}\widehat{Y}_0 - Y_0 &= \boldsymbol{a}_0^\mathsf{T}\boldsymbol{Y} - Y_0 && (1.98)\\ &= \boldsymbol{a}_0^\mathsf{T}(\mathbf{F}\boldsymbol{\beta} + \boldsymbol{Z}) - (\boldsymbol{f}_0^\mathsf{T}\boldsymbol{\beta} + Z_0)) && (1.99)\\ &= \boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z} - Z_0 + (\boldsymbol{a}_0^\mathsf{T}\mathbf{F} - \boldsymbol{f}_0^\mathsf{T})\boldsymbol{\beta}. && (1.100)\end{aligned}$$

The unbiasedness property requires that:

$$\begin{aligned}\mathbb{E}\left[\widehat{Y}_0 - Y_0\right] &= \mathbb{E}\left[\boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z} - Z_0 + (\boldsymbol{a}_0^\mathsf{T}\mathbf{F} - \boldsymbol{f}_0^\mathsf{T})\boldsymbol{\beta}\right]\\ &= \mathbb{E}\left[\boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z} - Z_0\right] + (\boldsymbol{a}_0^\mathsf{T}\mathbf{F} - \boldsymbol{f}_0^\mathsf{T})\boldsymbol{\beta} = 0, && (1.101)\end{aligned}$$

where the first expected term is zero because of the zero-mean assumption for the underlying Gaussian process Z:

$$\mathbb{E}\left[\boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z} - Z_0\right] \equiv 0, \tag{1.102}$$

so that the non-bias constraint reduces to:

$$\boldsymbol{a}_0^\mathsf{T}\mathbf{F} - \boldsymbol{f}_0^\mathsf{T} = 0. \tag{1.103}$$

Using the above equation together with Eq. (1.100), the mean-squared prediction error further reads:

$$\begin{aligned}\mathbb{E}\left[\left(\widehat{Y}_0 - Y_0\right)^2\right] &= \mathbb{E}\left[\left(\boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z} - Z_0\right)^2\right] && (1.104)\\ &= \mathbb{E}\left[\boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z}\boldsymbol{Z}^\mathsf{T}\boldsymbol{a}_0 + Z_0^2 - 2\,\boldsymbol{a}_0^\mathsf{T}\boldsymbol{Z}\,Z_0\right] && (1.105)\\ &= \boldsymbol{a}_0^\mathsf{T}\,\mathbb{E}\left[\boldsymbol{Z}\boldsymbol{Z}^\mathsf{T}\right]\,\boldsymbol{a}_0 + \mathbb{E}\left[Z_0^2\right] - 2\,\boldsymbol{a}_0^\mathsf{T}\,\mathbb{E}\left[\boldsymbol{Z}\,Z_0\right] && (1.106)\\ &= \boldsymbol{a}_0^\mathsf{T}\,\sigma^2\mathbf{R}\,\boldsymbol{a}_0 + \sigma^2 - 2\,\boldsymbol{a}_0^\mathsf{T}\,\sigma^2\,\boldsymbol{r}_0 && (1.107)\end{aligned}$$

since $\sigma^2\mathbf{R} \equiv \mathbb{E}\left[\boldsymbol{Z}\boldsymbol{Z}^\mathsf{T}\right]$, $\sigma^2 \equiv \mathbb{E}\left[Z_0^2\right]$ and $\sigma^2\boldsymbol{r}(\boldsymbol{x}) \equiv \mathbb{E}\left[\boldsymbol{Z}\,Z_0\right]$ by definition. Thus the mean squared prediction error eventually reads:

$$\mathbb{E}\left[\left(\widehat{Y}_0 - Y_0\right)^2\right] = \sigma^2\left(1 + \boldsymbol{a}_0^\mathsf{T}\,\left(\mathbf{R}\boldsymbol{a}_0 - 2\,\boldsymbol{r}_0\right)\right). \tag{1.108}$$

Eq. (1.97) states that the best linear unbiased predictor $\widehat{Y}_0$ is the solution of an equality constrained optimization problem. Let us introduce a vector of Lagrange multipliers $\boldsymbol{\lambda}_0 \equiv \boldsymbol{\lambda}(\boldsymbol{x}^{(0)})$ to enforce the simplified equality constraint in Eq. (1.103) during the minimization of the mean squared prediction error in Eq. (1.108). The Lagrangian reads:

$$L(\boldsymbol{a}_0, \boldsymbol{\lambda}_0) = \sigma^2\left(1 + \boldsymbol{a}_0^\mathsf{T}\,\left(\mathbf{R}\boldsymbol{a}_0 - 2\,\boldsymbol{r}_0\right)\right) + \boldsymbol{\lambda}_0^\mathsf{T}\,\left(\boldsymbol{a}_0^\mathsf{T}\mathbf{F} - \boldsymbol{f}_0^\mathsf{T}\right). \tag{1.109}$$

Hence the associated Lagrange optimality conditions read as follows:

$$\left\{\begin{array}{rcl}\nabla_{\boldsymbol{a}_0} L &=& 2\,\sigma^2\left(\mathbf{R}\boldsymbol{a}_0 - \boldsymbol{r}_0\right) + \mathbf{F}\boldsymbol{\lambda}_0 = 0\\ \nabla_{\boldsymbol{\lambda}_0} L &=& \mathbf{F}^\mathsf{T}\boldsymbol{a}_0 - \boldsymbol{f}_0 = 0\end{array}\right. \tag{1.110}$$

The latter equation is indeed a linear system in the unknowns $\boldsymbol{a}_0$ and $\boldsymbol{\lambda}_0$ so that it may be rewritten in the following more convenient matrix form:

$$\left[\begin{array}{cc}\mathbf{R} & \mathbf{F}\\ \mathbf{F}^\mathsf{T} & \mathbf{0}\end{array}\right]\left\{\begin{array}{c}\boldsymbol{a}_0\\ \widetilde{\boldsymbol{\lambda}}_0\end{array}\right\} = \left\{\begin{array}{c}\boldsymbol{r}_0\\ \boldsymbol{f}_0\end{array}\right\} \quad \textit{with} \quad \widetilde{\boldsymbol{\lambda}}_0 = \frac{\boldsymbol{\lambda}_0}{2\,\sigma^2}. \tag{1.111}$$

Multiplying the first line of this system by $-\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}$ *on the left and adding the resulting quantity to the second line turns out to give a new equation in* $\widetilde{\boldsymbol{\lambda}}_0$ *only whose solution reads:*

$$\widetilde{\boldsymbol{\lambda}}_0^* = \left(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\left(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 - \boldsymbol{f}_0\right). \tag{1.112}$$

Substituting this solution in the first line of the linear system in Eq. (1.111) gives the sought optimal vector of weights $\boldsymbol{a}_0^*$ *which reads:*

$$\boldsymbol{a}_0^* = \mathbf{R}^{-1}\left(\boldsymbol{r}_0 - \mathbf{F}\left(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\left(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 - \boldsymbol{f}_0\right)\right). \tag{1.113}$$

Finally, the expression for the mean of the best linear unbiased predictor $\widehat{Y}_0$ *is obtained by replacing the sought weight vector* $\boldsymbol{a}_0$ *by this above solution in Eq. (1.94) and by using the symmetry of* $\mathbf{R}$*:*

$$\begin{aligned}
\mu_{\widehat{Y}_0} &= \boldsymbol{a}_0^{*\mathsf{T}}\boldsymbol{y} \\
&= [\boldsymbol{r}_0 - \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 - \boldsymbol{f}_0)]^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y} \\
&= \boldsymbol{r}_0^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y} - [(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 + (\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{f}_0]^{\mathsf{T}}\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y} \\
&= \boldsymbol{f}_0^{\mathsf{T}}\underbrace{(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y}}_{\widehat{\boldsymbol{\beta}}} + \boldsymbol{r}_0^{\mathsf{T}}\mathbf{R}^{-1}(\boldsymbol{y} - \mathbf{F}\underbrace{(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{y}}_{\widehat{\boldsymbol{\beta}}}).
\end{aligned}$$

The expression for its variance is similarly obtained by introducing $\boldsymbol{u}_0 \equiv \mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 - \boldsymbol{f}_0$ *in Eq. (1.113) for the sake of brevity, and by replacing the obtained expression for* $\boldsymbol{a}_0^*$ *in Eq. (1.108):*

$$\begin{aligned}
\sigma^2_{\widehat{Y}_0} &= \sigma^2[1 + \boldsymbol{a}_0^{*\mathsf{T}}(\mathbf{R}\boldsymbol{a}_0^* - 2\boldsymbol{r}_0)] \\
&= \sigma^2[1 + (\boldsymbol{r}_0 - \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0)^{\mathsf{T}}\mathbf{R}^{-1}((\boldsymbol{r}_0 - \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0) - 2\boldsymbol{r}_0)] \\
&= \sigma^2[1 - \underbrace{(\boldsymbol{r}_0 - \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0)^{\mathsf{T}}\mathbf{R}^{-1}(\boldsymbol{r}_0 + \mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0)}_{\text{“}(\boldsymbol{a}-\boldsymbol{b})^{\mathsf{T}}\mathbf{K}(\boldsymbol{a}+\boldsymbol{b}) = \boldsymbol{a}^{\mathsf{T}}\mathbf{K}\boldsymbol{a} - \boldsymbol{b}^{\mathsf{T}}\mathbf{K}\boldsymbol{b}\text{” because } \mathbf{K} \text{ is symmetric}}] \\
&= \sigma^2[1 - (\boldsymbol{r}_0^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 - (\mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0)^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0)] \\
&= \sigma^2[1 - \boldsymbol{r}_0^{\mathsf{T}}\mathbf{R}^{-1}\boldsymbol{r}_0 + \boldsymbol{u}_0^{\mathsf{T}}\underbrace{(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})}_{\mathbf{I}}(\mathbf{F}^{\mathsf{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\boldsymbol{u}_0].
\end{aligned}$$

1.4.4 Properties of the best linear unbiased predictor

1.4.4.1 Interpolation

The best linear unbiased predictor *interpolates* the observations in the dataset such as illustrated in Figure 1.9.

To prove it, let choose some $i \in [\![1;\, m]\!]$ and define $\boldsymbol{f}_i \equiv \boldsymbol{f}(\boldsymbol{x}^{(i)})$ which is also the i-th row of the regression matrix $\mathbf{F}$, and $\boldsymbol{r}_i \equiv \boldsymbol{r}\left(\boldsymbol{x}^{(i)}\right)$ which is also the i-th row of the correlation matrix $\mathbf{R}$. Let also point out that the following relationship holds:

$$\mathbf{R}^{-1}\boldsymbol{r}_i = \boldsymbol{e}_i, \tag{1.114}$$

where $\boldsymbol{e}_i$ is the i-th basis vector of $\mathbb{R}^m$ which has all-zero components except its i-th component equal to one, because the latter dot product is also the i-th column of $\mathbf{R}^{-1}\mathbf{R} = \mathbf{I}$. Using Eq. (1.90) with $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(i)}$ and replacing $\mathbf{R}^{-1}\boldsymbol{r}_i$ with $\boldsymbol{e}_i$ leads to:

$$\mu_{\widehat{Y}_i} = \boldsymbol{f}_i^{\mathsf{T}}\widehat{\boldsymbol{\beta}} + \boldsymbol{e}_i^{\mathsf{T}}\left(\boldsymbol{y} - \mathbf{F}\widehat{\boldsymbol{\beta}}\right) = y_i, \tag{1.115}$$

which proves that the best linear unbiased predictor interpolates the observations in the dataset.

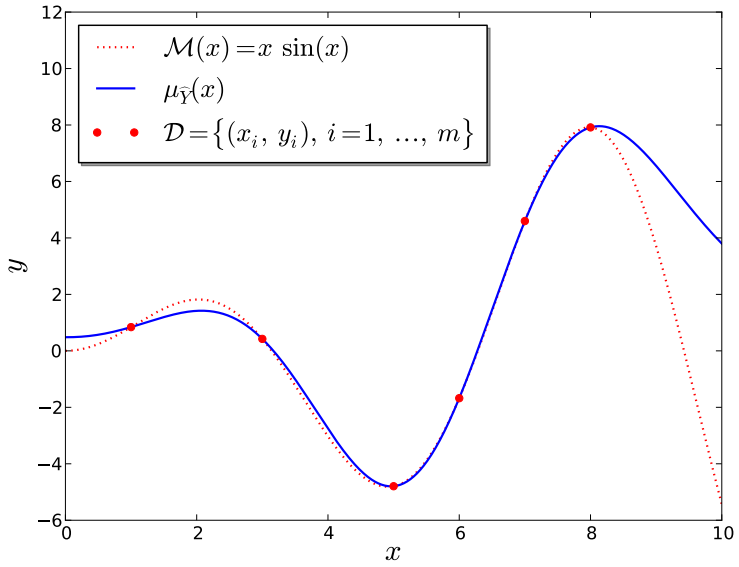


Figure 1.9: Interpolation property of the best linear unbiased predictor for the one-dimensional function $\mathscr{M}(x) = x\ \sin(x)$.

1.4.4.2 Asymptotic consistency

Vazquez (2005, pp. 132–156) points out that the universal Kriging predictor is asymptotically consistent when the actual autocovariance $\sigma^2 R(\bullet, \boldsymbol{\theta})$ of the underlying Gaussian process is *regular*.

The proof consists in showing that the mean squared prediction error of the best linear unbiased predictor tends to zero everywhere as the dataset $\mathscr{D} = \left\{(\boldsymbol{x}^{(i)}, \boldsymbol{y}_i),\ i = 1, \ldots, m\right\}$ becomes dense in $\mathbb{X}$ (meaning that all $\boldsymbol{x}^{(i)}$'s are unique elements of $\mathbb{X}$):

$$\mathbb{E}\left[(\widehat{Y}_0 - Y_0)^2\right] \underset{m \to \infty}{\longrightarrow} 0 \quad \text{with} \quad \boldsymbol{x}^{(i)} \neq \boldsymbol{x}^{(j)}\ \forall\, (i, j) \in [\![1;\ m]\!]^2. \tag{1.116}$$

Indeed, when considering Eq. (1.91) for $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(i)}$, one proves that the term $\boldsymbol{u}_i$ defined in Eq. (1.93) vanishes since:

$$\boldsymbol{u}_i = \mathbf{F}^{\mathsf{T}}\, \boldsymbol{e}_i - \boldsymbol{f}_i = \boldsymbol{f}_i - \boldsymbol{f}_i = \mathbf{0}. \tag{1.117}$$

Thus,

$$\sigma^2_{\widehat{Y}_i} = \sigma^2 \left(1 - \boldsymbol{r}_i^{\mathsf{T}}\, \boldsymbol{e}_i + 0\right). \tag{1.118}$$

This means that the Kriging variance is zero for all $\boldsymbol{x}^{(i)}$'s (which are assumed to be dense in $\mathbb{X}$) if and only if $\boldsymbol{r}_i = \boldsymbol{e}_i$ which is only true for *regular* autocorrelation function for which $R(|\boldsymbol{x} - \boldsymbol{x}|) = R(\mathbf{0}) = 1$.

As a consequence this consistency property might not hold in some practical applications where the degree of regularity is empirically chosen or estimated. As a matter of fact, it makes the use of the forthcoming *empirical* best linear unbiased predictors as deterministic surrogates for some experiment $\mathcal{M}$ a *heuristic* approach. It is worth mentioning though that Vazquez (2005) also demonstrated through numerical experiments on a selected set of examples that empirical best linear unbiased predictors tend to reach asymptotic consistency when the *regularity* of the autocovariance is accurately estimated.

This property is illustrated in Figure 1.10. It can be seen that the predictor variance reduces to zero at all observations so that it tends to be zero everywhere when the design of experiments $\mathcal{X} = \left\{\boldsymbol{x}^{(i)},\, i = 1, \ldots, m\right\}$ becomes dense in $\mathbb{X}$.

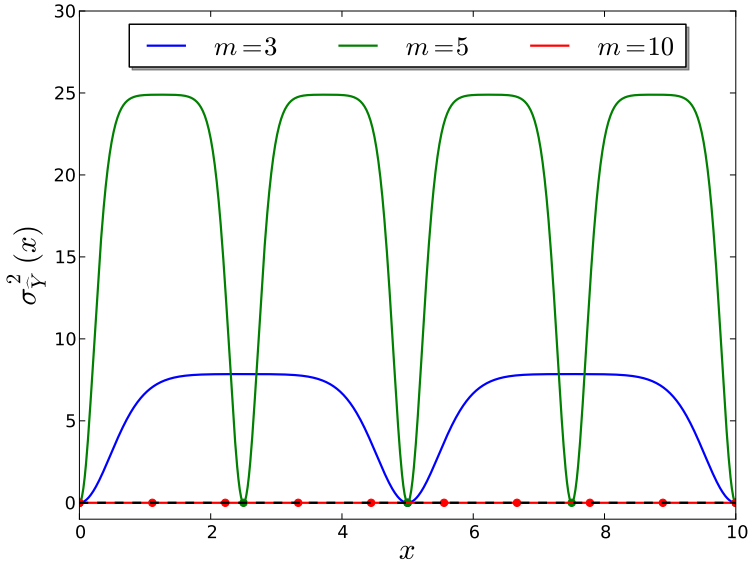


Figure 1.10: Asymptotic consistency of the best linear unbiased predictor for the one-dimensional function $\mathcal{M}(x) = x\sin(x)$. The scale parameter ℓ in the chosen squared exponential autocovariance model has been estimated by means of the maximum likelihood estimation technique (see Section 1.4.5.2) for each m.

1.4.4.3 Gaussianity

As opposed to its forthcoming Bayesian counterpart, the frequentist interpretation of the Kriging prediction methodology does not necessarily require the Gaussian assumption for the underlying stochastic process. Indeed, it only needs the assumptions regarding its stationary second-order statistics and the proof makes only use of the linearity property of the expectation operator. However the Gaussian assumption is transmitted to the predictor $\widehat{Y}_0$ because the Kriging predictor is a linear combination of the Gaussian observations $\boldsymbol{Y}$:

$$\widehat{Y}_0 = \boldsymbol{a}_0^{\mathsf{T}}\,\boldsymbol{Y} \sim \mathcal{N}_1(\mu_{\widehat{Y}_0},\, \sigma^2_{\widehat{Y}_0}). \tag{1.119}$$

Thus it is a convenient property to compute:

- exceedance probabilities:

$$\mathscr{P}\left[\widehat{Y}_0 \leq t\right] = \Phi\left(\frac{t-\mu_{\widehat{Y}_0}}{\sigma_{\widehat{Y}_0}}\right), \tag{1.120}$$

  where Φ denotes the CDF of the standard normal distribution (see Eq. (1.20));

- confidence intervals:

$$Y_0 \in \left[\mu_{\widehat{Y}_0} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sigma_{\widehat{Y}_0};\ \mu_{\widehat{Y}_0} + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sigma_{\widehat{Y}_0}\right] \text{ with prob. } 1-\alpha, \tag{1.121}$$

  where Φ^{-1} denotes the inverse CDF of the standard normal distribution. For instance, the value $\alpha = 5\%$ implies that $\Phi^{-1}(1-0.05/2) = 1.96$, and corresponds to a $1-\alpha = 95\%$ confidence interval. This value will be widely used throughout this manuscript.

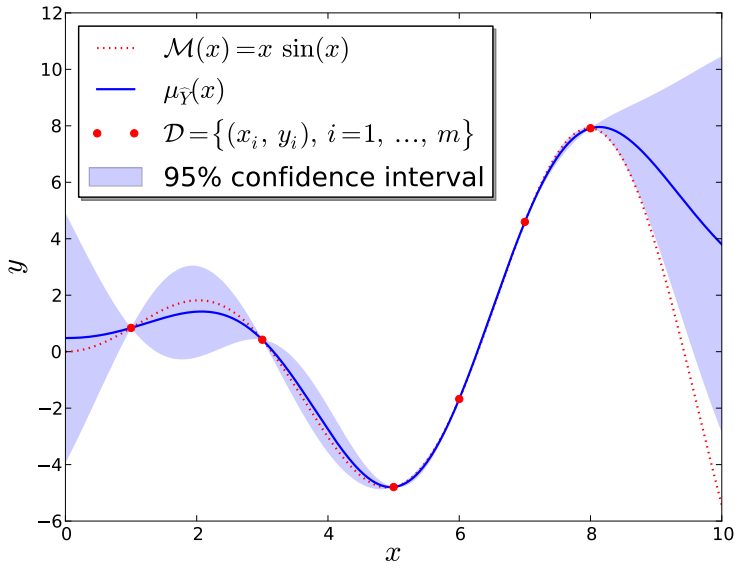


Figure 1.11: Gaussianity of the best linear unbiased predictor for the one-dimensional function $\mathcal{M}(x) = x\ \sin(x)$.

1.4.5 Empirical best linear unbiased predictors

It was previously assumed, at the beginning of Section 1.4.3, that the autocovariance function $\sigma^2 R(\bullet, \boldsymbol{\theta})$ was known. In the present context of computer experiments though it is never the case, so that the user has to *(i)* choose a family of autocorrelation functions *e.g.* amongst the one depicted in Section 1.3, and then *(ii)* estimate the unknown hyper-parameters $\boldsymbol{\theta}$ and the variance σ^2 from the dataset $\mathscr{D} = \left\{(\boldsymbol{x}^{(i)}, y_i),\ i = 1, \ldots, m\right\}$. The resulting best linear unbiased predictors are called *empirical best linear unbiased predictors* in the book by Santner et al. (2003) because they result of an empirical (although motivated) choice. Some of these estimation techniques are reviewed in the sequel.

1.4.5.1 Variogram estimation

The use of this technique is strictly limited to the field of geostatistics where it was initially proposed (see *e.g.* Cressie, 1993, for a review). It consists in computing an empirical (binned) estimate of the autocovariance function from the dataset $\mathscr{D}$ assuming the underlying stationary Gaussian process Z is ergodic. Loosely speaking, the procedure is alike the histogram estimation of probability density functions which is widely used in statistical inference. However it is not applicable to computer experiments because it is hardly generalizable in dimensions greater than $n = 3$ when accounting for anisotropy in the autocorrelation function, and it does not permit to account for the non-stationary trend $\boldsymbol{f}(\boldsymbol{x})^{\mathsf{T}}\boldsymbol{\beta}$. The limitation with respect to the dimension is not a problem in geostatistics because the Kriging prediction methodology is used to estimate the spatial variability of properties in a 2- or 3-dimensional space.

It is also important to note that the property of ergodicity for a stationary Gaussian process is only true if the dataset $\mathscr{D}$ covers a domain $\mathbb{X}$ that is sufficiently large when compared to its correlation lengths. According to the author's experience, in a one-dimensional input space $\mathbb{X} = [0;\, L]$, the true correlation length ℓ of the Gaussian process must be less than $L/30$ on purpose to yield satisfactory estimates (provided the sample path is observed on a sufficiently dense partition of $\mathbb{X}$). This is rarely the case for smooth and slowly-varying computer experiments. Strictly speaking, ergodicity is rigorously defined with respect to the *covariance integral range* which depends on the autocorrelation function as a whole and not only on the correlation length ℓ (see *e.g.* Chilès and Delfiner, 1999, Chapters 1 and 2).

1.4.5.2 Maximum likelihood estimation

The *maximum likelihood estimation* (MLE) technique is better suited to computer experiments because it does not depend on the dimension n of the input space $\mathbb{X}$.

Recall that the likelihood of the observations $\boldsymbol{y}$ is defined with respect to its multivariate normal distribution:

$$\mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}\right) = \frac{1}{\left((2\,\pi\,\sigma^2)^m\,[\det \mathbf{R}(\boldsymbol{\theta})]\right)^{1/2}} \exp\left[-\frac{1}{2\,\sigma^2}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta})^{\mathsf{T}}\mathbf{R}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta})\right], \tag{1.122}$$

which depends on $\boldsymbol{\beta}$, σ^2 and $\boldsymbol{\theta}$ (through the correlation matrix $\mathbf{R}$).

The corresponding opposite log-likelihood then reads:

$$\begin{aligned} -\log \mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}\right) = {} & \frac{1}{2\,\sigma^2}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta})^{\mathsf{T}}\mathbf{R}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta}) + \frac{m}{2}\log(2\,\pi) \\ & + \frac{m}{2}\log\left(\sigma^2\right) + \frac{1}{2}\log\left([\det \mathbf{R}(\boldsymbol{\theta})]\right). \end{aligned} \tag{1.123}$$

The maximum likelihood estimates of $\boldsymbol{\beta}$ and σ^2 were already derived in Section 1.2.3. These two solutions are the so-called *generalized least squares estimates* and depend both here on $\boldsymbol{\theta}$:

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\mathbf{F}^{\mathsf{T}}\mathbf{R}(\boldsymbol{\theta})^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathsf{T}}\mathbf{R}(\boldsymbol{\theta})^{-1}\boldsymbol{y}, \tag{1.124}$$

$$\widehat{\sigma^2}(\boldsymbol{\theta}) = \frac{1}{m}(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}})^{\mathsf{T}}\mathbf{R}(\boldsymbol{\theta})^{-1}(\boldsymbol{y}-\mathbf{F}\widehat{\boldsymbol{\beta}}). \tag{1.125}$$

Plugging these two solutions in Eq. (1.123) leads to a new expression that depends on $\boldsymbol{\theta}$ only:

$$\begin{aligned} -\log \mathsf{L}\left(\boldsymbol{y} \mid \boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}\right) &= \frac{m}{2} + \frac{m}{2}\log(2\pi) + \frac{m}{2}\log\left(\widehat{\sigma^2}(\boldsymbol{\theta})\right) + \frac{1}{2}\log([\det \mathbf{R}(\boldsymbol{\theta})]) \\ &= \frac{m}{2}\log(\psi(\boldsymbol{\theta})) + \frac{m}{2}\left(\log(2\pi)+1\right), \end{aligned} \tag{1.126}$$

where the so-called *reduced likelihood function* has been introduced:

$$\psi(\boldsymbol{\theta}) = \widehat{\sigma^2}(\boldsymbol{\theta})\,[\det \mathbf{R}(\boldsymbol{\theta})]^{1/m}. \tag{1.127}$$

The maximum likelihood estimate of $\boldsymbol{\theta}$ is eventually defined as the following global minimizer:

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \quad \psi(\boldsymbol{\theta}). \tag{1.128}$$

because minimizing ψ is equivalent to minimizing the negative log-likelihood as the natural logarithm and the affine functions are strictly increasing.

The global optimization problem in Eq. (1.128) cannot be solved analytically and one usually resorts to numerical global optimization techniques. For instance, the DACE Matlab toolbox by Lophaven et al. (2002) uses the BOXMIN algorithm which is a sort of multivariate dichotomy algorithm while the DiceKriging R package by Roustant et al. (2010) resorts to a gradient-based genetic algorithm by Sekhon and Mebane (2011).

It is worth mentioning here the work by Marrel (2005, 2008); Marrel et al. (2008) who investigated the numerical difficulties related to this global optimization problem by studying the properties of the reduced likelihood function depending on the chosen autocorrelation function and the size m of the dataset $\mathscr{D}$. The authors pointed out that the problem is particularly ill-posed when both the squared exponential correlation function is used and the dataset $\mathscr{X} = \left\{\boldsymbol{x}^{(i)}, i = 1, \ldots, m\right\}$ becomes dense in $\mathbb{X}$. The author explains this phenomenon by observing that the squared exponential autocorrelation assumes infinite regularity (degree of differentiability) for the experiment $\mathscr{M}$ so that the unavoidable noise in the observations $\boldsymbol{y}$ raises some numerical inconsistencies with respect to that former assumption.

This observation is also mentioned by Vazquez (2005) who recommends the use of the Matérn autocorrelation function instead of the squared-exponential because it allows one to explicitly control the finite regularity of the experiment through one of its parameters (see Section 1.3.4).

Another workaround which is widely implemented in many Kriging softwares and packages consists in raising the diagonal of the correlation matrix $\mathbf{R}$ with a quantity ε that is negligible in front of the variance of the underlying Gaussian process. This improves the conditioning of the correlation matrix and makes its Cholesky decomposition possible. However Vazquez (2005) warns that the interpolating and consistency properties enunciated in Section 1.4.4 do not hold anymore when this technique is used, although it almost holds when ε is sufficiently small. The quantity ε is often referred to as a *nugget variance* because of its historical background: this technique was initially proposed by geostatisticians that had to deal with *in situ* experiments containing irregularities often due to the presence of singularities in the soil (nuggets).

The MLE algorithm used for the applications presented in this thesis is the one implemented in the DACE toolbox by Lophaven et al. (2002) where the first guess is derived under an isotropic assumption following the idea developed in Marrel et al. (2008) based on the work by Welch et al. (1992).

1.4.5.3 Cross-validated empirical best linear unbiased predictors

The cross-validation technique is a very popular tool for *model selection* in the supervised learning related literature. The practice is even so large that it is hard to say who proposed it in the first place, although Allen (1971) and Stone (1974) are often cited as the earliest references.

The general principle of cross-validation techniques consists in splitting the whole dataset $\mathscr{D}$ in K mutually exclusive and collectively exhaustive subsets $\{\mathscr{D}_k, k = 1, \ldots, K\}$ such that:

$$\mathscr{D}_i \cap \mathscr{D}_j = \emptyset \quad \forall (i, j) \in [\![1; K]\!]^2 \quad \text{and} \quad \bigcup_{k=1}^{K} \mathscr{D}_k = \mathscr{D}. \tag{1.129}$$

The k-th ($k \in [\![1; K]\!]$) set of cross-validated predictions is obtained by fitting the model using all the subsets but the k-th one $\mathscr{D} \setminus \mathscr{D}_k$ and predicting it on that specific k-th fold that was left apart. The leave-one-out cross-validation procedure corresponds to the special case for which $K = m$.

Dubrule (1983) showed that leave-one-out predictions for the universal Kriging predictor can be efficiently obtained by inverting the matrix involved in Eq. (1.111) only once as for ordinary least-squares regression (see *e.g.* Saporta, 2006, Chapter 17). The i-th leave-one-out Kriging prediction built from the reduced dataset $\mathscr{D}_{-i} = \mathscr{D} \setminus (\boldsymbol{x}^{(i)}, y_i)$ and evaluated at point $\boldsymbol{x}^{(i)}$ is denoted by $\widehat{Y}_{-i}$ in the sequel.

Indeed, by introducing the matrix $\mathbf{B}$ whose terms read:

$$B_{ij} = S^{-1}_{ij}, \quad i, j = 1, \ldots, m, \tag{1.130}$$

where S^{-1}_{ij} denotes the terms of the inverse of the matrix:

$$\mathbf{S} = \begin{bmatrix} \sigma^2 \mathbf{R} & \mathbf{F} \\ \mathbf{F}^{\mathsf{T}} & \mathbf{0} \end{bmatrix}, \tag{1.131}$$

Dubrule (1983) showed that the m leave-one-out predictions of the universal Kriging mean and variance respectively read:

$$\mu_{\widehat{Y}_{-i}} = -\sum_{\substack{j=1 \\ j\neq i}}^{m} \frac{B_{ij}}{B_{ii}} y_j, \quad i = 1, \ldots, m, \tag{1.132}$$

and

$$\sigma^2_{\widehat{Y}_{-i}} = \frac{1}{B_{ii}}, \quad i = 1, \ldots, m. \tag{1.133}$$

The paper also generalizes the use of this technique to any K-folds cross-validation of the universal Kriging predictor.

Then leave-one-out predictions allow one to estimate a meaningful coefficient of determination even when the model is interpolating as for the Kriging predictor:

$$Q^2 = 1 - \frac{1}{m}\sum_{i=1}^{m}\left(\frac{\mu_{\widehat{Y}_{-i}} - y_i}{\sigma_{\widehat{Y}_{-i}}}\right)^2, \tag{1.134}$$

The definition of this *generalization score* varies in the Kriging-related literature depending on the context. Some researchers use the prediction variance in the denominator as in the latter equation while others use the empirical variance of the observations in the dataset. As for the usual coefficient of determination, the higher the generalization score, the better the predictor. The generalization score cannot exceed 1.

1.4.6 Bayesian predictors

It has already been shown in Section 1.2.4 that the Bayesian framework allows one to compute posterior distributions for the unknown parameters of a given model and that this posterior distribution is conditioned on both the prior knowledge (or belief) expressed in terms of a prior distribution and the data. In the present context, Santner et al. (2003, amongst others) show that it is possible to propagate the uncertainty carried by the posterior distributions of the unknown parameters down to the final *predictive distribution* Y_0 of the quantity of interest. This section reviews the fundamental concepts of this prediction methodology and some reference results available in the literature.

1.4.6.1 Outline of the Bayesian prediction methodology

First, let us group the unknown parameters of the two-stage Gaussian process model of Section 1.4.2 in the vector:

$$\boldsymbol{\omega} \equiv \left(\boldsymbol{\beta}^\mathsf{T}, \sigma^2, \boldsymbol{\theta}^\mathsf{T}\right)^\mathsf{T}, \tag{1.135}$$

and assume it has prior distribution $p(\boldsymbol{\omega})$. Note that practical priors often have the following form (Santner et al., 2003):

$$p(\boldsymbol{\omega}) = p(\boldsymbol{\beta} \mid \sigma^2)\, p(\sigma^2)\, p(\boldsymbol{\theta}), \tag{1.136}$$

because it makes the approach tractable, and because it is reasonable to assume that the parameters in the autocorrelation function are independent of $[\boldsymbol{\beta}, \sigma^2]$ *a priori*.

Then, the posterior distribution of these unknown parameters (conditional on the data $\boldsymbol{y}$) can be obtained by applying Bayes' theorem:

$$p(\boldsymbol{\omega} \mid \boldsymbol{y}) \propto p(\boldsymbol{y} \mid \boldsymbol{\omega})\, p(\boldsymbol{\omega}), \tag{1.137}$$

as in Section 1.2.4 for the specific case of a non-informative Jeffreys prior on $\boldsymbol{\beta}$ and σ^2. Deriving this posterior distribution for any given prior distribution is a first computational challenge, even though the literature contains many reference analytical results (see *e.g.* Zellner, 1971).

The joint posterior distribution of $[Y_0, \boldsymbol{\omega} \mid \boldsymbol{Y}]$ is defined as follows:

$$p(y_0, \boldsymbol{\omega} \mid \boldsymbol{y}) \equiv p(\widehat{y}_0 \mid \boldsymbol{y}, \boldsymbol{\omega})\, p(\boldsymbol{\omega} \mid \boldsymbol{y}), \tag{1.138}$$

where the first term corresponds to the posterior distribution of the prediction given the parameters $\boldsymbol{\omega}$. Deriving the posterior $p(\widehat{y}_0 \mid \boldsymbol{y}, \boldsymbol{\omega})$ for any given multivariate second-order probability distribution is not straightforward.

For the multivariate Gaussian distribution though, Theorem A.4.2 of Appendix A states that this conditional distribution is another Gaussian distribution:

$$[Y_0 \mid \boldsymbol{Y}, \boldsymbol{\omega}] \sim \mathcal{N}_1 \left(\boldsymbol{f}_0^{\mathsf{T}} \boldsymbol{\beta} + \boldsymbol{r}_0^{\mathsf{T}} \mathbf{R}^{-1} (\boldsymbol{Y} - \mathbf{F}\boldsymbol{\beta}),\, \sigma^2 (1 - \boldsymbol{r}_0^{\mathsf{T}} \mathbf{R}^{-1} \boldsymbol{r}_0) \right). \tag{1.139}$$

This theorem is the reason why the Gaussian process prior assumption formulated in Section 1.4.2 is so popular in the literature on Bayesian prediction. Note that the mean of this conditional distribution already corresponds to the best linear unbiased predictor $\mu_{\widehat{Y}_0}$ of Theorem 1.4.2 while the variance lacks an additive term compared to Eq. (1.91). The reason for this is that the posterior distribution in Eq. (1.139) assumes the location parameters $\boldsymbol{\beta}$ are known whereas the distribution of the universal Kriging predictor of Theorem 1.4.2 sums up the uncertainty induced by a non-informative prior on the unknown parameters $\boldsymbol{\beta}$.

The last step of the Bayesian prediction methodology consists in marginalizing the joint posterior distribution by integrating $\boldsymbol{\omega}$ out. Indeed the posterior predictive distribution is defined as follows:

$$p(\widehat{y}_0) \equiv p(y_0 \mid \boldsymbol{y}) \equiv \int p(y_0, \boldsymbol{\omega} \mid \boldsymbol{y})\, \mathrm{d}\boldsymbol{\omega}. \tag{1.140}$$

This very last step constitutes the second computational challenge.

1.4.6.2 Some reference results available in the literature on Gaussian process predictors

For instance, Santner et al. (2003, pp. 116–118) show that the universal Kriging predictor of Theorem 1.4.2 corresponds to the case where no prior knowledge for $\boldsymbol{\beta}$ is available while σ^2 and $\boldsymbol{\theta}$ are known. The corresponding non-informative prior distribution thus reads:

$$p(\boldsymbol{\omega}) = p(\boldsymbol{\beta}) \propto 1, \tag{1.141}$$

and the posterior predictive distribution is:

$$\widehat{Y}_0 \sim \mathcal{N}_1\left(\boldsymbol{f}_0^{\mathsf{T}}\,\widehat{\boldsymbol{\beta}} + \boldsymbol{r}_0^{\mathsf{T}}\,\mathbf{R}^{-1}(\boldsymbol{Y} - \mathbf{F}\,\widehat{\boldsymbol{\beta}}),\, \sigma^2\,(1 - \boldsymbol{r}_0^{\mathsf{T}}\,\mathbf{R}^{-1}\,\boldsymbol{r}_0 + \boldsymbol{u}_0^{\mathsf{T}}\,\mathbf{F}^{\mathsf{T}}\,\mathbf{R}^{-1}\,\mathbf{F}\,\boldsymbol{u}_0)\right), \tag{1.142}$$

where $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{u}_0$ were already defined in Eq. (1.92) and Eq. (1.93), respectively.

Table 1.1 summarizes the results available in the literature for the posterior predictive distribution $\widehat{Y}_0$ depending on the available prior knowledge. From the first to the last row, the computational complexity considerably increases, and so does the variance of the posterior predictive distribution as reported by Handcock and Stein (1993).

Prior			**Posterior**
$[\boldsymbol{\beta} \mid \sigma^2]$	$[\sigma^2]$	$[\boldsymbol{\theta}]$	$\widehat{Y}_0 \equiv [Y_0 \mid \boldsymbol{Y}]$
Known	*Known*	*Known*	Gaussian as in Eq. (1.139)
Non-informative	*Known*	*Known*	Gaussian as in Eq. (1.142)
Non-informative or Gaussian	Non-informative or Inv.-Gamma	*Known*	Shifted Student
Non-informative or Gaussian	Non-informative or Inv.-Gamma	informative	Non-analytical posterior distribution

Table 1.1: Nature of the predictive posterior distribution depending on the prior knowledge available expressed in terms of a joint prior distribution.

The first row is the simplest form of Bayesian Kriging. It could be referred to as *simple Kriging*, even though the simple Kriging predictor does not involve any regression stage (see *e.g.* Ginsbourger, 2009, pp. 75–77). The second row is the universal Kriging case of Theorem 1.4.2. The third row corresponds to Theorem 4.1.2 in the book of Santner et al. (2003). The authors point out that the number of degrees of freedom in the shifted Student distribution interestingly increases with the quantity of prior knowledge available. As a consequence, the variance of the posterior predictive distribution decreases and tends towards the universal Kriging case.

The last row, which corresponds to the case where one wants to propagate the uncertainty in all the parameters including $\boldsymbol{\theta}$, is greatly complicated by the way these parameters enter in the likelihood (recall that $\boldsymbol{\theta}$ defines $\boldsymbol{r}_0$ and $\mathbf{R}$). It involves the use of numerical integration techniques to compute the posterior predictive distribution (see *e.g.* Handcock and Stein, 1993), and/or Markov chain Monte Carlo (MCMC) to sample from the posterior predictive distribution. Regarding the MCMC sampling of two-stage Gaussian process models, the interested reader is referred to the PyMC Python package by Patil et al. (2010) and its Gaussian process subpackage. The latter most comprehensive case is rather cumbersome to implement and much more computationally expensive than the other cases which are analytical.

As a conclusion, it has been considered that for the present application of Gaussian process predictors as surrogates for expensive-to-evaluate black-box functions, the universal Kriging model offers a sufficient degree of fidelity even if its variance underestimates the

real one associated with our prior knowledge of σ^2 and $\boldsymbol{\theta}$. In another identification context though (such as the identification of random fields), the complete Bayesian approach corresponding to the last row of Table 1.1 may provide more meaningful results.

1.5 Illustration

1.5.1 Least-squares linear regression

This illustration is provided so as to show the interrelation between Gaussian process regression and ordinary least-squares regression through the use of the singular nugget autocorrelation function which assumes spatial independence of the observations as ordinary least-squares regression does. The proposed regression problem consists in fitting the slope and the intercept of a line from the dataset $\mathscr{D} = \left\{(x^{(i)}, y_i), i = 1, \ldots, 20\right\}$ where the $x^{(i)}$'s split the interval $\mathbb{X} = [-4;\ 4]$ in equal subintervals, and the y_i's are computed from the following noisy linear experiment:

$$\mathscr{M}(x) = a\,x + b + \varepsilon \tag{1.143}$$

where $a = 1$, $b = 3$ and $\varepsilon \sim \mathscr{N}_1(0,\, 1)$.

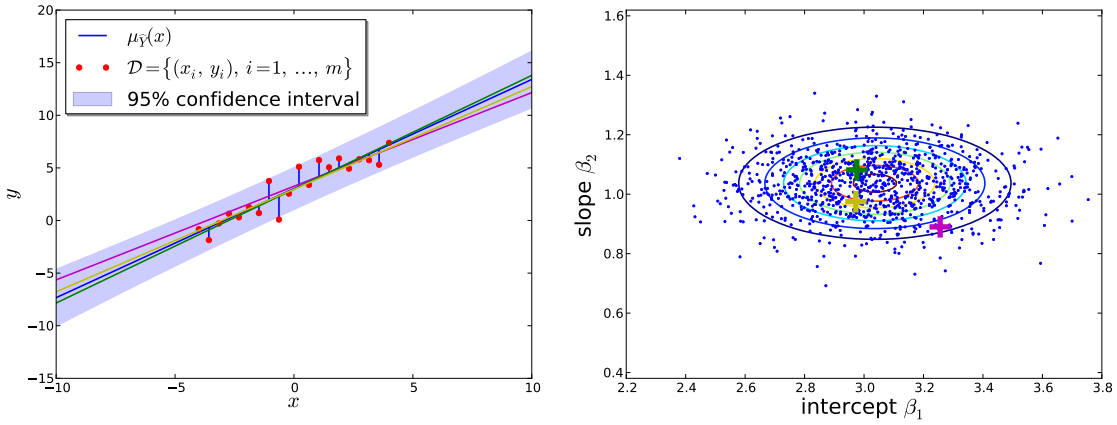
The fitted Gaussian process model assumes a linear trend $\boldsymbol{f} = \{1,\, x\}$, and a parameter-free nugget autocorrelation function (as defined in Eq. (1.74)). This means that the observations falls under the ordinary least-squares form:

$$Y_i = \beta_1 + \beta_2\, x^{(i)} + Z_i, \quad i = 1, \ldots, 20 \tag{1.144}$$

where $Z_i = Z(x^{(i)})$ assumes spatial independence of the observations. The problem is reduced to the inference of the regression weights: the intercept β_1 and the slope β_2.

It can be seen from Figure 1.12 that the best linear unbiased predictor still interpolates the data as the Dirac term from the nugget autocorrelation function forces the prediction to fetch the observations. Practical implementation of the best linear unbiased predictor consists in using an autocorrelation function that introduces a non-zero correlation length to smooth the prediction.

The variance is assumed to be known even if it has been estimated from the observations in the dataset using the maximum likelihood principle, so that the posterior distribution of the regression weights is Gaussian and the final predictive distribution is also Gaussian (universal Kriging). The 95% confidence interval represented by the shaded area is computed from this convenient assumption. In this particular case, the posterior distribution for the intercept β_1 and the slope β_2 seem uncorrelated, and the uncertainty in the intercept is greater than the one in the slope.



(a) $p(\widehat{y} \mid \sigma^2)$ (b) $p(\boldsymbol{\beta} \mid \boldsymbol{y}, \sigma^2)$

Figure 1.12: Interrelation between Gaussian process regression and ordinary least-squares regression. Panel (a) shows the resulting best linear unbiased predictor with 95% confidence intervals. Panel (b) shows the posterior predictive distribution of the regression weights (assuming the variance σ^2 is known) together with a few random samples (the points), three of which (the "+") are used to plot samples of the fitted linear model on panel (a).

1.5.2 Gaussian process regression

This section provides an empirical comparison of different empirical best linear unbiased predictors fitted to a unique sparse set of observations of the one-dimensional model $\mathscr{M}(x) = x\ \sin(x)$. All predictors assumed a regression model reduced to a constant (*i.e.* $\boldsymbol{f} = \{1\}$ and $\boldsymbol{\beta} = \beta_1$), and only the autocorrelation functions are varied in order to illustrate their properties. The scale parameter ℓ of each correlation model is fitted using the maximum likelihood principle which consists in finding the global minimum of the reduced likelihood function ψ in Eq. (1.128). The reduced likelihood function ψ is plotted together with the generalization score Q^2 in Eq. (1.134) as functions of the inverse scale parameter $1/\ell$.

A first observation from this experimental study is that the interpolation and consistency properties enunciated in Section 1.4.4 hold for all these usual regular autocorrelation functions as illustrated on the left-hand panels of Figures 1.14 and 1.15. It is also important to note that the prediction variance $\sigma^2_{\widehat{Y}}$ does not depend on the observations as it can be seen from its definition in Eq. (1.91). However, it clearly depends on both the design of experiment $\mathscr{X}$ and the autocorrelation function as illustrated in Figure 1.13.

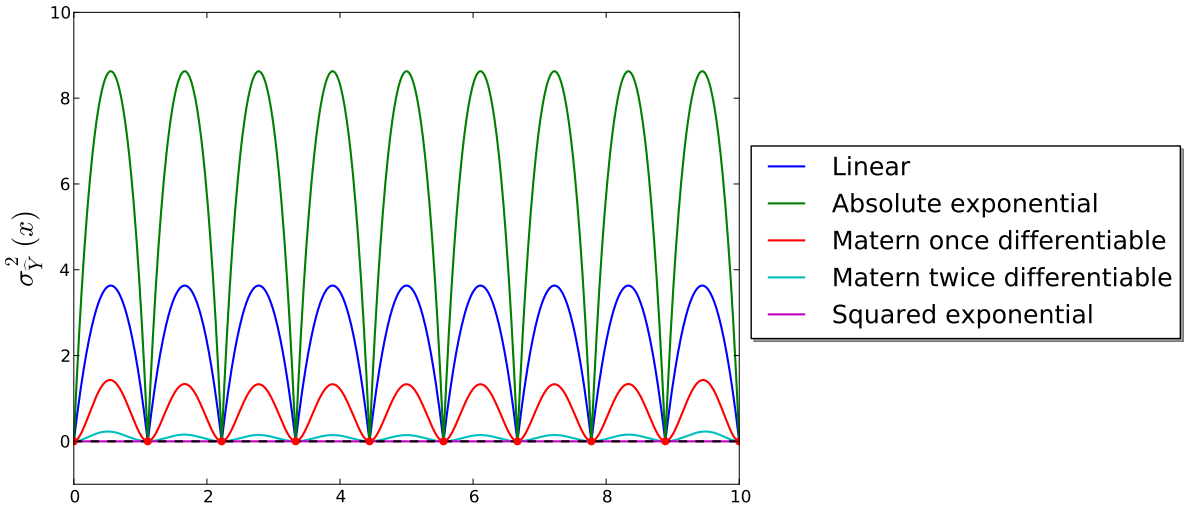
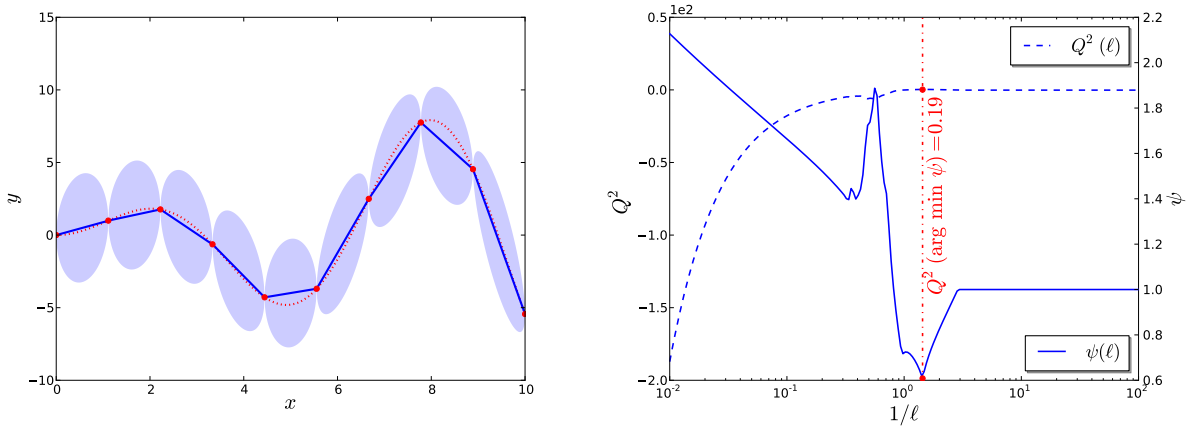
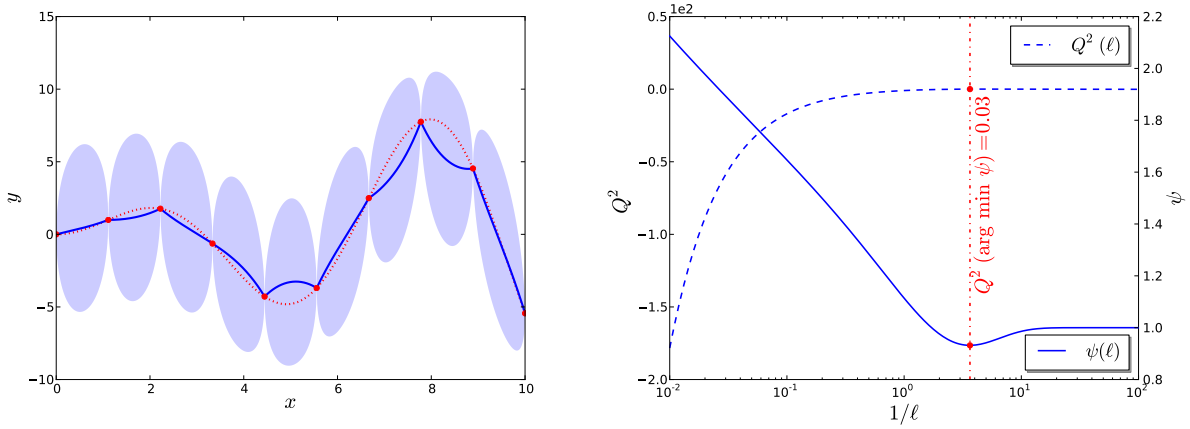


Figure 1.13: Prediction variance depending on the autocorrelation function for the approximation of the model $\mathscr{M}(x) = x\ \sin(x)$.

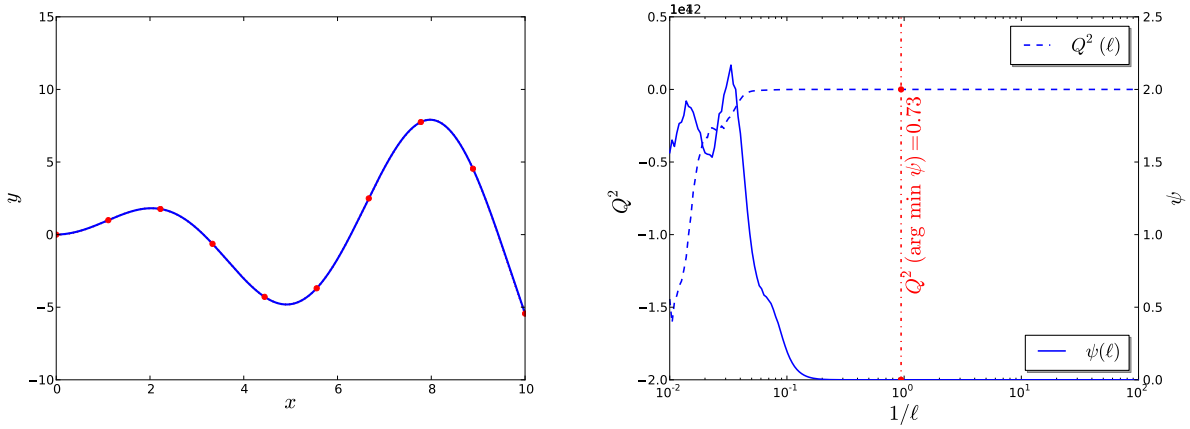
Another important observation concerns the conditioning of the maximum likelihood estimation problem. It can be seen from the right-hand panels of Figures 1.14 and 1.15 that the sharper the autocorrelation function, the easier the minimization of the reduced likelihood function ψ. The linear correlation appears as an exception as it features several local *extrema* although it has a clear global *minimum*. The squared exponential autocorrelation is clearly the worst case for a tractable inference, because the reduced likelihood function features several local *extrema* and a large plateau containing the sought optimal solution. On the contrary, the more regular the autocorrelation function, the higher the generalization score Q^2. Eventually, it can be seen from Figure 1.15(a) and 1.15(b) that the finite differentiable Matérn autocorrelation function offers an interesting trade-off between inference complexity and prediction quality.



(a) Linear autocorrelation function (Eq. (1.75))

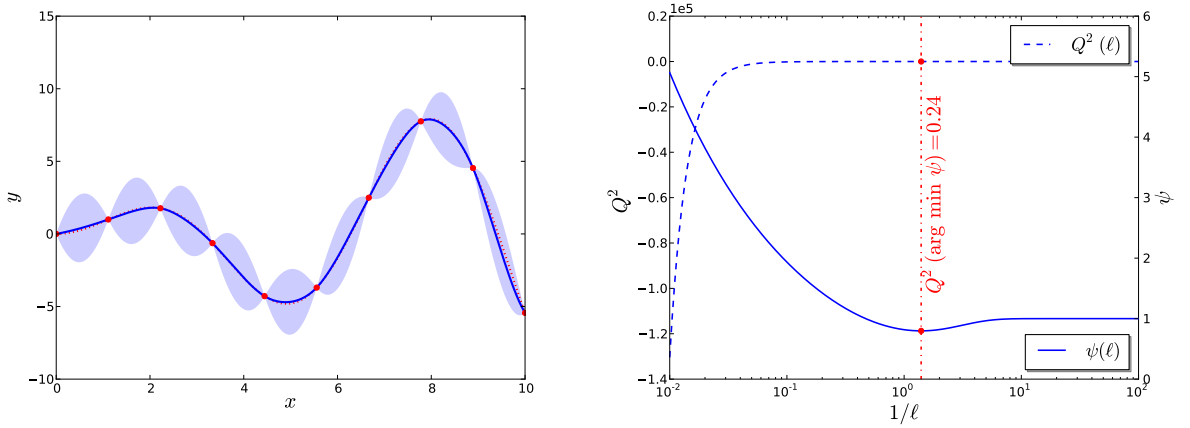


(b) Exponential autocorrelation function (Eq. (1.76))

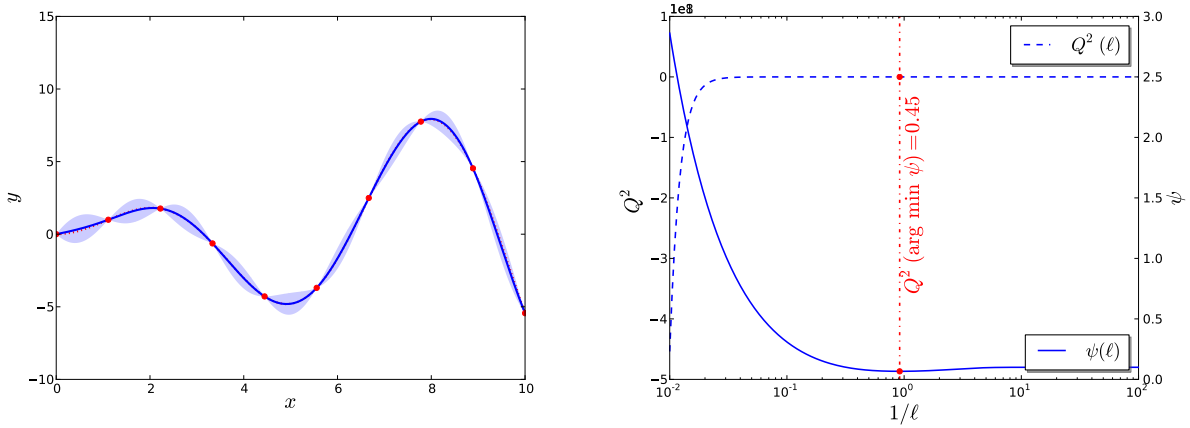


(c) Squared exponential autocorrelation function (Eq. (1.77))

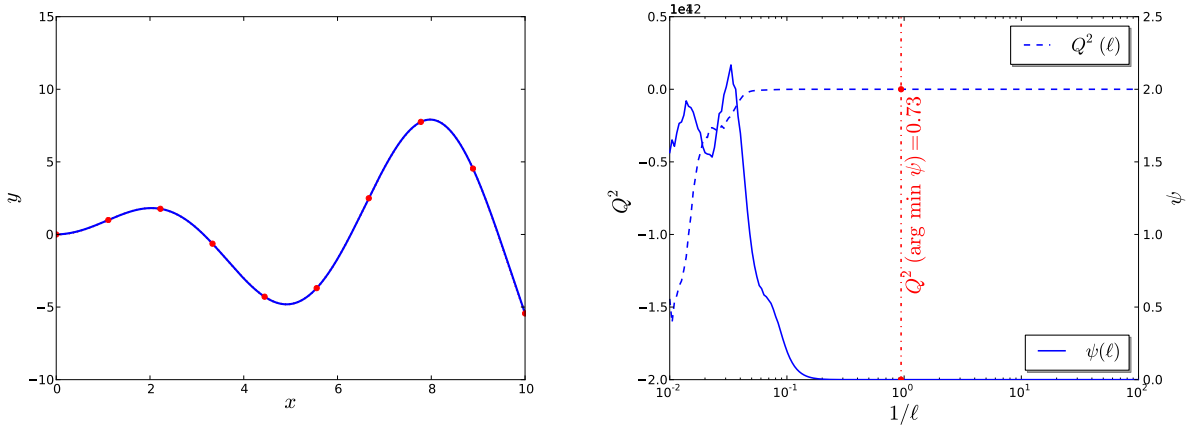
Figure 1.14: Universal Kriging models for $\mathcal{M}(x) = x\ \sin(x)$ and various autocorrelation functions. The left-hand panels represent the observations and the probabilistic prediction as in Figure 1.11. The right-hand panels depict the reduced likelihood function ψ and the generalization score Q^2 as functions of the inverse scale parameter $(1/\ell)$.



(a) Once differentiable Matérn autocorrelation function (Eq. (1.78) with $\nu = 3/2$)



(b) Twice differentiable Matérn autocorrelation function (Eq. (1.78) with $\nu = 5/2$)



(c) Infinitely differentiable autocorrelation function (Eq. (1.78) with $\nu \to \infty$)

Figure 1.15: *(continued from Figure 1.14)* Universal Kriging models for $\mathcal{M}(x) = x\ \sin(x)$ and various autocorrelation functions. The left-hand panels represent the observations and the probabilistic prediction as in Figure 1.11. The right-hand panels depict the reduced likelihood function ψ and the generalization score Q^2 as functions of the inverse scale parameter $(1/\ell)$.

1.5.3 From Gaussian process regression to probabilistic classification

This last illustration is essential for the understanding of the ideas developed throughout this manuscript. Let us consider the following two-dimensional *limit-state* function from Der Kiureghian and Dakessian (1998):

$$\mathfrak{g}(\boldsymbol{x}) = b - x_2 - \kappa\,(x_1 - e)^2 \tag{1.145}$$

where $b = 5$, $\kappa = 0.5$ and $e = 0.1$. In a reliability analysis (or equivalently in a constrained optimization problem), the purpose of such a limit-state function is to split the input space $\mathbb{X}$ into two complementary subsets, namely:

- the *failure domain* canonically defined as:

$$\mathbb{F} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) \leq 0\}, \tag{1.146}$$

- the *safe domain* canonically defined as:

$$\overline{\mathbb{F}} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) > 0\}. \tag{1.147}$$

In order to solve this *classification* problem, it is proposed to use a Gaussian process predictor $\widehat{Y}$ for the model $\mathfrak{g}$. It makes use of a constant regression trend ($\boldsymbol{f} = \{1\}$ and $\boldsymbol{\beta} = \beta_1$) and a squared exponential autocorrelation function. The two scale parameters in the autocorrelation function are arbitrarily set to $\ell_1 = \ell_2 = 2$ rather than being estimated from the data. The reason for this is that the Kriging predictor is already fully consistent when the scale parameters are correctly estimated so that it does not deserve the present illustration as discussed hereafter. Also, the experimental design $\mathscr{X} = \{\boldsymbol{x}^{(i)}, i = 1, \ldots, 8\}$ is deliberately chosen sparse and not uniformly distributed over the square input space $\mathbb{X} = [-8;\, 8]^2$ for the same reason.

In Figure 1.16, the original limit-state surface $\mathscr{S}_0 = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) = 0\}$ is represented by the dashed black line. The red "$-$" and the blue "$+$" represent the initial dataset $\mathscr{D}$ from which the Kriging meta-model is built. The symbol is related to the sign of the real-valued observations $\boldsymbol{y}$ used to fit the Kriging predictor $\widehat{Y}$. The mean prediction of the limit-state surface, which is defined as $\widehat{\mathscr{S}_0} = \{\boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) = 0\}$, is represented by the solid black line.

1.5.3.1 The probabilistic classification function

Using the Gaussianity of the Kriging predictor $\widehat{Y}$, the *probabilistic classification function* is introduced:

$$\begin{aligned}\pi(\boldsymbol{x}) &= \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq t\right] \\ &= \Phi\left(\frac{t - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \quad t \in \mathbb{Y}.\end{aligned} \tag{1.148}$$

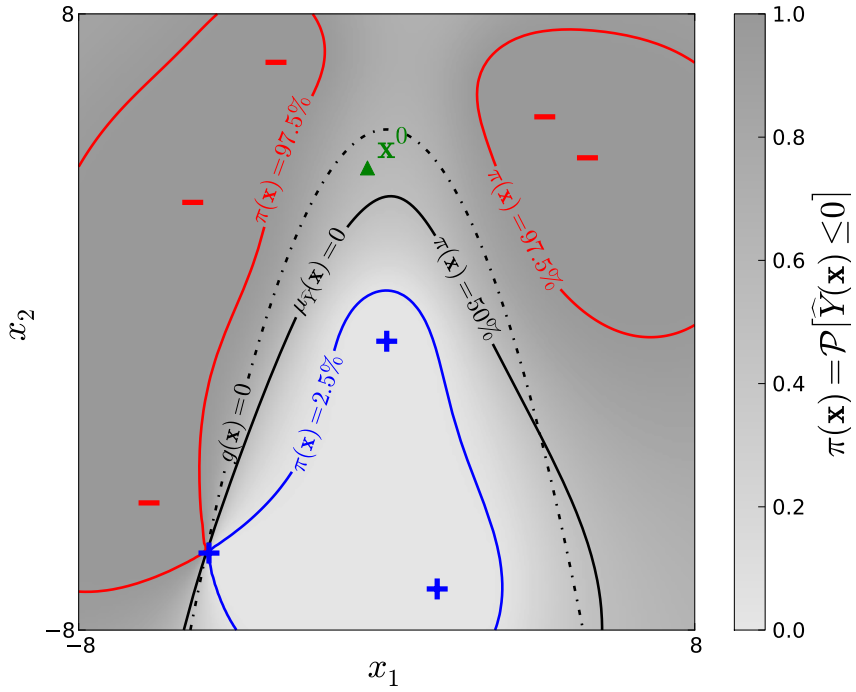


Figure 1.16: Probabilistic classification function on a simple two-dimensional computer experiment from Der Kiureghian and Dakessian (1998). The probabilistic classification function π at $\boldsymbol{x}^{(0)}$ approximately equals 0.6.

where Φ denotes the CDF of the standard normal distribution (see Eq. (1.20)).

In Figure 1.16, the grey shade represents the values of the probabilistic classification function π. The red (resp. the blue) solid line corresponds to the contour of the probabilistic classification function at level $\pi(\boldsymbol{x}) = 0.975$ (resp. $\pi(\boldsymbol{x}) = 0.025$). In other words, these contours bound the 95% confidence region around the mean prediction of the limit-state surface. The probabilistic classification equals one (resp. zero) at the red "−" (resp. at the blue "+") because the Kriging prediction is interpolating ($\mu_{\widehat{Y}}(\boldsymbol{x}^{(i)}) = y_i$) and consistent ($\sigma_{\widehat{Y}}(\boldsymbol{x}^{(i)}) = 0$). Indeed, with slight abuse of notation:

$$\pi(\boldsymbol{x}^{(i)}) = \Phi\left(\frac{t - y_i}{0}\right) = \begin{cases} \Phi(-\infty) = 0 & \text{if } y_i > t \\ \Phi(+\infty) = 1 & \text{if } y_i \leq t \end{cases}, \quad i = 1, \ldots, m, \quad t \in \mathbb{Y}. \tag{1.149}$$

The green triangle $\boldsymbol{x}^0$ in Figure 1.16 is quite interesting because two different decisions can be made depending on the *decision function* that is used. Indeed, the original limit-state function $\mathfrak{g}$ is positive at $\boldsymbol{x}^0$ so that $\boldsymbol{x}^0 \in \overline{\mathbb{F}}$. However, according to the mean prediction $\mu_{\widehat{Y}}(\boldsymbol{x}^0)$, which is predicted to be negative, $\boldsymbol{x}^0$ would belong to $\mathbb{F}$. This problem is referred to as a *classification error*. Note that the previously defined probabilistic classification function, represented in grey shade, allows a smoother decision. Indeed, $\mathfrak{g}(\boldsymbol{x}^0)$ is predicted to be negative with a 60% probability *w.r.t.* the epistemic uncertainty in the random prediction $\widehat{Y}(\boldsymbol{x}^0) \sim \mathcal{N}_1(\mu_{\widehat{Y}}(\boldsymbol{x}^0), \sigma^2_{\widehat{Y}}(\boldsymbol{x}^0))$. This probabilistic information is richer than the simpler binary classification based on the sign of the mean prediction and might be exploited in a reliability analysis.

1.5.3.2 Approximate failure domains and margin of uncertainty

Let us introduce the following three *approximate failure domains*:

$$\widehat{\mathbb{F}}_{1-\alpha}^{\,i} \equiv \left\{ \boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) \leq t + i\, k_{1-\alpha}\, \sigma_{\widehat{Y}}(\boldsymbol{x}) \right\}, \quad i = -1,\ 0, +1, \quad t \in \mathbb{Y}, \tag{1.150}$$

where $1-\alpha$ is the selected confidence-level so that $k_{1-\alpha} = \Phi^{-1}(1-\alpha/2)$. Recall that for the applications in this manuscript $1-\alpha = 95\%$, thus $k_{1-\alpha} = 1.96$. These three subsets are equivalently bounded by the contour of the probabilistic classification function at level 0.975, 0.5 and 0.025 for $\widehat{\mathbb{F}}_{1-\alpha}^{-1}$, $\widehat{\mathbb{F}}_{1-\alpha}^{\,0}$ and $\widehat{\mathbb{F}}_{1-\alpha}^{+1}$, respectively. However, the definition in Eq. (1.150) is easier to evaluate because it does not require the evaluation of the standard normal CDF Φ.

The three approximate failure domains are such that:

- a point $\boldsymbol{x} \in \widehat{\mathbb{F}}_{1-\alpha}^{\,0}$ has fifty percent chances to belong to the failure domain, making $\widehat{\mathbb{F}}_{1-\alpha}^{\,0}$ the most uncertain prediction of the actual failure domain $\mathbb{F}$;
- a point $\boldsymbol{x} \in \widehat{\mathbb{F}}_{1-\alpha}^{-1}$ belongs to the failure domain with high confidence $1-\alpha/2$, making $\widehat{\mathbb{F}}_{1-\alpha}^{-1}$ an *optimistic* prediction of the actual failure domain $\mathbb{F}$;
- a point $\boldsymbol{x} \in \widehat{\mathbb{F}}_{1-\alpha}^{+1}$ belongs to the failure domain with (lower) confidence $\alpha/2$, making $\widehat{\mathbb{F}}_{1-\alpha}^{+1}$ a *pessimistic* (*conservative*) prediction of the actual failure domain $\mathbb{F}$.

In addition, thanks to the positiveness of the Kriging standard deviation $\sigma_{\widehat{Y}}$, these three subsets satisfy the following property:

$$\widehat{\mathbb{F}}_{1-\alpha}^{-1} \subseteq \widehat{\mathbb{F}}_{1-\alpha}^{\,0} \subseteq \widehat{\mathbb{F}}_{1-\alpha}^{+1}. \tag{1.151}$$

These three approximate failure domains $\widehat{\mathbb{F}}_{1-\alpha}^{-1}$ (red), $\widehat{\mathbb{F}}_{1-\alpha}^{\,0}$ (grey) and $\widehat{\mathbb{F}}_{1-\alpha}^{+1}$ (blue) are illustrated in Figure 1.17 for the limit-state function in Eq. (1.145) and $1-\alpha = 95\%$.

The *margin of uncertainty* at confidence level $1-\alpha$ associated with a Kriging approximation of some contour $\mathscr{S}_t, t \in \mathbb{Y}$ is defined as follows:

$$\begin{aligned} \mathbb{M}_{1-\alpha} &\equiv \widehat{\mathbb{F}}_{1-\alpha}^{+1} \setminus \widehat{\mathbb{F}}_{1-\alpha}^{-1} \\ &= \left\{ \boldsymbol{x} \in \mathbb{X} :\ t - k_{1-\alpha}\, \sigma_{\widehat{Y}}(\boldsymbol{x}) \leq \mu_{\widehat{Y}}(\boldsymbol{x}) \leq t + k_{1-\alpha}\, \sigma_{\widehat{Y}}(\boldsymbol{x}) \right\}. \end{aligned} \tag{1.152}$$

It is the *relative complement* of $\widehat{\mathbb{F}}_{1-\alpha}^{-1}$ in $\widehat{\mathbb{F}}_{1-\alpha}^{+1}$.

In other words, $\mathbb{M}_{1-\alpha}$ is the $(1-\alpha)$-confidence region for the Kriging approximation of the contour of interest. The real limit-state surface $\mathscr{S}_t$ has $1-\alpha$ percent chances to be located in this margin *w.r.t.* the Kriging epistemic uncertainty. In addition, the spread of this margin turns out to be a useful indication on the quality of the approximation $\widehat{\mathbb{F}}^{\,0}$ for $\mathbb{F}$. The 95%-confidence margin of uncertainty in Figure 1.17 for the limit-state function in Eq. (1.145) corresponds to the relative complement of the red set in the blue set.

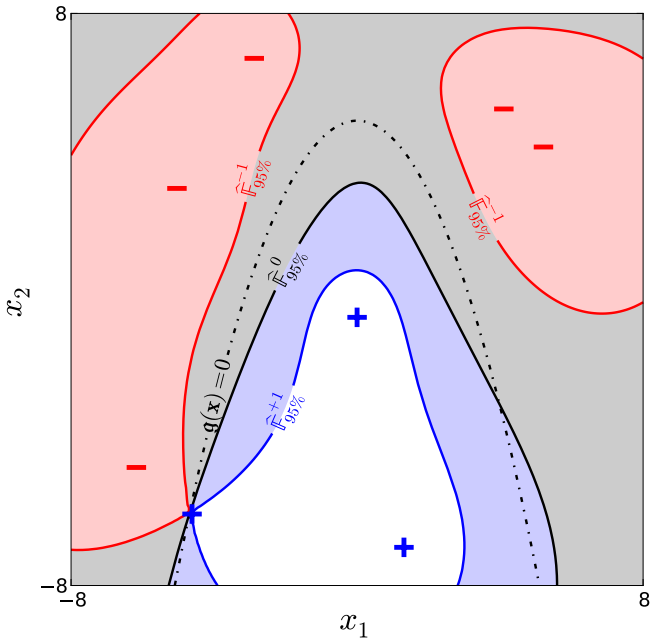


Figure 1.17: Approximate failure domains $\widehat{\mathbb{F}}^{-1}_{1-\alpha}$ (red), $\widehat{\mathbb{F}}^{0}_{1-\alpha}$ (grey) and $\widehat{\mathbb{F}}^{+1}_{1-\alpha}$ (blue) for the example from Section 1.5.3. The corresponding margin of uncertainty is defined as the relative complement of the red set in the blue set: $\mathbb{M}_{95\%} = \widehat{\mathbb{F}}^{+1}_{95\%} \setminus \widehat{\mathbb{F}}^{-1}_{95\%}$.

1.6 Conclusion

This chapter has reviewed some state-of-the-art meta-modelling techniques. Considering that the kernel-based approaches are better suited to the classification problem underlying a reliability analysis than other general linear models, it has been decided to use Gaussian process predictors in order to circumvent the computational burden induced by the resolution of the reliability-based design optimization problem.

Despite its Gaussianity may seem heuristic, the universal Kriging predictor turns out to be an interesting tool to surrogate a computer experiment in the context of reliability analysis because it allows one to genuinely quantify the error induced by this substitution (up to the empirically chosen autocovariance model). This uncertainty is only due to the sparsity of the dataset $\mathscr{D}$ and it is therefore reducible by adding new observations corresponding to additional points in the experimental design $\mathscr{X}$. Such a reducible uncertainty is referred to as an *epistemic* uncertainty as opposed to the *aleatoric* uncertainty modelled by the random distribution of the input $\boldsymbol{X}$ in an uncertainty quantification analysis. For a further discussion about these two types of uncertainty and how they might be involved in uncertainty quantification, the interested reader is referred to the article by Der Kiureghian and Ditlevsen (2009).

2

Adaptive designs of experiments

Contents

2.1 Introduction

In the previous chapter, supervised learning was presented as the discipline of statistical learning that aims at building an emulator for a given model $\mathscr{M}$ from a dataset $\mathscr{D} = \left\{(\boldsymbol{x}^{(i)}, y_i), i = 1, \dots, m\right\}$. This chapter is concerned with the problem of selecting the best *design of experiments* (DOE) $\mathscr{X} = \left\{\boldsymbol{x}^{(i)}, i = 1, \dots, m\right\}$ in order to ensure the accuracy of the emulator over the whole input space $\mathbb{X}$ or some part of it.

Such a DOE should be parsimonious (*i.e.* it should contain the least number m of experiments) for two reasons. First, the emulator is built to surrogate an expensive-to-evaluate computer model $\mathscr{M}$. The cost to build this emulator is expected to be far smaller than the one induced by using the original computer model $\mathscr{M}$ to perform the analysis of interest (*e.g.* global optimization, reliability analysis, global sensitivity analysis, or reliability-based design optimization). Second, most emulators (including Gaussian processes) lose computational efficiency when the size m of the dataset tends to be large, and becomes intractable above a certain level (say $m > 10^4$ for the present applications of Gaussian processes). Note that this phenomenon gets even worse for anisotropic Gaussian process meta-models when the dimension n of the input space $\mathbb{X}$ tends to be large.

This chapter contains three main sections. Section 2.2 reviews state-of-the-art methods to choose a finite set of experiments in the continuous input space $\mathbb{X}$ from a fresh start where no observations are available yet. It also proposes a heuristic shape-filling procedure that was used in this manuscript. Sections 2.3 and 2.4 are concerned with the *adaptive refinement* of an existing DOE for two different purposes, namely: *global optimization* or *contour approximation*. Section 2.3 reviews state-of-the-art strategies which are based on the resolution of a *global optimization problem* while Section 2.4 explores a *sampling alternative*.

2.2 Initial designs of experiments

In this section various state-of-the art techniques to select the experiments are reviewed. These techniques may then be applied to construct an initial DOE before switching to an adaptive refinement strategy as discussed in Section 2.3 if one can afford additional runs of the computational model $\mathscr{M}$.

2.2.1 Space-filling designs of experiments

The so-called *space-filling DOEs* attempt to fill the input space $\mathbb{X}$ with a finite number m of computer experiments in order to capture the largest amount of information to emulate the model $\mathscr{M}$. The only information that is required in order to use these techniques is the shape of the input space $\mathbb{X}$ (when it is bounded) and the size m of the sought experimental design.

2.2.1.1 State-of-the-art techniques

State-of-the art space-filling experimental designs attempt to fill the n-dimensional unit hypercube $[0;\ 1]^n$ according to some metric.

Deterministic designs

Full factorial designs are certainly the most popular DOEs for *in situ* experiments (see *e.g.* Montgomery, 2004). Indeed a full factorial design is simply defined as the nodes of an L-level grid ($L > 1$) defined over the n margin input variables (or *factors*) composing the input $\boldsymbol{x} = (x_1, \ldots, x_n)^\mathsf{T}$. Formally, a full factorial design is obtained as the following tensor product:

$$\mathscr{X}_{\text{FF}} = \left\{ \frac{i-1}{L-1},\ i = 1, \ldots, L \right\}^n . \tag{2.1}$$

Practical use of full factorial designs does not involve more than 2 or 3 qualitative levels which makes them applicable. However they do not suit to the present purpose because their size ($m = L^n$) drastically increases with both the dimension n and the number of levels L which should be greater than 2 in the directions where the model $\mathscr{M}$ is highly nonlinear.

Fractional factorial designs (see *e.g.* Montgomery, 2004) were designed to make sparser designs as subset of full factorial designs. They are based on the *sparsity-of-effect principle* according to which the model $\mathscr{M}$ is primarily driven by margin effects (functions of only one factor x_i, $i \in [\![1;\ n]\!]$) and low-order interactions (functions of $k < n$ factors only). As a consequence, the experiments in a full factorial design involving the variation of more than k variables at a time can be omitted thus resulting in a fractional factorial design. However such designs are somewhat *model-specific* because they require some prior knowledge about the interactions that can be omitted.

Random designs

Uniform Monte Carlo sampling is another technique that is rather easy to implement. Indeed a uniform Monte Carlo sample is simply defined as:

$$\mathscr{X}_{\text{MCS}} = \left\{ x_j^{(i)},\ i = 1, \ldots, m,\ j = 1, \ldots, n \right\}, \tag{2.2}$$

where each $x_j^{(i)}$ is a realization of a uniform random number in $[0;\ 1]$. However a high number of experiments m is required to guarantee that any given sample truly fills the hypercube and this does not suit the parsimony constraint.

Latin hypercube sampling (LHS, McKay et al., 1979) is an alternative to uniform Monte Carlo sampling that ensures uniformity of the sample on the margin input variables $\left\{ x_j,\ j = 1, \ldots, n \right\}$. In addition to that interesting property, it has a rather simple definition that makes it a very popular technique to build space-filling design of experiments.

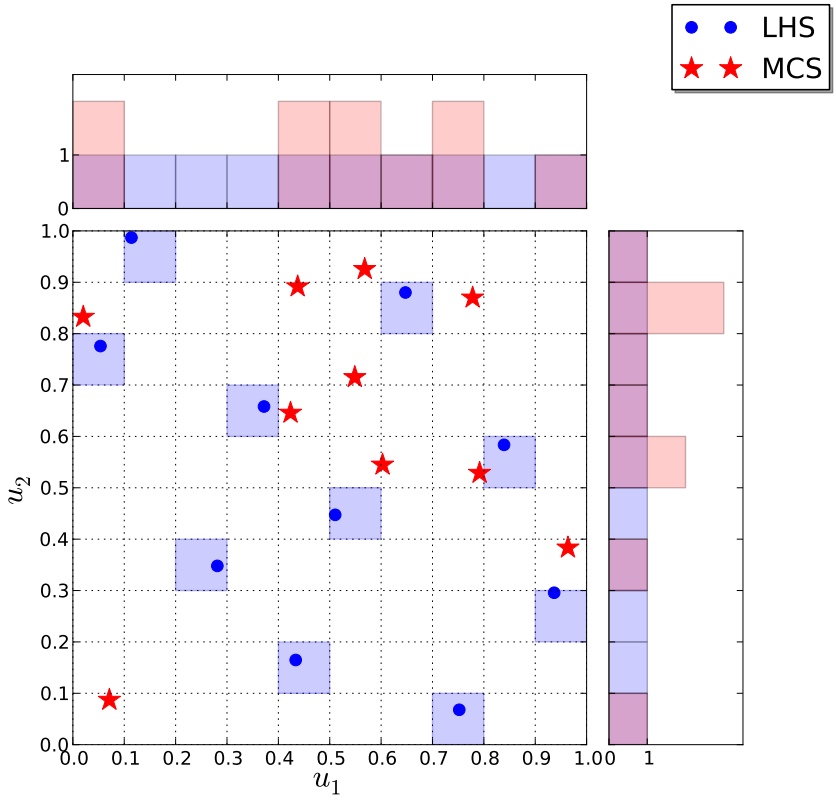


Figure 2.1: A two-dimensional Latin hypercube sample (LHS) with $m = 10$ elements compared to a Monte Carlo sample (MCS) with identical size. The Monte Carlo sample does not feature uniform margins whereas the Latin hypercube sample does.

A Latin hypercube sample of size m over the n-dimensional unit hypercube $[0;\ 1]^n$ is defined as follows:

$$\mathscr{X}_{\text{LHS}} = \left\{ x_j^{(i)} = \frac{\pi_j^{(i)} - 1 + u_j^{(i)}}{m}, \quad i = 1, \dots, m, j = 1, \dots, n \right\}, \tag{2.3}$$

where:

$\pi_j^{(i)}$ is the i-th element of the j-th random permutation of the sequence $\{1, \dots, m\}$,

$u_j^{(i)}$ is a realization of a random variable U that is uniformly distributed over $[0;\ 1]$.

Note that the j-th random permutation of the sequence $\{1, \dots, m\}$ can simply be obtained as the list of integers that sorts an m-sample $\{e_i,\ i = 1, \dots, m\}$ of a uniformly distributed random variable E over $[0;\ 1]$ in its ascending order:

$$\pi_j = \arg\operatorname{sort} \{e_i,\ i = 1, \dots, m\}. \tag{2.4}$$

In order to understand this definition, consider that each axis is divided into m subintervals of equal width $1/m$. By making a tensor product of the n meshed axes, one obtains

the grid illustrated in Figure 2.1 (for $n = 2$ and $m = 10$). Note that the nodes (edges) of this grid constitute an m-level full factorial design. Then, m cells are picked at random in the grid so as to ensure uniformity on the axes. This is achieved by computing the n permutations of the sequence $\{1, \dots, m\}$. Indeed, by doing so each margin cell is only picked once thus ensuring uniformity on the axes as illustrated in Figure 2.1. Eventually, a vector is picked randomly in each selected cell by adding a uniform random number in $[0;\ 1/m]$ to each lower bound of each selected cell.

Even if Latin hypercube sampling ensures uniformity on the axes, it does not ensure the uniformity in the higher-dimensional projections and this may lead to inaccurate results when emulating models involving strong interactions. Post-processing techniques of Latin hypercube samples have been proposed in an attempt to solve this problem, one of which consists in maximizing the minimum Euclidean distance (in $\mathbb{R}^n$) between each couple of points (*maximin LHS*, see *e.g.* Stein, 1987; Johnson et al., 1990).

Quasi-random designs

Niederreiter (1992) introduces the concept of *discrepancy* which is defined as the deviation of a given sequence $\mathscr{X}$ from the uniform distribution in the unit hypercube $[0;\ 1]^n$. The discrepancy is defined as follows:

$$D_m(\mathscr{X}) = \sup_{\mathscr{J} \in \mathfrak{J}} \Delta(\mathscr{J}, \mathscr{X}), \tag{2.5}$$

where $\mathscr{J}$ is one subset amongst all the rectangular subsets $\mathfrak{J}$ of $[0;\ 1]^n$, and:

$$\Delta(\mathscr{J}, \mathscr{X}) = \frac{\text{Card}\left(\{\boldsymbol{x} \in \mathscr{X} \cap \mathscr{J}\}\right)}{m} - \mathscr{V}(\mathscr{J}) \tag{2.6}$$

where $\text{Card}(\bullet)$ denotes the number of elements in $\bullet$, and $\mathscr{V}(\bullet)$ denotes its volume. The definition varies depending on how $\mathscr{J}$ is defined (see Franco, 2008, for a deeper review).

In order to provide a deeper insight on this abstract definition, Franco (2008) points out that in one dimension, the discrepancy is nothing but the statistic of a Kolmogorov goodness-of-fit test applied to the uniform distribution over $[0;\ 1]$:

$$D_m(\mathscr{X}) = \sup_{x \in [0;\ 1]} \left| \widehat{F}_m(x) - x \right|, \tag{2.7}$$

where $\widehat{F}_m(x)$ is the empirical cumulative distribution of the sequence $\mathscr{X}$ defined as follows:

$$\widehat{F}_m(x) = \frac{1}{m} \sum_{i=1}^{m} \mathbb{1}(x^{(i)} \leq x). \tag{2.8}$$

Its explicit computation in dimensions higher than 2 becomes intractable so that one resorts to computationally expensive approximations (see *e.g.* Franco, 2008). As a reference however, one may remember the following results.

- A sequence of points in the hypercube is uniformly distributed if its discrepancy $D(m)$ tends to zero as m tends to infinity.
- The discrepancy of a uniform Monte Carlo sample in the hypercube $[0;\ 1]^n$ of size m equals $1/\sqrt{m}$.

These results motivated the use of the so-called *low-discrepancy sequences* (Niederreiter, 1992) to build experimental designs because they feature lower discrepancies than a uniform Monte Carlo sample and thus ensure a better coverage of the unit hypercube. It can also be proven that despite their conceptual simplicity, full factorial designs yield a higher discrepancy ($1/m$) than the uniform random sequence (Franco, 2008) thus making uniform Monte Carlo sampling more efficient with respect to that metric. In addition to their nice space-filling property low-discrepancy sequences enables to augment the size of the experimental design to refine the emulator without altering the discrepancy of the augmented sequence (see *e.g.* Blatman et al., 2007, for such an application of Sobol' sequences).

The so-called *Halton*, *Faure* and *Sobol'* low-discrepancy sequences are indeed specific n-dimensional generalizations of a reference one-dimensional sequence known as the *van der Corput* (VDC) sequence. The latter sequence is defined as follows:

$$\mathscr{X}_{\text{VDC}} = \left\{ x^{(i)} = \sum_{k=0}^{d} \frac{b_k}{2^{k+1}}, \quad i = 1, \dots, m \right\} \tag{2.9}$$

where b_k is the k-th digit in the base 2 (or binary) representation of the integer i (*e.g.* if $i = (12)_{\text{base }10}$, then $i = (1\times 2^3 + 1\times 2^2 + 0\times 2^1 + 0\times 2^0)_{\text{base }10} = (1100)_{\text{base }2}$ meaning that $d = 3$, $b_3 = 1$, $b_2 = 1$, $b_1 = 0$ and $b_0 = 0$). Halton's sequence is obtained by incrementing the base for each dimension (starting from 2). The way the other two sequences are built is beyond the scope of this thesis and the interested reader is referred to the reference book by Niederreiter (1992) or the Ph.D. thesis of Franco (2008). These two references also present a larger set of space-filling DOE classes including *distance*-based (*e.g. maximin* and *minimax* designs) and *entropy*-based designs.

2.2.1.2 A heuristic hypersphere-filling design of experiments

State-of-the-art space-filling experimental designs are preliminary designed to fill the unit hypercube. Their use may possibly be extended to the simulation of any random vector with independent margins by means of the usual inverse transform technique:

$$\mathscr{X} = \left\{ x_j^{(i)} = F_{X_j}^{-1}(u_j^{(i)}), \quad i = 1, \dots, m,\ j = 1, \dots, n \right\}, \tag{2.10}$$

where $F_{X_j}^{-1}$ denotes the inverse cumulative density function for the j-th random margin, and $\{\boldsymbol{u}^{(i)},\ i = 1, \dots, m\}$ is a DOE that fills the unit hypercube $[0;\ 1]^n$. However, to the author's knowledge, the literature provides few techniques to sample in an hypersphere (Nie and Ellingwood, 2000, 2004).

Nevertheless sampling in an hypersphere might be relevant when the importance of the behaviour of the original simulator $\mathscr{M}$ vanishes with the distance to the origin. This section reviews an original technique that was initially proposed by Deheeger (2008); Bourinet et al. (2011) and that was used to construct the Kriging predictors in this thesis.

First the inverse transform method formulated in Eq. (2.10) is not applicable there because one does not know the marginal distributions for a random vector that is uniformly distributed in the hypersphere. One first intuitive approach would thus consist in sampling in the smallest hypercube containing the hypersphere using any of the previously introduced state-of-the-art techniques and reject the points that are not inside the hypersphere. In practice though, this approach becomes intractable for dimensions higher than say $n = 5$ because the rejection rate dramatically increases with the dimension. The reason for this is that the ratio between the volume of an hypersphere and the volume of its circumscribed hypercube is shrinking exponentially with the dimension. Instead, it is proposed to use a direct Monte Carlo method to generate random vectors in the hypersphere without any rejection. This direct sampling technique is based on the following theorem (see *e.g.* Rubinstein and Kroese, 2008, p. 69).

Theorem 2.2.1. ***Uniform random sampling within the unit hypersphere***
Let $\boldsymbol{Y}$ be a multivariate normal random vector with n independent margins, zero mean and unit variance, let $\|\boldsymbol{Y}\|_2$ denote its $\mathscr{L}_2$ norm in $\mathbb{R}^n$ and U be another random variable that is uniformly distributed over $[0;\ 1]$, then the random vector:

$$\boldsymbol{X} \equiv U^{1/n} \frac{\boldsymbol{Y}}{\|\boldsymbol{Y}\|_2} \tag{2.11}$$

is uniformly distributed within the unit-hypersphere.

The proof of this theorem is two-fold.

(i) First consider the random vector $\boldsymbol{Y}/\|\boldsymbol{Y}\|_2$. The random vector $\boldsymbol{Y}$ is uniformly distributed in direction because the multivariate normal density is spherically symmetrical so that it does not favor any direction. Uniformity on the surface of the hypersphere is guaranteed by the bijective nature of the projection $\boldsymbol{Y}/\|\boldsymbol{Y}\|_2$. For a more rigorous elaboration of these intuitions see *e.g.* Rubinstein (1982).

(ii) Then one can prove (see *e.g.* Cumbus et al., 1996) that the distribution of the norm R of a random vector that is uniformly distributed over the unit hypersphere satisfies $\mathbb{P}[R \leq r] = r^n$, $r \in [0;\ 1]$. So, using the usual inverse transform technique, the random radius R can be efficiently simulated as $U^{1/n}$ with U uniformly distributed over $[0;\ 1]$.

A uniform random sample in the unit 2-dimensional hypersphere (*i.e.* the unit disk) is illustrated in Figure 2.2 (blue dots). It can be seen that the required uniformity of the sample within the disk is fulfilled although a large sample size is required to guarantee that it is truly space-filling (here $m = 1,000$).

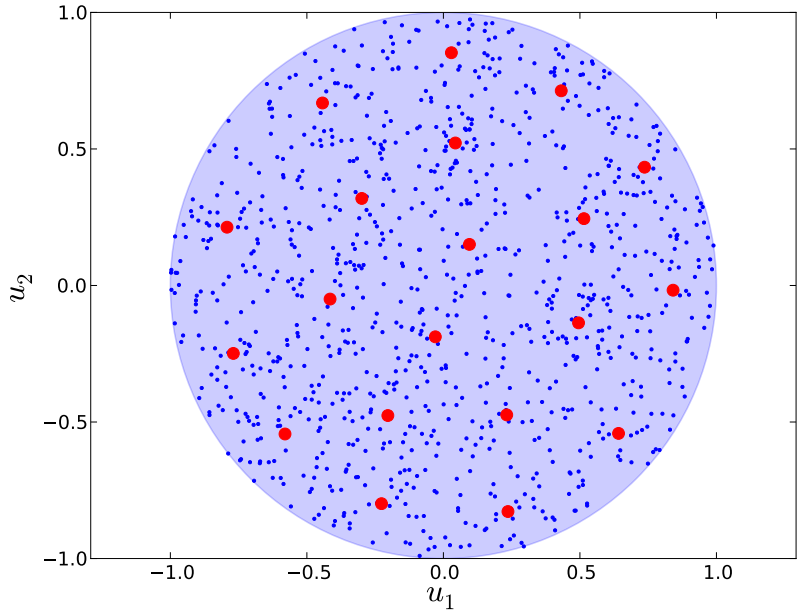


Figure 2.2: A heuristic hypersphere-filling experimental design. The blue dots constitute a uniform sample within the unit disk while the red points form the proposed hypersphere-filling experimental design with size $m = 20$ obtained by K-means clustering.

However, the problem is that one cannot afford such a large sample size in the present context. It is thus proposed to reduce this large population to its K cluster centres using K-means clustering where $K = m$ is the size of the experimental design. The basic idea underlying K-means clustering is presented in Section 2.2.1.3. An example of a hypersphere-filling experimental design obtained by this heuristic approach is illustrated in Figure 2.2 (red points) for the unit disk. The $m = 20$ cluster centres of the larger candidate population (blue points) uniformly span the hypersphere as desired.

2.2.1.3 K-means clustering

The clustering problem is an *unsupervised classification* problem. Considering a sample $\mathfrak{C}$ with size N as well as a given number of classes $K > 1$, the clustering problem consists in allocating a class (a label) amongst $\{k, k = 1, \ldots, K\}$ to the individuals in $\mathfrak{C}$ based on some metric to be defined. Here the samples in $\mathfrak{C}$ are points of the Euclidean space $\mathbb{R}^n$ so that the Euclidean distance seems to be an appropriate metric, but the clustering problem can also be considered for a collection of functions (see *e.g.* Auder, 2011).

The *K-means clustering* problem (Steinhaus, 1956; MacQueen, 1967; Lloyd, 1982) makes use of the usual $\mathscr{L}_2$ distance in $\mathbb{R}^n$ (denoted by $\| \bullet \|_2$) in order to assign each point in $\mathfrak{C}$ to their closest *cluster centre* (or *centroid*). More specifically, it relies on the so-called *inertia* which is defined as follows:

$$I_K(\mathscr{O}, \mathfrak{C}) \equiv \sum_{i=1}^{N} \min_{k=1,\ldots,K} \left\| \boldsymbol{o}^{(k)} - \boldsymbol{c}^{(i)} \right\|_2^2, \tag{2.12}$$

where $\mathscr{O} = \{\boldsymbol{o}^{(k)} \in \mathbb{R}^n, k = 1, \dots, K\}$ is the set of centres of each class. The best set of centres $\mathscr{O}^*$ is then defined as the following minimizer:

$$\mathscr{O}^* = \arg \min_{\mathscr{O} \in \mathfrak{O}(\mathbb{R}^n, K)} I_K(\mathscr{O}, \mathfrak{C}), \tag{2.13}$$

where $\mathfrak{O}(\mathbb{R}^n, K)$ denotes all the subsets of $\mathbb{R}^n$ with size K.

The solution of the optimization problem in Eq. (2.13) is not analytical so that practical implementations of the K-means clustering resorts to dedicated iterative procedures. For instance the so-called *expectation-maximization* (EM) algorithm starts from a set of *initial centroids* $\mathscr{O}^{[0]}$ and proceeds iteratively on $[j]$ in two steps:

(E) Compute the inertia $I_K(\mathscr{O}^{[j]}, \mathfrak{C})$ for the current set of centroids $\mathscr{O}^{[j]}$. This computation requires the allocation of a label to each point in the sample $\mathfrak{C}$ and is defined as follows:

$$l^{(i)[j]} = \arg \min_{k=1,\dots,K} \left\| \boldsymbol{o}^{(k)[j]} - \boldsymbol{c}^{(i)} \right\|_2^2, \quad i = 1, \dots, N. \tag{2.14}$$

(M) Define the next set of centroids $\mathscr{O}^{[j+1]}$ as the barycentre of each subset obtained from the expectation step. These K subsets are defined as follows:

$$\mathfrak{C}^{(k)[j]} = \left\{ \boldsymbol{c}^{(t)} \in \mathfrak{C} : l^{(t)[j]} = k \right\}, \quad k = 1, \dots, K, \tag{2.15}$$

yielding the following expression for the k-th barycentre:

$$\boldsymbol{o}^{(k)[j+1]} = \frac{1}{\mathrm{Card}\left(\mathfrak{C}^{(k)[j]}\right)} \sum_{\boldsymbol{c}^{(t)} \in \mathfrak{C}^{(k)[j]}} \boldsymbol{c}^{(t)}, \tag{2.16}$$

where $\mathrm{Card}(\bullet)$ denotes the number of elements in $\bullet$.

The algorithm stops when the latest improvement brought by the maximization step in terms of *inertia* is below a given threshold ϵ_0. For instance it stops at iteration i if:

$$0 < \frac{\left| I_K\left(\mathscr{O}^{(i-1)}, \mathfrak{C}\right) - I_K\left(\mathscr{O}^{(i)}, \mathfrak{C}\right) \right|}{I_K\left(\mathscr{O}^{(i-1)}, \mathfrak{C}\right)} \leq \epsilon_0. \tag{2.17}$$

However it is worth mentioning that the K-means problem does not have a unique solution, so that the EM algorithm should be restarted from different initial centroids. For the sake of numerical efficiency though, it rather resorts to heuristic initialization procedures like the *K-means++* procedure proposed by Arthur and Vassilvitskii (2007) which yields satisfying results in practice (at least for the present purpose of reducing $\mathfrak{C}$ to a smaller population with similar statistical properties).

2.2.2 Model-specific designs of experiments

Model-specific DOEs differ from the previously introduced space-filling experimental designs in the sense that one already knows what kind of emulator will be used and may incorporate this additional information to construct the best-suited experimental design.

2.2.2.1 Experimental designs for general linear models

In Chapter 1, it has been shown that general linear models are solely determined from the choice of the functional basis $\boldsymbol{f} = \{f_i, i = 1, \ldots, p\}$, $p < m$. The optimal designs reviewed in the sequel are constructed from this additional information by deriving specific criteria ensuring various properties. All these criteria are implementable from a very fresh start as they do not depend on the observations $\boldsymbol{y}$ (unavailable yet).

A-optimal designs attempt to minimize the averaged variance of the posterior distribution of the regression weights $\widehat{\boldsymbol{\beta}}$. Recalling that $\mathbf{F}$ denotes the regression matrix (see Eq. (1.33)), *A*-optimality is achieved by minimizing the sum of the diagonal elements of the regression weights' covariance matrix (*i.e.* its trace) defined as $\sigma^2(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F})^{-1}$ (see Eq. (1.53)). Since the variance σ^2 depends on the observations and is thus not available yet, the minimization is performed on $(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F})^{-1}$ only, leading to the following definition:

$$\mathcal{X}_\mathrm{A} = \arg\min_{\mathcal{X}\in\mathbb{X}} \quad \mathrm{trace}\left(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}. \tag{2.18}$$

In a similar fashion *D-optimal designs* attempt to minimize the determinant of the covariance matrix which is another invariant summary statistic for square matrices. To avoid the inversion though one maximizes the determinant of the *information matrix* $\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F}$, neglecting the unavailable multiplicative variance σ^2 and thus leading to the following definition:

$$\mathcal{X}_\mathrm{D} = \arg\max_{\mathcal{X}\in\mathbb{X}} \quad \det\left(\mathbf{F}^\mathsf{T}\mathbf{R}^{-1}\mathbf{F}\right). \tag{2.19}$$

Eventually *G-optimal designs* try to minimize the maximum variance of the predicted value so as to augment the leave-one-out estimate of the coefficient of determination previously defined in Section 1.4.5.3 for the Kriging predictor and denoted by Q^2. The interested reader may refer to the thesis by Micol (2007) who uses the *G*-criterion to build and enrich a DOE to perform a reliability analysis with a polynomial model. Note however that all these optimal designs require the resolution of non-trivial and high dimensional optimization problems (find the $m \times n$ scalar values composing $\mathcal{X}$).

2.2.2.2 Experimental designs for Gaussian process predictors

Inspired by the optimal designs for general linear models, Sacks et al. (1989a) proposed to extend their use to Gaussian process predictors by exploiting their interrelation with generalized least-squares models. This extension is essentially based on the fact that the prediction variance of the best linear unbiased predictor previously denoted by $\sigma^2_{\widehat{Y}}$ does not depend on the observations $\boldsymbol{y}$ so that it can be computed *a priori* alike the other *A*- *D*- or *G*-criteria.

The *integrated mean-squared error* (IMSE) criterion is defined as the following weighted integral of the mean-squared prediction error associated to the best linear unbiased predictor:

$$\mathrm{IMSE}(\mathcal{X}) = \int_{\mathbb{X}} \sigma^2_{\widehat{Y}}(\boldsymbol{x})\, w(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \tag{2.20}$$

where w is a probability density function, meaning it has unit integral over $\mathbb{X}$. Note that Sacks et al. (1989a,b) provide simplified expression for this integral when both the regression and autocorrelation model are tensor products of one-dimensional models, and when w is the uniform density over some bounded subset of $\mathbb{X}$. I-optimal designs for Gaussian process predictors aim at minimizing this new criterion and are consequently defined as follows:

$$\mathscr{X}_{\mathrm{I}} = \arg\min_{\mathscr{X}\in\mathbb{X}} \quad \mathrm{IMSE}(\mathscr{X}). \tag{2.21}$$

The *maximum mean-squared error* (MMSE) criterion is defined as the maximum value of the mean-squared prediction error over the input space:

$$\mathrm{MMSE}(\mathscr{X}) = \arg\max_{\boldsymbol{x}\in\mathbb{X}} \quad \sigma^2_{\widehat{Y}}(\boldsymbol{x}). \tag{2.22}$$

The latter should also be minimized, thus leading to the following definition for M-optimal designs:

$$\mathscr{X}_{\mathrm{M}} = \arg\min_{\mathscr{X}\in\mathbb{X}} \quad \mathrm{MMSE}(\mathscr{X}). \tag{2.23}$$

However a practical implementation of these criteria raises a fundamental problem to construct DOEs from scratch. The reason for this is that best linear unbiased predictors cannot be used in practice due to the unknown autocorrelation function $R(\bullet, \boldsymbol{\theta})$. Empirical best linear unbiased predictors are used instead, as introduced in Section 1.4.5. The parameters $\boldsymbol{\theta}$ must be estimated from the unavailable observations $\boldsymbol{y}$ thus making I- and M-optimal designs non-implementable from an empty dataset $\mathscr{D}$. However Sacks et al. (1989a,b), among others, propose a heuristic approach to circumvent this problem.

2.3 Adaptive designs of experiments for Gaussian process predictors

In this section it is now assumed that the initial experimental design $\mathscr{X} = \{\boldsymbol{x}^{(i)}, i = 1, \ldots, m\}$ was built and that a first Gaussian process predictor $\widehat{Y}(\boldsymbol{x}) \sim \mathscr{N}_1(\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma^2_{\widehat{Y}}(\boldsymbol{x}))$ is available. It is also assumed that one can afford new evaluations of the original model $\mathscr{M}$ in order to enrich the prediction. The refinement of the prediction may now rely on the mean prediction $\mu_{\widehat{Y}}$ based on the available observations $\boldsymbol{y}$. This additional information is of utmost importance in global optimization or for the present problem of refining an initial contour approximation.

Alike model-specific DOEs, state-of-the-art *adaptive* DOEs for Gaussian process predictors all rely on the definition of a criterion that depends on the input $\boldsymbol{x}$ and which is maximum or minimum in a region of interest: the neighbourhood of the global optimum of the emulated model $\mathscr{M}$ or of a specific contour of interest. Given such a criterion, they attempt to find *the best next point* $\boldsymbol{x}^{(m+1)}$ that will significantly improve the prediction by solving a global optimization subproblem on the *inexpensive* predictor. The required accuracy about the quantity of interest is achieved by involving this global optimization subproblem into an iterative scheme.

2.3.1 Historical premise: efficient global optimization

Adaptive designs of experiments for Gaussian process predictors first became popular in the 90's due to the work by Mockus (1994) and Jones et al. (1998). Their purpose was to solve a *global optimization problem* involving an expensive-to-evaluate objective function $\mathscr{M}$. The global optimization problem takes the following canonical form:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathbb{X}} \quad \mathscr{M}(\boldsymbol{x}). \tag{2.24}$$

While most conventional optimization algorithms use the local information provided by the partial derivatives of the function $\mathscr{M}$ with respect to $\boldsymbol{x}$, the so-called *efficient global optimization* (EGO) algorithm exploits the probabilistic information provided by a Gaussian process predictor built from a more global experimental design $\mathscr{X}$ to select the next evaluation point $\boldsymbol{x}^{(m+1)}$.

2.3.1.1 The expected improvement

Considering the initial dataset $\mathscr{D} = \{(\boldsymbol{x}^{(i)}, y_i), i = 1, \dots, m\}$ and an initial Gaussian process predictor $\widehat{Y}$, Jones et al. (1998) defines the so-called *improvement* as follows:

$$\mathrm{I}(\boldsymbol{x}) \equiv \begin{cases} y_{\min} - \widehat{Y}(\boldsymbol{x}) & \text{if } \widehat{Y}(\boldsymbol{x}) \leq y_{\min} \\ 0 & \text{otherwise} \end{cases} = \max\left\{y_{\min} - \widehat{Y}(\boldsymbol{x}),\, 0\right\}, \tag{2.25}$$

where $y_{\min} = \min\{y_i, i = 1, \dots, m\}$ is the minimum value of $\mathscr{M}$ in the available dataset $\mathscr{D}$. As it depends on the Gaussian prediction $\widehat{Y}(\boldsymbol{x})$, the improvement at $\boldsymbol{x}$ is also a Gaussian variable truncated to positive values only (as a negative improvement is a decline).

So as to summarize this random variable, Jones et al. (1998) propose to use its expectation thus leading to the so-called *expected improvement*:

$$\mathrm{EI}(\boldsymbol{x}) \equiv \mathbb{E}\left[\mathrm{I}(\boldsymbol{x})\right] = \int_{-\infty}^{y_{\min}} (y_{\min} - \widehat{y})\, \varphi\left(\frac{\widehat{y} - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) \mathrm{d}\widehat{y} \tag{2.26}$$

which turns out after integrating by parts (see *e.g.* Bichon, 2010, p. 143–145) to the following closed-form expression:

$$\mathrm{EI}(\boldsymbol{x}) = \left(y_{\min} - \mu_{\widehat{Y}}(\boldsymbol{x})\right) \Phi\left(\frac{y_{\min} - \mu_{\widehat{Y}(\boldsymbol{x})}}{\sigma_{\widehat{Y}(\boldsymbol{x})}}\right) + \sigma_{\widehat{Y}}(\boldsymbol{x})\, \varphi\left(\frac{y_{\min} - \mu_{\widehat{Y}(\boldsymbol{x})}}{\sigma_{\widehat{Y}(\boldsymbol{x})}}\right), \tag{2.27}$$

where φ and Φ denotes the standard Gaussian PDF and CDF respectively. Note that this definition still holds for the points in the DOE by taking the limit value and leads to a zero expected improvement as intuited.

The next point that should be added to the DOE is defined as the maximizer of the expected improvement, and leads to the elaboration of the following iteration rule:

$$\boldsymbol{x}^{(m+1)} = \arg\max_{\boldsymbol{x}\in\mathbb{X}} \quad \mathrm{EI}(\boldsymbol{x}). \tag{2.28}$$

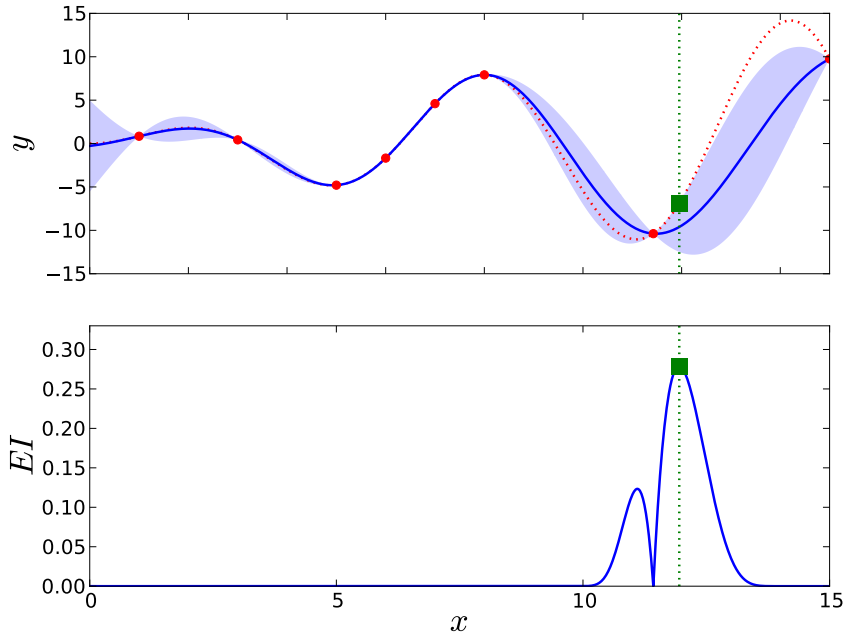


Figure 2.3: Implementation of the efficient global optimization technique for the minimization of $\mathscr{M}(x) = x\sin(x)$ on the interval $\mathbb{X} = [0; 15]$. The upper panel illustrates the original function (dotted line) together with its Gaussian process approximation (continuous line) built from the available dataset $\mathscr{D}$ (red points). The lower panel illustrates the expected improvement and its maximizer which corresponds to the next evaluation required by EGO.

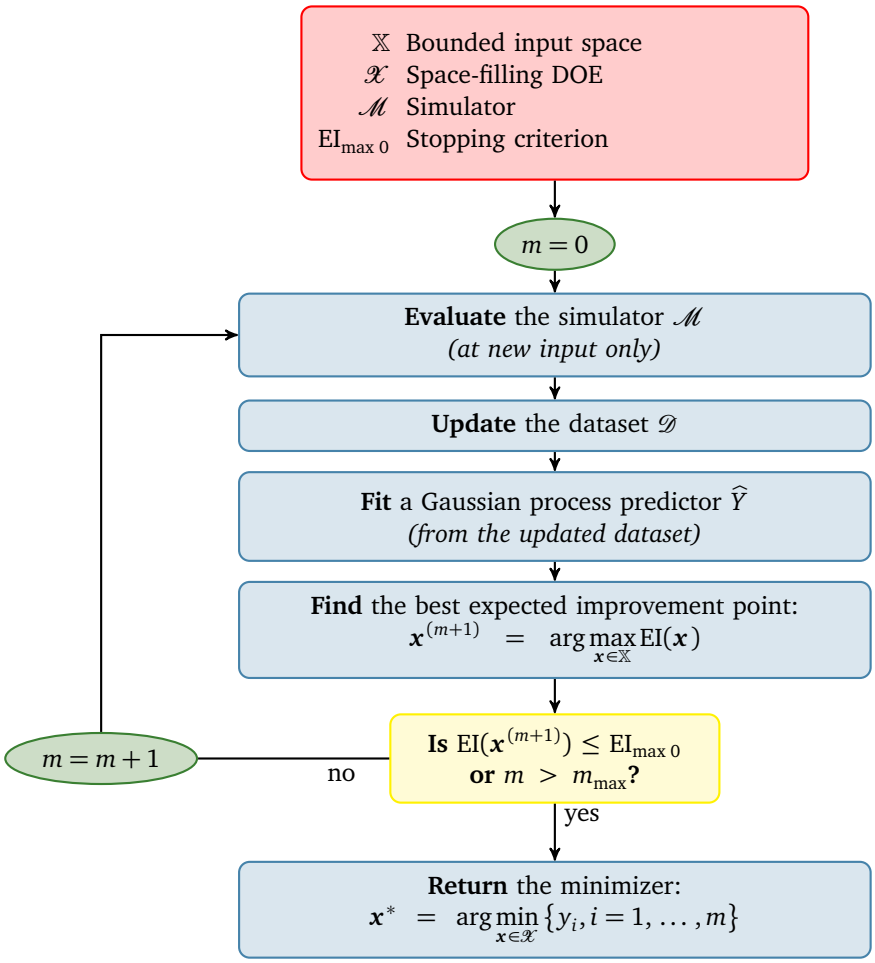
The *expected improvement* is depicted in Figure 2.3 for the one-dimensional objective function $\mathscr{M}(x) = x\sin(x)$ used in Section 1.5.2. This figure illustrates the state of the *efficient global optimizer* at some point where the DOE contains $m = 8$ points (in red). The maximizer of the expected improvement is represented as the green square and corresponds to the best current guess about the location of the global minimizer of $\mathscr{M}$ on the bounded interval $\mathbb{X} = [0; 15]$. Note that the "best guess" is wrong in this case although the second EI-extremum was right. This remark partly motivates the approach proposed in Section 2.4.

2.3.1.2 The efficient global optimizer

Eventually, based on the expected improvement, Jones et al. (1998) came up with the *efficient global optimizer* detailed in Algorithm 2.1.

The algorithm requires a bounded admissible input space $\mathbb{X}$, an initial space-filling experimental design $\mathscr{X}$ as well as the ability to sequentially ask for new evaluations of the original model $\mathscr{M}$ at most $m_{\max}$ times. It begins with the computation of the model response on the initial DOE. Then it fits a Gaussian process model using *e.g.* maximum likelihood estimation as detailed in Section 1.4.5.2. The next point $\boldsymbol{x}^{(m+1)}$ is the global maximizer of the expected improvement computed from the available Gaussian process predictor $\widehat{Y}$. Note that any brute force optimization algorithm (*e.g.* sampling-based algorithms) can be used here because the expected improvement is based on the inexpensive predictor $\widehat{Y}$.

Algorithm 2.1 Efficient Global Optimizer



The overall procedure is repeated until the maximum expected improvement falls under some given threshold $\mathrm{EI}_{\max 0}$ unless the model evaluation budget $m_{\max}$ is exceeded before convergence. Jones et al. (1998) recommends $\mathrm{EI}_{\max 0} = 10^{-2}$, although this value should be carefully selected with respect to the order of magnitude of the model response y. The algorithm eventually returns the best minimizer in the final DOE.

The convergence of this optimizer is not proved though and mostly depends on the ability of the Gaussian process model to capture the behaviour of the original function $\mathscr{M}$. In an attempt to improve its convergence properties, various modifications were brought to that pioneering algorithm, most of them concentrating on its criterion: the *expected improvement*. For instance, Schonlau (1997) proposed the *generalized expected improvement* criterion which features an additional parameter to allow a more global search than its standard counterpart. Villemonteix et al. (2009) uses an *entropy* criterion summarizing the amount of uncertainty about the location of the sought minimizer $\boldsymbol{x}^*$ expressed in the form of a probability distribution. The latter interpretation of these criteria will be developed later on in Section 2.4.

2.3.2 Refinement criteria for contour approximation

Another local refinement of particular interest throughout this manuscript concerns the accurate approximation of a specific contour of the original computational model $\mathscr{M}$. Recall that a contour at level $t \in \mathbb{Y}$ is defined as $\mathscr{S}_t = \{\boldsymbol{x} \in \mathbb{X} : \mathscr{M}(\boldsymbol{x}) = t\}$. The correct approximation of this contour is of utmost importance in at least two kinds of problems: for constrained optimization and for reliability analysis where the constraint or the limit-state function (respectively) is approximated with a Gaussian process predictor. This section reviews state-of-the-art criteria available in the literature in order to adapt the efficient global optimization meta-heuristic to the contour refinement problem.

2.3.2.1 Simple criteria

The simplest criteria presented in this section are based on the following premise. In Section 1.5.3, it was emphasized that the class (failed or safe) of the points $\boldsymbol{x}$ located in the margin of uncertainty $\mathbb{M}_{1-\alpha}$ (see Eq. (1.152)) is not predicted with a sufficient confidence level. Its sign is indeed guaranteed at a lesser probability level than $1-\alpha/2$. It is argued here that the refinement of the contour of interest can be achieved by reducing the width of this margin of uncertainty. Indeed, adding points in $\mathbb{M}_{1-\alpha}$ will significantly improve the prediction and reduce its spread thanks to the consistency of the Kriging predictor (see Section 1.4.4).

Note that a similar idea for *support vector classifiers* is referred to as the *margin shrinking concept* in the book by Hurtado (2004b). According to that concept, only the points located in the margin of the classifier can correct its shape because they are most likely to become *support vectors*. This concept was successfully applied to structural reliability problems by Deheeger (2008); Bourinet et al. (2011) and Basudhar and Missoum (2010).

The margin indicator function

This idea eventually leads to the elaboration of the first simplest statistic named the *margin indicator* function defined as:

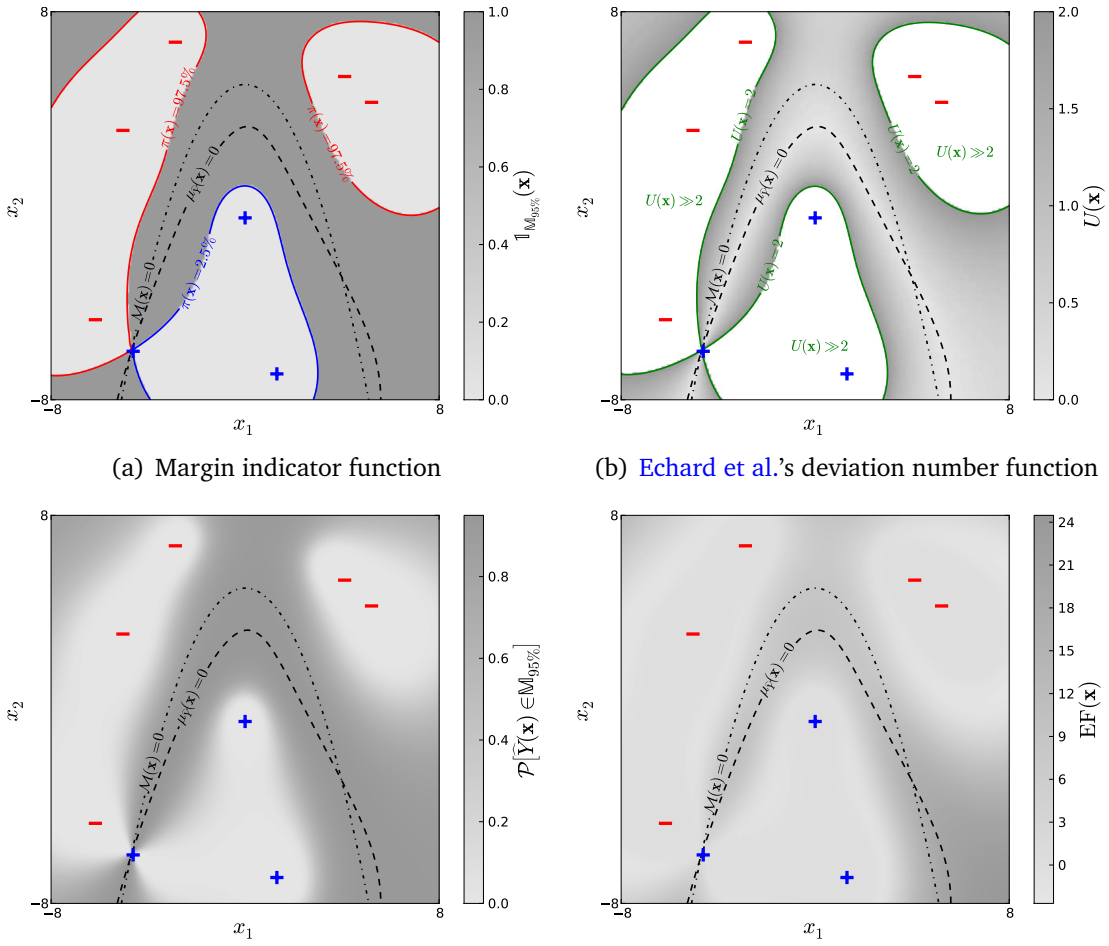
$$\mathbb{1}_{\mathbb{M}_{1-\alpha}}(\boldsymbol{x}) = \begin{cases} 1 & \text{if} \quad \mu_{\widehat{Y}}(\boldsymbol{x}) \in \mathbb{M}_{1-\alpha} \\ 0 & \text{otherwise} \end{cases} . \tag{2.29}$$

The deviation number function

Echard et al. (2011) proposed an equivalent criterion measuring how close is the mean prediction $\mu_{\widehat{Y}}$ to the contour of interest in terms of number of Kriging standard deviations $\sigma_{\widehat{Y}}$. This lead to the definition of the following *deviation number* function:

$$U(\boldsymbol{x}) = \frac{\left|t - \mu_{\widehat{Y}}(\boldsymbol{x})\right|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}, \tag{2.30}$$

$t \in \mathbb{Y}$ being the level of interest. Echard et al. (2011) also state that a point $\boldsymbol{x}$ for which $U(\boldsymbol{x}) > 2$ is classified with a sufficient confidence level. Actually, this arbitrary value comes from a confidence level close to 95% (remember $k_{95\%} = 1.96$) so as to compare with the margin of uncertainty $\mathbb{M}_{1-\alpha}$. This comparison is illustrated in Figure 2.4. In Figure 2.4(b), only the values of the deviation number function less than 2 were represented because its gradient is then too large.



(a) Margin indicator function

(b) Echard et al.'s deviation number function

(c) Margin probability function

(d) Bichon et al.'s expected feasibility function

Figure 2.4: Representation of the four simple criteria on the example from Section 1.5.3. They all have their extrema in the vicinity of the identified contour.

The margin probability function

The smoother *margin probability* function which is used in this thesis is defined as follows:

$$\mathrm{MP}(\boldsymbol{x}) \equiv \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \in \mathbb{M}_{1-\alpha}\right] = \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \in \left(\widehat{\mathbb{F}}^{+1}_{1-\alpha} \setminus \widehat{\mathbb{F}}^{-1}_{1-\alpha}\right)\right]. \tag{2.31}$$

It further simplifies in application of the inclusion in Eq. (1.151). Indeed:

$$\begin{aligned}\mathrm{MP}(\boldsymbol{x}) &= \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \in \widehat{\mathbb{F}}^{+1}_{1-\alpha}\right] - \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \in \widehat{\mathbb{F}}^{-1}_{1-\alpha}\right] \\ &= \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq t + k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x})\right] - \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq t - k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x})\right] \\ &= \Phi\left(\frac{t + k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x}) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) - \Phi\left(\frac{t - k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x}) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right). \end{aligned} \tag{2.32}$$

This statistic was already mentioned by Picheny et al. (2010a), although it is used in a more elaborated *one-step-look-ahead strategy* as evoked hereafter.

In this manuscript, it is defended that this criterion is sufficiently objective to efficiently apply the *margin shrinking concept* for the refinement of a contour approximated by a Kriging prediction. It has a spread comparable with the margin of uncertainty $\mathbb{M}_\alpha$ and vanishes far from the contour of interest.

The expected feasibility function

Eventually, based on the *expected improvement* statistic, Bichon et al. (2008) proposed the *expected feasibility* function which measures the expected *proximity* of the Kriging prediction to the threshold t of interest. The name "feasibility" is actually due to its optimization background: it was first conceived to deal with equality constrained optimization problems for which a *feasible* solution lies on a specific contour. The definition of the so-called *feasibility* function requires the choice of a proximity parameter $\varepsilon \in \mathbb{Y}$ and is as follows:

$$\mathrm{F}(\boldsymbol{x}) \equiv \begin{cases} 0 & \text{if } \widehat{Y}(\boldsymbol{x}) \leq t - \varepsilon \\ \varepsilon - \left|\widehat{Y}(\boldsymbol{x}) - t\right| & \text{if } t - \varepsilon \leq \widehat{Y}(\boldsymbol{x}) \leq t + \varepsilon \\ 0 & \text{if } t + \varepsilon \leq \widehat{Y}(\boldsymbol{x}) \end{cases} = \max\left\{\varepsilon - \left|\widehat{Y}(\boldsymbol{x}) - t\right|, 0\right\}. \tag{2.33}$$

Its relationship with the improvement in Eq. (2.25) is clear from its definition. Indeed, the feasibility is nothing but a bilateral extension of the improvement because the progress now consists in finding a point $\boldsymbol{x} \in \mathbb{X}$ for which the Gaussian prediction $\widehat{Y}(\boldsymbol{x})$ is closer to threshold t than the given proximity $\varepsilon \in \mathbb{Y}$. Although that proximity could be set to the smallest deviation in the experimental design, Bichon et al. (2008) propose instead to make it vary in the input space $\mathbb{X}$ by setting it proportional to the Kriging standard deviation: $\varepsilon(\boldsymbol{x}) \propto \sigma_{\widehat{Y}}(\boldsymbol{x})$. This somehow results in measuring the spread of the margin of uncertainty $\mathbb{M}_{1-\alpha}$.

As for the improvement $\mathrm{I}(\boldsymbol{x})$, the feasibility is a truncated Gaussian random variate so that it has to be summarized with a constant value to be used as a refinement statistic. This lead Bichon et al. (2008) to propose the *expected feasibility* function defined as:

$$\mathrm{EF}(\boldsymbol{x}) \equiv \mathbb{E}\left[\mathrm{F}(\boldsymbol{x})\right] = \int_{t-\varepsilon(\boldsymbol{x})}^{t+\varepsilon(\boldsymbol{x})} \left(\varepsilon(\boldsymbol{x}) - \left|\widehat{y} - t\right|\right) \varphi\left(\frac{\widehat{y} - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) \mathrm{d}\widehat{y}, \tag{2.34}$$

which they show (see Bichon, 2010, p. 145–147) to have the following closed-form expression:

$$\begin{aligned}\mathrm{EF}(\boldsymbol{x}) = &\left(\mu_{\widehat{Y}}(\boldsymbol{x}) - t\right)\left[2\,\Phi\left(z^{\,0}(\boldsymbol{x})\right) - \Phi\left(z^{-1}(\boldsymbol{x})\right) - \Phi\left(z^{+1}(\boldsymbol{x})\right)\right] \\ &+ \sigma_{\widehat{Y}}(\boldsymbol{x})\left[2\,\varphi\left(z^{\,0}(\boldsymbol{x})\right) - \varphi\left(z^{-1}(\boldsymbol{x})\right) - \varphi\left(z^{+1}(\boldsymbol{x})\right)\right] \\ &+ \varepsilon(\boldsymbol{x})\left[\Phi\left(z^{+1}(\boldsymbol{x})\right) - \Phi\left(z^{-1}(\boldsymbol{x})\right)\right],\end{aligned} \tag{2.35}$$

where we have introduced:

$$z^{\,i}(\boldsymbol{x}) = \frac{t + i\,\varepsilon(\boldsymbol{x}) - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}, \quad i = -1,\ 0,\ +1. \tag{2.36}$$

The expected feasibility function is always positive and it vanishes (at the limit) for the points in the DOE $\mathcal{X}$ alike the margin probability and the expected improvement functions.

Illustration and conclusion

These simple contour refinement criteria are represented in Figure 2.4 on the example from Section 1.5.3. It can be seen that they all have their *extrema* in the vicinity of the identified limit-state (the margin of uncertainty) where the sign of the prediction is the most uncertain. From that premise the criterion maximization principle of the efficient global optimization algorithm still holds, so that *the next best point* to add to the DOE is:

$$\boldsymbol{x}^{(m+1)} = \arg \max_{\boldsymbol{x}\in\mathbb{X}} \ \mathscr{C}(\boldsymbol{x}) \tag{2.37}$$

where $\mathscr{C}$ might be set equal to either the *margin indicator* function in Eq. (2.29), the *margin probability* function in Eq. (2.32) or the *expected feasibility* function in Eq. (2.35). The *deviation number* function excepts this rule as it should be minimized, so that:

$$\boldsymbol{x}^{(m+1)} = \arg \min_{\boldsymbol{x}\in\mathbb{X}} \ U(\boldsymbol{x}) \tag{2.38}$$

It can be seen from Figures 2.4(c) and 2.4(d) that the other two criteria are highly multi-modal – as the *expected improvement* is in Figure 2.3, so that the global optimization problem in Eq. (2.37) is still non-trivial to solve. This partly motivates the sampling alternative introduced in Section 2.4.

The criteria presented here are qualified as *simple* because they all have closed-form expressions and are thus both easy to implement and fast to evaluate in contrast with the others forthcoming *one-step-look-ahead criteria*.

2.3.2.2 One-step-look-ahead criteria

The *one-step-look-ahead criteria* introduced in this section attempts to quantify the improvement of the Kriging approximation one could expect when adding a specific point $\tilde{\boldsymbol{x}}^{(m+1)} \in \mathbb{X}$ in the DOE $\mathcal{X}$. In other words, they try to anticipate the state of the refinement procedure one step ahead. Therefore the strategies implementing such criteria were named *one-step-look-ahead strategies* by Bect et al. (2011).

The weighted integrated mean-squared error

Picheny et al.'s *weighted integrated mean-squared error* (2010a) is an extension of Sacks et al.'s *integrated mean-squared error* (1989a) for the contour refinement problem. It consists in integrating the prediction variance $\sigma^2_{\tilde{Y}}$ of a *virtual* Kriging predictor $\tilde{Y}$ built from the augmented DOE $\mathscr{X} \cup \left\{\tilde{\boldsymbol{x}}^{(m+1)}\right\}$ over a vicinity of the available approximation of the contour of interest $\left\{\boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) = t\right\}$ where $\mu_{\widehat{Y}}$ is the mean of the Kriging predictor $\widehat{Y}$ built from the available dataset $\mathscr{D}$.

More specifically, its definition is the same as the integrated mean-squared error in Eq. (2.20):

$$\text{W-IMSE}(\mathscr{X} \cup \{\tilde{\boldsymbol{x}}^{(m+1)}\}) \equiv \int_{\mathbb{X}} \sigma^2_{\tilde{Y}}(\boldsymbol{x})\, w\left(\widehat{Y}(\boldsymbol{x})\right) \mathrm{d}\boldsymbol{x}, \tag{2.39}$$

except that the weighting density $w(\widehat{Y}(\bullet))$ is selected so as to emphasize the mean-squared error of the closest points to the identified contour. For instance Picheny et al. (2010a) proposes to use the margin indicator statistic in Eq. (2.29) or the margin probability statistic in Eq. (2.32) as such weighting densities.

Unlike the integrated mean-squared error statistic evoked earlier in Section 2.2.2, the weighted integrated mean-squared error proposed here is practically implementable because the m first elements of the dataset $\mathscr{D}$ are available and thus allow one to compute:

- the prediction variance $\sigma^2_{\tilde{Y}}$ of the virtual Kriging predictor $\tilde{Y}$, because it does not depend on the unavailable observation $y_{m+1} = \mathscr{M}(\boldsymbol{x}^{(m+1)})$ and because the autocorrelation $R(\bullet, \boldsymbol{\theta})$ was already determined from the available observations $\boldsymbol{y}$;
- and, the weighting density $w(\widehat{Y}(\bullet))$ solely depends on the available observations $\boldsymbol{y}$.

The best candidate point that should be added to the DOE according to that novel metric is defined so as to minimize it:

$$\boldsymbol{x}^{(m+1)} = \arg \min_{\tilde{\boldsymbol{x}}^{(m+1)} \in \mathbb{X}} \text{W-IMSE}(\tilde{\boldsymbol{x}}^{(m+1)}). \tag{2.40}$$

However the repeated estimations of the integral in Eq. (2.39) require a larger computational effort than the other simpler criteria because they resort to numerical integration techniques. Picheny (2009, pp. 97–98) provides more computational details on both the integration and optimization problems, as well as heuristic approximations to speed up the decision process.

Stepwise uncertainty reduction strategies

The *stepwise uncertainty reduction* criteria from Bect et al. (2011) are specifically designed for the estimation of rare event probabilities, so that we need to introduce a few more facts about the surrogate-based probability estimation problem. These facts will be

made short here and discussed into more details in Chapter 3. Consider that the input $\boldsymbol{x}$ is the realization of a random vector $\boldsymbol{X}$ with known PDF $f_{\boldsymbol{X}}$ (and probability measure $\mathbb{P}$). The probability estimation problem that is being dealt with by Bect et al. (2011) consists in estimating the following failure probability defined as:

$$p_f \equiv \mathbb{P}\left[\mathscr{M}(\boldsymbol{X}) \leq t\right] = \int_{\mathbb{F}=\{\boldsymbol{x}\in\mathbb{X}:\,\mathscr{M}(\boldsymbol{x})\leq t\}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \mathbb{E}_{f_{\boldsymbol{X}}}\left[\mathbb{1}_{\mathbb{F}}(\boldsymbol{X})\right], \tag{2.41}$$

where we have introduced the failure indicator function $\mathbb{1}_{\mathbb{F}}$ which is equal to one if $\boldsymbol{x} \in \mathbb{F}$ and 0 otherwise.

Starting from the usual premise that $\mathscr{M}$ is too expensive-to-evaluate to accurately estimate the integral defining p_f on $\mathscr{M}$, Bect et al. (2011) resort to a Kriging approximation $\widehat{Y}$ for $\mathscr{M}$, and the failure indicator function $\mathbb{1}_{\mathbb{F}}$ is approximated with the *probabilistic classification function* π already defined in Eq. (1.148). Eventually, they propose to estimate the probability in Eq. (2.41) with the following plug-in estimator:

$$\tilde{p}_f \equiv \mathbb{P}\left[\mathscr{P}\left[\widehat{Y}(\boldsymbol{X}) \leq t\right]\right] = \int_{\mathbb{X}} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \mathbb{E}_{f_{\boldsymbol{X}}}\left[\pi(\boldsymbol{X})\right]. \tag{2.42}$$

This surrogate-based estimator being defined, the purpose of Bect et al. (2011) is to design a refinement criterion that will summarize the improvement brought by a new observation at the candidate location $\tilde{\boldsymbol{x}}^{(m+1)}$ on the final probability of interest. They propose the following mean-squared error metric:

$$J_{\mathrm{SUR}}(\tilde{\boldsymbol{x}}^{(m+1)}) = \mathbb{E}_{\tilde{Y}}\left[(p_f - \tilde{p}_f)^2\right], \tag{2.43}$$

where $\mathbb{E}_{\tilde{Y}}\left[\bullet\right]$ denotes (with slight abuse of notation) the mathematical expectation with respect to the virtual Kriging predictor $\tilde{Y}$ built from the augmented dataset $\mathscr{D} \cup (\tilde{\boldsymbol{x}}^{(m+1)}, \tilde{y}_{m+1})$.

However, once again, this statistic is impractical because one does not know the sought failure probability p_f and the future observation $\tilde{y}_{m+1}$. As a matter of fact, Bect et al. (2011) derived two upper bounds for this statistic using Minkowski and Cauchy-Schwartz inequalities that fortunately do not depend on the former unavailable quantities.

The evaluation of these upper bounds require a double numerical integration over $\mathbb{X}\times\mathbb{Y}$ which is admitted to be quite expensive. The inner integration over $\mathbb{Y}$ resorts to Gauss quadrature while the outer one resorts to Monte Carlo simulation. The next best point to add to the DOE minimizes the upper bound $\overline{J}_{\mathrm{SUR}}$ of the *stepwise uncertainty reduction* statistic:

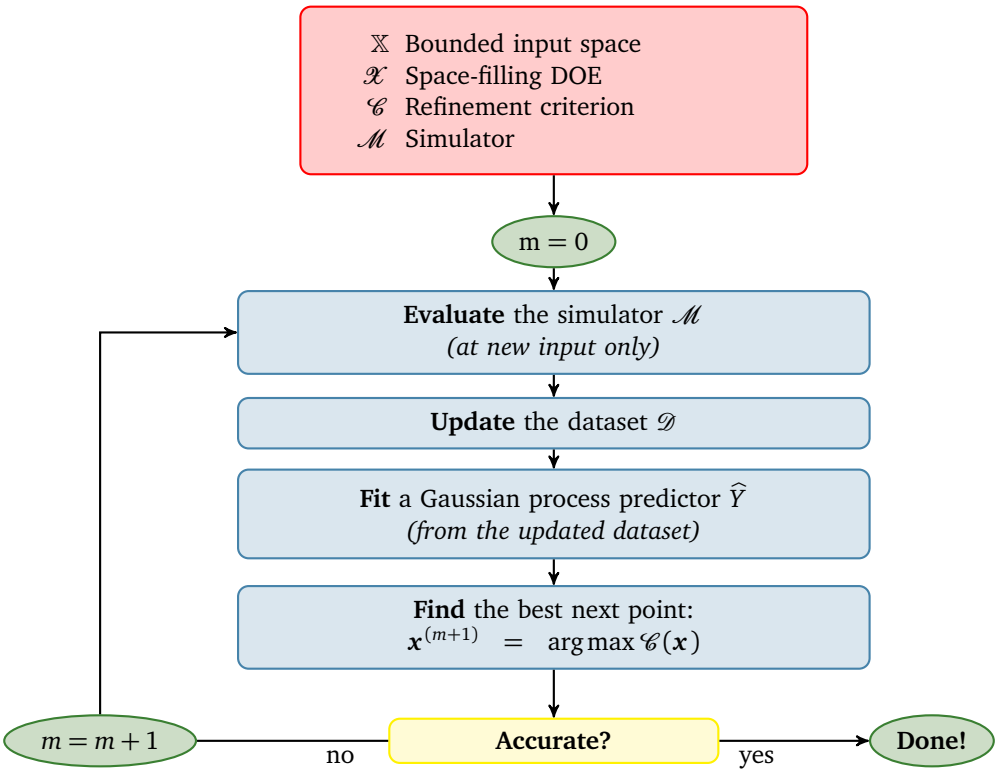
$$\boldsymbol{x}^{(m+1)} = \arg \min_{\tilde{\boldsymbol{x}}^{(m+1)}\in\mathbb{X}} \overline{J}_{\mathrm{SUR}}(\tilde{\boldsymbol{x}}^{(m+1)}). \tag{2.44}$$

To conclude, it is worth pointing out that Bect et al. (2011) benchmarked most of the statistics introduced in this section and concluded that they all asymptotically lead to a good refinement of the contour of interest even though their *stepwise uncertainty reduction* statistics perform even better (according to the metric and the examples used to assess this performance). It is argued here that the *margin shrinking concept* is a sufficient heuristic to achieve an accurate contour approximation at a more reasonable computational expense.

2.3.3 Standard adaptive refinement strategy

The standard refinement strategy is based on the efficient global optimization meta-heuristic. It is summarized in Algorithm 2.2. The procedure requires an initial dataset $\mathscr{D}$ and assumes that one can afford new evaluations of the computational model $\mathscr{M}$. First, a Gaussian process predictor $\widehat{Y}$ is fitted to the dataset. The selected refinement criterion is defined from this first prediction and a global optimization algorithm is used to find its optimizer $\boldsymbol{x}^{(m+1)}$. The response value of the simulator is evaluated at $\boldsymbol{x}^{(m+1)}$, and the dataset $\mathscr{D}$ is updated. The procedure is repeated until some accuracy criteria are met. These accuracy criteria depends on the context and on the refinement criterion which is used.

Algorithm 2.2 Standard adaptive refinement strategy



It should be pointed out that the strategy proposed by Echard et al. (2011) is slightly different. Before applying the refinement procedure, the input space $\mathbb{X}$ is discretized with a large sample $\mathfrak{X}$ of size $N = 10^{5-6}$ and their purpose is to ensure that the points in this sample are accurately classified. Thus they apply the same procedure as in Algorithm 2.2 except that the optimization is discrete. The discrete optimization problem is solved by means of an exhaustive search amongst the $N - m$ points in the initial sample $\mathfrak{X}$ that are not in the experimental design $\mathscr{X}$. The procedure stops when there is no more points $\boldsymbol{x}$ in $\mathfrak{X}$ such that their statistic $U(\boldsymbol{x})$ is less than 2. As a result, all the points in the initial sample $\mathfrak{X}$ are accurately classified with a confidence level equal to $\Phi(2) > 97\%$ *w.r.t.* the Kriging epistemic uncertainty.

2.4 Sampling-based adaptive designs of experiments

2.4.1 Motivation

It has been shown in the previous section (see Figures 2.3 and 2.4) that the refinement criteria proposed in the literature are all highly multi-modal and sometimes even feature a *large plateau*. The existence of such plateau is especially true for the contour refinement criteria as the approximate contour $\widehat{\mathscr{S}_t} = \{\boldsymbol{x} \in \mathbb{X} : \widehat{Y}(\boldsymbol{x}) = t\}$ is composed with an infinity of points.

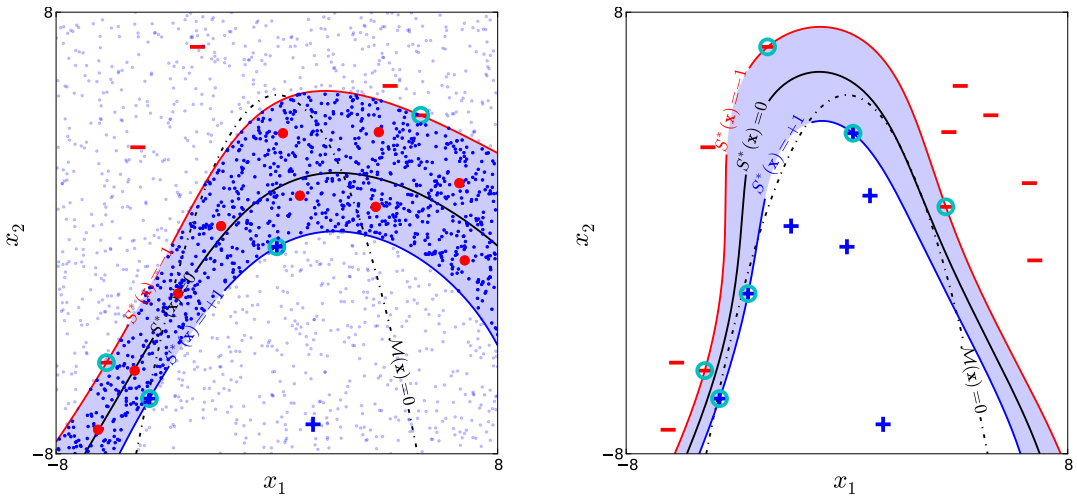
This multi-modality thus makes the global optimization problem involved in Algorithm 2.2 rather difficult to solve. Moreover if it is hard to solve, then it is most probably ill-posed. This would mean that there does not exist *one single* best point or at least that one does not have enough knowledge yet to say that such a point even exists. Nevertheless, this refinement strategy usually ends up giving an accurate contour approximation after a sufficient number of iterations. Mockus (1994) explains that "there is no need for *exact* minimization of the risk function[1]" because the global optimization of the criterion is involved in an *iterative* procedure.

On the other hand, the margin shrinking concept (Hurtado, 2004b) which consists in adding points located in the margin of a support vector classifier does not benefit from a real-valued refinement criterion. Indeed, the refinement statistic is alike the margin indicator function $\mathbb{1}_{\mathbb{M}_{1-\alpha}}$ defined in Eq. (2.29) and only takes binary values. Global optimization of such a criterion is definitely ill-posed and most authors involved in the adaptive refinement of such classifiers came up with the conclusion that *any* point in the margin significantly improves the prediction.

Points in the margin are classically obtained by means of an accept-reject Monte Carlo simulation in the input space $\mathbb{X}$. However, it can be seen from Figure 2.5 that a large number of points is required to ensure that the Monte Carlo sample truly fills the margin. Again, in the present context such a high number is not affordable so that Deheeger (2008); Bourinet et al. (2011) proposed to reduce this large candidate population to its $K > 1$ cluster centres. Such centres are obtained by means of a state-of-the-art K-means clustering algorithm reviewed in Section 2.2.1.3. The overall procedure is illustrated in Figure 2.5. It can be seen that 4 out of the 5 support vectors (circled in cyan in Figure 2.5(b)) are amongst the $K = 10$ cluster centres proposed by the margin shrinking procedure. This is the key argument of the margin shrinking concept: only the points located in the margin can significantly change the shape of the contour and make it tend toward the real separator, *i.e.* here $\mathscr{S}_0 = \{\boldsymbol{x} \in \mathbb{X} : \mathscr{M}(\boldsymbol{x}) = 0\}$.

As opposed to the standard refinement strategy for Gaussian process predictors, the margin shrinking concept delivers a set of points instead of a single one. This interesting feature enables the use of distributed computing platforms, *i.e.* the simulator $\mathscr{M}$ may possibly be evaluated in parallel on a set of available CPUs.

[1] Mockus' refinement criterion.



(a) Initial margin filled with $K = 10$ points

(b) Updated (shrinked) margin

Figure 2.5: Application of the margin shrinking concept to a support vector classifier. The $K = 10$ points (red dots in the left-hand panel) selected for the refinement were clustered from a larger population of size $N = 1\,000$ (blue points) generated by an accept-reject Monte Carlo simulation in the input space $\mathbb{X} = [-8;\ 8]$.

It is also worth mentioning that this limitation of Gaussian process refinement strategies has already been investigated by Ginsbourger et al. (2010) in a global optimization context. To achieve the parallelization of the efficient global optimizer on a cluster of PCs, Ginsbourger et al. (2010) uses a modification of the expected improvement in Eq. (2.27) that attempts to account for the joint improvement of q points picked simultaneously. The so-called *q-point expected improvement* is defined as follows:

$$\mathrm{EI}(\boldsymbol{x}^{(m+1)}, \ldots, \boldsymbol{x}^{(m+q)}) \equiv \mathbb{E}\left[\max_{i=1,\ldots,q}\left(\max\left(\widehat{Y}(\boldsymbol{x}^{(m+i)}) - y_{\min}, 0\right)\right)\right]. \tag{2.45}$$

The motivation for picking the maximum of the q marginal improvements is not clear though. The full closed-form expression for the 2-points expected improvement is derived in the original book chapter as well as in Ginsbourger's Ph.D. thesis (2009). It involves the CDF of the bivariate Gaussian distribution Φ_2. Ginsbourger et al. (2010) show on an example that the two points yielded by this criterion approximately correspond to the two main modes of the usual 1-point expected improvement (see *e.g.* Figure 2.3 for a bi-modal expected improvement). However Ginsbourger et al. (2010) also admit that "this is not a general fact" . Higher order q-point expected improvements involve the CDF of the q-variate ($q > 2$) Gaussian distribution Φ_q which can be approximated numerically (see *e.g.* Genz, 1992, 1993).

From the above remarks it has been decided to propose an adaptive improvement that is based on sampling instead of the optimization of a specific refinement criterion.

2.4.2 Proposed adaptive refinement strategy

The proposed adaptive refinement strategy is inspired by the margin shrinking concept for support vector classifiers applied here to Gaussian process predictors. It consists in two key steps which are reviewed in the sequel: a *sampling* step and a *clustering* step.

2.4.2.1 Sampling of the refinement criterion

Let $\mathscr{C}$ denote any of the criteria introduced in the previous section. Let us additionally require that *(i)* the "*best*" points *maximize* this quantity, and that *(ii)* $\mathscr{C}$ is positive (possibly zero) everywhere in the input space $\mathbb{X}$. Note that even if all the criteria available in the literature are positive, a negative criterion might be translated up to the positive real line by means of the following mapping:

$$\mathscr{C}^+(\boldsymbol{x}) = \exp(\pm\mathscr{C}(\boldsymbol{x})), \quad \boldsymbol{x} \in \mathbb{X}, \tag{2.46}$$

which is commonly used in the *simulated annealing* optimization meta-heuristic. The exponential function is strictly increasing so that it does not change the sense of variation of the original criterion. However the sign in the exponential possibly allows one to alter the change of variation to match the first requirement *(i)*.

The proposed refinement procedure relies on the interpretation of the refinement criterion $\mathscr{C}$ as a probability density function for the "*best*" points. In other words, the "best" point is indeed a random vector $\boldsymbol{C}$ with PDF:

$$f_{\boldsymbol{C}}(\boldsymbol{c}) \propto \mathscr{C}(\boldsymbol{c}), \quad \boldsymbol{c} \in \mathbb{R}^n. \tag{2.47}$$

Such an interpretation raises a fundamental problem though. Indeed, all the refinement criteria are based on the Kriging prediction variance which is known to increase "far" from the experimental design $\mathscr{X}$. Thus, the refinement criteria does not have a finite integral on $\mathbb{R}^n$:

$$\int_{\mathbb{R}^n} \mathscr{C}(\boldsymbol{c})\,\mathrm{d}\boldsymbol{c} \to \infty. \tag{2.48}$$

A simple workaround to the latter problem is considered. Recall that in the standard refinement strategy, the global optimization algorithm usually requires that the input space $\mathbb{X}$ is bounded so that the optimization does not get lost far from the DOE $\mathscr{X}$. A direct consequence for the PDF of $\boldsymbol{C}$ is that the criterion might indeed be weighted as follows:

$$f_{\boldsymbol{C}}(\boldsymbol{c}) \propto \mathscr{C}(\boldsymbol{c})\,w(\boldsymbol{c}), \quad \boldsymbol{c} \in \mathbb{R}^n. \tag{2.49}$$

where w is the weighting pseudo-PDF (not necessarily normalized) that might either be set to a simple *input space indicator function*:

$$\mathbb{1}_{\mathbb{X}}(\boldsymbol{c}) = \begin{cases} 1 & \text{if } \boldsymbol{c} \in \mathbb{X} \\ 0 & \text{otherwise} \end{cases}, \quad \boldsymbol{c} \in \mathbb{R}^n. \tag{2.50}$$

or a real PDF f_X for the input $\boldsymbol{x}$ when appropriate, *e.g.* in an uncertainty quantification context. Note that a real PDF f_X may also be used to express a prior belief about the location of the global optimizer in a global optimization context.

Thanks to that weighting function w, the PDF of the "best" point is now *proper* since:

$$\int_{\mathbb{R}^n} \mathscr{C}(\boldsymbol{c}) w(\boldsymbol{c}) \, \mathrm{d}\boldsymbol{c} < \infty, \tag{2.51}$$

and it is thus well-suited to sample from. Sampling from f_C is not direct though. It resorts to *Markov chain Monte Carlo* (MCMC, see Appendix B) techniques which are specifically designed to sample from distributions that are only defined with a PDF (possibly non-normalized). Let us denote by $\mathfrak{C} = \{\boldsymbol{c}^{(i)}, i = 1, \ldots, N\}$ the sample of "candidates" generated from f_C.

According to the author's experience the *slice sampling* technique proposed by Neal (2003) and shortly reviewed in Section B.3.3 of Appendix B outperforms the other more conventional Metropolis-Hastings algorithm in this context because of the high number of modes featured by f_C. Indeed it can be seen from Figure 2.4 that the refinement statistics are all narrow-banded in the vicinity of the approximated contour. Moreover this spread will vanish along with the refinement of the contour of interest resulting in highly skewed multimodals PDFs for which slice sampling performs much better than Metropolis-Hastings.

Note that MCMC techniques require a first point that is distributed according to the targeted PDF (here f_C). Such a point is referred to as a *seed* denoted by $\boldsymbol{c}^{(0)}$ and it is such that:

$$f_C(\boldsymbol{c}^{(0)}) > 0. \tag{2.52}$$

Such a seed may easily be obtained by an accept-reject Monte Carlo simulation using the weighting PDF w as a proposal PDF. This method may prove inefficient though as the PDF f_C gets skewed in the vicinity of the approximate contour $\widehat{\mathscr{S}_t}$ (in $\mathbb{M}_{1-\alpha}$) along with the refinement procedure. But if the problem of finding the seed becomes difficult, then it certainly means that the approximate contour $\widehat{\mathscr{S}_t}$ is sufficiently accurate (and that the refinement should stop).

Figure 2.6 illustrates one sample generated according to the weighted margin probability function in Eq. (2.32). Sampling is restricted to the bounded input space $\mathbb{X} = [-8; 8]^2$ by means of the indicator function in Eq. (2.50). It can be seen that the samples are denser in the vicinity of the modes of f_C than anywhere else as requested.

2.4.2.2 Statistical reduction of the candidates

As for the margin shrinking concept, a large sample size is required in order to ensure that the N-sample $\mathfrak{C}$ truly fills the region of interest (say $N = 10^{3-4}$). Again, one cannot afford evaluating a few thousand points with respect to the limited computational budget (*i.e.* limited number of calls to the simulator $\mathscr{M}$). Hence, it is proposed to reduce the large sample $\mathfrak{C}$ to a smaller subset that exhibits the same statistical properties, *i.e.* it uniformly spans the region of interest. As proposed by Deheeger (2008); Bourinet et al. (2011) such

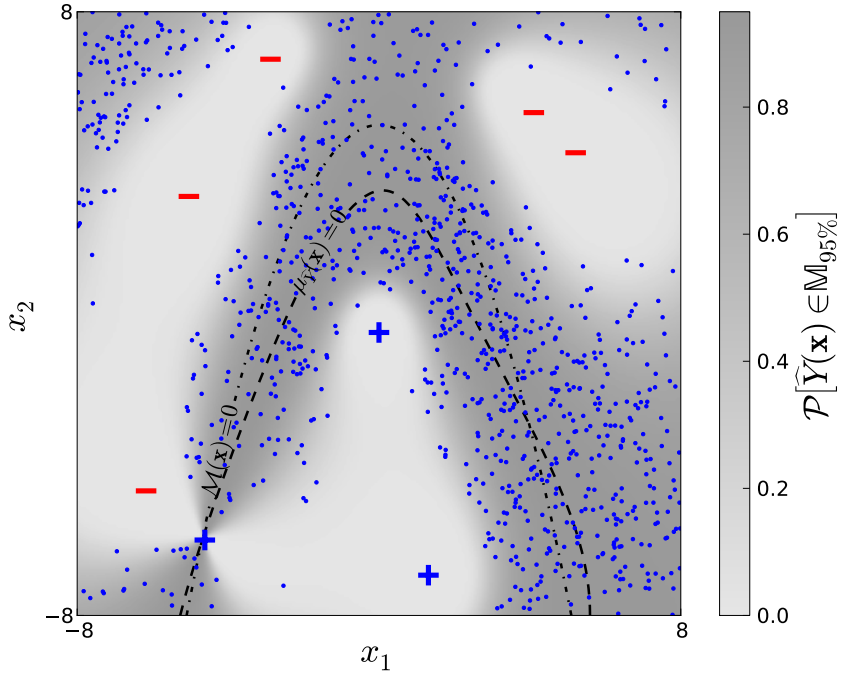


Figure 2.6: A few candidates distributed according to the weighted margin probability function defined in Eq. (2.32). The samples are denser in the vicinity of the identified contour as expected.

a *reduction* may be carried out by means of the K-means clustering algorithm detailed in Section 2.2.1.3.

However, since the final set of cluster centres $\mathcal{O}^*$ does not belong to the initial sample $\mathfrak{C}$ and that some of the proposed centres may be associated with low values of f_C, it is proposed to pick the closest points in $\mathfrak{C}$ to the cluster centres provided by the K-means EM algorithm detailed in Section 2.2.1.3. The new set of points yielded by this procedure is thus defined as follows:

$$\boldsymbol{x}^{(m+k)} = \arg \min_{\boldsymbol{c} \in \mathfrak{C} \setminus \{\boldsymbol{x}^{(m+i)},\, i=1,\dots,k\}} \|\boldsymbol{o}^{*(k)} - \boldsymbol{c}\|_2^2, \quad k = 1, \dots, K, \tag{2.53}$$

where each point is picked iteratively one after the other. The uniqueness of each point is ensured by successively removing them from the initial sample $\mathfrak{C}$.

The points obtained with this heuristic procedure are illustrated as the red dots in Figure 2.7. It can be seen that they uniformly span the candidate population and thus the region of interest as required.

Note that the present section has not provided any rule to choose the number of cluster centres K yet. There are mainly two reasons for this. First, from a pragmatic viewpoint, K is set equal to the number of available CPUs (to run $\mathscr{M}$ in parallel) for the applications presented in this manuscript. Second, choosing K objectively is a non-trivial problem. A common practice consists in comparing the different minimal inertia yielded by the EM procedure, and picking the centres associated with the *optimum optimorum* but this technique is more a rule of thumb than a real criterion.

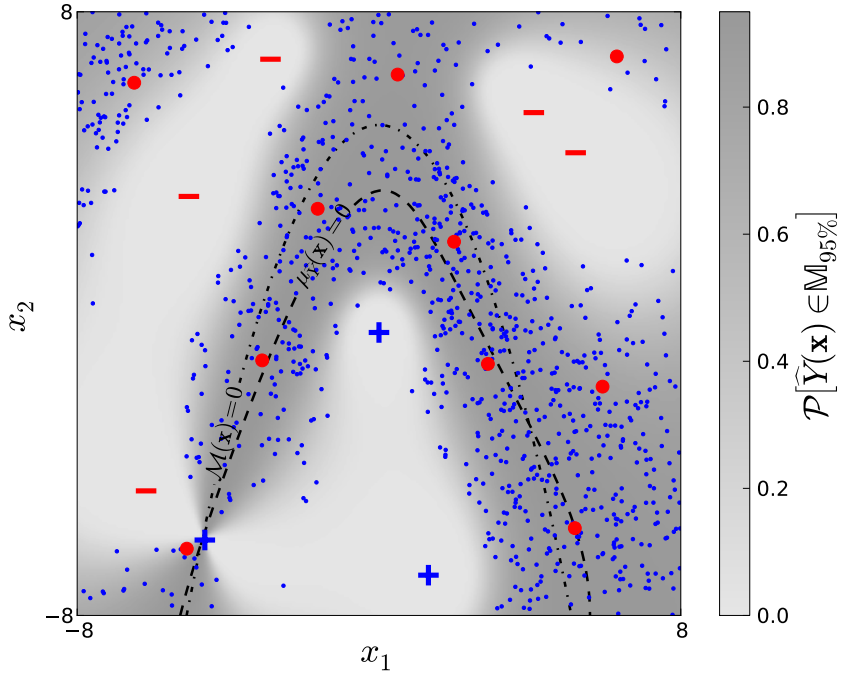


Figure 2.7: Cluster centres obtained by K-means clustering ($K = 10$) of the candidate population distributed according to the weighted margin probability function defined in Eq. (2.32).

2.4.2.3 Implementation

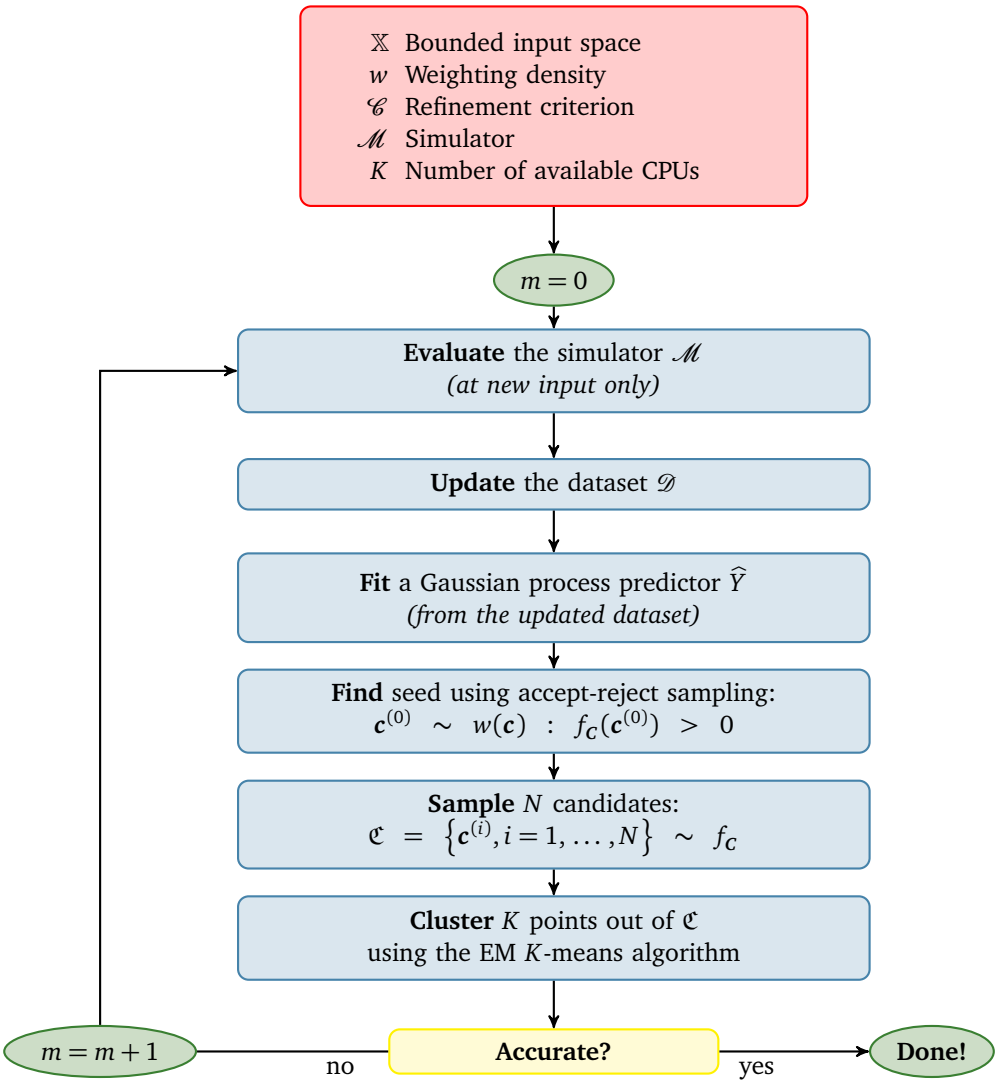
The overall sampling-based refinement procedure proposed in this manuscript is summarized in Algorithm 2.3. It requires the choice of a refinement criterion $\mathscr{C}$ together with a weighting PDF w and assumes that one can afford evaluating K new points in parallel within a finite number of runs $m_{\max}$. The procedure interestingly extends to the initialization from an empty DOE by considering that the criterion $\mathscr{C}$ is proportional to one everywhere in the bounded input space $\mathbb{X}$ for this first iteration. Accuracy criteria depends on the context and will be discussed later in this manuscript along with the application of this procedure. Recall that if the seed becomes hard to find, then it certainly means that the approximate contour $\widehat{\mathscr{S}_t}$ is sufficiently accurate.

2.5 Illustration

In this section, the proposed adaptive refinement strategy is illustrated on a structural reliability example from Waarts (2000). The limit-state function is defined as follows:

$$\mathfrak{g}(\boldsymbol{x}) = \min \begin{pmatrix} 3 + (x_1 - x_2)^2/10 - (x_1 + x_2)/\sqrt{2} \\ 3 + (x_1 - x_2)^2/10 + (x_1 + x_2)/\sqrt{2} \\ x_1 - x_2 + 7/\sqrt{2} \\ x_2 - x_1 + 7/\sqrt{2} \end{pmatrix}, \quad \boldsymbol{x} \in \mathbb{X}, \tag{2.54}$$

Algorithm 2.3 Proposed adaptive refinement strategy



and the associated limit-state (*i.e.* the contour of interest) is defined as $\mathscr{S}_0 = \{\boldsymbol{x} \in \mathbb{X} : g(\boldsymbol{x}) = 0\}$. It is illustrated in Figures 2.8 and 2.9 as the dotted black line. The input space $\mathbb{X} = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 \leq 8\}$ is the disk of radius $R = 8$ centred at the origin. This contour features 4 sharp corners which are difficult to approximate from a single space-filling DOE, so that it deserves the illustration of the proposed refinement strategy.

The Gaussian process predictor is fitted using maximum likelihood estimation at each refinement iteration assuming a constant regression trend and a squared exponential autocorrelation function. The refinement criterion $\mathscr{C}$ is chosen as the margin probability function in Eq. (2.32) as for all the other applications presented in this manuscript. The refinement procedure was ran two times assuming different weighting PDFs w. The first run whose iterations are illustrated in Figure 2.8 uses the input space indicator function in Eq. (2.50), *i.e.* $w(\boldsymbol{c}) \propto \mathbb{1}_{\|\boldsymbol{c}\|_2 \leq 8}(\boldsymbol{c})$.

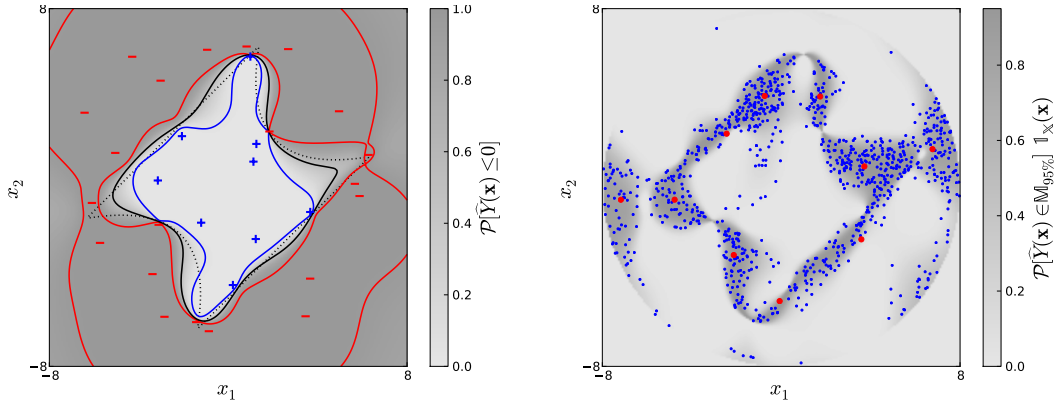
The second run, whose iterations are illustrated in Figure 2.9, uses the standard bivariate normal PDF (*i.e.* $w(\boldsymbol{c}) \propto e^{-\|\boldsymbol{c}\|_2^2/2}$) as appropriate in an uncertainty quantification context. In both cases K is set equal to 10.

The second run results in a more conservative experimental design which is dense near the origin in the first iterations and which tend to explore further in the "tails" of the weighting normal PDF along with the increasing confidence provided by the refined Kriging approximation. However it will hardly reach the four sharp corners because the standard bivariate normal PDF vanishes in this area assuming that a classification error in this area is less significant than a classification error closer to the origin. It is argued that in an uncertainty quantification context where the PDF $f_{\boldsymbol{X}}$ of the random input $\boldsymbol{X}$ is known, $f_{\boldsymbol{X}}$ is a more relevant choice for the weighting PDF as it ensures conservative predictions scaled with the relative importance of the input realizations $\boldsymbol{x}$.

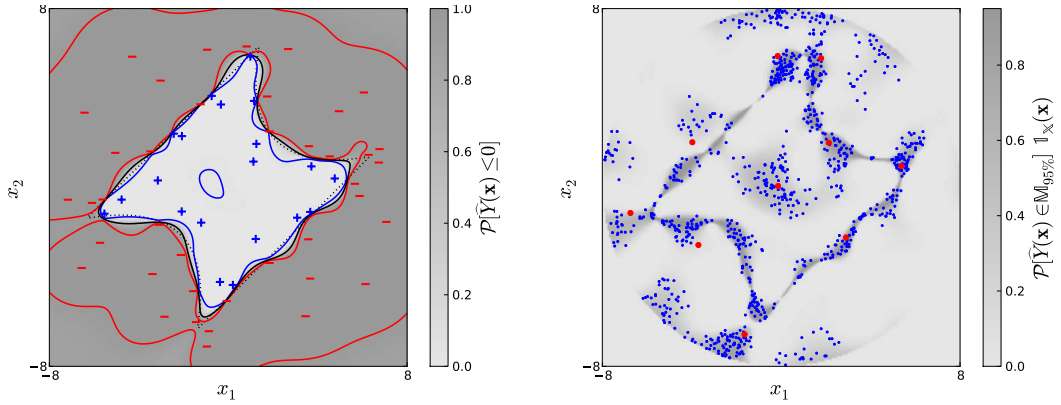
2.6 Conclusion

This chapter has reviewed state-of-the-art techniques to build experimental designs for computational models. A first part was devoted to space-filling DOEs whose common purpose is to initialize the refinement procedure from an empty dataset. In particular, a heuristic hypersphere-filling DOE was presented and will be widely used in the application presented in this manuscript.

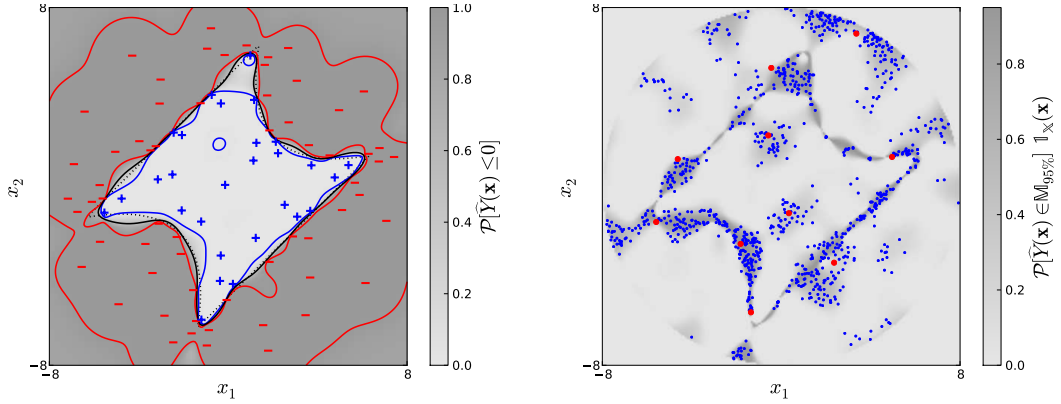
The other two sections focused on the iterative refinement of an initial Kriging approximation for two different purposes, namely: global optimization and contour approximation. Starting from the premise that most of these heuristics involve an ill-posed optimization problem, a sampling-based alternative is proposed. It basically consists in interpreting the usual refinement criteria as a PDF for the location of the points that will significantly improve the Kriging approximation. A MCMC algorithm is used to sample a large candidate population from these refinement criteria which is then reduced to a smaller one using K-means clustering for the sake of parsimony. The K cluster centres obtained with this approach are expected to uniformly span the vicinity of the modes of the refinement criteria. The number of cluster centres K is set equal to the number of available CPUs to run the simulator $\mathscr{M}$ in parallel. Note that one may also pick the best point in the large candidate population if the simulator can only be run sequentially.



(a) Iteration #3 – 3×10 points in the DOE

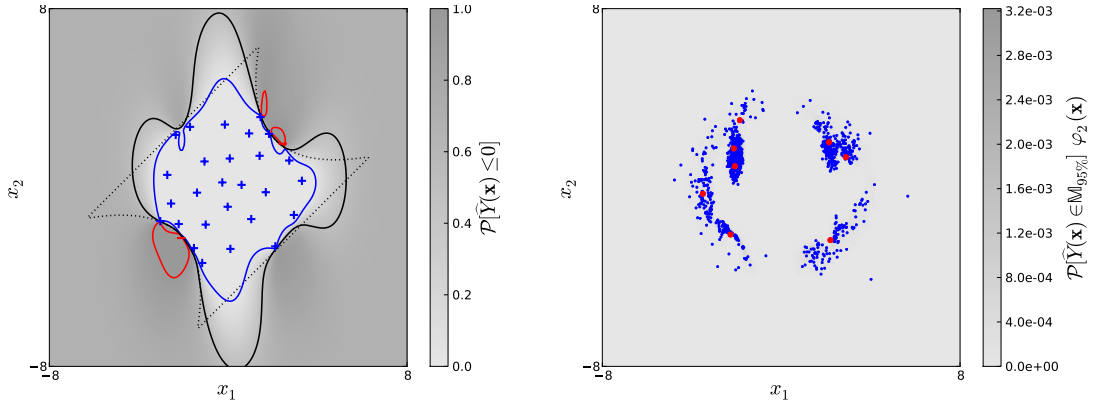


(b) Iteration #6 – 6×10 points in the DOE

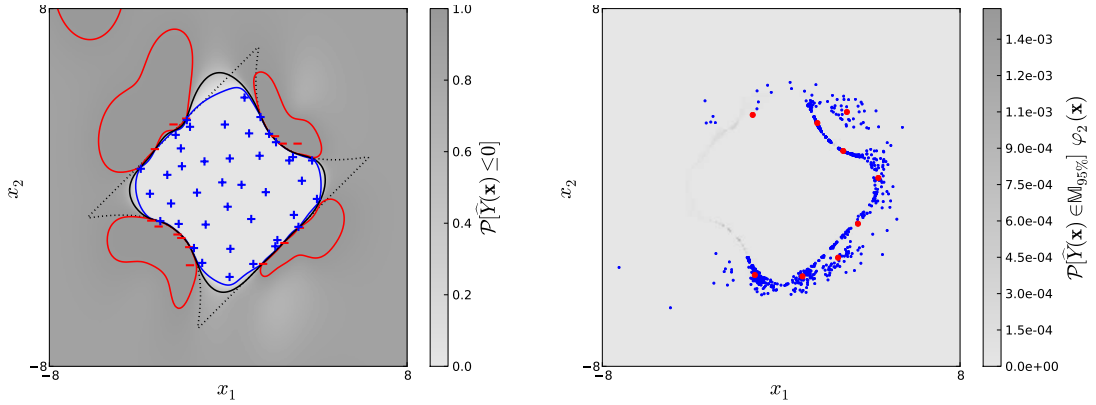


(c) Iteration #9 – 9×10 points in the DOE

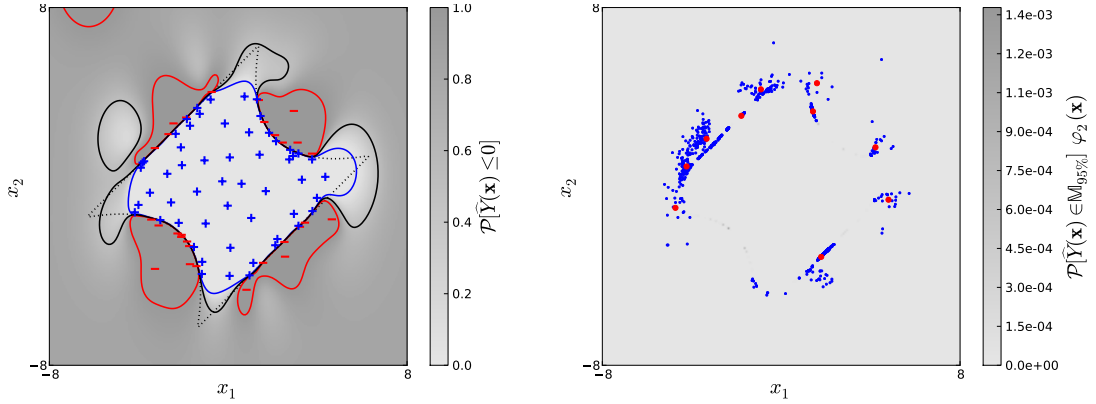
Figure 2.8: A few iterations of the refinement procedure applied to the limit-state surface defined in Eq. (2.54) using a uniform weighting PDF on the disk of radius $R = 8$. The left-hand panels illustrate the current approximation as in Figure 1.16 whereas the right hand-panel depicts the cluster centres (red dots) of the candidate population (blue dots) that will be evaluated at the next iteration.



(a) Iteration #3 – 3×10 points in the DOE



(b) Iteration #6 – 6×10 points in the DOE



(c) Iteration #9 – 9×10 points in the DOE

Figure 2.9: A few iterations of the refinement procedure applied to the limit-state surface defined in Eq. (2.54) using the standard multivariate normal as weighting PDF. The left-hand panels illustrate the current approximation as in Figure 1.16 whereas the right hand-panel depicts the cluster centres (red dots) of the candidate population (blue dots) that will be evaluated at the next iteration.

3

Reliability analysis

Contents

3.1 Introduction

Uncertainties arise in many engineering disciplines thus giving much importance to their *quantification* for the sake of safety and/or optimal performance. The sources of uncertainty in a physical system are multiple but they are often classified into two basic categories.

First, *aleatoric uncertainty* is associated with natural variability of a quantity at a large scale of observation. For instance, in fatigue design, the number of cycles before a component fails is random from a macroscopic point of view, although it could be (theoretically) deterministically predicted from a fine description of the part's micro-structure. Hence, at a large scale of observation, the components' properties can only be modelled in the form of probability distributions.

Second, *epistemic uncertainty* is an artificial source of uncertainty which is caused by a lack of knowledge and which is deliberately introduced as an assumption for computational reasons. For instance, most models arising in physical sciences rely on simplifying assumptions such as the discretization of a set of partial differential equations (*i.e.* the finite element method) or the choice for the boundary conditions. Such uncertainties are generally not considered though because they are hard to model in a probabilistic framework and negligible with respect to the other aleatoric uncertainties. However, other non-negligible sources of epistemic uncertainty will arise along with the forthcoming presentation of the so-called reliability methods. For example, in this manuscript, the uncertainty induced by the Gaussian process meta-modelling of an expensive-to-evaluate computational model (see Chapter 1) will be propagated down to the final quantity of interest.

From now on, it is assumed that a physical system (at least at its design stage) can be reduced to a *deterministic* model describing its performance. This *performance* is mathematically represented by a scalar function denoted by $\mathfrak{g}$. It takes n inputs grouped in the vector $\boldsymbol{x}$ which is assumed to be a realization of a random vector $\boldsymbol{X}$ with probability measure $\mathbb{P}(\mathrm{d}\boldsymbol{x}) = f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and support $\mathbb{X}$.

Recall that $\mathscr{P}$ denotes the probability measure associated with the meta-modelling epistemic *uncertainty, whereas $\mathbb{P}$ denotes the one corresponding to the* aleatoric *uncertainty in the random input $\boldsymbol{X}$.*

In this context, a reliability analysis aims at estimating the *probability of failure* of the system. Such a probability is conveniently defined as follows:

$$p_f \equiv \mathbb{P}\left[G \leq 0\right] = \mathbb{P}\left[\mathfrak{g}(\boldsymbol{X}) \leq 0\right], \tag{3.1}$$

where $G \equiv \mathfrak{g}(\boldsymbol{X})$ is the performance variable now random, and 0 is the *critical threshold* yielding failure. Thanks to the orientation of the inequality, this first definition implies that:

$$p_f = F_G(0), \tag{3.2}$$

where F_G is the cumulative distribution function of G.

Eventually, since F_G is generally unknown, the computational definition of the failure probability takes the form of the following integral:

$$p_f = \int_{\mathbb{F}=\{\boldsymbol{x}\in\mathbb{X}:\ \mathfrak{g}(\boldsymbol{x})\leq 0\}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \tag{3.3}$$

where $\mathbb{F} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) \leq 0\}$ is the *failure domain* and $\mathscr{S}_0 = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) = 0\}$ is the *limit-state surface*.

This chapter is concerned with the estimation of the failure probability in an industrial context. This means that the estimation should be possible within a reasonable computational time when both $\mathfrak{g}$ is expensive-to-evaluate and p_f has a low order of magnitude. Failure is indeed an undesired event, so that the failure probability is often required to be as low as possible, or at least lower than a prescribed acceptable level chosen by the stakeholders.

Section 3.2 introduces the basics of the *joint* probabilistic modelling of $\boldsymbol{X}$ that is adopted in this thesis. It also presents the so-called *isoprobabilistic transform*. In Section 3.3 the best established techniques for estimating such low failure probabilities are revisited from an importance sampling viewpoint. This original interpretation matters for the understanding of the contributions introduced in the next two sections. Section 3.4 examines the substitution of a meta-model for the performance function $\mathfrak{g}$ in order to speed up the estimation of the failure probability. An important part of this section is devoted to the *efficient* quantification of the *substitution error*. At last, an original reliability estimation technique is proposed in Section 3.5. The purpose of the latter contribution is to extend the use of meta-modelling techniques to reliability problems featuring a large number n of random variables.

3.2 Joint probabilistic modelling

3.2.1 The copula formalism

In this thesis, it is assumed that the probability distribution of $\boldsymbol{X}$ is completely known and described as follows. It is composed with n parametric marginal distributions, and a stochastic dependence structure expressed in the form of a *copula*.

The formal definition of a copula C is given in the book by Nelsen (1999). It is not given in this thesis for the reason that it does not provide any insight about what it is used for. Instead, it is informally defined as *what remains of a joint probability distribution once the effect of the marginal distributions has been removed*. In other words, it is a probability distribution defined on the n following uniform margins:

$$u_i = F_{X_i}(x_i), \quad i = 1, \ldots, n. \tag{3.4}$$

This less formal definition becomes clearer from the application of the following theorem due to Sklar (1959).

Theorem 3.2.1. ***Sklar's theorem*** *(see* e.g. *Nelsen, 1999, p. 41)*
Let $F_{\boldsymbol{X}}$ be a joint CDF with margins $\{F_{X_i},\, i = 1, \dots, n\}$. Then there exists a copula C such that:

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = C\left(F_{X_1}(x_1), \dots, F_{X_n}(x_n)\right). \tag{3.5}$$

In addition, if the margins are continuous, then C is unique.

Corollary 3.2.1. *(see* e.g. *Nelsen, 1999, p. 41)*
Let $F_{\boldsymbol{X}}$, $\{F_{X_i},\, i = 1, \dots, n\}$ and C be as in Theorem 3.2.1, and let $\{F_{X_i}^{-1},\, i = 1, \dots, n\}$ be the inverse CDFs of the n margins. Then for any $\boldsymbol{u} \in [0;\, 1]^n$:

$$C(u_1, \dots, u_n) = F_{\boldsymbol{X}}\left(F_{X_1}^{-1}(u_1), \dots, F_{X_n}^{-1}(u_n)\right). \tag{3.6}$$

This theorem justifies the use of copulas to specify a joint probability distribution. Its corollary provides a first mean to construct copulas from the joint and marginal CDFs of a given probability distribution. In addition, in the present case of continuous margin distributions, this theorem enables the explicit computation of the joint PDF $f_{\boldsymbol{X}}$ which is as follows:

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}) &= \frac{\partial^n F_{\boldsymbol{X}}(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n} \\ &= \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \prod_{i=1}^{n} \frac{\partial F_{X_i}(x_i)}{\partial x_i} \\ &= c\left(F_1(x_1), \dots, F_n(x_n)\right) \prod_{i=1}^{n} f_{X_i}(x_i), \end{aligned} \tag{3.7}$$

where the density of the copula c and the PDFs of the margins $\{f_{X_i},\, i = 1, \dots, n\}$ have been introduced.

In particular, it can be seen from Eq. (3.5) (resp. Eq. (3.7)) that this formalism extends to the classical case of independence through the use of the so-called *independent copula* whose CDF (resp. PDF) reads:

$$C_{\mathrm{Id}}(u_1, \dots, u_n) = \prod_{i=1}^{n} u_i, \quad \boldsymbol{u} \in [0;\, 1]^n, \tag{3.8}$$

$$c_{\mathrm{Id}}(u_1, \dots, u_n) = 1, \quad \boldsymbol{u} \in [0;\, 1]^n. \tag{3.9}$$

3.2.2 The Nataf distribution and the normal copula

There exists a large variety of copulas dictating various types of dependence, although only the normal copula is considered in this manuscript. The reader is referred to the thesis by Charpentier (2006) for a more comprehensive review of joint stochastic modelling using copulas (from the theoretical background to the practical problem of inference from data).

The so-called *Nataf distribution* (1962) is widely used in the structural reliability literature due to the pioneering article by Liu and Der Kiureghian (1986). The reason why it

became popular in this field (at least in the two decades following this article) is that it was the first proposition in the reliability engineering literature to define a joint probabilistic model in the form of a collection of marginal distributions and a correlation structure expressed in the form of a (Pearson) correlation matrix. This particular joint distribution was recently reviewed and criticized by Lebrun and Dutfoy (2009b,a,c). In particular Lebrun and Dutfoy warn the probabilistic model builders about the assumptions underlying the choice for the Nataf distribution by pointing out that it implicitly assumes a normal copula which is not all-purpose. The Nataf distribution is now reviewed from the (normal) copula viewpoint as a particular and convenient means to describe a joint probabilistic model assuming that the normal copula fits to the description of the input $\boldsymbol{X}$.

The *normal copula* is defined from the *standard multivariate Gaussian distribution* by means of Corollary 3.2.1. Indeed, its copula reads:

$$C_\Phi(u_1, \dots, u_n;\, \mathbf{R}_0) = \Phi_n\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n);\, \mathbf{R}_0\right), \quad \boldsymbol{u} \in [0;\, 1]^n, \tag{3.10}$$

where Φ_n is the standard multivariate Gaussian CDF:

$$\Phi_n(x_1, \dots, x_n) = \int_{[-\infty;\, x_1]\times\ldots\times[-\infty;\, x_n]} (2\,\pi)^{-n/2} \exp\left[-\frac{1}{2}\,\boldsymbol{t}^{\mathsf{T}}\,\mathbf{R}_0^{-1}\,\boldsymbol{t}\right] \mathrm{d}\boldsymbol{t}, \tag{3.11}$$

and Φ^{-1} is the inverse of the standard univariate Gaussian CDF (see Eq. (1.20)). Both Φ_n and Φ^{-1} and therefore C_Φ can only be approximated numerically (see *e.g.* Genz, 1992, 1993, for the numerical evaluation of Φ_n).

Moreover, the PDF associated with this copula can be elaborated from Eq. (3.10) as follows:

$$\begin{aligned} c_\Phi(u_1, \dots, u_n;\, \mathbf{R}_0) &= \frac{\partial^n \Phi_n\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n);\, \mathbf{R}_0\right)}{\partial u_1 \dots \partial u_n} \\ &= \left.\frac{\partial^n \Phi_n\left(z_1, \dots, z_n;\, \mathbf{R}_0\right)}{\partial z_1 \dots \partial z_n}\right|_{z_i = \Phi^{-1}(u_i)} \prod_{i=1}^{n} \frac{\partial \Phi^{-1}(u_i)}{\partial u_i} \\ &= \frac{\varphi_n\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n);\, \mathbf{R}_0\right)}{\prod_{i=1}^{n} \varphi\left(\Phi^{-1}(u_i)\right)}. \end{aligned} \tag{3.12}$$

This first definition further simplifies by introducing $\boldsymbol{z} = (\Phi^{-1}(u_i),\, i = 1, \dots, n)^{\mathsf{T}}$ and by replacing the multivariate and univariate Gaussian PDFs φ_n and φ with their respective expressions (see Eq. (A.6) and (1.19)):

$$\begin{aligned} c_\Phi(u_1, \dots, u_n;\, \mathbf{R}_0) &= \frac{(2\,\pi)^{-n/2}\,\left[\det \mathbf{R}_0\right]^{-1/2} \exp\left[-\frac{1}{2}\,\boldsymbol{z}^{\mathsf{T}}\,\mathbf{R}_0^{-1}\,\boldsymbol{z}\right]}{(2\,\pi)^{-n/2} \exp\left[-\frac{1}{2}\,\boldsymbol{z}^{\mathsf{T}}\,\boldsymbol{z}\right]} \\ &= \left[\det \mathbf{R}_0\right]^{-1/2} \exp\left[-\frac{1}{2}\,\boldsymbol{z}^{\mathsf{T}}\left(\mathbf{R}_0^{-1} - \mathbf{I}\right)\boldsymbol{z}\right]. \end{aligned} \tag{3.13}$$

It can be seen from Eq. (3.10) that the copula C_Φ is simply defined by a symmetric positive definite matrix $\mathbf{R}_0 = [\rho_{0\,ij}, i, j = 1, \dots, n]$. Even though the off-diagonal terms in

this matrix are comprised in $]-1;\ 1[$ and its diagonal terms are equal to 1, it shall not be confused with the more usual linear correlation matrix $\mathbf{R} = [\rho_{ij}, i,\ j = 1, \ldots, n]$. It should instead be remembered that their terms are related by the following integral:

$$\rho_{ij} = \frac{1}{\sigma_i\,\sigma_j} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(F_{X_i}^{-1}\left(\Phi(z_i)\right) - \mu_{X_i}\right) \left(F_{X_j}^{-1}\left(\Phi(z_j)\right) - \mu_{X_j}\right) \times \ldots \\ \ldots \times \varphi_2(z_i, z_j;\ \rho_{0ij}) \mathrm{d}z_i\,\mathrm{d}z_j, \quad i,\ j = 1, \ldots, n, \tag{3.14}$$

provided it exists. Indeed, Bourinet and Lemaire (2008) and Lebrun and Dutfoy (2009a) pointed out that a high Pearson correlation coefficient is incompatible with highly heteroscedastic marginal distributions (*i.e.* if $\sigma_1 = 1$ and $\sigma_2 \to \infty$ then the admissible range for ρ_{12} vanishes to a short interval around 0). However, when the Pearson correlation coefficient is reasonably low (in magnitude), the integral in Eq. (3.14) can be inverted numerically using Gaussian quadrature and an optimization algorithm.

More importantly, Lebrun and Dutfoy (2009a) also point out that the Spearman (or rank) correlation coefficient is better suited to parametrize a copula because it leads to a simpler closed-form expression for ρ_{ij}. This simpler expression is due to the fact that Spearman's correlation coefficient is independent of the marginal distributions[1]. More specifically, denoting by $\mathbf{R}_S = [\rho_{Sij},\ i,\ j = 1, \ldots, n]$ the given symmetric positive definite matrix of Spearman correlation coefficients, Lebrun and Dutfoy prove that:

$$\rho_{0ij} = 2 \sin\left(\frac{\pi}{6}\rho_{Sij}\right), \quad i,\ j = 1, \ldots, n, \tag{3.15}$$

This expression is by far simpler than Eq. (3.14). Hence, it is argued that Spearman's correlation coefficient should be preferred to Pearson's when building a probabilistic model from experimental data.

Eventually, note that the *independent copula* is obtained from the *normal copula* when the parameters matrix $\mathbf{R}_0$ reduces to the identity matrix $\mathbf{I}$, indeed:

$$C_\Phi(u_1, \ldots, u_n;\ \mathbf{I}) = \prod_{i=1}^{n} \Phi(\Phi^{-1}(u_i)) = \prod_{i=1}^{n} u_i, \quad \boldsymbol{u} \in [0;\ 1]^n, \tag{3.16}$$

$$c_\Phi(u_1, \ldots, u_n;\ \mathbf{I}) = \frac{1}{[\det \mathbf{I}]^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{z}^{\mathsf{T}}\left(\mathbf{I}^{-1} - \mathbf{I}\right)\boldsymbol{z}\right] = 1, \quad \boldsymbol{u} \in [0;\ 1]^n. \tag{3.17}$$

3.2.3 Isoprobabilistic transforms

The transform T exposed in this section is a *one-to-one differentiable* and *bijective* mapping (*i.e.* a diffeomorphism), such that the probability distribution of $\boldsymbol{U} = T(\boldsymbol{X})$ is invariant by rotation. It was named *isoprobabilistic* transform because it also conserves probabilities:

$$F_X(\boldsymbol{x}) = F_U(\boldsymbol{u}). \tag{3.18}$$

[1] It is defined as the usual linear correlation coefficient of the transformed variables $\{F_{X_i}(X_i),\ i = 1, \ldots, n\}$.

Its use is widely spread in the structural reliability literature. The historical motivation for this transform is that it allows tractable analytical approximations of the failure probability when the first- or second- order reliability method is used (see Section 3.3.4.2). In addition, it then revealed a fundamental prior step to many methods arising in the field of uncertainty quantification. For instance, this transform eases the choice of the polynomial basis in a polynomial chaos expansion (see *e.g.* Berveiller et al., 2006) and it makes the tuning of the reliability methods evoked in this chapter robust to the possible heterogeneity of scale in the components of $\boldsymbol{X}$.

3.2.3.1 The Nataf transform

In particular, the Nataf transform maps any random vector $\boldsymbol{X}$ having a Nataf distribution (*e.g.* a normal copula) into a standard Gaussian random vector $\boldsymbol{U}$ (*i.e.* a random vector whose components are independent and normally distributed with zero mean and unit variance). It was first introduced in the structural reliability literature by Liu and Der Kiureghian (1986) and it is explained in reference books (see *e.g.* Ditlevsen and Madsen, 1996; Lemaire, 2009).

It is defined as the following composed application:

$$T : \left| \begin{array}{ccc} \mathbb{X} & \mapsto & \mathbb{R}^n \\ \boldsymbol{x} & \to & \boldsymbol{u} = (T_3 \circ T_2 \circ T_1)(\boldsymbol{x}) \end{array} \right. , \tag{3.19}$$

where each sub-transform is defined as follows:

T_1 the margin variables are first mapped into uniform margins using their respective CDFs:

$$T_1 : \left| \begin{array}{ccc} \mathbb{X} & \mapsto & [0;\, 1]^n \\ \boldsymbol{x} & \to & \boldsymbol{w} = \left(F_{X_i}(x_i),\, i = 1, \ldots, n\right)^{\mathsf{T}} \end{array} \right. , \tag{3.20}$$

T_2 these uniform margins are then mapped into correlated standard Gaussian variables using the inverse CDF of the Gaussian distribution:

$$T_2 : \left| \begin{array}{ccc} [0;\, 1]^n & \mapsto & \mathbb{R}^n \\ \boldsymbol{w} & \to & \boldsymbol{v} = \left(\Phi^{-1}(w_i),\, i = 1, \ldots, n\right)^{\mathsf{T}} \end{array} \right. , \tag{3.21}$$

T_3 eventually, the correlated standard Gaussian variables are "de-correlated" through the following linear transform:

$$T_3 : \left| \begin{array}{ccc} [0;\, 1]^n & \mapsto & \mathbb{R}^n \\ \boldsymbol{v} & \to & \boldsymbol{u} = \mathbf{L}_0^{-1}\, \boldsymbol{v} \end{array} \right. , \tag{3.22}$$

where $\mathbf{L}_0$ is the lower Cholesky decomposition of the matrix $\mathbf{R}_0$ that parametrizes the normal copula (*i.e.* such that $\mathbf{R}_0 = \mathbf{L}_0\, \mathbf{L}_0{}^{\mathsf{T}}$).

Since T is a one-to-one differentiable mapping, its Jacobian exists and is computed by differentiating the composed application in Eq. (3.19):

$$\mathbf{J}_{\boldsymbol{u},\boldsymbol{x}} = \left[\frac{\partial u_i}{\partial x_j},\, i,\, j = 1, \ldots, n \right] = \mathbf{L}_0^{-1}\, \mathbf{J}_{\boldsymbol{v},\boldsymbol{x}} \tag{3.23}$$

where $\mathbf{J}_{\boldsymbol{v},\boldsymbol{x}}$ is a diagonal matrix whose terms are defined as follows:

$$J_{v_i,x_i} = \frac{\partial \Phi^{-1}\left(F_{X_i}(x_i)\right)}{\partial x_i} = \frac{f_{X_i}(x_i)}{\varphi(v_i)}, \quad i = 1, \ldots, n, \tag{3.24}$$

where $v_i = \Phi^{-1}(F_{X_i}(x_i))$. The latter equation further simplifies depending on the marginal distributions.

Since T is a one-to-one bijective mapping, the Jacobian of the inverse mapping T^{-1} exists and is simply defined as follows:

$$\mathbf{J}_{\boldsymbol{x},\boldsymbol{u}} = \mathbf{J}_{\boldsymbol{u},\boldsymbol{x}}^{-1}. \tag{3.25}$$

Moreover in the sequel, the performance function in the standard space is denoted by $\mathfrak{g}^\circ$ and is defined as the following composed application:

$$\mathfrak{g}^\circ \equiv (\mathfrak{g} \circ T^{-1}) : \left| \begin{array}{ccc} \mathbb{R}^n & \mapsto & \mathbb{Y} \\ \boldsymbol{u} & \to & \mathfrak{g}\left(T^{-1}(\boldsymbol{u})\right) \end{array} \right. , \tag{3.26}$$

where $\mathbb{Y} \subseteq \mathbb{R}$ is the original output space of $\mathfrak{g}$. Its gradient is related with that of $\mathfrak{g}$ as the gradient of a composed application and it reads:

$$\boldsymbol{\nabla}_{\boldsymbol{u}} \mathfrak{g}^\circ(\boldsymbol{u}) = \mathbf{J}_{\boldsymbol{x},\boldsymbol{u}}^{\mathsf{T}}\, \boldsymbol{\nabla}_{\boldsymbol{x}} \mathfrak{g}(\boldsymbol{x}) = \mathbf{J}_{\boldsymbol{u},\boldsymbol{x}}^{-1\,\mathsf{T}}\, \boldsymbol{\nabla}_{\boldsymbol{x}} \mathfrak{g}(\boldsymbol{x}). \tag{3.27}$$

The transformed failure domain is denoted by $\mathbb{F}^\circ \equiv \{\boldsymbol{u} \in \mathbb{R}^n : \mathfrak{g}^\circ(\boldsymbol{u}) \leq 0\}$, and the transformed limit-state by $\mathscr{S}_0^\circ \equiv \{\boldsymbol{u} \in \mathbb{R}^n : \mathfrak{g}^\circ(\boldsymbol{u}) = 0\}$.

3.2.3.2 The Rosenblatt transform

This alternative transform can be applied to a larger class of multivariate probability distributions for mapping them into a standard Gaussian random vector. It basically consists in applying a one-dimensional transform to each component of the random vector $\boldsymbol{X}$ recursively. Namely, the Rosenblatt transform is defined as follows:

$$T_{\text{Rosenblatt}} : \left| \begin{array}{ccc} \mathbb{X} & \mapsto & \mathbb{R}^n \\ \boldsymbol{x} & \to & \boldsymbol{u} = \left\{ \begin{array}{c} \Phi^{-1}\left(F_{X_1}(x_1)\right) \\ \Phi^{-1}\left(F_{X_2|X_1}(x_2)\right) \\ \vdots \\ \Phi^{-1}\left(F_{X_n|X_1,\ldots,X_n}(x_n)\right) \end{array} \right\} \end{array} \right. , \tag{3.28}$$

This transform is more computationally involved than the Nataf transform as it requires the evaluation of conditional CDFs of $\boldsymbol{X}$. Nevertheless, when the joint distribution is defined as a collection of margin distributions and a copula (possibly composed), this transform is computationally tractable (Lebrun and Dutfoy, 2009b). Eventually, note that despite this transform preserves the probability measure $\mathbb{P}$, it is not unique as it depends on the conditioning order. A direct consequence of this non-unicity is that the (exact) failure probability in Eq. (3.3) is left unchanged although some of its (geometrical) approximations reviewed in the sequel may change.

3.2.4 Illustration: capacity and demand

Let us consider the following basic reliability problem along with the methods reviewed in this chapter. The performance function reads:

$$g(r, s) = r - s, \tag{3.29}$$

where r is the outcome of the random *capacity* R, and s is the outcome of the random *demand* S. Note that in this context, $G \equiv g(R, S)$ is referred to as the *safety margin* for obvious reasons. Capacity and demand are grouped in the input random vector $\boldsymbol{X} = (R, S)^{\mathsf{T}}$ and a Nataf joint distribution is assumed. It is parametrized as follows.

- R is lognormally distributed with mean $\mu_R = 7$ and standard deviation $\sigma_R = 0.5$;
- S is lognormally distributed with mean $\mu_S = 1$ and standard deviation $\sigma_S = 0.5$;
- the underlying normal copula is parametrized with a Pearson's correlation matrix $\mathbf{R}$ whose off-diagonal term is set equal to $\rho = 0.5$.

The normal copula parameter defined in Eq. (3.14) can be computed analytically in this specific case (Liu and Der Kiureghian, 1986):

$$\rho_0 = \frac{\log\left(1 + \rho\, \delta_R\, \delta_S\right)}{\sqrt{\log\left(1 + \rho\, \delta_R^2\right) \log\left(1 + \rho\, \delta_S^2\right)}} \tag{3.30}$$

where $\delta_\bullet = \sigma_\bullet / \mu_\bullet$ are the coefficients of variation. The latter formula leads to $\rho_0 \approx 0.525$. The contours of the PDF of $\boldsymbol{X}$, the limit-state surface as well as the failure domain are illustrated in Figure 3.1.

The Nataf transform and its inverse also come analytically in this simple case, so that the transformed performance function $\mathfrak{g}^\circ$ has a closed-form expression. Actually, the three steps of the inverse Nataf transform are derived as follows.

T_3^{-1} The first step consists in "re-correlating" the standard Gaussian random vector $\boldsymbol{u}$. The normal copula parameters matrix admits the following Cholesky decomposition:

$$\mathbf{R}_0 = \begin{bmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ \rho_0 & \sqrt{1-\rho_0^2} \end{bmatrix}}_{\mathbf{L}_0} \underbrace{\begin{bmatrix} 1 & \rho_0 \\ 0 & \sqrt{1-\rho_0^2} \end{bmatrix}}_{\mathbf{L}_0^{\mathsf{T}}}. \tag{3.31}$$

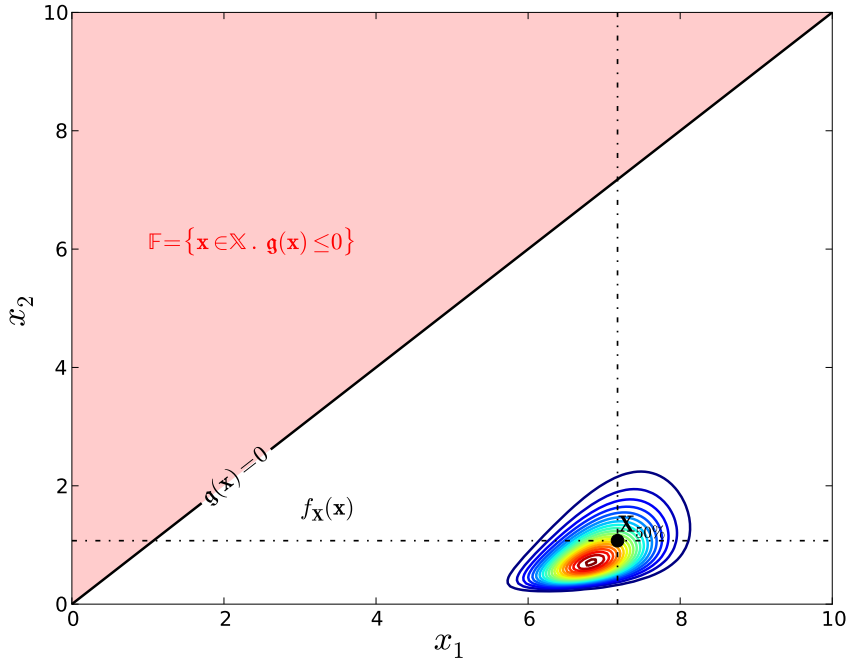


Figure 3.1: Physical space spanned by $\boldsymbol{X}$ for the basic structural reliability problem defined in Section 3.2.4.

Hence,

$$\boldsymbol{v} = T_3^{-1}(\boldsymbol{u}) = \mathbf{L}_0\,\boldsymbol{u} = \begin{pmatrix} u_R \\ \rho_0\,u_R + \sqrt{1-\rho_0^2}\,u_S \end{pmatrix}. \tag{3.32}$$

T_2^{-1} Then, the correlated standard Gaussian random variates are mapped into uniform margins by means of the standard Gaussian CDF:

$$\boldsymbol{w} = T_2^{-1}(\boldsymbol{v}) = \left(\Phi(v_R),\,\Phi(v_S)\right)^{\mathsf{T}} = \begin{pmatrix} \Phi\left(u_R\right) \\ \Phi\left(\rho_0\,u_R + \sqrt{1-\rho_0^2}\,u_S\right) \end{pmatrix}. \tag{3.33}$$

T_1^{-1} Finally, recall that the lognormal distribution is parametrized with its scale and location parameters respectively defined as follows:

$$\zeta_\bullet = \sqrt{\log\left(1+\delta_\bullet^2\right)}, \tag{3.34}$$

$$\lambda_\bullet = \log(\mu_\bullet) - \frac{1}{2}\,\zeta_\bullet^2, \tag{3.35}$$

where $\delta_\bullet = \sigma_\bullet/\mu_\bullet$ is the coefficient of variation. Its inverse CDF reads:

$$F_\bullet^{-1}(p) = \exp\left(\lambda_\bullet + \zeta_\bullet\,\Phi^{-1}(p)\right), \quad i = 1,\,2. \tag{3.36}$$

Hence, the physical vector reads as follows:

$$\begin{aligned} \boldsymbol{x} &= T_1^{-1}(\boldsymbol{w}) = \left(F_R^{-1}(w_R),\,F_S^{-1}(w_S)\right)^{\mathsf{T}} \\ &= \begin{pmatrix} \exp\left(\lambda_R + \zeta_R\,u_R\right) \\ \exp\left(\lambda_S + \zeta_S\left(\rho_0\,u_R + \sqrt{1-\rho_0^2}\,u_S\right)\right) \end{pmatrix}. \end{aligned} \tag{3.37}$$

This eventually yields the expression of the performance function in the standard space:

$$\mathfrak{g}^\circ(\boldsymbol{u}) = \exp\left(\lambda_R + \zeta_R\, u_R\right) - \exp\left(\lambda_S + \zeta_S\, \left(\rho_0\, u_R + \sqrt{1-\rho_0^2}\, u_S\right)\right). \tag{3.38}$$

Note that the transformed performance function is not linear anymore although the transformed limit-state surface $\mathscr{S}_0^\circ$ is still an hyperplane in the standard space for this very particular example. This can be proved as follows:

$$\begin{aligned}
\mathscr{S}_0^\circ &= \left\{\boldsymbol{u} \in \mathbb{R}^2 : \mathfrak{g}^\circ(\boldsymbol{u}) = 0\right\} \\
&= \left\{\boldsymbol{u} \in \mathbb{R}^2 : 1 - \frac{\exp\left(\lambda_S + \zeta_S\, \left(\rho_0\, u_R + \sqrt{1-\rho_0^2}\, u_S\right)\right)}{\exp\left(\lambda_R + \zeta_R\, u_R\right)} = 0\right\} \\
&= \left\{\boldsymbol{u} \in \mathbb{R}^2 : \left(\zeta_R - \rho_0\, \zeta_S\right) u_R - \zeta_S \sqrt{1-\rho_0^2}\, u_S + \left(\lambda_R - \lambda_S\right) = 0\right\}.
\end{aligned} \tag{3.39}$$

The contours of the standard Gaussian PDF of $\boldsymbol{U}$, the transformed limit-state surface as well as the transformed failure domain are illustrated in Figure 3.2. Note that the origin of the so-called *standard space* spanned by $\boldsymbol{U}$ is the image of the median of the physical distribution (denoted by $\boldsymbol{X}_{50\%}$ in Figure 3.1) through the isoprobabilistic transform T.

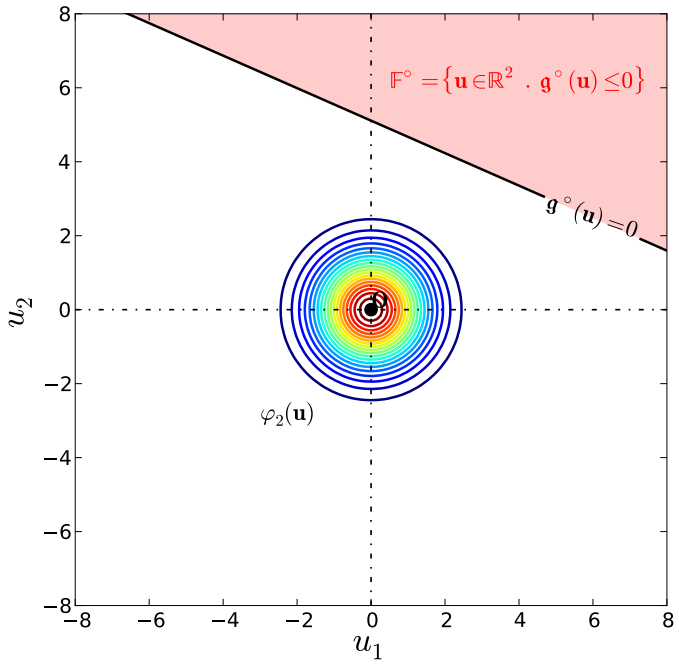


Figure 3.2: Standard space spanned by $\boldsymbol{U}$ for the basic structural reliability problem defined in Section 3.2.4.

The probability of failure $p_f = \mathbb{P}[R \leq S]$ associated with this simple example will be computed with the reliability methods reviewed in this chapter for the sake of illustration.

3.3 State-of-the-art reliability methods

This section reviews some state-of-the-art methods for estimating the failure probability defined in Eq. (3.3). Monte Carlo sampling is first reviewed as a motivation for introducing more advanced techniques. Then specific emphasis is put on the central role of the so-called *importance sampling* technique. A more conventional presentation of these methods can be found in the reference books by Ditlevsen and Madsen (1996) and Lemaire (2009). The example of Section 3.2.4 will be used for illustration purposes.

3.3.1 Monte Carlo sampling

3.3.1.1 Principle

First, let us define the *failure indicator function* as follows:

$$\mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) = \begin{cases} 1 & \text{if} \quad \mathfrak{g}(\boldsymbol{x}) \leq 0 \\ 0 & \text{otherwise} \end{cases} . \tag{3.40}$$

By introducing this failure indicator function in Eq. (3.3), the failure probability can be recast as:

$$p_f = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_X(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \equiv \mathbb{E}\left[\mathbb{1}_{\mathbb{F}}(\boldsymbol{X})\right]. \tag{3.41}$$

Hence, this quantity can be *estimated* as follows:

$$\widehat{p}_{f\,\mathrm{MCS}} \equiv \widehat{\mathbb{E}}\left[\mathbb{1}_{\mathbb{F}}(\boldsymbol{X})\right] = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\mathbb{F}}(\boldsymbol{X}^{(i)}), \tag{3.42}$$

where $\{\boldsymbol{X}^{(i)}, i = 1, \ldots, N\}$ is a sample of N independent copies of the random vector $\boldsymbol{X}$.

This estimator is itself a random variable since it is defined as the sum of N independent and identically distributed Bernoulli random variables $\{\mathbb{1}_{\mathbb{F}}(\boldsymbol{X}^{(i)}), i = 1, \ldots, N\}$. According to the central limit theorem, provided N is sufficiently large, this estimator is unbiased and normally distributed meaning that:

$$\left[\widehat{p}_{f\,\mathrm{MCS}} - p_f\right] \underset{N\to\infty}{\hookrightarrow} \mathcal{N}_1(0, \sigma^2_{\mathrm{MCS}}). \tag{3.43}$$

Its variance is easily proved (see *e.g.* Lemaire, 2009, pp. 250–251) to read as follows:

$$\sigma^2_{\mathrm{MCS}} = \frac{1}{N} p_f \, (1 - p_f). \tag{3.44}$$

It can be seen from the above equation that this variance decays with the number of samples used for estimating the failure probability, meaning that the uncertainty in the estimation is purely epistemic (reducible). This is often referred to as the *statistical error*.

In order to assess the accuracy of a sampling run, when $p_f \leq 50\%$ one usually resorts to the *coefficient of variation* defined as follows:

$$\delta_{\text{MCS}} = \frac{\sigma_{\text{MCS}}}{p_f} = \sqrt{\frac{1-p_f}{N\,p_f}}. \tag{3.45}$$

The lower δ_{MCS} the more confidence in $\widehat{p}_{f\,\text{MCS}}$ with respect to the residual epistemic uncertainty in the estimation.

3.3.1.2 Limitations and properties

The fundamental problem of this simple estimator is that the required sample size N to ensure that a failure probability is estimated with a given coefficient of variation dramatically increases as the probability gets low. For instance, the estimation of a probability of 10^{-k} (with $k < 2$) and a 10% coefficient of variation requires about 10^{k+2} evaluations of the performance function. Note that if p_f is greater than 50% (close to 1), one should instead compute the coefficient of variation of the complementary failure probability $1 - p_f$ which exhibits the same property (the variance of $1 - p_f$ is indeed the same as p_f's in Eq. (3.44)). This premise motivates the forthcoming reliability methods.

Despite its computational cost, it is worth mentioning that Monte Carlo sampling is highly distributable (*i.e.* the N independent calls to the performance function $\mathfrak{g}$ can be run in parallel). It is also robust with respect to both the input probabilistic model (its size $n = \dim(\boldsymbol{X})$) and the performance function (its regularity). Indeed, Monte Carlo sampling simply requires the availability of a random generator for the random performance G.

Eventually, Monte Carlo sampling enables the computation of the full empirical CDF of the random performance G which might also be of interest. This computation uses the collected sample of $\mathfrak{g}$-values $\mathscr{G} = \{\mathfrak{g}(\boldsymbol{x}^{(i)}), i = 1, \dots, n\}$. Indeed, the empirical CDF is a staircase function whose usual definition closely resembles the Monte Carlo estimator in Eq. (3.42):

$$\widehat{F}_G(g) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\left(\mathfrak{g}\left(\boldsymbol{x}^{(i)}\right) \leq g\right). \tag{3.46}$$

Its estimation variance is defined by analogy with Eq. (3.44) and it reads as follows:

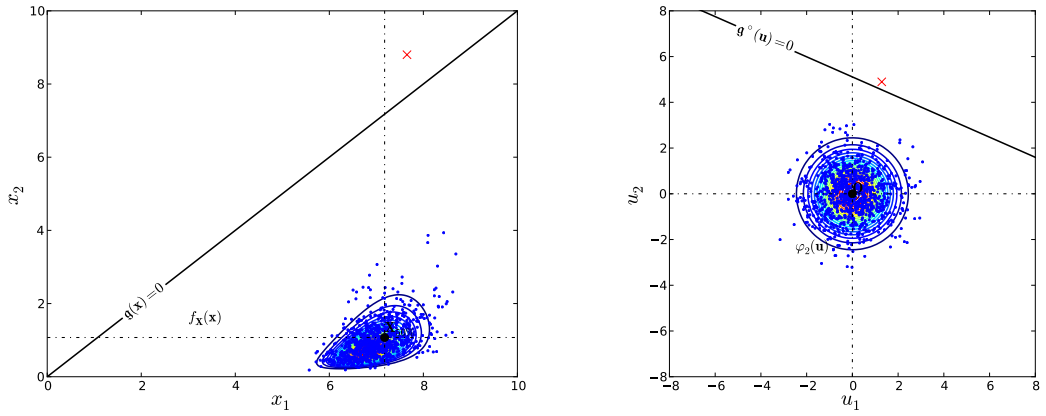
$$\sigma^2_{\widehat{F}_G}(g) = \frac{1}{N}\,\widehat{F}_G(g)\left(1 - \widehat{F}_G(g)\right). \tag{3.47}$$

This variance in turns enables a fair approximation of the 95% confidence interval bounds on the empirical CDF thanks to the central limit theorem, provided N is sufficiently large. These two bounds read as follows:

$$\widehat{F}^{\pm}_G(g) = \widehat{F}_G(g) \pm 1.96\sqrt{\sigma^2_{\widehat{F}_G}(g)}. \tag{3.48}$$

3.3.1.3 Illustration

Monte Carlo sampling is applied to the basic structural reliability problem introduced in Section 3.2.4. The computation is performed using a limited sample size set equal to 5×10^6. A sub-sample containing 10^3 points is illustrated in both the standard and physical spaces in Figure 3.3 (even if the isoprobabilistic transform is not a prerequisite for this simple technique). The blue dots are such that $\mathfrak{g}$ is positive (safe points) and the red crosses are such that $\mathfrak{g}$ is negative (failure points). Since failure is rare under the chosen probabilistic model, it is hard to draw a sub-sample of size $N = 10^3$ containing at least one failure point. The sub-sample plotted in Figure 3.3 has been specifically isolated during the sampling procedure for the sake of illustration.



(a) A batch sample in the physical space

(b) A batch sample in the standard space

Figure 3.3: Monte Carlo sampling applied to the basic structural reliability problem defined in Section 3.2.4.

Once the analysis is done, the empirical CDF is estimated from the sample $\mathscr{G}$ using Eq. (3.46) and it is illustrated in Figure 3.4 as a staircase plot. The red and blue curves represent the 95% confidence interval bounds computed from Eq. (3.48). In particular the failure probability estimate can be read on the vertical axis of Figure 3.4 as $\widehat{p}_{f\,\mathrm{MCS}} = \widehat{F}_G(0) \approx 1.60\times10^{-6}$. The coefficient of variation is approximately equal to 35% and yields the following 95% confidence interval: $[5.02\times10^{-7};\ 2.70\times10^{-6}]$. Another Monte Carlo sampling driven by a target coefficient of variation on p_f equal to $\delta_0 = 10\%$ yields a finer estimate $\widehat{p}_{f\,\mathrm{MCS}} \approx 1.51\times10^{-6}$ at the expense of approximately 10^8 calls to the performance function.

3.3.2 Statistical inference of the output distribution

The methods reviewed in this section attempt to approximate the failure probability from a set of assumptions regarding the unknown probability distribution of G.

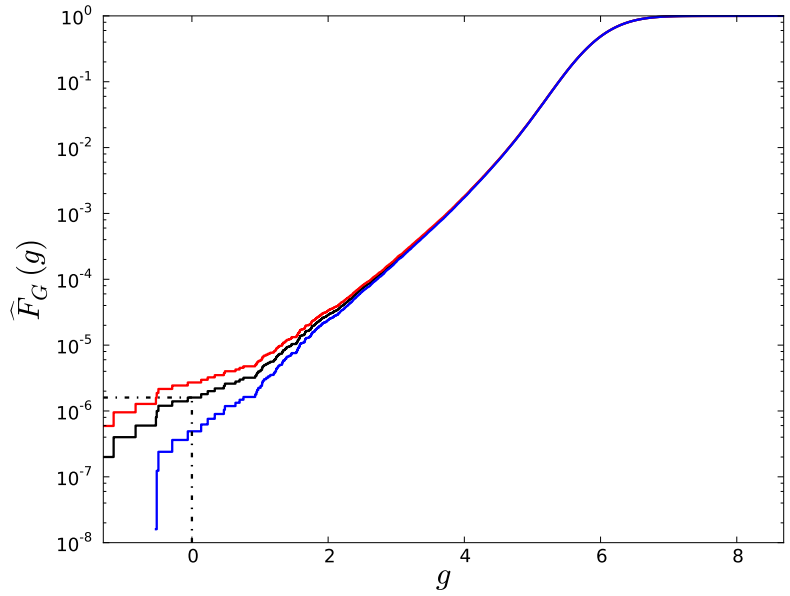


Figure 3.4: Empirical CDF of $G \equiv \mathfrak{g}(\boldsymbol{X}) = \mathfrak{g}^\circ(\boldsymbol{U})$ estimated by Monte Carlo sampling of the basic structural reliability problem defined in Section 3.2.4.

3.3.2.1 The median value first-order second-moment method

The *median value first-order second-moment* method (MV-FOSM) consists in linearizing the transformed performance function at the origin of the standard space (which is the image of the median of the original distribution $f_{\boldsymbol{X}}$ through the isoprobabilistic transform). The first-order Taylor series expansion of the transformed performance function at the origin reads as follows:

$$\mathfrak{g}^\circ(\boldsymbol{u}) = \mathfrak{g}^\circ_{1,\mathbf{0}}(\boldsymbol{u}) + o\left(\|\boldsymbol{u}\|_2^2\right), \tag{3.49}$$

where:

$$\mathfrak{g}^\circ_{1,\mathbf{0}}(\boldsymbol{u}) = \mathfrak{g}^\circ(\mathbf{0}) + \nabla_{\boldsymbol{u}}\mathfrak{g}^\circ(\mathbf{0})^{\mathsf{T}}\,\boldsymbol{u}. \tag{3.50}$$

Thanks to the linearity of the approximation, the propagation of the standard Gaussian random vector $\boldsymbol{U}$ through the latter first-order approximation yields a Gaussian variable:

$$G_{1,\mathbf{0}} \sim \mathscr{N}_1\left(\mu_{G_{1,\mathbf{0}}}, \sigma_{G_{1,\mathbf{0}}}\right), \tag{3.51}$$

with mean (resp. standard deviation):

$$\mu_{G_{1,\mathbf{0}}} = \mathbb{E}\left[\mathfrak{g}^\circ_{1,\mathbf{0}}(\boldsymbol{U})\right] = \mathfrak{g}^\circ(\mathbf{0}), \tag{3.52}$$

$$\sigma_{G_{1,\mathbf{0}}} = \sqrt{\mathrm{Var}\left[\mathfrak{g}^\circ_{1,\mathbf{0}}(\boldsymbol{U})\right]} = \left\|\nabla_{\boldsymbol{u}}\mathfrak{g}^\circ(\mathbf{0})\right\|_2, \tag{3.53}$$

$$\tag{3.54}$$

where $\|\bullet\|_2$ denotes the usual $\mathscr{L}_2$ norm in $\mathbb{R}^n$.

This approximation of G yields the following estimate of the failure probability:

$$p_{f\,\mathrm{MV-FOSM}} = \mathbb{P}\left[G_{1,\mathbf{0}} \leq 0\right] = F_{G_{1,\mathbf{0}}}(0) = \Phi\left(\frac{0 - \mu_{G_{1,\mathbf{0}}}}{\sigma_{G_{1,\mathbf{0}}}}\right). \tag{3.55}$$

Despite this method is rather inexpensive, it often yields inaccurate results due to the strong nonlinearities in $\mathfrak{g}^\circ$. In addition, this approximation does not come with any error measure so that it is impossible to check if the linear approximation is appropriate. It can be seen from Figure 3.5 that the MV-FOSM approximation of the failure probability is strongly inaccurate and that it does not provide any quantification of the error.

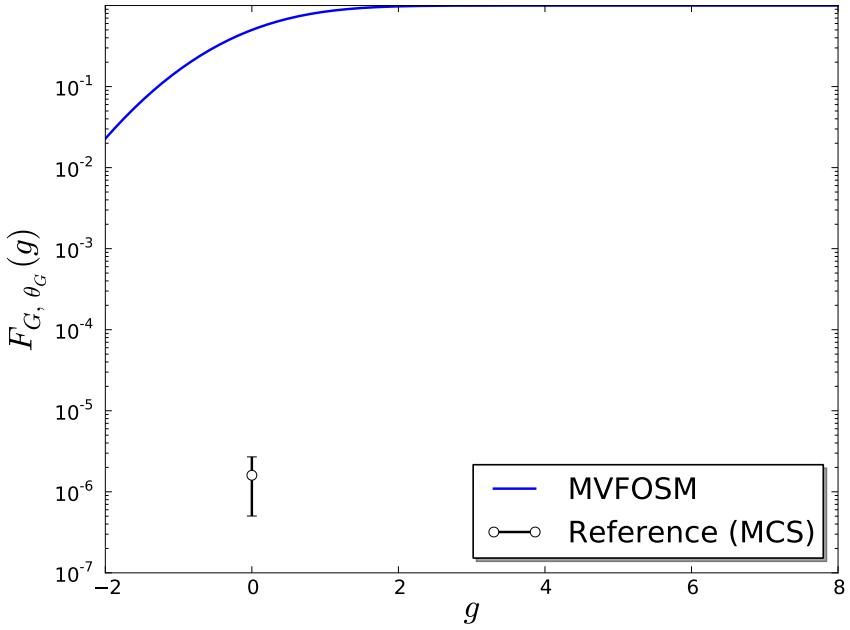


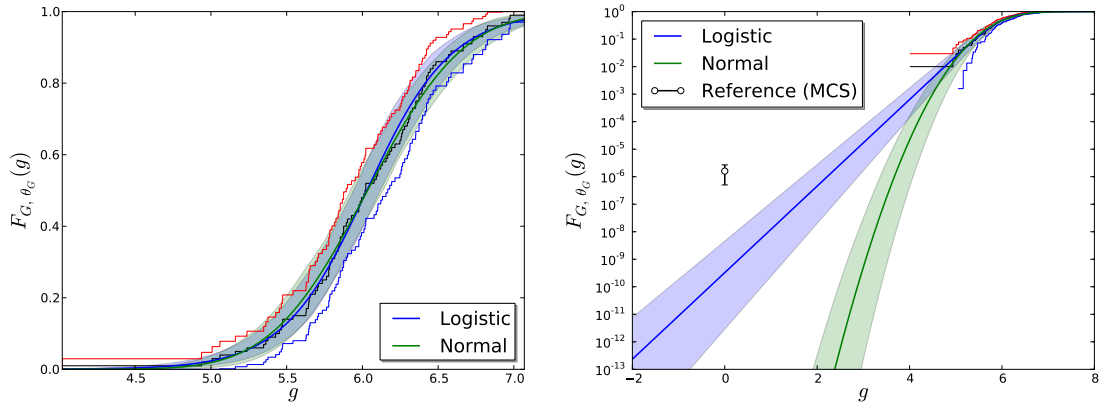
Figure 3.5: Median value first-order second-moment approximation of the random performance CDF for the basic structural reliability problem defined in Section 3.2.4.

3.3.2.2 Inference of the performance distribution

Another obvious technique consists in building a probabilistic distribution for the unknown random variable G using the usual statistical inference techniques. Starting from a sample $\mathscr{G} = \{g(\boldsymbol{x}^{(i)}), i = 1, \dots, n\}$ of reasonable size $30 \leq N \leq 100$ (obtained by Monte Carlo sampling), the empirical CDF $\widehat{F}_G$ is computed using its definition in Eq. (3.46). Then, one or several parametric probability distribution(s) $F_{G\,\boldsymbol{\theta}_G}$ may be proposed based on the shape of the empirical CDF. The parameters $\boldsymbol{\theta}_G$ of these distributions are estimated from the sample $\mathscr{G}$ using *e.g.* maximum likelihood or Bayesian estimation (provided a prior distribution $p(\boldsymbol{\theta}_G)$ for the parameters is available). Goodness-of-fit tests such as the one-sample Kolmogorov test may then be used to accept or reject the proposed distribution(s). Eventually, the failure probability is computed using the CDF of the retained best-fit parametric distribution.

Again, this technique is rather inexpensive but it is never used in practice for the following reasons. First it might be difficult to find a parametric distribution that fits to G. More importantly, the inference techniques and goodness-of-fit tests for parametric distributions lack objectivity far in the tail where failure occurs.

The technique is applied to the basic structural reliability problem defined in Section 3.2.4 and illustrated in Figure 3.6. The left-hand panel illustrates the perfect fit of the two proposed distributions (normal and logistic) that cannot be rejected by a Kolmogorov-Smirnov test (at the limit risk of $\alpha = 5\%$). The shaded areas represent the 95% confidence interval obtained by simulation when the epistemic uncertainty in the parameters $\boldsymbol{\theta}_G$ (built from the Fisher information matrix and a Gaussian assumption) is propagated through the CDF equations. Despite the central tendency seems perfectly fitted, the right-hand panel demonstrate how bad these two inferred distributions are with respect to the reference solution obtained by Monte Carlo sampling. It can also be seen that the solution strongly depends on the proposed distribution and that the spread of the 95% confidence interval on the CDF is not an objective error measure.



(a) Central tendency (b) Lower tail

Figure 3.6: Attempts to infer the distribution of the random performance G for the basic structural reliability problem defined in Section 3.2.4.

3.3.2.3 Inference of the tail of the performance distribution

The idea

Based on the premises exposed in the previous section, some authors proposed to infer only *the tail* of the distribution using appropriate parametric models (see *e.g.* Bucher, 2009; Naess et al., 2009; Nishijima et al., 2010; Sichani et al., 2011).

First, the empirical CDF is estimated using Eq. (3.46) and a Monte Carlo sample $\mathscr{G}$ of G. The lower and upper bounds of the 95% confidence interval are also computed using Eq. (3.48). These three empirical CDFs are then truncated to the tail of interest.

Assuming the lower tail is of interest, the empirical CDF is truncated to probability levels that are less than p_0 (say $p_0 = 10^{-2}$) by retaining only the samples that are less than the p_0-quantile, namely:

$$\mathscr{G}_{p_0} = \left\{ g \in \mathscr{G} : g \leq \widehat{G}_{p_0} \right\}, \tag{3.56}$$

where $\widehat{G}_{p_0}$ is the p_0-quantile approximated from the whole sample $\mathscr{G}$ using *e.g.* the following estimate:

$$\widehat{G}_{p_0} = \arg\max \left\{ g \in \mathscr{G} : \widehat{F}_G(g) \leq p_0 \right\}. \tag{3.57}$$

The lower p_0, the more samples N in $\mathscr{G}$ will be required so that the number of elements N_{p_0} in $\mathscr{G}_{p_0}$ enables an accurate estimate further in the tail (say up to 10^{-4} assuming a much lower failure probability is sought).

Bucher (2009), Naess et al. (2009) and Nishijima et al. (2010) then propose different parametric models $F_{G,\boldsymbol{\theta}_G}(\bullet \mid g \leq \widehat{G}_{p_0})$ for the lower tail distribution. Naess et al. (2009) determine the parameters $\boldsymbol{\theta}_G$ in this tail model using the following weighted least-squares regression formulation:

$$\boldsymbol{\theta}_G^* = \arg\min_{\boldsymbol{\theta}_G} \sum_{i=1}^{N_{p_0}} \left(\frac{\log F_{G,\boldsymbol{\theta}_G}\left(g^{(i)};\, \boldsymbol{\theta}_G \,\middle|\, \widehat{G}_{p_0}\right) - \log \widehat{F}_G(g^{(i)})}{\log \widehat{F}_G^+(g^{(i)}) - \log \widehat{F}_G^-(g^{(i)})} \right)^k, \tag{3.58}$$

where $k = 1$ or 2. The logarithmic spread of the 95% confidence interval is used in the denominator so as to put more weight on the more confident estimates of the CDF and less on the others. The logarithmic scale is used to ease the numerical resolution of this nonlinear regression problem. According to the author's experience, the most appropriate numerical optimizer for this task is the Levenberg-Marquardt algorithm (Marquardt, 1963).

Illustration of the concept

The concept is now applied to the basic structural reliability example defined in Section 3.2.4 for the sake of illustration. First, Monte Carlo sampling is used in order to generate a sample $\mathscr{G}$ of 10^5 performance function values. The probability level p_0 delimiting the tail is arbitrarily set to 10^{-2} and the corresponding quantile is found equal to $\widehat{G}_{p_0} = 4.65$.

For the present illustration, the *generalized extreme value distribution* is then proposed as a parametric model for the CDF tail. It reads as follows:

$$F_{\text{GEV}}\left(g;\, (\mu, \sigma, \xi)^\intercal \,\middle|\, \widehat{G}_{p_0}\right) = \begin{cases} \exp\left[-\left(1 + \xi\left(\frac{g-\mu}{\sigma}\right)\right)^{-1/\xi} \right] & \text{if } g \leq \widehat{G}_{p_0} \\ 0 & \text{otherwise} \end{cases}, \tag{3.59}$$

where $\mu < \widehat{G}_{p_0}$, $\sigma > 0$ and $\xi \in \mathbb{R}$ are the location, scale and shape parameters respectively. Note that Bucher (2009), Naess et al. (2009) and Nishijima et al. (2010) resort to a different parametric model that is inspired from the asymptotic approximation of the standard Gaussian CDF originally proposed in the book by Abramowitz and Stegun (1970). However, both models demonstrate a similar degree of freedom to fit most CDF tails.

The regression of the proposed model to the empirical CDF tail is truncated to the part that is estimated with a coefficient of variation that is less than 40%. The resulting model is illustrated in Figure 3.7. It can be seen that the predicted value for the failure probability is in reasonable agreement with the reference solution obtained by Monte Carlo sampling. The confidence bounds delimiting the shaded area in Figure 3.7 are obtained by fitting the

boundaries $\widehat{F}_G^-$ and $\widehat{F}_G^+$ instead of the mean estimate $\widehat{F}_G$. The overall approach is more accurate than the MV-FOSM method and the conventional inference of the overall distribution at a greater computational expense though. However, it is clear from the spread of the confidence bounds that the required sample size still depends on the order of magnitude of the failure probability. As intuited from Figure 3.7 and the applications presented in the original papers, interpolating the CDF tail too far from the empirical estimate does not seem reasonable.

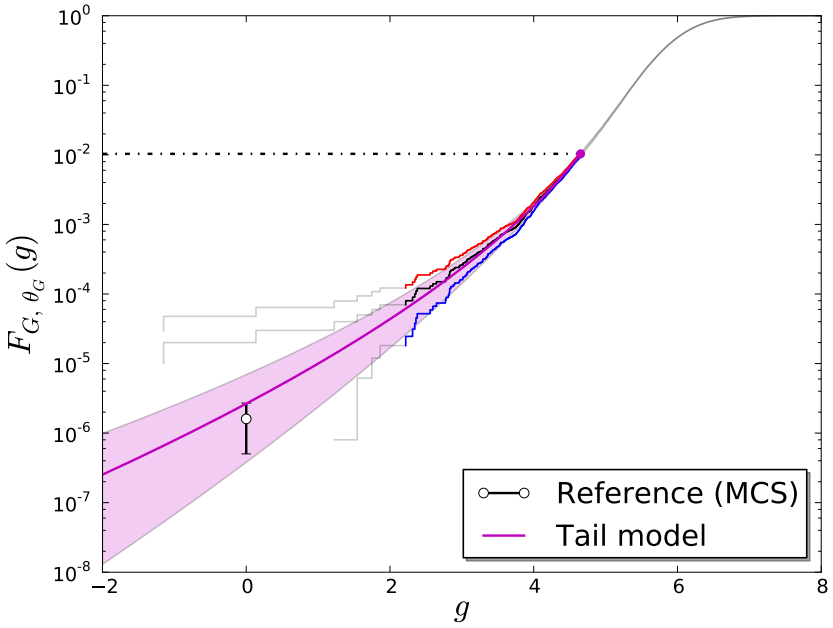


Figure 3.7: Inference of the tail of the distribution of the random performance G for the basic structural reliability problem defined in Section 3.2.4.

3.3.3 Importance sampling

In statistics, the so-called *variance reduction techniques* aim at deriving estimators of a quantity of interest (such as the failure probability in Eq. (3.3)) featuring a variance of estimation that is much less than that of the Monte Carlo estimator. The book by Rubinstein and Kroese (2008) reviews a large number of such techniques amongst which are: control variates, stratified sampling[2], subset sampling and cross-entropy methods. In this thesis, it is argued that most of these techniques (apart from stratified sampling) are provided as practical approximations of an optimal although impracticable meta-technique known as *importance sampling*.

[2]The Latin hypercube sampling technique exposed in Chapter 2 (Page 54) is a particular application of stratified sampling. See the article by Owen (1992) for the properties of an LHS-based estimator.

3.3.3.1 Principle

Let us first review this technique in a more general framework than the present reliability analysis context. Assume the following integral has to be approximated by means of Monte Carlo sampling:

$$\overline{\phi} = \int_{\mathbb{X}} \phi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \mathbb{E}_{\boldsymbol{X}} \left[\phi(\boldsymbol{X}) \right], \tag{3.60}$$

where ϕ is a real-valued function defined over $\mathbb{X}$. Note that this integral is nothing but a generalization of the definition of the failure probability given in Eq. (3.41) which can be obtained by substituting the failure indicator function $\mathbb{1}_{\mathbb{F}}$ for ϕ and the failure probability p_f for $\overline{\phi}$.

Importance sampling starts from the premise that ϕ may be close to zero in regions where the PDF $f_{\boldsymbol{X}}$ is large whereas it may take more significant values in regions where $f_{\boldsymbol{X}}$ tends to zero so that it makes the averaging of ϕ by Monte Carlo sampling of $\boldsymbol{X}$ rather inefficient. Hence, importance sampling consists in replacing the original PDF $f_{\boldsymbol{X}}$ with another PDF h to make the Monte Carlo approximation of Eq. (3.60) more efficient.

Let h denote such an *instrumental PDF* such that it dominates $\phi \, f_{\boldsymbol{X}}$, meaning that:

$$\begin{array}{rcl} h(\boldsymbol{x}) = 0 & \Rightarrow & \phi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) = 0 \\ \Leftrightarrow \quad \phi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \neq 0 & \Rightarrow & h(\boldsymbol{x}) \neq 0. \end{array} \tag{3.61}$$

The quantity of interest $\overline{\phi}$ rewrites as follows:

$$\begin{aligned} \overline{\phi} &= \int_{\mathbb{X}} \phi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \frac{h(\boldsymbol{x})}{h(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{X}} \frac{\phi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})}{h(\boldsymbol{x})} h(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \qquad (3.62) \\ &\equiv \mathbb{E}_{\boldsymbol{Z}} \left[\frac{\phi(\boldsymbol{Z}) f_{\boldsymbol{X}}(\boldsymbol{Z})}{h(\boldsymbol{Z})} \right], \qquad (3.63) \end{aligned}$$

where $\boldsymbol{Z}$ is the random vector distributed according to h. $\overline{\phi}$ may then be approximated by its Monte Carlo estimator which reads as follows:

$$\widehat{\overline{\phi}}_{\mathrm{IS}} = \widehat{\mathbb{E}}_{\boldsymbol{Z}} \left[\frac{\phi(\boldsymbol{Z}) f_{\boldsymbol{X}}(\boldsymbol{Z})}{h(\boldsymbol{Z})} \right] = \frac{1}{N} \sum_{i=1}^{N} \frac{\phi(\boldsymbol{Z}^{(i)}) f_{\boldsymbol{X}}(\boldsymbol{Z}^{(i)})}{h(\boldsymbol{Z}^{(i)})}, \tag{3.64}$$

where $\mathscr{Z} = \{\boldsymbol{Z}^{(i)}, i = 1, \ldots, N\}$ is a sample containing N independent copies of the random vector $\boldsymbol{Z}$ with PDF h. Thanks to the central limit theorem, provided N is sufficiently large, the latter quantity is an unbiased estimator of the quantity of interest:

$$\left[\widehat{\overline{\phi}}_{\mathrm{IS}} - \overline{\phi} \right] \underset{N \to \infty}{\hookrightarrow} \mathscr{N}_1(0, \sigma^2_{\mathrm{IS}}). \tag{3.65}$$

where the variance of estimation σ^2_{IS} is derived as follows:

$$\begin{aligned}\sigma^2_{\text{IS}} &= \text{Var}_Z\left[\widehat{\overline{\phi}}_{\text{IS}}\right] \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\text{Var}_{\boldsymbol{Z}^{(i)}}\left[\frac{\phi(\boldsymbol{Z}^{(i)})f_X(\boldsymbol{Z}^{(i)})}{h(\boldsymbol{Z}^{(i)})}\right] \\ &= \frac{1}{N^2}N\,\text{Var}_Z\left[\frac{\phi(\boldsymbol{Z})f_X(\boldsymbol{Z})}{h(\boldsymbol{Z})}\right],\end{aligned}$$

since $\mathscr{Z}$ contains N independent copies of $\boldsymbol{Z}$. At last, the variance further reduces to:

$$\begin{aligned}\sigma^2_{\text{IS}} &= \frac{1}{N}\left[\mathbb{E}_Z\left[\frac{\phi(\boldsymbol{Z})^2 f_X(\boldsymbol{Z})^2}{h(\boldsymbol{Z})^2}\right] - \mathbb{E}_Z\left[\frac{\phi(\boldsymbol{Z})f_X(\boldsymbol{Z})}{h(\boldsymbol{Z})}\right]^2\right] \\ &= \frac{1}{N}\left[\int_{\mathbb{X}}\frac{\phi(\boldsymbol{x})^2 f_X(\boldsymbol{x})^2}{h(\boldsymbol{x})^2}h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \left(\int_{\mathbb{X}}\frac{\phi(\boldsymbol{x})f_X(\boldsymbol{x})}{h(\boldsymbol{x})}h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)^2\right] \\ &= \frac{1}{N}\left[\int_{\mathbb{X}}\frac{\phi(\boldsymbol{x})^2 f_X(\boldsymbol{x})^2}{h(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} - \overline{\phi}^2\right]. \qquad (3.66)\end{aligned}$$

The following theorem from the book by Rubinstein and Kroese (2008, Chapter 5) provides the *optimal instrumental density* h^* that minimizes the latter variance of estimation.

Theorem 3.3.1. ***Optimal instrumental PDF***
The instrumental density that minimizes the variance of estimation in Eq. (3.66) reads:

$$h^*(\boldsymbol{x}) = \frac{\left|\phi(\boldsymbol{x})\right| f_X(\boldsymbol{x})}{\int_{\mathbb{X}}\left|\phi(\boldsymbol{x})\right| f_X(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}}. \qquad (3.67)$$

In addition, provided $\phi(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathbb{X}$, the optimal variance of estimation is zero.

Proof 3.3.1. *The proof is two-fold.*

(i) *Let us consider the variance of estimation yielded by any instrumental PDF h in Eq. (3.66). The first term (in the brackets) of this equation rewrites as follows:*

$$\int_{\mathbb{X}}\frac{\phi(\boldsymbol{x})^2 f_X(\boldsymbol{x})^2}{h(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{X}}\frac{\phi(\boldsymbol{x})^2 f_X(\boldsymbol{x})^2}{h(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x}\int_{\mathbb{X}}h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \qquad (3.68)$$

since the new term on the right is equal to one by definition of a PDF. Then the Cauchy-Schwartz inequality ensures that:

$$\int_{\mathbb{X}}\frac{\phi(\boldsymbol{x})^2 f_X(\boldsymbol{x})^2}{h(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x}\int_{\mathbb{X}}h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \left(\int_{\mathbb{X}}\frac{\left|\phi(\boldsymbol{x})\right| f_X(\boldsymbol{x})}{h(\boldsymbol{x})^{1/2}}h(\boldsymbol{x})^{1/2}\,\mathrm{d}\boldsymbol{x}\right)^2. \qquad (3.69)$$

Hence the following lower bound holds for the variance of estimation in Eq. (3.66):

$$\sigma^2_{\text{IS}} \geq \sigma^2_{\text{IS,LB}} \equiv \frac{1}{N}\left[\left(\int_{\mathbb{X}}\left|\phi(\boldsymbol{x})\right| f_X(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)^2 - \overline{\phi}^2\right]. \qquad (3.70)$$

Yet, according to Eq. (3.66), the variance of estimation yielded by the optimal instrumental density h^* reads as follows:

$$\mathrm{Var}_{\boldsymbol{Z}^*}\left[\widehat{\overline{\phi}}_{\mathrm{IS}}\right] = \frac{1}{N}\left[\int_{\mathbb{X}} \frac{\phi(\boldsymbol{x})^2 f_{\boldsymbol{X}}(\boldsymbol{x})^2}{h^*(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} - \overline{\phi}^2\right], \tag{3.71}$$

and it can be further elicited by substituting the quantity in Eq. (3.67) for h^*:

$$\begin{aligned}\mathrm{Var}_{\boldsymbol{Z}^*}\left[\widehat{\overline{\phi}}_{\mathrm{IS}}\right] &= \frac{1}{N}\left[\int_{\mathbb{X}} \frac{\phi(\boldsymbol{x})^2 f_{\boldsymbol{X}}(\boldsymbol{x})^2}{\frac{\left|\phi(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\mathbb{X}}\left|\phi(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}}}\,\mathrm{d}\boldsymbol{x} - \overline{\phi}^2\right] \\ &= \frac{1}{N}\left[\int_{\mathbb{X}} \frac{\phi(\boldsymbol{x})^2 f_{\boldsymbol{X}}(\boldsymbol{x})^2}{\left|\phi(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} \int_{\mathbb{X}}\left|\phi(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \overline{\phi}^2\right] \\ &= \frac{1}{N}\left[\left(\int_{\mathbb{X}}\left|\phi(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)^2 - \overline{\phi}^2\right]. \end{aligned} \tag{3.72}$$

By comparing Eq. (3.70) and Eq. (3.72), it appears that h^* realizes the lower bound of the variance of estimation, i.e. $\sigma^{2\,*}_{\mathrm{IS}} = \sigma^2_{\mathrm{IS,LB}}$.

(ii) *Now assume that ϕ takes only positive values, then the minimal variance of estimation in Eq. (3.72) reduces to:*

$$\sigma^{2\,*}_{\mathrm{IS}} = \frac{1}{N}\left[\left(\int_{\mathbb{X}} \phi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)^2 - \overline{\phi}^2\right] = \frac{1}{N}\left[\overline{\phi}^2 - \overline{\phi}^2\right] = 0. \tag{3.73}$$

3.3.3.2 Impracticality of optimal importance sampling

Applying Theorem 3.3.1 to the problem of estimating the failure probability from its definition in Eq. (3.41) yields the following optimal instrumental density:

$$h^*(\boldsymbol{x}) = \frac{\left|\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\mathbb{X}}\left|\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\right| f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}} = \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})}{p_f}. \tag{3.74}$$

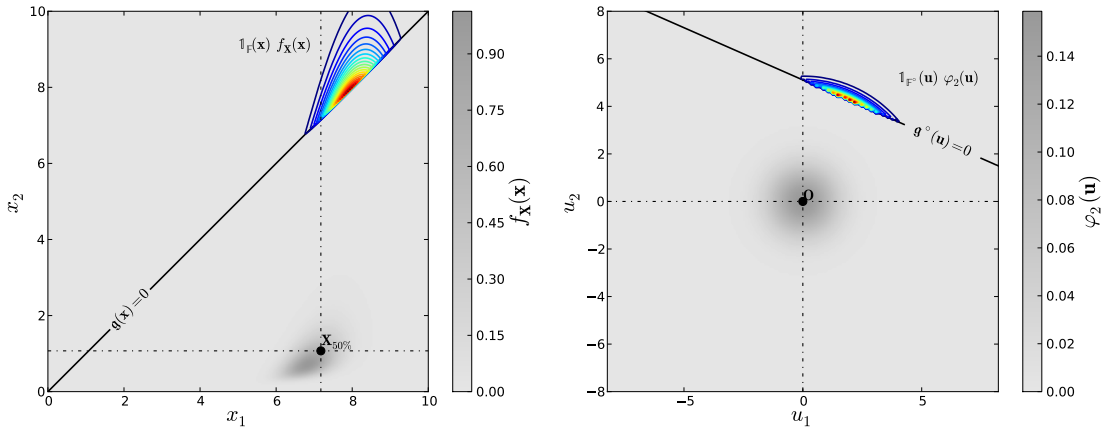
In addition, it ensures that the variance of estimation is zero because the failure indicator function takes only positive values (0 or 1). Equivalently, the optimal instrumental density in the standard space reads:

$$h^{\circ\,*}(\boldsymbol{u}) = \frac{\mathbb{1}_{\mathbb{F}^\circ}(\boldsymbol{u})\,\varphi_n(\boldsymbol{u})}{p_f}. \tag{3.75}$$

Unfortunately, this optimal density is intractable because it involves the unknown quantity of interest p_f in its denominator. Figure 3.8 represents the unnormalized optimal PDFs in both standard and physical spaces for the basic structural reliability problem defined in Section 3.2.4 in order to provide some more insight on Theorem 3.3.1. It is clear from this illustration that the optimal instrumental PDF is equal to the PDF of $\boldsymbol{X}$ (resp. $\boldsymbol{U}$) truncated

to the failure domain $\mathbb{F}$ (resp. $\mathbb{F}^\circ$). In other words the optimal instrumental PDF is nothing but the probability distribution for the failure points, and it is clear that such a perfect knowledge of the failure region is not available in practice.

Thus *the art of importance sampling* consists in building a quasi-optimal instrumental density that significantly reduces the variance of estimation without reducing it to zero though. The remaining of this section explains how such a variance reduction can be achieved by means of the conventional reliability methods.



(a) In the physical space spanned by $\boldsymbol{X}$ (b) In the standard space spanned by $\boldsymbol{U}$

Figure 3.8: Optimal (although impracticable) instrumental PDFs for the basic structural reliability problem defined in Section 3.2.4.

3.3.4 Most-probable-failure-point-based approaches

The approaches introduced in this section have been a reference in the structural reliability literature for the past decades. Despite their limitations, they remain quite popular because they provide satisfactory results when the assumptions they are built on hold. They also involve an optimization problem that yields interesting by-products in addition to the failure probability approximation at a reasonable computational cost (see Lemaire, 2009, Chapter 6).

3.3.4.1 The most probable failure point

It is clear from Figure 3.8 that a large amount of probability condensates in a region close to the median point in the physical space or to the origin in the standard space. This remark naturally leads to the definition of *the most probable failure point* (MPFP) as the maximizer (or *mode*) of the optimal instrumental density in the standard space:

$$\boldsymbol{u}^* = \arg \max_{\boldsymbol{u}\in\mathbb{R}^n} \frac{\mathbb{1}_{\mathbb{F}^\circ}(\boldsymbol{u})\,\varphi_n(\boldsymbol{u})}{p_f}. \tag{3.76}$$

The unknown normalizing constant p_f does not alter the variations of the objective function so that the initial definition rewrites as follows:

$$\begin{aligned} \boldsymbol{u}^* &= \arg \max_{\boldsymbol{u}\in\mathbb{R}^n} \mathbb{1}_{\mathbb{F}^\circ}(\boldsymbol{u})\,\varphi_n(\boldsymbol{u}) \\ &= \arg \max_{\boldsymbol{u}\in\mathbb{R}^n} \frac{1}{(2\,\pi)^{n/2}} \exp\left[-\frac{1}{2}\,\boldsymbol{u}^\mathsf{T}\,\boldsymbol{u}\right] \quad \text{s.t.} \quad \boldsymbol{u}\in\mathbb{F}^\circ \\ &= \arg \min_{\boldsymbol{u}\in\mathbb{R}^n} \boldsymbol{u}^\mathsf{T}\,\boldsymbol{u} \quad \text{s.t.} \quad \mathfrak{g}^\circ(\boldsymbol{u}) \leq 0. \end{aligned} \tag{3.77}$$

This is the usual definition of the MPFP as defined within the context of the first order reliability method (FORM). See the reference books by Ditlevsen and Madsen (1996) and Lemaire (2009). The MPFP is *the closest failure point to the origin of the standard space*. It is sometimes referred to as the *design point* as it corresponds to a sort of optimal deterministic design under the non-linear constraint $\mathfrak{g}^\circ(\boldsymbol{u}) \leq 0$.

The problem in Eq. (3.77) is a quadratic optimization problem under non-linear constraint. Thus it can be efficiently solved by means of the usual sequential quadratic programming algorithm or other more specific algorithms such as the Abdo-Rackwitz algorithm (Abdo and Rackwitz, 1990), or the improved Hasofer-Lind-Rackwitz-Fiessler algorithm (iHLRF, Zhang and Der Kiureghian, 1995).

3.3.4.2 First-order reliability approximation

Assuming that the MPFP is unique, its coordinates enable an analytical approximation for the failure probability. This approximation is obtained by linearizing the performance function in the vicinity of the latter point. The first-order Taylor series expansion at the MPFP reads as follows:

$$\mathfrak{g}^\circ(\boldsymbol{u}) = \mathfrak{g}^\circ_{1,\boldsymbol{u}^*}(\boldsymbol{u}) + o\left(\|\boldsymbol{u}-\boldsymbol{u}^*\|_2^2\right), \tag{3.78}$$

where:

$$\mathfrak{g}^\circ_{1,\boldsymbol{u}^*}(\boldsymbol{u}) = \mathfrak{g}^\circ(\boldsymbol{u}^*) + \nabla_u \mathfrak{g}^\circ(\boldsymbol{u}^*)^\mathsf{T}\,(\boldsymbol{u}-\boldsymbol{u}^*). \tag{3.79}$$

Since $\boldsymbol{u}^*$ is on the limit-state surface, $\mathfrak{g}^\circ(\boldsymbol{u}^*)$ is equal to zero. By introducing, the opposite normalized gradient of the performance function at the MPFP:

$$\boldsymbol{\alpha} = -\frac{\nabla_u \mathfrak{g}^\circ(\boldsymbol{u}^*)}{\left\|\nabla_u \mathfrak{g}^\circ(\boldsymbol{u}^*)\right\|_2}, \tag{3.80}$$

the expansion in Eq. (3.79) further reads after normalization:

$$\mathfrak{g}^\circ_{1,\boldsymbol{u}^*}(\boldsymbol{u}) = \left\|\nabla_u \mathfrak{g}^\circ(\boldsymbol{u}^*)\right\|_2 \left(\boldsymbol{\alpha}^\mathsf{T}\boldsymbol{u}^* - \boldsymbol{\alpha}^\mathsf{T}\boldsymbol{u}\right) \tag{3.81}$$

Thus the first-order failure probability approximation is obtained as follows:

$$p_{f,1} = \mathbb{P}\left[\mathfrak{g}^\circ_{1,\boldsymbol{u}^*}(\boldsymbol{U}) \leq 0\right] = \mathbb{P}\left[-\boldsymbol{\alpha}^\mathsf{T}\boldsymbol{U} \leq -\boldsymbol{\alpha}^\mathsf{T}\boldsymbol{u}^*\right]. \tag{3.82}$$

Recall that $\boldsymbol{\alpha}$ has its Euclidean norm equal to one by definition (see Eq. (3.80)) and that $\boldsymbol{U}$ is a standard Gaussian random vector, so that $\boldsymbol{\alpha}^\mathsf{T}\boldsymbol{U}$ is a standard Gaussian random variate. In

addition, let us introduce the so-called *Hasofer-Lind reliability index* defined as the following algebraic (signed) distance in $\mathbb{R}^n$:

$$\beta_{\mathrm{HL}} \equiv \boldsymbol{\alpha}^{\mathsf{T}} \boldsymbol{u}^{*}. \tag{3.83}$$

Eventually, the first-order failure probability approximation becomes:

$$p_{f,1} = \Phi(-\beta_{\mathrm{HL}}). \tag{3.84}$$

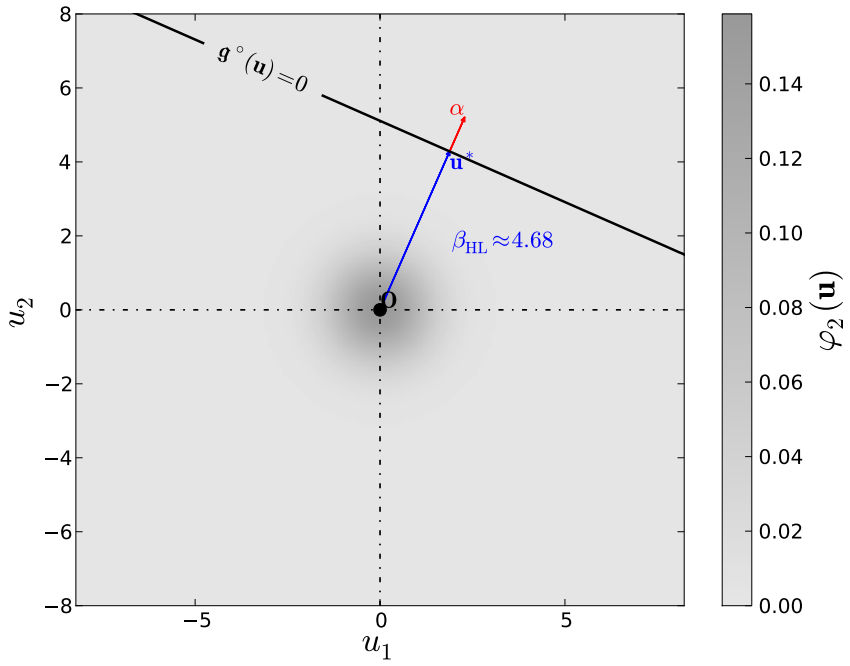


Figure 3.9: First-order reliability approximation for the basic structural reliability problem defined in Section 3.2.4.

Both $\boldsymbol{\alpha}$ and $\boldsymbol{u}^*$ are illustrated in Figure 3.9 for the basic reliability example defined in Section 3.2.4. Since the limit-state surface $\mathscr{S}_0$ is an hyperplane with closed-form expression (see Eq. (3.39)), the Hasofer-Lind reliability index comes analytically as the distance from the origin of the standard space to the latter hyperplane:

$$\beta_{\mathrm{HL}} = \mathrm{dist}\left(\boldsymbol{O}, \mathscr{S}_0\right) = \frac{\left|\lambda_R - \lambda_S\right|}{\sqrt{\left(\zeta_S \rho_0 - \zeta_R\right)^2 + \zeta_S^2 \left(1 - \rho_0^2\right)}}. \tag{3.85}$$

The numerical evaluation gives $\beta_{\mathrm{HL}} \approx 4.68$ which in turns leads to $p_{f,1} \approx 1.44 \times 10^{-6}$ as the first-order approximation of the failure probability (see Eq. (3.84)). Since the limit-state surface is linear, this first-order approximation is equal to the exact failure probability in this simple example although this is not a general fact. Indeed the Hasofer-Lind reliability index is usually estimated from Eq. (3.83) and $p_{f,1}$ is then only an approximation of the exact failure probability.

3.3.4.3 Corrections for non-linear limit-state surfaces

For limit-state surfaces featuring curvature at the design point, Breitung (1984) proposed an analytical second-order approximation for the failure probability. The approximation holds only asymptotically and is as follows:

$$p_{f,\text{Breitung}} \underset{\beta_{\text{HL}}\to\infty}{\sim} \Phi(-\beta_{\text{HL}}) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1+\beta_{\text{HL}}\,\kappa_i}}, \tag{3.86}$$

where $\{\kappa_i,\, i = 1, \ldots, n-1\}$ are the principal curvatures of the limit-state surface (*i.e.* the curvatures of the limit-state surface in the rotated basis that admits $\boldsymbol{\alpha}$ as a principal axis). The second-order approximation of a limit-state surface around the design point requires computing the Hessian matrix $\nabla^2_{\boldsymbol{u},\boldsymbol{u}}\mathfrak{g}^\circ(\boldsymbol{u}^*)$ (see Hohenbichler et al., 1987; Hohenbichler and Rackwitz, 1988). This computation might be non-trivial though. It is especially true when the performance function involves the output of a noisy computational model for which a finite forward difference scheme is hard to tune. Der Kiureghian et al. (1987) propose to use a composite experimental design centered at the design point in order to approximate the limit-state surface with a quadratic response surface. This approach uses larger perturbations and is thus more robust regarding the possible noise in the performance function.

For general non-linear limit-state surfaces, Melchers (1989) proposed to implement an importance sampling scheme with the following instrumental PDF:

$$h^\circ_{\boldsymbol{u}^*}(\boldsymbol{u}) = \varphi_n(\boldsymbol{u} - \boldsymbol{u}^*). \tag{3.87}$$

Even if this instrumental PDF is not optimal, it may yield a significant variance reduction with respect to the standard Monte Carlo estimator.

Melchers' instrumental PDF is illustrated in Figure 3.10 for the basic reliability example defined in Section 3.2.4. An importance sampling scheme using this instrumental PDF is implemented. It is driven with a target coefficient of variation $\delta_0 = 10\%$ and the corresponding probability estimate $\widehat{p}_{f\,\boldsymbol{u}^*\text{IS}} \approx 1.49 \times 10^{-6}$ is obtained at the expense of only 600 calls[3] to the performance function.

3.3.4.4 Limitations

Despite all these MPFP-based approaches provide fairly accurate approximations of the failure probability when the MPFP is unique, they may lead to very inaccurate results when it is not.

To illustrate this point, Der Kiureghian and Dakessian (1998) provided the following example. The performance function is directly defined in the standard space and it reads as follows:

$$\mathfrak{g}^\circ(\boldsymbol{u}) = b - u_2 - \kappa\,(u_1 - e)^2 \tag{3.88}$$

[3] Up to 50 runs because the sampling was run by batches of 50.

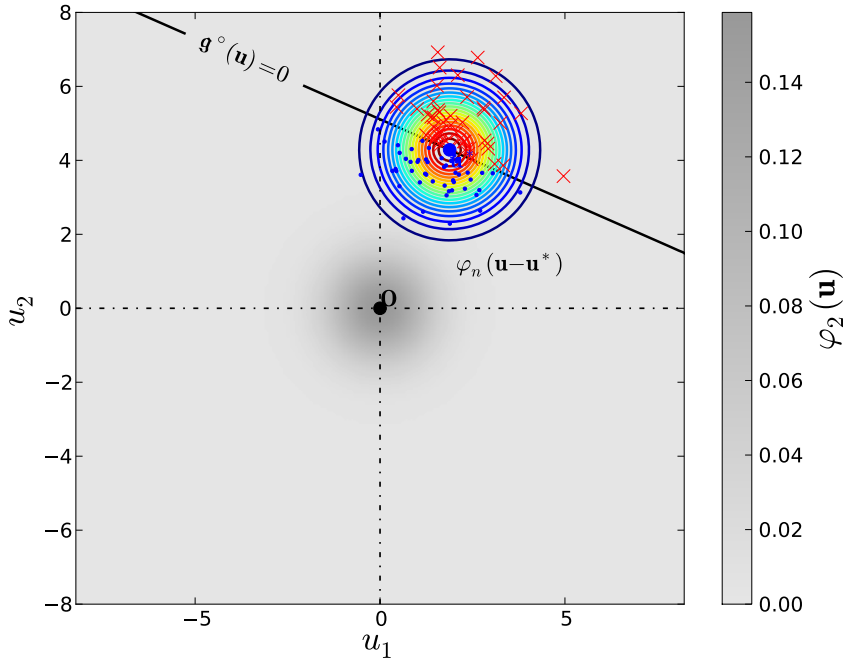


Figure 3.10: MPFP-based importance sampling for the basic structural reliability problem defined in Section 3.2.4.

where $\boldsymbol{u}$ is a realization of the standard Gaussian random vector $\boldsymbol{U}$, $b = 5$, $\kappa = 0.5$ and $e = 0.1$. The limit-state surface is the parabola shown in Figure 3.11. The curvature κ and the eccentricity e control the severity of the example for MPFP-based approaches. The true failure probability is sufficiently high to be estimated by means of crude Monte Carlo sampling and it is found equal to $\widehat{p}_{f\,\mathrm{MCS}} \approx 2.99 \times 10^{-3}$ (with a coefficient of variation that is less than 1%).

The iHLRF algorithm (Zhang and Der Kiureghian, 1995) is used to find the MPFP. The corresponding Hasofer-Lind reliability index is $\beta_{\mathrm{HL}} \approx 3.09$ so that the first-order approximation of the failure probability is $p_{f,1} \approx 9.87 \times 10^{-4}$. Then, Melchers' instrumental PDF is used to check if the first-order reliability method assumptions hold. The importance sampling procedure is driven with a target coefficient of variation $\delta_0 = 10\%$, and it provides the following failure probability estimate $p_{f\,\boldsymbol{u}^*\mathrm{IS}} \approx 1.06 \times 10^{-3}$ (with a coefficient of variation that is less than 10%).

Despite all accuracy criteria seems fulfilled, both the first-order approximation $p_{f,1}$ and the importance sampling estimate $p_{f\,\boldsymbol{u}^*\mathrm{IS}}$ are biased with respect to the Monte Carlo reference estimate due to the existence of a second design point. Both points are illustrated in Figure 3.11 together with Melchers' instrumental PDF centered at the first design point.

This example also illustrates the dangers underlying the choice of the instrumental PDF. The instrumental density expresses a sort of prior knowledge about the failure region. If this information is strongly biased, then the central limit theorem and its asymptotic variance of estimation do not hold. Note also that in this case, Melchers' instrumental PDF barely fulfils the domination requirement formulated in Eq. (3.61). Indeed, despite it has infinite

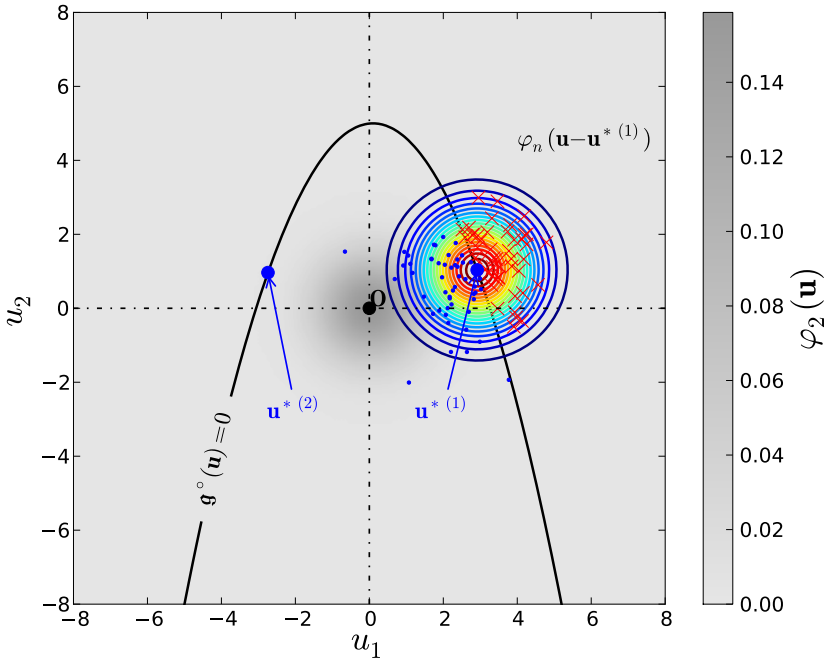


Figure 3.11: Limitation of MPFP-based approaches on the example from Der Kiureghian and Dakessian (1998).

support, Melchers' instrumental PDF takes very low values in the vicinity of the second design point and there is very little chance to sample in this area. As a consequence, the true variance of estimation is greater than that provided by the central limit theorem.

3.3.4.5 Turnaround for multiple design points

For cases featuring multiple design points, Der Kiureghian and Dakessian (1998) proposed the *restarted iHLRF* algorithm as a heuristic turnaround to find several design points. This algorithm proceeds as follows. First, the usual iHLRF algorithm is used in order to find a first design point $\boldsymbol{u}^{*(1)}$. Once a design point is found, the algorithm is restarted from a different starting point on a modified performance function that excludes an hypersphere (a *bulge*) centred at the previous design point. The modified performance function reads as follows:

$$\mathfrak{g}^\circ_{\text{Bulge}}(\boldsymbol{u}) = \mathfrak{g}^\circ(\boldsymbol{u}) + s\,\mathbb{1}_{\mathscr{B}^{(1)}}(\boldsymbol{u})\left(r^{(1)2} - \left\|\boldsymbol{u}^{*(1)}\right\|_2^2\right)^2, \tag{3.89}$$

where $\mathbb{1}_{\mathscr{B}^{(1)}}$ is the *bulge indicator function* defined as follows:

$$\mathbb{1}_{\mathscr{B}^{(1)}}(\boldsymbol{u}) = \begin{cases} 1 & \text{if } \left\|\boldsymbol{u}-\boldsymbol{u}^{*(1)}\right\|_2 \leq r^{(1)} \\ 0 & \text{otherwise} \end{cases}. \tag{3.90}$$

and s controls the height of the bulge. The approach is repeated incrementally by adding new bulges centred on all the previously found design points so that the modified perfor-

mance function excluding the Q already found design points reads as follows:

$$\mathfrak{g}^{\circ}_{\text{Bulge}}(\boldsymbol{u}) = \mathfrak{g}^{\circ}(\boldsymbol{u}) + \sum_{q=1}^{Q} s\, \mathbb{1}_{\mathscr{B}^{(q)}}(\boldsymbol{u}) \left(r^{(q)^2} - \left\| \boldsymbol{u}^{*(q)} \right\|_2^2 \right)^2, \tag{3.91}$$

where $\{\mathbb{1}_{\mathscr{B}^{(q)}}, q = 1, \dots, Q\}$ are the bulge indicator functions for all the design points (defined as in Eq. (3.90)).

If the algorithm converges to a point that is located on the foot of one of the bulges, it is discarded. At this point, the algorithm requires an intervention from the user in order to decide whether to perform a new search or not because the modified performance function may feature too many bulges. In addition to this first limitation, the algorithm features different parameters such as the bulges' radii $\{r^{(i)}, i = 1, \dots, n_{\text{MPFP}}\}$ and their height s that are rather hard to tune. A more detailed description and illustrations of this algorithm can be found in the original article of Der Kiureghian and Dakessian (1998) as well as in FERUM v4.0's documentation (Bourinet et al., 2009).

A collection of approximating hyperplanes can then be built from the collection of design points $\{\boldsymbol{u}^{*(i)}, i = 1, \dots, n_{\text{MPFP}}\}$ and unit normal vectors $\{\boldsymbol{\alpha}^{(i)}, i = 1, \dots, n_{\text{MPFP}}\}$. The linear approximations of the n_{MPFP} failure domains read as follows:

$$\widetilde{\mathbb{F}}^{\circ\,(i)} = \left\{ \boldsymbol{u} \in \mathbb{R}^n : \beta^{(i)}_{\text{HL}} - \boldsymbol{\alpha}^{(i)^{\mathsf{T}}} \boldsymbol{u} \leq 0 \right\}, \quad i = 1, \dots, n_{\text{MPFP}}. \tag{3.92}$$

The corresponding multi-linear approximation of the overall failure region is defined as the union of all these marginal failure domains:

$$\widetilde{\mathbb{F}}^{\circ} = \bigcup_{i=1}^{n_{\text{MPFP}}} \widetilde{\mathbb{F}}^{\circ\,(i)}. \tag{3.93}$$

It can then be proved (see *e.g.* Lemaire, 2009, Chapter 9) that the first-order approximation of this *series system* failure probability is given by the centred (zero-mean) multivariate Gaussian CDF:

$$p_{f\,1\,\Sigma} = \mathbb{P}\left[\boldsymbol{U} \in \widetilde{\mathbb{F}}^{\circ} \right] = 1 - \Phi_{n_{\text{MPFP}}}\left(\boldsymbol{\beta}_{\text{HL}} \middle| \mathbf{0}, \boldsymbol{\rho} \right), \tag{3.94}$$

where $\boldsymbol{\beta}_{\text{HL}} = (\beta^{(i)}_{\text{HL}}, i = 1, \dots, n_{\text{MPFP}})^{\mathsf{T}}$ is the vector of Hasofer-Lind reliability indices for each MPFP, and $\boldsymbol{\rho}$ is the correlation matrix between each linear approximation whose terms read as follows:

$$\rho_{ij} = \boldsymbol{\alpha}^{(i)^{\mathsf{T}}} \boldsymbol{\alpha}^{(j)}, \quad i, j = 1, \dots, n_{\text{MPFP}}. \tag{3.95}$$

Before the work by Genz (1992, 1993) on the computation of the multivariate Gaussian CDF $\Phi_{n_{\text{MPFP}}}$ defined in Eq. (3.11), Ditlevsen (1979b) proposed the so-called bi-modal bounds obtained by considering only the second-order interactions between each linear limit-state which only involve the bivariate Gaussian CDF Φ_2 (see *e.g.* Lemaire, 2009, Chapter 9).

The restarted iHLRF algorithm was applied to the multiple failure modes example from Der Kiureghian and Dakessian (1998). The marginal Hasofer-Lind reliability indices $\beta^{(1)}_{\text{HL}} \approx 3.09$ and $\beta^{(2)}_{\text{HL}} \approx 2.91$ are almost equivalent and the correlation between the two limit-state hyperplanes $\rho_{1,2} \approx -0.78$ is quite significant. Indeed it can be seen on Figure 3.12 that

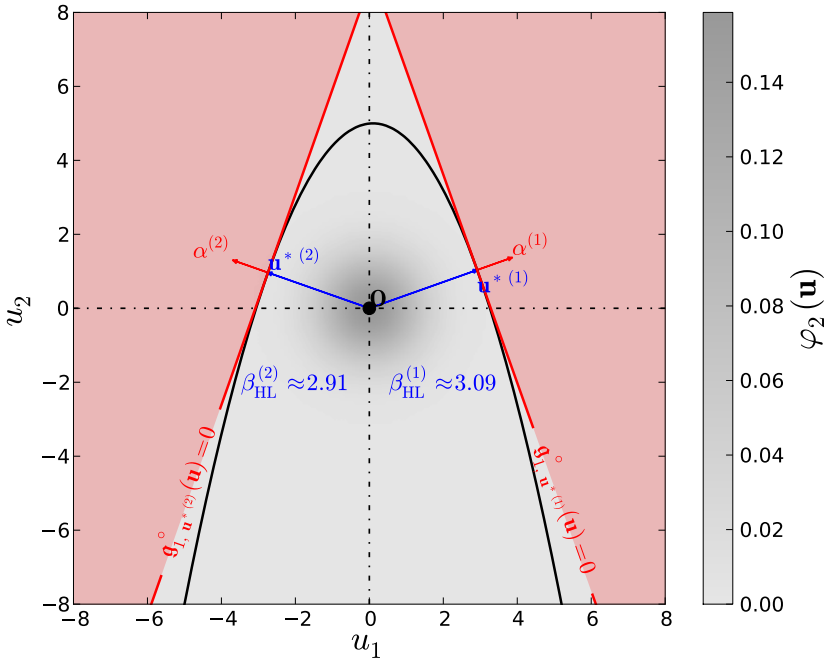


Figure 3.12: First-order approximate of a *serial system* failure probability on the example from Der Kiureghian and Dakessian (1998).

the two linear approximations are almost parallel and that the corresponding approximate failure domains are facing each other. Eventually, the first-order approximation of the serial system failure probability is $p_{f\,1\,\Sigma} \approx 2.82 \times 10^{-3}$. The latter approximation is fairly more accurate than the marginal first-order approximation $p_{f,1}^{(1)} \approx 9.87 \times 10^{-4}$ and $p_{f,1}^{(2)} \approx 1.83 \times 10^{-3}$ even though it is still biased with respect to the Monte Carlo reference solution $\widehat{p}_{f\,\mathrm{MCS}} \approx 2.99 \times 10^{-3}$. Another more robust workaround that was not implemented in this thesis would consist in using the importance sampling technique proposed by Cambier et al. (2002, Section 3.5) which is the generalization of the work by Melchers (1989) to multiple design points (*i.e.* Cambier et al.'s instrumental PDF is a mixture of Melchers' unimodal PDFs centred at each design point).

The pathological example proposed by Der Kiureghian and Dakessian (1998) may seem quite artificial but it is not. For instance, multiple failure modes may occur if the limit-state surface is defined from a black-box computer model featuring a built-in serial system (*e.g.* with "*if-then-else*" statements). Structural reliability examples involving stochastic processes (or random fields) are also known to feature several *most probable failure configurations* (see *e.g.* Verhoosel and Gutiérrez, 2009; Dubourg et al., 2009b).

3.3.5 Subset sampling

Subset sampling is a well-known variance reduction technique amongst statisticians (see *e.g.* Glasserman et al., 1999; Garnier and Del Moral, 2006; Cérou and Guyader, 2007; L'Ecuyer et al., 2009). It was brought to the structural reliability community by Au and Beck (2001);

Ching et al. (2005b,a); Au et al. (2007); Katafygiotis and Cheung (2005, 2007). It is sometimes also referred to as *splitting simulation* for reasons that will become clearer in the sequel.

3.3.5.1 Principle

In this section, the key idea of subset sampling is reviewed from the optimal importance sampling viewpoint. For a more conventional introduction, the reader is referred to the reference article by Au and Beck (2001).

First, recall that the optimal instrumental PDF for estimating the failure probability reads as follows:

$$h^*(\boldsymbol{x}) = \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_X(\boldsymbol{x})}{\int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_X(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}} = \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_X(\boldsymbol{x})}{p_f}. \tag{3.96}$$

It was earlier argued that this instrumental PDF is impracticable because it involves the unknown quantity of interest p_f in its denominator. This argument is reconsidered here. In fact, p_f could be estimated using crude Monte Carlo sampling although this would be rather inefficient if the failure probability is low. Hence, subset sampling aims at *splitting* the optimal instrumental PDF in a collection of *sub-optimal* instrumental PDFs whose normalizing constants are sufficiently high to be accurately estimated by means of crude Monte Carlo sampling. This collection of sub-optimal instrumental PDFs is defined as follows.

Let us consider a sequence of s reals such that $q_1 > \ldots > q_s = 0$. Since the performance function $\mathfrak{g}$ is assumed to take its values on the whole real line, this sequence may then be used to define the following subsets:

$$\mathbb{F}_i = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \mathfrak{g}(\boldsymbol{x}) \leq q_i \right\}, \quad i = 1, \ldots, s. \tag{3.97}$$

A collection of such subsets is illustrated in Figure 3.13 on the example defined in Section 3.2.4.

In addition, let us define their associated indicator functions:

$$\mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x}) = \left\{ \begin{array}{ll} 1 & \text{if} \quad \mathfrak{g}(\boldsymbol{x}) \leq q_i \\ 0 & \text{otherwise} \end{array} \right. , \quad i = 1, \ldots, s. \tag{3.98}$$

Due to the required ordering of the sequence $\{q_i, i = 1, \ldots, s\}$, the following relation holds:

$$\mathbb{F} = \mathbb{F}_s \subset \ldots \subset \mathbb{F}_1. \tag{3.99}$$

Hence, the corresponding indicator functions satisfy the following implications:

$$\mathbb{1}_{\mathbb{F}_1}(\boldsymbol{x}) = 0 \quad \Rightarrow \ldots \Rightarrow \quad \mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x}) = 0 \quad \Rightarrow \quad \mathbb{1}_{\mathbb{F}_s}(\boldsymbol{x}) = \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) = 0. \tag{3.100}$$

The collection of sub-optimal instrumental PDFs is now defined as follows:

$$h_i^*(\boldsymbol{x}) = \frac{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) f_X(\boldsymbol{x})}{p_{i-1}}, \quad i = 2, \ldots, s, \tag{3.101}$$

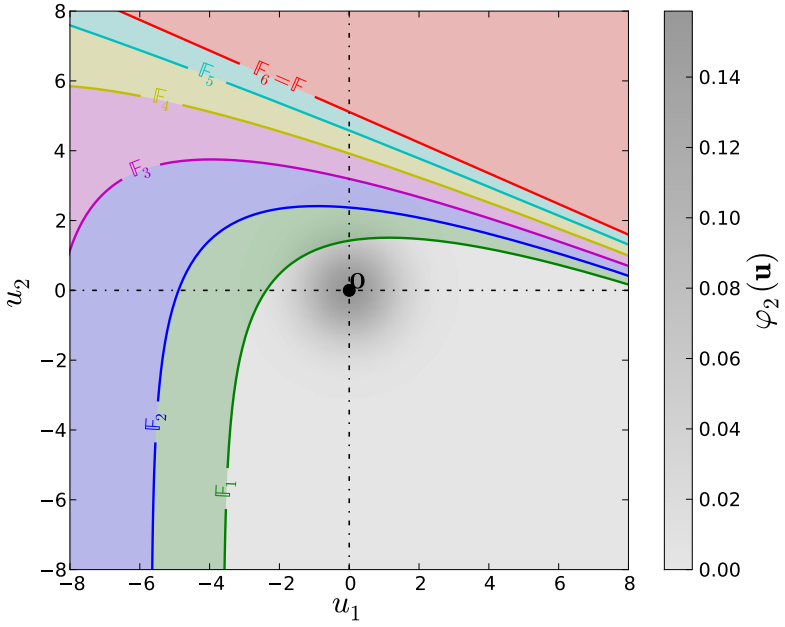


Figure 3.13: A collection of subsets defined from a strictly decreasing sequence of 6 positive reals in the standard space for the basic structural reliability example defined in Section 3.2.4.

where their normalizing constants read:

$$p_{i-1} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \quad i = 2, \ldots, s. \tag{3.102}$$

Note that Eq. (3.100) ensures that this collection of instrumental PDFs dominate each others and eventually dominate $\mathbb{1}_{\mathbb{F}} f_{\boldsymbol{X}}$:

$$h_2^*(\boldsymbol{x}) = 0 \quad \Rightarrow \ldots \Rightarrow \quad h_{s-1}^*(\boldsymbol{x}) = 0 \quad \Rightarrow \quad h_s^*(\boldsymbol{x}) = 0 \quad \Rightarrow \quad \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) = 0. \tag{3.103}$$

Let us now use the last instrumental PDF h_s^* in order to rewrite the failure probability defined in Eq. (3.41) as follows:

$$\begin{aligned} p_f &= \int_{\{\boldsymbol{x} \in \mathbb{X}: \, h_s^*(\boldsymbol{x}) > 0\}} \mathbb{1}_{\mathbb{F}_s}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \frac{h_s^*(\boldsymbol{x})}{h_s^*(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x} \\ &= \int_{\{\boldsymbol{x} \in \mathbb{X}: \, h_s^*(\boldsymbol{x}) > 0\}} \mathbb{1}_{\mathbb{F}_s}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \frac{p_{s-1}}{\mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})} h_s^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\ &= \int_{\{\boldsymbol{x} \in \mathbb{X}: \, h_s^*(\boldsymbol{x}) > 0\}} \frac{\mathbb{1}_{\mathbb{F}_s}(\boldsymbol{x})}{\mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x})} h_s^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \, p_{s-1} \\ &= p_{s|s-1} \, p_{s-1}, \end{aligned} \tag{3.104}$$

where:

$$p_{s|s-1} = \int_{\{\boldsymbol{x} \in \mathbb{X}: \, h_s^*(\boldsymbol{x}) > 0\}} \frac{\mathbb{1}_{\mathbb{F}_s}(\boldsymbol{x})}{\mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x})} h_s^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \tag{3.105}$$

The first splitting of the optimal instrumental PDF is now completed. However, in order to implement an importance sampling scheme that uses h_s^* as instrumental PDF one first needs to compute its normalizing constant p_{s-1} in Eq. (3.102) which is another possibly low probability. Hence, a Monte Carlo estimation of p_{s-1} might again be rather inefficient. This is the reason why it is proposed to substitute Eq. (3.102) for p_{s-1} in Eq. (3.104) so that it comes:

$$\begin{aligned} p_f &= p_{s|s-1} \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \\ &= p_{s|s-1} \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\, h_{s-1}^*(\boldsymbol{x})>0\right\}} \mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \frac{h_{s-1}^*(\boldsymbol{x})}{h_{s-1}^*(\boldsymbol{x})} \,\mathrm{d}\boldsymbol{x} \\ &= p_{s|s-1} \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\, h_{s-1}^*(\boldsymbol{x})>0\right\}} \mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \frac{p_{s-2}}{\mathbb{1}_{\mathbb{F}_{s-2}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})} h_{s-1}^*(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \\ &= p_{s|s-1} \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\, h_{s-1}^*(\boldsymbol{x})>0\right\}} \frac{\mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x})}{\mathbb{1}_{\mathbb{F}_{s-2}}(\boldsymbol{x})} h_{s-1}^*(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}\, p_{s-2} \\ &= p_{s|s-1}\, p_{s-1|s-2}\, p_{s-2}, \end{aligned} \tag{3.106}$$

where:

$$p_{s-1|s-2} = \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\, h_{s-1}^*(\boldsymbol{x})>0\right\}} \frac{\mathbb{1}_{\mathbb{F}_{s-1}}(\boldsymbol{x})}{\mathbb{1}_{\mathbb{F}_{s-2}}(\boldsymbol{x})} h_{s-1}^*(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}. \tag{3.107}$$

The normalizing constant p_{s-2} might again be too low to be efficiently estimated using Monte Carlo sampling so that it should be split again. Thus after $s-1$ more splits the initial failure probability rewrites as follows:

$$p_f = p_1 \prod_{i=2}^{s} p_{i|i-1}, \tag{3.108}$$

where the intermediate probabilities read as follows:

$$p_1 = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_1}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \equiv \mathbb{P}\left[\boldsymbol{X} \in \mathbb{F}_1\right], \tag{3.109}$$

$$p_{i|i-1} = \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\, h_i^*(\boldsymbol{x})>0\right\}} \frac{\mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x})}{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})} h_i^*(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}, \quad i = 2, \ldots, s. \tag{3.110}$$

The latter equation further reduces to the following simpler expression:

$$p_{i|i-1} = \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\, h_i^*(\boldsymbol{x})>0\right\}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x}) h_i^*(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \equiv \mathbb{P}\left[\boldsymbol{X} \in \mathbb{F}_i \,\middle|\, \boldsymbol{X} \in \mathbb{F}_{i-1}\right], \quad i = 2, \ldots, s, \tag{3.111}$$

by pointing out that:

$$h_i^*(\boldsymbol{x}) \neq 0 \quad \Rightarrow \quad \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) \neq 0 \quad \Leftrightarrow \quad \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) = 1, \quad i = 2, \ldots, s, \tag{3.112}$$

by definition (see Eq. (3.98) and (3.101)).

3.3.5.2 Adaptive splitting

The initial problem of estimating the failure probability p_f now consists in finding an *optimal* sequence of positive reals $q_1 > \ldots > q_s = 0$ such that the intermediate probabilities can be efficiently estimated by means of Monte Carlo sampling. Note that the q_i's are basically quantiles of the random variable G. Even though these quantiles could be specified, a better practice consists in estimating them by fixing the intermediate probabilities to a reasonably high value p_0 (*e.g.* $p_0 = 10\%$). Amongst statisticians, this technique is known as *adaptive splitting* whereas it was called *subset sampling* in the structural reliability community. It proceeds as follows.

First step

The first p_0-quantile q_1 is estimated from a sample of performance function values $\mathscr{G}^{[1]} = \{\mathfrak{g}(\boldsymbol{X}^{(k)}),\, k = 1, \ldots, N\}$ computed from a sample $\mathscr{X}^{[1]} = \{\boldsymbol{X}^{(k)},\, k = 1, \ldots, N\}$ containing N independent copies of the random vector $\boldsymbol{X}$. To be more specific, it is the $\lceil p_0 N \rceil$-th value of $\mathscr{G}^{[1]}$ sorted in ascending order, where $\lceil \bullet \rceil$ denotes the ceiling function (*i.e.* $\lceil x \rceil$ is the smallest integer greater than or equal to $x \in \mathbb{R}$). Once the first quantile q_1 is estimated, the corresponding probability and its coefficient of variation are computed as follows:

$$\widehat{p}_1 = \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{\mathbb{F}_1} \left(\boldsymbol{X}^{(k)} \right), \quad \mathbb{E}\left[\widehat{p}_1 \right] = p_0, \tag{3.113}$$

$$\delta_1 = \sqrt{\frac{1 - p_1}{N\, p_1}}. \tag{3.114}$$

i-th step ($i = 2, \ldots, s$)

The other p_0-quantiles $q_2 > \ldots > q_s = 0$ are estimated in a similar fashion from conditional samples that are distributed according to the sub-optimal instrumental PDFs $\{h_i^*,\, i = 2, \ldots, s\}$. Let $\boldsymbol{X}^{[i]}$ denote the i-th random vector that is distributed according to the i-th instrumental PDF h_i^*, and $\mathscr{X}^{[i]} = \{\boldsymbol{X}^{[i](k)},\, k = 1, \ldots, N\}$ denote a sample of N copies of the random vector $\boldsymbol{X}^{[i]}$.

Note that sampling from the i-th instrumental PDF h_i^* is not straightforward due to its definition. Indeed, recall that it is defined as follows:

$$h_i^*(\boldsymbol{x}) \propto \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}), \tag{3.115}$$

where the normalizing constant p_{i-1} is omitted as it is useless in the present sampling context. Despite an accept-reject sampling technique could be considered with $f_{\boldsymbol{X}}$ as a proposal PDF, it would be particularly inefficient as there is little chance to draw a sample from $f_{\boldsymbol{X}}$ such that the i-th indicator function $\mathbb{1}_{\mathbb{F}_i}$ is not zero. Chances even decrease along with the steps of the algorithm as the subsets move towards regions with lesser probability content.

Thus Au and Beck (2001) propose instead to resort to a Markov Chain Monte Carlo (MCMC) sampling technique. This technique, known as *modified Metropolis-Hastings* (Au and Beck, 2001) is further detailed in Section B.3.2 of Appendix B. For now, assume that a sample $\mathscr{X}^{[i]}$ has been generated from a sample of seeds, *i.e.* from a sample that is distributed according to the target PDF h_i^*. In practice, the MCMC sampler is initiated with the $\lceil p_0 N \rceil$ seeds contained in the previous sample $\mathscr{X}^{[i-1]}$ for which the indicator function $\mathbb{1}_{\mathbb{F}_{i-1}} = 1$, hence for which h_i^* is not zero.

A set of performance function values $\mathscr{G}^{[i]} = \{\mathfrak{g}(\boldsymbol{X}^{[i](k)}),\, k = 1, \ldots, N\}$ is computed from the sample $\mathscr{X}^{[i]}$. Note that the performance function values are already available as they were computed during the sampling of $\boldsymbol{X}^{[i]}$ in order to evaluate the indicator function $\mathbb{1}_{\mathbb{F}_{i-1}}$ defining the i-th instrumental PDF h_i^*. The i-th quantile q_i is then estimated as the $\lceil p_0 N \rceil$-th value of $\mathscr{G}^{[i]}$ sorted in ascending order. The corresponding i-th probability and its coefficient of variation are computed as follows:

$$\widehat{p}_{i|i-1} = \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\mathbb{F}_i}\left(\boldsymbol{X}^{[i](k)}\right), \quad \mathbb{E}\left[\widehat{p}_{i|i-1}\right] = p_0, \tag{3.116}$$

$$\delta_i = \sqrt{\frac{1-p_{i|i-1}}{N\, p_{i|i-1}}\,(1+\gamma_i)}. \tag{3.117}$$

where γ_i is a coefficient that accounts for the correlation between the N random vectors that constitute the i-th sample when it is generated by means of an MCMC sampler as detailed in Section B.3.2 of Appendix B. The reader is referred to this appendix for the expression of the coefficient γ_i together with the proof of the latter equation.

Last step

The last step is determined depending on the sign of the quantile obtained at the i-th step ($i = 1, \ldots, s$). Indeed the procedure is stopped when a quantile q_i is negative, then $s = i$ and q_s is replaced by the critical threshold: $q_s \equiv 0$. The corresponding probability and its coefficient of variation are computed as follows:

$$\widehat{p}_{s|s-1} = \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\mathbb{F}_s}\left(\boldsymbol{X}^{[s](k)}\right), \quad \mathbb{E}\left[\widehat{p}_{s|s-1}\right] \geq p_0, \tag{3.118}$$

$$\delta_{s|s-1} = \sqrt{\frac{1-p_{s|s-1}}{N\, p_{s|s-1}}\,(1+\gamma_s)}. \tag{3.119}$$

Final estimator

Eventually the estimator of the failure probability defined in Eq. (3.108) is obtained as follows:

$$\widehat{p}_{f\,\mathrm{SS}} = \widehat{p}_1 \prod_{i=2}^{s} \widehat{p}_{i|i-1}. \tag{3.120}$$

Au and Beck (2001) point out that this estimator is biased for every N but *asymptotically unbiased*. This bias comes from the way the modified Metropolis-Hastings algorithm is seeded at each step. Indeed, the procedure resorts to the $\lceil p_0 N \rceil$ lowest values in the $(i-1)$-th sample $\mathscr{X}^{[i-1]}$ in order to seed the next i-th sample $\mathscr{X}^{[i]}$ so that the successive conditional samples are dependent. As a consequence, the intermediate probability estimators are also dependent and the estimator in Eq. (3.120) is biased. However the longer the chains, the smaller the bias.

In order to assess convergence (bias and variance reduction), Au and Beck (2001) provide the following bounds on the coefficient of variation:

$$\sqrt{\sum_{i=1}^{s} \delta_i^2} \leq \delta_{\text{SS}} \leq \sqrt{\sum_{i=1}^{s} \sum_{j=1}^{s} \delta_i \, \delta_j}. \tag{3.121}$$

The lower (resp. upper) bound is derived under the assumption that the intermediate probability estimates are independent (resp. fully correlated). The reader is referred to the original article by Au and Beck (2001) for the two propositions and the proof of Eq. (3.121). It is worth mentioning here that Au et al. (2007) showed on several examples that the *empirical coefficient of variation*[4] is usually closer to the lower bound than the upper one.

Algorithm 3.1 is provided so as to summarize the adaptive splitting algorithm detailed above. The computation of the coefficient of variation is not mentioned for the sake of clarity.

3.3.5.3 Tuning of the parameters

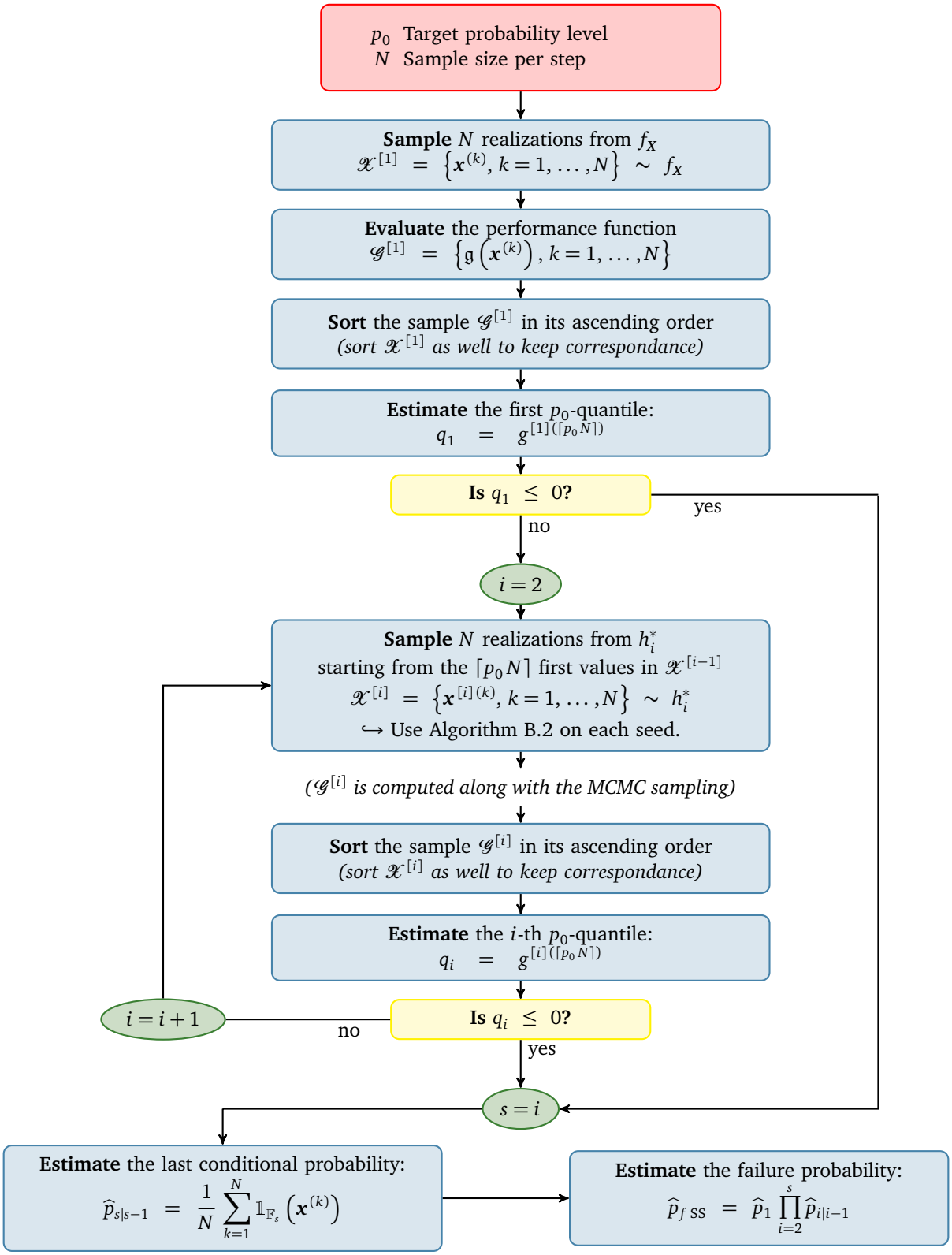
The problem of tuning the parameters involved in the subset sampling technique was investigated by Au and Beck (2001) and Au et al. (2007).

Target probability level p_0

First, the algorithm requires the choice of a target probability level p_0 to define the conditional quantiles. The optimal choice for this value is essentially a compromise between the sample size per step N and the number of steps s. Indeed, if p_0 is too low then the algorithm will converge within a small number of steps s although it will require a large sample size per step N to ensure that the intermediate coefficient of variations δ_i are sufficiently low. On the contrary, if p_0 is too large then the algorithm will require a smaller sample size per step N but more steps to converge. Hence the lower bound on the coefficient of variation in Eq. (3.121) will suffer from the total number of steps due to the correlation between the intermediate probabilities. Au and Beck (2001) found that $p_0 = 0.1$ yields good efficiency in practice. As a conclusion, $p_0 = 0.1$ is the value that is most currently used in practice.

[4] The coefficient of variation estimated on a sample of failure probabilities obtained from several independent runs of the present subset sampling algorithm.

Algorithm 3.1 Adaptive splitting (or subset sampling)



Minimum sample size and optimal intermediate probabilities

The choice for the minimum sample size per step, say $N_{\min}$, is also non-trivial and it actually depends on the target probability level p_0. Assume that the value of the coefficients $\{\gamma_i, i = 2, \dots, s\}$ is equal to a constant value γ_0. Thus the average lower bound on the squared coefficient of variation becomes:

$$\delta_{\text{SS}}^2 \geq \frac{1-p_0}{N\,p_0} + \sum_{i=2}^{s} (1+\gamma_0)\,\frac{1-p_0}{N\,p_0} = \frac{1-p_0}{N\,p_0}\left[1+(s-1)\,(1+\gamma_0)\right]. \tag{3.122}$$

Hence, if a coefficient of variation δ_0 is targeted, the minimum sample size per step is:

$$N_{\min} = \frac{1-p_0}{\delta_0^2\,p_0}\left[1+(s-1)\,(1+\gamma_0)\right]. \tag{3.123}$$

Assume also that the failure probability is equal to the product of s intermediate probabilities all equal to p_0, *i.e.* $p_f = p_0^s$ so that one can substitute $\log(p_f)/\log(p_0)$ for s in the latter equation:

$$N_{\min} = \frac{1-p_0}{\delta_0^2\,p_0}\left[1+\left(\frac{\log(p_f)}{\log(p_0)}-1\right)(1+\gamma_0)\right]. \tag{3.124}$$

Using the common values $p_0 = 0.1$, $\delta_0 = 0.1$ and $\gamma_0 = 3$ yields a fair estimate of the minimum sample size per step depending on the order of magnitude of the failure probability.

Figure 3.14 depicts the one-to-one relationship between the total number of simulations $N_{\text{tot}} = \log(p_f)/\log(p_0)N_{\min}$ and p_0 for different orders of magnitude of the failure probability (10^{-s}, $s = 3, \dots, 10$) when $\gamma_0 = 3$ and $\delta_0 = 0.1$. It is clear from this illustration that $p_0 \in [0.1;\ 0.2]$ is a good compromise for efficiently estimating a wide range of low failure probabilities.

MCMC sampler parameters

Finally, the MCMC sampler used to sample the sub-optimal instrumental PDFs also involves a set of parameters that are rather difficult to tune in the physical space. Au and Beck (2001) found out that the modified Metropolis-Hastings sampler is a lot simpler to tune in the standard space. Indeed, the margin variables in the random vector $\boldsymbol{U}$ are identically distributed standard Gaussian random variates. Therefore they all have the same order of magnitude so that it eases the choice of the *proposal PDF* (see Section B.3.2 of Appendix B). Moreover, these variables are also independent. The independence of the margins is actually the key feature of the so-called modified Metropolis-Hastings algorithm which is nothing but a *two-step componentwise* Metropolis-Hastings algorithm.

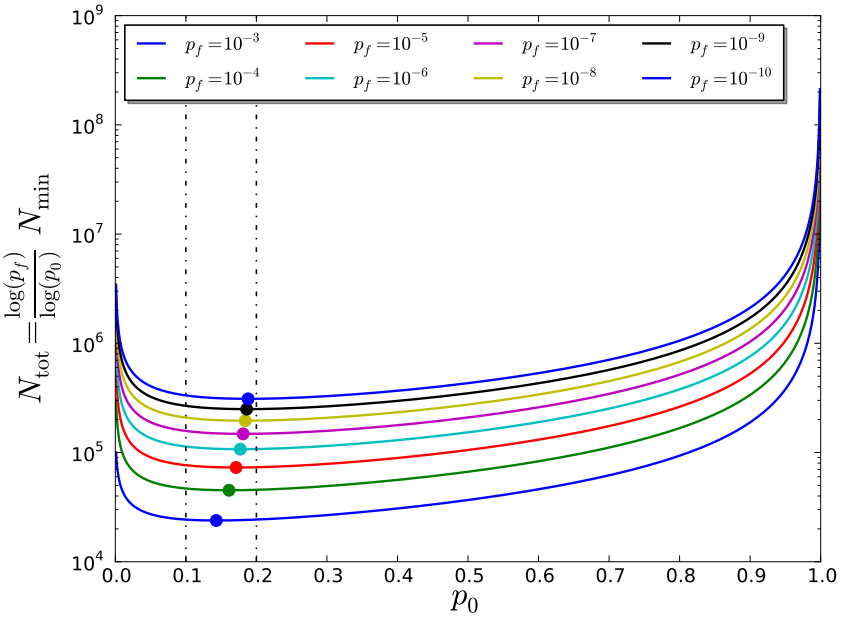


Figure 3.14: One-to-one relationship between the total number of simulations $N_{\text{tot}} = \log(p_f)/\log(p_0) N_{\text{min}}$ and the intermediate probability level p_0 computed by means of Eq. (3.124) for $\gamma_0 = 3$ and $\delta_0 = 0.1$. The dots represent the minimum of each curve. Note that all minima belong to the optimal range of p_0 values provided by Au and Beck (2001).

3.3.5.4 Illustration

For illustration purposes, the subset sampling technique is applied to the basic structural reliability example defined in Section 3.2.4. Since the order of magnitude of the failure probability is known from the previous reliability analyses ($p_f \approx 10^{-6}$), Eq. (3.124) is used to estimate the minimum sample size per step $N_{\text{min}} \approx 20,000$.

Actually, the overall subset sampling procedure required only 112, 317 calls to the performance function. It is less than the $6 \times 20,000$ samples required due to the present implementation of the MCMC sampler. It performs 20, 000 propositions per step, some of which are rejected as they do not seem to be distributed according to f_X (see Section B.3.2 of Appendix B for further details). The failure probability estimate is finally equal to $\widehat{p}_{f\,\text{SS}} \approx 1.45 \times 10^{-6}$ and the coefficient of variation is comprised between 10% and 22% (see Eq. (3.121)). Repeating the estimation 100 times with the same setup leads to an empirical coefficient of variation $\widehat{\delta}_{\text{SS}} \approx 13\%$ which is closer to the lower bound than the upper one. Despite it cannot be formally proved (see *e.g.* Bourinet et al., 2011, for an example where the empirical coefficient of variation is closer to the upper bound than the lower one), the applications presented in this manuscript will always be justified with that lower bound.

The variance reduction offered by subset sampling is really significant on this example. Indeed, for the same coefficient of variation (10%), crude Monte Carlo sampling requires 1, 000 times more calls to the performance function ($\approx 10^8$) than subset sampling ($\approx 10^5$).

It is more computationally demanding than the MPFP-based approaches, but it has a larger field of application as it does not make any assumption regarding the properties of the performance function $\mathfrak{g}$.

Figure 3.16 depicts the 6 conditional samples together with the intermediate limit-states corresponding to the sequence of quantiles $q_1 > \ldots > q_6 = 0$ in the standard space so as to illustrate how subset sampling works. The sample size per step has been reduced to $N = 1,000$ for the sake of clarity. Note that h_6^* in Figure 3.16(f) is close (although not equal) to the optimal instrumental PDF illustrated in Figure 3.8(b).

Figure 3.15 represents the empirical CDF of the random variable G reconstructed from the 6 conditional samples (of size $N = 20,000$) together with its confidence bounds at ± 2 standard deviations. The computation of this empirical CDF is defined *piecewise* from the empirical CDFs of all the 6 sub-samples. The first CDF is truncated below the first quantile q_1. The i-th empirical CDF is truncated above by the previous quantile $q - 1$ by definition. It is weighted by the normalizing constant $\widehat{p}_{i-1} = \widehat{p}_1 \widehat{p}_{2|1} \ldots \widehat{p}_{i-1|i-2}$ and truncated below the next quantile q_i. The last empirical CDF is not truncated below $q_s = 0$. The availability of this empirical CDF is another interesting feature of subset sampling.

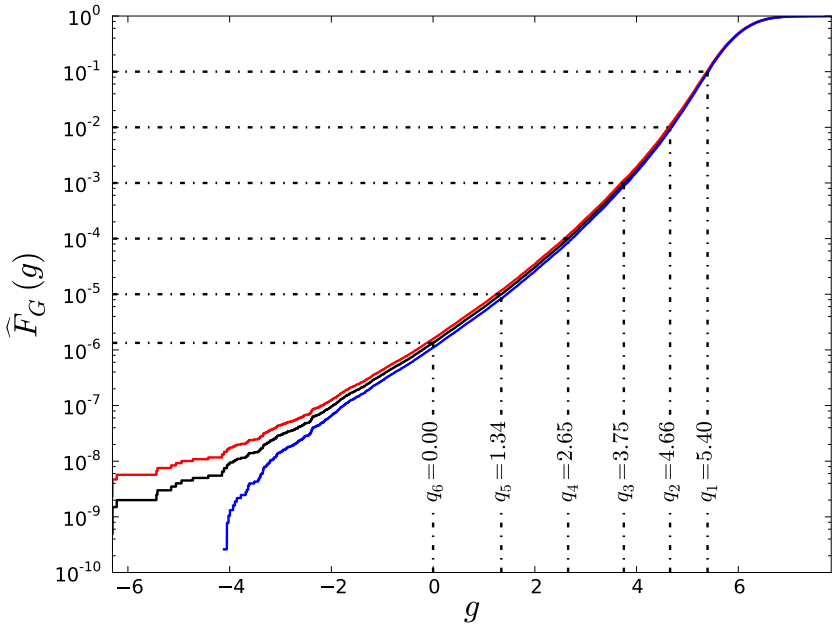
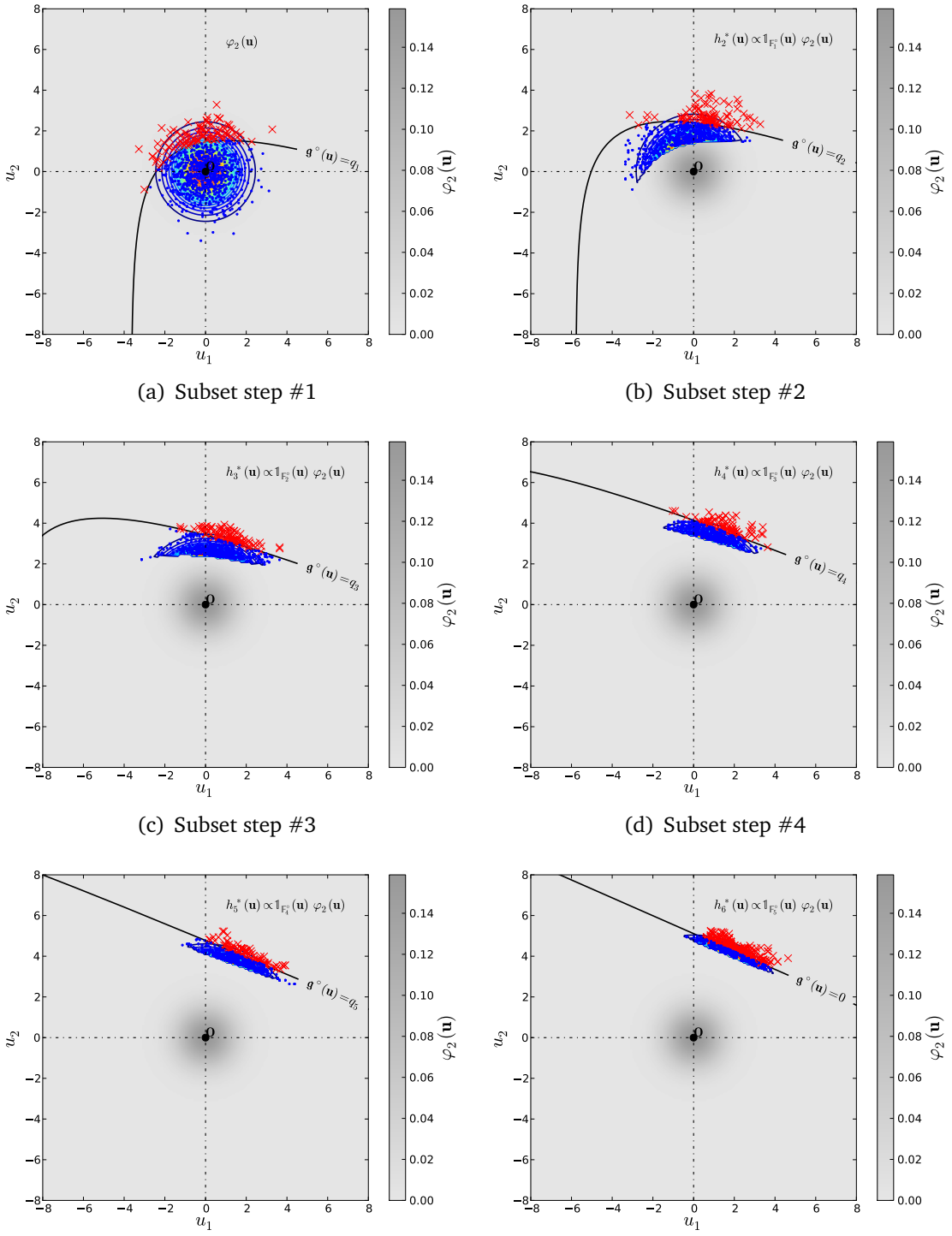


Figure 3.15: Empirical CDF reconstructed from all the sub-samples obtained by subset sampling for the basic structural reliability example defined in Section 3.2.4.



(a) Subset step #1

(b) Subset step #2

(c) Subset step #3

(d) Subset step #4

(e) Subset step #5

(f) Subset step #6

Figure 3.16: Subset sampling applied to the basic structural reliability example defined in Section 3.2.4. Each panel represents a step of the algorithm. The very first step uses a sample generated from the original PDF φ_2. The other steps use the sub-optimal instrumental PDFs h_i^* to generate more samples from the seeds obtained in the previous step (red crosses). The procedure is stopped at the sixth step because the corresponding quantile q_6 is negative.

3.3.6 Conclusion

Before concluding, it is worth mentioning that there exists many other techniques for estimating rare event probabilities in addition to the ones reviewed in this section such as directional sampling (Bjerager, 1988), line sampling (Koutsourelakis et al., 2004; Schuëller et al., 2004; Schuëller and Pradlwarter, 2007), cross-entropy methods (see *e.g.* Rubinstein and Kroese, 2008, Chapter 8). Note also that the MPFP-based approaches as well as the subset sampling technique have been revisited here from the importance sampling viewpoint. This original interpretation inspired the techniques exposed in the two remaining sections of this chapter.

3.4 Surrogate-based reliability analysis

In the previous section, it has been shown that MPFP-based strategies may lead to inaccurate results in the case of multiple failure modes. It was also emphasized that these strategies do not provide any self-quantification of the possible error. In addition, the gradient-based optimization algorithm used to find the MPFP might be hard to tune when the gradients are approximated by means of a finite difference scheme (due to the possible noise in the performance function).

Despite the more robust variance reduction strategies (especially subset sampling) offer a significant reduction of the computational cost induced by a reliability analysis, they still require a few hundreds or thousands of calls to the performance function. Hence, sampling-based reliability analyses remain intractable when the performance function involves the output of an expensive-to-evaluate computational model.

Starting from these two observations, the so-called *surrogate-based strategies* are now reviewed.

3.4.1 Principle

In essence, surrogate-based strategies consist in using a surrogate $\widetilde{\mathfrak{g}}$ for the performance function $\mathfrak{g}$ that is much faster to evaluate. Such a surrogate might be the analytical counterpart of a high-fidelity numerical model derived on simplifying assumptions. Nevertheless, it is not always available or it may introduce arbitrarily large conservatism that would eventually lead to an overestimation of the actual failure probability. Hence, meta-modelling (see Chapter 1) reveals a powerful tool in this context.

A surrogate-based estimator of the failure probability is simply defined as the following plug-in estimator:

$$\widetilde{p}_f \equiv \mathbb{P}\left[\boldsymbol{X} \in \widetilde{\mathbb{F}}\right] = \int_{\widetilde{\mathbb{F}}=\{\boldsymbol{x}\in\mathbb{X}:\,\widetilde{\mathfrak{g}}(\boldsymbol{x})\leq 0\}} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \tag{3.125}$$

where $\widetilde{\mathbb{F}} = \{\boldsymbol{x} \in \mathbb{X} : \widetilde{\mathfrak{g}}(\boldsymbol{x}) \leq 0\}$ is the surrogate-based prediction of the actual failure domain $\mathbb{F}$.

Meta-model-based reliability analysis has gained a large attention amongst researchers in the past two decades. So far, the number of meta-model-based strategies is as large as the number of meta-models and conventional reliability methods available in the literature so that it is hard to make an exhaustive state-of-the-art. To cite a few though, Bucher and Bourgund (1990); Kim and Na (1997); Das and Zheng (2000); Gayton et al. (2003); Micol (2007) investigate the use of *quadratic response surfaces* whose coefficients are obtained by a least-square fit as detailed in Chapter 1. *Sparse polynomial chaos expansions* were also proved to provide significant results for reliability analysis (Blatman and Sudret, 2008a; Blatman, 2009; Blatman and Sudret, 2010b).

Hurtado and Alvarez (2000); Papadrakakis and Lagaros (2002) resort to *neural networks* which is another sort of kernel-based meta-model that was not considered in this thesis. Hurtado (2004a,b, 2007); Deheeger and Lemaire (2007); Deheeger (2008); Piera-Martìnez et al. (2007); Piera-Martìnez (2008); Basudhar and Missoum (2010); Bourinet et al. (2011) have recourse to *support vector margin classifiers*. Eventually, the use of *Gaussian process (or Kriging) predictors* has been explored by Kaymaz (2005); Bichon et al. (2008); Bichon (2010); Picheny (2009); Picheny et al. (2010a,b); Echard et al. (2011); Bect et al. (2011) and is further dealt with in this manuscript.

A large majority of these contributions makes use of Monte Carlo sampling on the surrogate in order to estimate the quantity in Eq. (3.125) although this can reveal rather inefficient. Indeed, recall that if the failure probability is in the order of magnitude of 10^{-k} ($2 \leq k \leq 12$), then the minimum sample size for a 10% coefficient of variation using Monte Carlo sampling is $N_{\min} = 10^{k+2}$. Such a high number of evaluations, even on a supposedly inexpensive-to-evaluate surrogate, is still quite computationally demanding. For this reason, Gayton et al. (2003); Kaymaz (2005) resort to MPFP-based approaches. Bichon et al. (2008); Bichon (2010) use an adaptive importance sampling scheme that accounts for multiple design points. Hurtado (2007); Piera-Martìnez et al. (2007); Piera-Martìnez (2008) use a filtered importance sampling scheme that will be further discussed in Section 3.5. Eventually, in this thesis as in Deheeger and Lemaire (2007); Deheeger (2008); Bourinet et al. (2011) the estimation makes use of subset sampling in order to make it efficient when dealing with low failure probabilities.

3.4.2 Error quantification

A fundamental problem of surrogate-based reliability analysis though is that it might introduce a *bias* in the failure probability estimate when the surrogate is not sufficiently accurate. Therefore, this section focuses on the specific problem of apprehending this bias when a kernel-based predictor is used to surrogate the performance function.

In Chapter 1, it has been shown that kernel-based predictors (namely *support vector margin classifiers* and *Kriging*) provide a measure of their local accuracy in addition to their prediction. Indeed, recall that support vector margin classifiers provide the boundaries of their *margin* while Kriging offers a full probabilistic prediction in the form of the *probabilistic classification function* π.

More specifically, the "confidence interval" measure provided by the margin of a support vector classifier can be mapped into a full probabilistic prediction. Indeed, probabilistic support vector margins (Platt, 1999) provides a probabilistic classification function that has similar properties to its Kriging counterpart. Conversely, the probabilistic response of the Kriging prediction can be used to construct confidence intervals on the limit-state surface as introduced in Section 1.5.3 of Chapter 1. In particular, the *margin of uncertainty* $\mathbb{M}_{1-\alpha}$ defined in Section 1.5.3 of Chapter 1 closely resembles the margin of a support vector classifier.

From now on, it is assumed that a Kriging predictor is used to surrogate the performance function $\mathfrak{g}$. It is denoted by $\widehat{Y}(\boldsymbol{x}) \sim \mathcal{N}_1(\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma^2_{\widehat{Y}}(\boldsymbol{x}))$ and has probability measure $\mathscr{P}$.

Recall that $\mathscr{P}$ denotes the probability measure associated with the meta-modelling epistemic *uncertainty, whereas $\mathbb{P}$ denotes the one corresponding to the* aleatoric *uncertainty in the random input $\boldsymbol{X}$.*

Its probabilistic classification function is defined as in Section 1.5.3 of Chapter 1 and it reads as follows:

$$\pi(\boldsymbol{x}) = \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq t\right] = \Phi\left(\frac{t - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \quad t \in \mathbb{Y}. \tag{3.126}$$

This definition is extended to the points in the dataset $\mathscr{D} = \{(\boldsymbol{x}^{(i)}, y_i), i = 1, \ldots, m\}$ for which the kriging variance equals zero by switching to the limit as follows:

$$\pi(\boldsymbol{x}^{(i)}) = \Phi\left(\frac{t - y_i}{0}\right) = \begin{cases} \Phi(-\infty) = 0 & \text{if } y_i > t \\ \Phi(+\infty) = 1 & \text{if } y_i \leq t \end{cases}, \quad i = 1, \ldots, m, \quad t \in \mathbb{Y}. \tag{3.127}$$

The Kriging-based prediction of the failure domain is denoted by $\widehat{\mathbb{F}}^{\,0}_{1-\alpha} = \{\boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) \leq 0\}$ and the corresponding *optimistic* (resp. *conservative*) prediction of the failure domain at confidence level $1-\alpha$ are defined as follows:

$$\widehat{\mathbb{F}}^{-1}_{1-\alpha} = \left\{\boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) \leq t - k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x})\right\}, \tag{3.128}$$

$$\widehat{\mathbb{F}}^{+1}_{1-\alpha} = \left\{\boldsymbol{x} \in \mathbb{X} : \mu_{\widehat{Y}}(\boldsymbol{x}) \leq t + k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x})\right\}, \quad t \in \mathbb{Y}, \tag{3.129}$$

where $k_{1-\alpha}$ is the number of Kriging standard deviations corresponding to the selected confidence level $1-\alpha$. Recall that $1-\alpha$ is set to 95% in this manuscript so that $k_{1-\alpha} = 1.96$.

3.4.2.1 Augmented failure probability

Definition

Picheny (2009); Picheny et al. (2010a); Bect et al. (2011) propose to use the probabilistic classification function π as a surrogate for the actual failure indicator function $\mathbb{1}_{\mathbb{F}}$

in Eq. (3.41). The corresponding plug-in estimator of the failure probability is therefore defined as follows:

$$p_{f\,\varepsilon} = \int_{\mathbb{X}} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \equiv \mathbb{E}_{\boldsymbol{X}}\left[\pi(\boldsymbol{X})\right]. \tag{3.130}$$

where the definition of the probabilistic classification function has been recalled in Eq. (3.126). This estimator accounts for the residual uncertainty in the Kriging prediction of the actual failure domain. However it condensates the whole information in a single value. Indeed $p_{f\,\varepsilon} = \mathbb{P}[\mathscr{P}[\widehat{Y}(\boldsymbol{X}) \leq 0]]$ somehow sums up both types of uncertainties so that it is hard to tell how far is this estimate from the real failure probability p_f. This is the reason why it will be referred to as the *augmented failure probability* in the sequel.

Estimation using crude Monte Carlo sampling

Picheny (2009); Picheny et al. (2010a); Bect et al. (2011) resort to Monte Carlo sampling of $\boldsymbol{X}$ to provide a fair estimation of the augmented failure probability. Given a sample $\mathscr{X} = \{\boldsymbol{X}^{(k)},\, k = 1, \ldots, N\}$ of N independent copies of the random vector $\boldsymbol{X}$, the Monte Carlo estimator reads as follows:

$$\widehat{p}_{f\,\varepsilon\,\mathrm{MCS}} = \widehat{\mathbb{E}}_{\boldsymbol{X}}\left[\pi(\boldsymbol{X})\right] = \frac{1}{N}\sum_{k=1}^{N} \pi\left(\boldsymbol{X}^{(k)}\right). \tag{3.131}$$

The central limit theorem ensures that this quantity is an unbiased Gaussian estimator of the augmented failure probability whose variance can be approximated as follows:

$$\sigma^2_{\varepsilon\,\mathrm{MCS}} \approx \frac{1}{N}\left(\frac{1}{N}\sum_{k=1}^{N} \pi\left(\boldsymbol{X}^{(k)}\right)^2 - p^2_{f\,\varepsilon}\right). \tag{3.132}$$

Estimation using subset sampling

As previously stated in this chapter, crude Monte Carlo sampling may reveal rather inefficient if the augmented failure probability is low. Instead, it is proposed to adapt the subset sampling scheme reviewed earlier in Section 3.3.5 to the problem of estimating the mathematical expectation of π under $f_{\boldsymbol{X}}$.

Recall that before the last simplification, the subset sampling definition of the failure probability in Eq. (3.108) still holds for the augmented failure probability. Actually, provided a strictly decreasing sequence of quantiles $q_1 > \ldots > q_s = 0$, Eq. (3.130) rewrites in the following splitted form:

$$p_{f\,\varepsilon} = p_{1\,\varepsilon} \prod_{i=2}^{s} p_{i|i-1\,\varepsilon}. \tag{3.133}$$

where the intermediate *augmented* probabilities read as follows:

$$p_{1\,\varepsilon} = \int_{\mathbb{X}} \pi_1(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \tag{3.134}$$

$$p_{i|i-1\,\varepsilon} = \int_{\mathbb{X}} \frac{\pi_i(\boldsymbol{x})}{\pi_{i-1}(\boldsymbol{x})}\, h^*_{i\,\varepsilon}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \quad i = 2, \ldots, s. \tag{3.135}$$

The latter expression involves the collection of *approximate* intermediate indicator functions that are defined as follows:

$$\pi_i(\boldsymbol{x}) = \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq q_i\right] = \Phi\left(\frac{q_i - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right), \quad i = 1, \ldots, s, \tag{3.136}$$

as well as the *approximate* sub-optimal instrumental PDFs that read:

$$h^*_{i\,\varepsilon}(\boldsymbol{x}) = \frac{\pi_{i-1}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\mathbb{X}} \pi_{i-1}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}}, \quad i = 2, \ldots, s. \tag{3.137}$$

Then, practical use of this splitting is as follows.

(i) The optimal decreasing sequence of quantiles $q_1 > \ldots > q_s = 0$ is determined by applying the adaptive splitting technique detailed in Section 3.3.5.2 on the mean prediction $\mu_{\widehat{Y}}$ of the Kriging predictor with intermediate probabilities set to p_0 (say $p_0 = 10\%$).

(ii) The first step of subset sampling is then performed without determining the first quantile q_1. It rather uses the one determined in the preliminary step *(i)* to compute:

$$\widehat{p}_{1\,\varepsilon} = \frac{1}{N} \sum_{k=1}^{N} \pi_1\left(\boldsymbol{X}^{(k)}\right), \tag{3.138}$$

$$\delta_{1\,\varepsilon} = \sqrt{\frac{1 - p_{1\,\varepsilon}}{N\, p_{1\,\varepsilon}}}, \tag{3.139}$$

where $\mathscr{X}^{[1]} = \{\boldsymbol{X}^{(k)},\, k = 1, \ldots, N\}$ is a sample of N independent copies of the random vector $\boldsymbol{X}$.

(iii) The other $s - 1$ steps are eventually performed by sampling N copies of the random vectors $\boldsymbol{X}_{\varepsilon}^{[i]}$ from their respective approximate sub-optimal instrumental PDFs $\{h^*_{i\,\varepsilon},\, i = 2, \ldots, s\}$. This sampling again resorts to the modified Metropolis-Hastings algorithm (see Section B.3.2 of Appendix B). Note that the Markov chains can be seeded with the points in the i-th sample generated in the preliminary step. Indeed, these samples are such that $h^*_{i\,\varepsilon}$ is non-zero because of the following implication:

$$\mu_{\widehat{Y}}(\boldsymbol{x}) \leq q_{i-1} \quad \Rightarrow \quad \pi_{i-1}(\boldsymbol{x}) = \Phi\left(\frac{q_{i-1} - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right) \geq 0.5, \quad i = 2, \ldots, s. \tag{3.140}$$

The estimation of the intermediate augmented probabilities uses the quantiles $\{q_i,\ i = 2,\dots,s\}$ determined in the preliminary step *(i)*. These probabilities (resp. their coefficient of variation) are computed as follows:

$$\widehat{p}_{i|i-1\,\varepsilon} = \frac{1}{N}\sum_{k=1}^{N}\frac{\pi_i\left(\boldsymbol{X}_{\varepsilon}^{[i](k)}\right)}{\pi_{i-1}\left(\boldsymbol{X}_{\varepsilon}^{[i](k)}\right)}, \tag{3.141}$$

$$\delta_{i|i-1\,\varepsilon} \approx \frac{1}{p_{i|i-1\,\varepsilon}}\sqrt{\frac{1}{N}\left[\frac{1}{N}\sum_{k=1}^{N}\frac{\pi_i\left(\boldsymbol{X}_{\varepsilon}^{[i](k)}\right)^2}{\pi_{i-1}\left(\boldsymbol{X}_{\varepsilon}^{[i](k)}\right)^2} - p_{i|i-1\,\varepsilon}^2\right]\left(1+\gamma_{i\,\varepsilon}\right)}, \tag{3.142}$$

where $\{\boldsymbol{X}_{\varepsilon}^{[i](k)},\ k = 1,\dots,N\}$ is a sample of N copies of the i-th instrumental random vector $\boldsymbol{X}^{[i]}$, and $\gamma_{i\,\varepsilon}$ is the coefficient that accounts for the correlation in the MCMC sample (see Section B.2.3 of Appendix B).

Eventually, the subset sampling estimator of the augmented failure probability is evaluated as follows:

$$\widehat{p}_{f\,\varepsilon\,\mathrm{SS}} = \widehat{p}_{1\,\varepsilon}\prod_{i=2}^{s}\widehat{p}_{i|i-1\,\varepsilon}, \tag{3.143}$$

and its coefficient of variation satisfies:

$$\sqrt{\sum_{i=1}^{s}\delta_{i\,\varepsilon}^2} \le \delta_{\varepsilon\,\mathrm{SS}} \le \sqrt{\sum_{i=1}^{s}\sum_{j=1}^{s}\delta_{i\,\varepsilon}\,\delta_{j\,\varepsilon}}. \tag{3.144}$$

3.4.2.2 Pseudo-confidence bounds

Definition

Another solution to quantify the error introduced by the use of a surrogate instead of the real performance function is inspired from the work by Deheeger and Lemaire (2007); Deheeger (2008) on support vector margin classifiers. It basically consists in estimating the three following probabilities:

$$p_{f\,(1-\alpha)}^{\,i} = \mathbb{P}\left[\boldsymbol{X} \in \widehat{\mathbb{F}}_{1-\alpha}^{\,i}\right], \quad i = +1,\ 0,\ -1. \tag{3.145}$$

Recall that thanks to the positiveness of the Kriging standard deviation $\sigma_{\widehat{Y}}$, the three approximate failure domains satisfy the following imbrication:

$$\widehat{\mathbb{F}}_{1-\alpha}^{-1} \subseteq \widehat{\mathbb{F}}_{1-\alpha}^{\,0} \subseteq \widehat{\mathbb{F}}_{1-\alpha}^{+1}, \tag{3.146}$$

which in turns ensures that:

$$p_{f\,(1-\alpha)}^{-1} \le p_{f\,(1-\alpha)}^{\,0} \le p_{f\,(1-\alpha)}^{+1}. \tag{3.147}$$

Therefore the spread of the interval $[p^{-1}_{f\,(1-\alpha)};\ p^{+1}_{f\,(1-\alpha)}]$ reveals a useful measure for assessing the accuracy of a Kriging surrogate in a reliability analysis.

However, it is worth mentioning here that there is no proof that the conservative failure domain $\widehat{\mathbb{F}}^{+1}_{1-\alpha}$ contains the actual failure domain $\mathbb{F}$, nor that the latter contains the optimistic failure domain $\widehat{\mathbb{F}}^{-1}_{1-\alpha}$. As a matter of fact, there is also no proof that the actual failure probability p_f belongs to the interval $[p^{-1}_{f\,(1-\alpha)};\ p^{+1}_{f\,(1-\alpha)}]$ even with 95% confidence. Nevertheless the metric defined in Eq. (3.157) reveals a satisfactory error measure in practice.

Estimation using *restarted* subset sampling

For the sake of efficiency it is proposed to use a *restarted* subset sampling algorithm for estimating the three probabilities in Eq. (3.147). Indeed provided the first (highest) probability $p^{+1}_{f\,(1-\alpha)}$ is known, Eq. (3.146) ensures that the other two probabilities rewrite as follows:

$$p^{\ 0}_{f\,(1-\alpha)} = \mathbb{P}\left[\boldsymbol{X} \in \mathbb{F}^{\ 0}_{1-\alpha} \,\middle|\, \boldsymbol{X} \in \mathbb{F}^{+1}_{1-\alpha}\right] p^{+1}_{f\,1-\alpha}, \tag{3.148}$$

$$p^{-1}_{f\,(1-\alpha)} = \mathbb{P}\left[\boldsymbol{X} \in \mathbb{F}^{-1}_{1-\alpha} \,\middle|\, \boldsymbol{X} \in \mathbb{F}^{\ 0}_{1-\alpha}\right] p^{\ 0}_{f\,1-\alpha}. \tag{3.149}$$

Hence, the three probabilities can be efficiently estimated by means of three successive runs of the adaptive splitting algorithm. The estimation proceeds as follows.

$\widehat{p}^{+1}_{f\,1-\alpha}$ is estimated by means of an adaptive splitting procedure on the following conservative prediction of the performance function:

$$\widetilde{\mathfrak{g}}^{+1}_{1-\alpha}(\boldsymbol{x}) = \mu_{\widehat{Y}}(\boldsymbol{x}) - k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x}), \tag{3.150}$$

where $k_{1-\alpha}$ is the number of Kriging standard deviations corresponding to the selected confidence level $1-\alpha$ (see Eq. (3.129)).

$\widehat{p}^{\ 0}_{f\,1-\alpha}$ is estimated by means of a *restarted* adaptive splitting procedure on the following prediction of the performance function:

$$\widetilde{\mathfrak{g}}^{\ 0}_{1-\alpha}(\boldsymbol{x}) = \mu_{\widehat{Y}}(\boldsymbol{x}). \tag{3.151}$$

The restarted adaptive splitting algorithm starts from the last conditional sample: *i.e.* $\mathscr{X}^{\ 0\,[1]} = \mathscr{X}^{+1\,[s]}$. The new performance function values are computed on this sample in order to estimate a new quantile $q_1^{\ 0} \geq q_s^{+1} = 0$. If this quantile is still positive the adaptive splitting algorithm is applied until a negative quantile is found. The final estimate of the quantity of interest is obtained as follows:

$$\widehat{p}^{\ 0}_{f\,(1-\alpha)} = \widehat{p}^{+1}_{\text{cond}} \prod_{i=2}^{s} \widehat{p}^{\ 0}_{i|i-1}, \tag{3.152}$$

where $\widehat{p}^{+1}_{\text{cond}}$ is the second last probability estimate obtained in the latter estimation of $\widehat{p}^{+1}_{f\,1-\alpha}$:

$$\widehat{p}^{+1}_{\text{cond}} = \widehat{p}^{+1}_{1} \prod_{i=2}^{s-1} \widehat{p}^{+1}_{i|i-1}. \tag{3.153}$$

$\widehat{p}_{f\,1-\alpha}^{-1}$ is estimated in a similar fashion. The *restarted* adaptive splitting procedure is performed on the following optimistic prediction of the performance function:

$$\widetilde{\mathfrak{g}}_{1-\alpha}^{-1}(\boldsymbol{x}) = \mu_{\widehat{Y}}(\boldsymbol{x}) + k_{1-\alpha}\,\sigma_{\widehat{Y}}(\boldsymbol{x}), \tag{3.154}$$

where $k_{1-\alpha}$ is the same number of Kriging standard deviations chosen in the very first step. The restarted adaptive splitting algorithm starts from the last conditional sample: *i.e.* $\mathscr{X}^{-1\,[1]} = \mathscr{X}^{\,0\,[s]}$. The new performance function values are computed on this sample in order to estimate a new quantile $q_1^{-1} \geq q_s{}^0 = 0$. If this quantile is still positive the adaptive splitting algorithm is applied until a negative quantile is found. The final estimate of the quantity of interest is obtained as follows:

$$\widehat{p}_{f\,(1-\alpha)}^{-1} = \widehat{p}_{\text{cond}}^{\,0} \prod_{i=2}^{s} \widehat{p}_{i|i-1}^{-1}, \tag{3.155}$$

where $\widehat{p}_{\text{cond}}^{\,0}$ is the second last probability estimate obtained in the latter estimation of $\widehat{p}_{f\,1-\alpha}^{\,0}$:

$$\widehat{p}_{\text{cond}}^{\,0} = \widehat{p}_{\text{cond}}^{+1} \prod_{i=2}^{s-1} \widehat{p}_{i|i-1}^{\,0}. \tag{3.156}$$

3.4.3 Stopping criterion for the refinement strategy

In Section 2.4 of Chapter 2, the proposed adaptive refinement strategy was left without any convergence criterion as it was argued that such a criterion depends on the final use of the Kriging surrogate. In the present context of surrogate-based reliability analysis, it is proposed to use one of the two error measures introduced above.

More specifically, the spread between the pseudo-confidence bounds $[p_{f\,(1-\alpha)}^{-1};\ p_{f\,(1-\alpha)}^{+1}]$ defined in Section 3.4.2.2 can be used in the form of the following metric:

$$\Delta p_{f\,(1-\alpha)} = \log_{10}\left(\frac{p_{f\,(1-\alpha)}^{+1}}{p_{f\,(1-\alpha)}^{-1}}\right) \geq 0. \tag{3.157}$$

Hence, the following criterion is used to stop the refinement strategy proposed in Section 2.4 of Chapter 2:

$$\Delta p_{f\,(1-\alpha)} \leq \Delta_0, \tag{3.158}$$

where Δ_0 is a specified threshold. Note that $\Delta_0 = 1$ is a minimum requirement to ensure that the surrogate-based failure probability is at least in the order of magnitude of the exact result.

3.4.4 Illustration

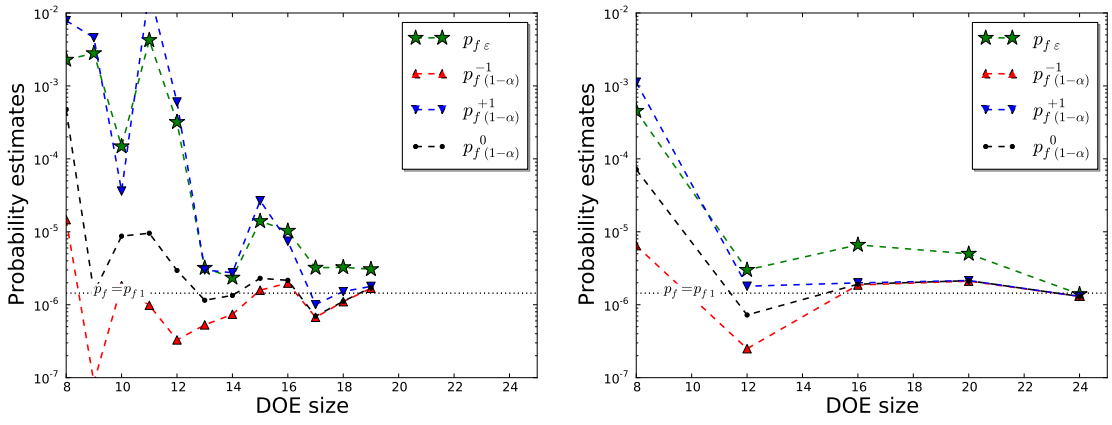
This section illustrates the behaviour of these heuristic error measures on the basic structural reliability example defined in Section 3.2.4. First, the performance function $\mathfrak{g}^\circ$ is

replaced with a Kriging meta-model built in the standard space. The first DOE uses 8 points obtained by the heuristic hypersphere-filling procedure detailed in Section 2.2.1.2 of Chapter 2 where the radius of the hypersphere is set equal to $R = 8$ because the multivariate standard Gaussian PDF mostly concentrates in this region.

This initial DOE is then enriched by means of the sampling-based refinement procedure detailed in Section 2.4 of Chapter 2. The refinement makes use of the margin probability refinement criterion which is weighted by the standard Gaussian PDF. The two heuristic error measures introduced in this section are estimated at each refinement iteration.

The refinement procedure was run twice. The first run uses the usual state-of-the-art refinement procedure where the points are added sequentially. In the second run, 4 points were added at each iteration. The heuristic error measures are plotted as functions of the total DOE size in Figure 3.17 for both runs. The algorithm is stopped when the logarithmic deviation $\Delta p_{f\,(1-\alpha)}$ between the two confidence bounds $\widehat{p}^{\,-1}_{f\,(1-\alpha)}$ and $\widehat{p}^{\,+1}_{f\,(1-\alpha)}$ remains less than $\Delta_0 = 0.25$ for three iterations.

It can be seen that the spread of the heuristic confidence interval $[\widehat{p}^{\,-1}_{f\,(1-\alpha)};\ \widehat{p}^{\,+1}_{f\,(1-\alpha)}]$ reduces as the DOE becomes denser in the vicinity of the limit-state surface and that $\widehat{p}^{\,0}_{f\,(1-\alpha)}$ tends towards the reference probability $p_{f,1} \approx 1.44 \times 10^{-6}$ given by FORM (see Section 3.3.4.2). This *empirically* demonstrates the consistency of the adaptive refinement strategies for reliability analysis. The successive estimates of the augmented failure probability $\widehat{p}_{f\,\varepsilon}$ also converge towards the reference probability although it is impossible to distinguish the surrogate error from the failure probability estimate. Finally, it should be pointed out that the second run shows a smoother convergence than the first one for the two error measures. This is because it adds several points in the DOE at each refinement iteration. Indeed, each batch of points uniformly spans the margin of uncertainty. Then, the spread of the latter decreases smoothly until it vanishes. The population-based refinement strategy does not require significantly more experiments than its sequential counterpart.



(a) Sequential refinement ($K = 1$) (b) Population-based refinement ($K = 4$)

Figure 3.17: Convergence of various approximations of the failure probability for the basic structural reliability problem defined in Section 3.2.4.

3.5 Meta-model-based importance sampling

In this section, a novel approach to surrogate-based reliability analysis is proposed. It combines Kriging meta-modelling together with importance sampling. This work has been originally published in Dubourg et al. (2011b,c) based on similar ideas developped in Dubourg and Deheeger (2010, in French).

3.5.1 Motivation

In the previous section, it was pointed out that the substitution error is rather hard to quantify. It is sometimes even impossible when the surrogate does not provide any quantification of its own accuracy (*e.g.* the Taylor expansion used in FORM does not provide such an information). Moreover, experience has shown that most meta-modelling techniques, including Kriging and support vector machines, lose efficiency in high dimension (*i.e.* when $n = \dim(\boldsymbol{X})$ gets large). For such problems a larger number of experiments is required in the DOE in order to ensure a significant reduction of the previously defined pseudo-confidence bounds on the failure probability. On the other hand, in highly sensitive applications such as in the nuclear industry, the use of surrogates for physical models is often rejected by the stakeholders. These pitfalls motivate the use of *hybrid approaches* that uses the information provided by a possibly coarse meta-model to sample smartly using either conditional or importance sampling.

Such hybrid approaches have not been given much attention in the literature though. Piera-Martìnez et al. (2007); Piera-Martìnez (2008) use the upper (conservative) bound of a support vector margin classifier in order to implement an efficient conditional sampling scheme on the original performance function $\mathfrak{g}$. Barbillon (2010, Chapter 6) exploits the same idea with Gaussian process predictors. In a similar fashion, Hurtado (2007) proposes a *filtered importance sampling scheme*, although the approach is limited to problems featuring a unique MPFP because it makes use of Melchers' instrumental PDF (see Section 3.3.3).

Cannamela et al. (2008) resort to a Gaussian process meta-model to implement three reduced variance estimators for low quantiles. The approaches are respectively based on importance sampling and two other fundamental concepts in the field of variance reduction techniques known as *control variates* and *stratified sampling* (see *e.g.* Rubinstein and Kroese, 2008, Chapter 5). The *parametric* importance-sampling-based approach is shown to yield the most accurate results. The strategy developed in the following sections explores the use of a *non-parametric* quasi-optimal instrumental PDF.

3.5.2 Approximation of the optimal instrumental PDF

In the sequel, it is proposed to build an approximation of the optimal instrumental PDF for the estimation of the failure probability in Eq. (3.3) based on a *probabilistic classification function* π that might either come from a Kriging predictor (see Eq. (3.126)) or from a probabilistic support vector margin classifier (Platt, 1999).

First, recall from Chapter 1 that in the case of a Kriging predictor denoted by $\widehat{Y}(\boldsymbol{x}) \sim \mathcal{N}_1(\mu_{\widehat{Y}}(\boldsymbol{x}), \sigma^2_{\widehat{Y}}(\boldsymbol{x}))$, the probabilistic classification function is defined as follows:

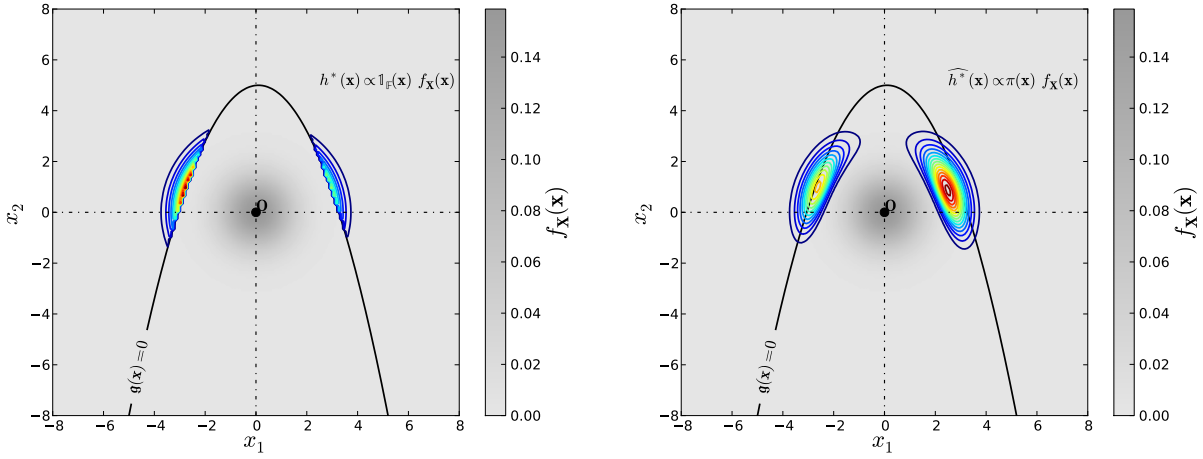
$$\pi(\boldsymbol{x}) = \mathscr{P}\left[\widehat{Y}(\boldsymbol{x}) \leq 0\right] = \Phi\left(\frac{0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right). \tag{3.159}$$

and that it can be extended to the points in the dataset for which the Kriging variance equals zero by switching to the limit (see Eq. (3.127)).

The proposed approximation of the optimal instrumental PDF h^* in Eq. (3.74) is obtained by substituting the probabilistic classification function π for the failure indicator function $\mathbb{1}_{\mathbb{F}}$ in its expression. Therefore, it reads as follows:

$$\widehat{h^*}(\boldsymbol{x}) = \frac{\pi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\mathbb{X}} \pi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}} = \frac{\pi(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})}{p_{f\varepsilon}}, \tag{3.160}$$

where the *augmented failure probability* $p_{f\varepsilon}$ defined in Eq. (3.130) has been introduced. Such a quasi-optimal instrumental PDF is illustrated on the example from Der Kiureghian and Dakessian (1998) introduced in Section 3.3.4.4 and already evoked in Section 1.5.3 of Chapter 1. It is compared to the optimal (although impracticable) instrumental PDF in Figure 3.18. It can be seen that, on this simple example, the proposed instrumental PDF is able to deal with multiple design points as opposed to Melchers' instrumental PDF (see Figure 3.10).



(a) Optimal instrumental PDF h^* (b) Proposed approximation $\widehat{h^*}$

Figure 3.18: Comparison of the optimal (although impracticable) instrumental PDF and its approximation on the two-dimensional example from Der Kiureghian and Dakessian (1998). The underlying Kriging predictor is built as described in Section 1.5.3 of Chapter 1.

By introducing the quasi-optimal instrumental PDF $\widehat{h^*}$ into Eq. (3.3), the failure probability rewrites as follows:

$$p_f = \int_{\{\boldsymbol{x} \in \mathbb{X}:\, \widehat{h^*}(\boldsymbol{x}) > 0\}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \frac{\widehat{h^*}(\boldsymbol{x})}{\widehat{h^*}(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x}. \tag{3.161}$$

This expression can be further simplified by replacing the instrumental PDF $\widehat{h}^*$ in the denominator on the right by its expression in Eq. (3.160):

$$\begin{aligned} p_f &= \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\,\widehat{h}^*(\boldsymbol{x})>0\right\}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \frac{p_{f\,\varepsilon}}{\pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}\, \widehat{h}^*(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\,\widehat{h}^*(\boldsymbol{x})>0\right\}} \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})}{\pi(\boldsymbol{x})}\, \widehat{h}^*(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\, p_{f\,\varepsilon} \\ &= \alpha_{\mathrm{corr}}\, p_{f\,\varepsilon}, \end{aligned} \tag{3.162}$$

where the so-called *correction factor*:

$$\alpha_{\mathrm{corr}} \equiv \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\,\widehat{h}^*(\boldsymbol{x})>0\right\}} \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})}{\pi(\boldsymbol{x})}\, \widehat{h}^*(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{3.163}$$

has been introduced. Both terms in Eq. (3.162) correspond to the following mathematical expectations:

$$p_{f\,\varepsilon} = \int_{\mathbb{X}} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \equiv \mathbb{E}_{\boldsymbol{X}}\left[\pi(\boldsymbol{X})\right], \tag{3.164}$$

$$\alpha_{\mathrm{corr}} = \int_{\left\{\boldsymbol{x}\in\mathbb{X}:\,\widehat{h}^*(\boldsymbol{x})>0\right\}} \frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x})}{\pi(\boldsymbol{x})}\, \widehat{h}^*(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \equiv \mathbb{E}_{\widehat{\boldsymbol{Z}}}\left[\frac{\mathbb{1}_{\mathbb{F}}(\widehat{\boldsymbol{Z}})}{\pi(\widehat{\boldsymbol{Z}})}\right], \tag{3.165}$$

where $\boldsymbol{X}$ (resp. $\widehat{\boldsymbol{Z}}$) is distributed according to $f_{\boldsymbol{X}}$ (resp. $\widehat{h}^*$).

The latter correction factor α_{corr} is thus defined as the expected ratio between the real indicator function $\mathbb{1}_{\mathbb{F}}$ and the probabilistic classification function π. Thus, if the Kriging prediction is fully accurate, the correction factor is equal to one and the failure probability is equal to the augmented failure probability (optimality of the proposed estimator). On the other hand, in the more general case where the Kriging prediction is not fully accurate (since it is obtained from a DOE $\mathscr{X}$ of finite size m), the correction factor modifies the augmented failure probability accounting for the residual epistemic uncertainty in the prediction.

3.5.3 Proposed estimator

3.5.3.1 Estimator of the augmented failure probability

The estimation of the augmented failure probability $p_{f\,\varepsilon}$ was already addressed in Section 3.4.2.1. Recall that this probability might be in the same order of magnitude than the failure probability, so that the modified subset sampling procedure is better suited than Monte Carlo sampling to provide an accurate estimate $\widehat{p}_{f\,\varepsilon}$ within a reasonable number of calls to the Kriging predictor, say N_ε. The corresponding variance of estimation is denoted by σ^2_ε. Note that it can be reduced to a negligible value because the augmented failure probability is estimated by means of the Kriging predictor which is a lot cheaper to evaluate than the original performance function $\mathfrak{g}$.

3.5.3.2 Estimator of the correction factor

The estimation of the correction factor α_{corr} resorts to Monte Carlo sampling. It is computed as follows:

$$\widehat{\alpha}_{\text{corr}} = \frac{1}{N} \sum_{k=1}^{N_{\text{corr}}} \frac{\mathbb{1}_{\mathbb{F}}\left(\widehat{\boldsymbol{Z}}^{(k)}\right)}{\pi\left(\widehat{\boldsymbol{Z}}^{(k)}\right)}, \tag{3.166}$$

where $\widehat{\mathscr{Z}} = \{\widehat{\boldsymbol{Z}}^{(k)}, k = 1, \dots, N_{\text{corr}}\}$ is a sample of N_{corr} samples of the random vector $\widehat{\boldsymbol{Z}}$ distributed according to $\widehat{h}^*$. Such samples can be obtained by means of any MCMC sampling algorithm such as the modified Metropolis-Hastings algorithm detailed in Section B.3.2 of Appendix B. Hence, the variance of estimation of this quantity reads as follows:

$$\sigma^2_{\text{corr}} = \frac{1}{N_{\text{corr}}} \left[\frac{1}{N_{\text{corr}}} \sum_{k=1}^{N_{\text{corr}}} \frac{\mathbb{1}_{\mathbb{F}}\left(\widehat{\boldsymbol{Z}}^{(k)}\right)^2}{\pi\left(\widehat{\boldsymbol{Z}}^{(k)}\right)^2} - \widehat{\alpha}^2_{\text{corr}} \right] \left(1 + \gamma_{\text{corr}}\right), \tag{3.167}$$

where γ_{corr} is the coefficient that accounts for the correlation in the sample $\mathscr{Z}$ generated by means of MCMC. The reader is referred to Section B.2.3 of Appendix B for the expression of the coefficient γ_{corr} together with the proof of the latter equation. Note that the coefficient γ_{corr} can be significantly reduced by means of the so-called *thinning* procedure detailed in Section B.2.4 of Appendix B because the instrumental PDF $\widehat{h}^*$ does not depend on the expensive-to-evaluate performance function $\mathfrak{g}$ as opposed to what happens in subset sampling. However, the estimation of the correction factor *does* depend on the performance function $\mathfrak{g}$ through the failure indicator function $\mathbb{1}_{\mathbb{F}}$ so that N_{corr} should be as small as possible for the sake of efficiency. N_{corr} depends on the order of magnitude of the correction factor α_{corr}. Indeed, the closer α_{corr} is to unity (optimality of the proposed estimator), the smaller N_{corr}.

3.5.3.3 Estimator of the failure probability

The final estimator of the failure probability is eventually defined as follows:

$$\widehat{p}_{f\,\text{metaIS}} = \widehat{p}_{f\,\varepsilon}\,\widehat{\alpha}_{\text{corr}}. \tag{3.168}$$

The two terms in this product can be estimated by means of two independent samples. Thus, $\widehat{p}_{f\,\varepsilon}$ and $\widehat{\alpha}_{\text{corr}}$ are independent as well. Based on this consideration, the variance of estimation of $\widehat{p}_{f\,\text{metaIS}}$ can be derived as follows:

$$\begin{aligned} \sigma^2_{\text{metaIS}} &\equiv \text{Var}\left[\widehat{p}_{f\,\text{metaIS}}\right] \\ &= \mathbb{E}\left[\widehat{p}^2_{f\,\varepsilon}\,\widehat{\alpha}^2_{\text{corr}}\right] - \mathbb{E}\left[\widehat{p}_{f\,\varepsilon}\,\widehat{\alpha}_{\text{corr}}\right]^2 \\ &= \mathbb{E}\left[\widehat{p}^2_{f\,\varepsilon}\right] \mathbb{E}\left[\widehat{\alpha}^2_{\text{corr}}\right] - \left(\mathbb{E}\left[\widehat{p}_{f\,\varepsilon}\right] \mathbb{E}\left[\widehat{\alpha}_{\text{corr}}\right]\right)^2. \end{aligned} \tag{3.169}$$

The first term may be further elaborated using the König-Huyghens theorem:

$$\sigma^2_{\text{metaIS}} = \left(\mathbb{E}\left[\widehat{p}_{f\,\varepsilon}\right]^2 + \sigma^2_{\varepsilon}\right) \left(\mathbb{E}\left[\widehat{\alpha}_{\text{corr}}\right]^2 + \sigma^2_{\text{corr}}\right) - \left(\mathbb{E}\left[\widehat{p}_{f\,\varepsilon}\right] \mathbb{E}\left[\widehat{\alpha}_{\text{corr}}\right]\right)^2. \tag{3.170}$$

Since the two estimators $\widehat{p}_{f\,\varepsilon}$ and $\widehat{\alpha}_{\text{corr}}$ are unbiased, the variance eventually reduces to:

$$\begin{aligned}\sigma^2_{\text{metaIS}} &= \left(p^2_{f\,\varepsilon}+\sigma^2_{\varepsilon}\right)\left(\alpha^2_{\text{corr}}+\sigma^2_{\text{corr}}\right)-\left(p_{f\,\varepsilon}\,\alpha_{\text{corr}}\right)^2 \\ &= \sigma^2_{\varepsilon}\,\sigma^2_{\text{corr}}+p^2_{f\,\varepsilon}\,\sigma^2_{\text{corr}}+\alpha^2_{\text{corr}}\,\sigma^2_{\varepsilon}. \end{aligned} \tag{3.171}$$

Finally, the coefficient of variation of the meta-model-based importance sampling estimator in Eq. (3.168) reads as follows:

$$\delta_{\text{metaIS}} \equiv \frac{\sigma_{\text{metaIS}}}{p_{f\,\varepsilon}\,\alpha_{\text{corr}}} = \sqrt{\delta^2_{\varepsilon}+\delta^2_{\text{corr}}+\delta^2_{\varepsilon}\,\delta^2_{\text{corr}}}, \tag{3.172}$$

where $\delta_{\varepsilon}=\sigma_{\varepsilon}/p_{f\,\varepsilon}$ (resp. $\delta_{\text{corr}}=\sigma_{\text{corr}}/\alpha_{\text{corr}}$) is the coefficient of variation of the augmented failure probability (resp. of the correction factor). In practice usual target coefficients of variation range from 1% to 10% so that:

$$\delta_{\text{metaIS}} \underset{\delta_{\varepsilon},\delta_{\text{corr}}\ll 1}{\approx} \sqrt{\delta^2_{\varepsilon}+\delta^2_{\text{corr}}} \tag{3.173}$$

is a fairly accurate approximation of the final coefficient of variation.

It is worth emphasizing again that the contribution of the coefficient of variation of the augmented failure probability δ_{ε} in that of the final estimator δ_{metaIS} can be significantly reduced at a low computational expense because the augmented failure probability $p_{f\,\varepsilon}$ is estimated by means of the Kriging predictor only.

3.5.4 Adaptive refinement of the proposed instrumental PDF

The significance of the variance reduction introduced by the proposed meta-model-based importance sampling technique mostly relies on the optimality of the proposed instrumental PDF $\widehat{h}^*$. Thus it is proposed here to adaptively refine the probabilistic classification function π so that it tends toward the failure indicator function $\mathbb{1}_{\mathbb{F}}$. Hence the quasi-optimal instrumental PDF $\widehat{h}^*$ will converge toward its optimal counterpart h^*, the correction factor α_{corr} will converge toward unity (σ^2_{corr} will tend to zero), and the variance of the failure probability in Eq. (3.171) will eventually reduce to that of the augmented failure probability σ^2_{ε} which can be reduced at a negligible computational cost.

The refinement of the probabilistic classification function is achieved by means of the sampling-based refinement procedure already detailed in Section 2.4 of Chapter 2. Indeed, this refinement procedure only requires a stopping criterion in order to decide whether the probabilistic classification function is sufficiently accurate for the proposed importance sampling scheme or not. The most appropriate metric to measure the optimality of the probabilistic classification function is nothing but the actual value of the correction factor α_{corr} although it is still unknown so far. Of course, it could be estimated at each refinement iteration. However this would be particularly inefficient because α_{corr} depends on the expensive-to-evaluate performance function $\mathfrak{g}$ so that it should be estimated only once, when the quasi-optimal instrumental density has converged.

As a matter of fact, it is proposed to resort to the following leave-one-out estimate of the correction factor:

$$\widehat{\alpha}_{\text{corr LOO}} = \frac{1}{m}\sum_{i=1}^{m}\frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{x}^{(i)})}{\mathscr{P}\left[\widehat{Y}_{-i}\leq 0\right]}, \tag{3.174}$$

where $\widehat{Y}_{-i} \sim \mathscr{N}_1(\mu_{\widehat{Y}_{-i}}, \sigma^2_{\widehat{Y}_{-i}})$ is the i-th *leave-one-out Kriging prediction* of the performance function $\mathfrak{g}$. The reader is referred to Section 1.4.5.3 of Chapter 1 for a more detailed introduction of this cross-validation procedure together with the analytical expression of the leave-one-out Kriging predictions $\{\widehat{Y}_{-i}, i = 1, \ldots, m\}$.

In practice, the iterative enrichment of the DOE $\mathscr{X}$ is stopped if its size is greater than say $N_{\min}$ and $\widehat{\alpha}_{\text{corr LOO}}$ is in the order of magnitude of unity. Usually, the minimum size $N_{\min}$ for averaging a quantity is given by the central limit theorem and $N_{\min} = 30$ yields satisfactory estimates. In the case of high dimensional and/or highly nonlinear problems, the size of the DOE is also limited to $m_{\max} = 1\,000$. One may also measure the improvement brought by the latest refinement iteration in terms of variation of $\widehat{\alpha}_{\text{corr LOO}}$ in order to decide whether to stop the refinement procedure or not.

Note that a leave-one-out estimate of the Kriging variance $\sigma^2_{\widehat{Y}_{-i}}$ might be rounded to zero due to machine precision. This is problematic because the ratio in Eq. (3.174) might not be defined in such cases. Indeed, if the corresponding mean prediction $\mu_{\widehat{Y}_{-i}}$ is positive, then the probability in the denominator equals zero. Hence, if the failure indicator function $\mathbb{1}_{\mathbb{F}}$ is also zero it may cause an exception because $0/0$ is undetermined. Recall that this variance should never be exactly zero. Indeed, the Kriging variance equals zero only at the points in the DOE (here $\mathscr{X}\setminus\boldsymbol{x}^{(i)}$). So, it is proposed to bound the leave-one-out probabilistic classification function in the denominator of Eq. (3.174) above a reasonably low value (say the machine precision, $\varepsilon_{\text{M}} \approx 10^{-16}$).

3.5.5 Implementation

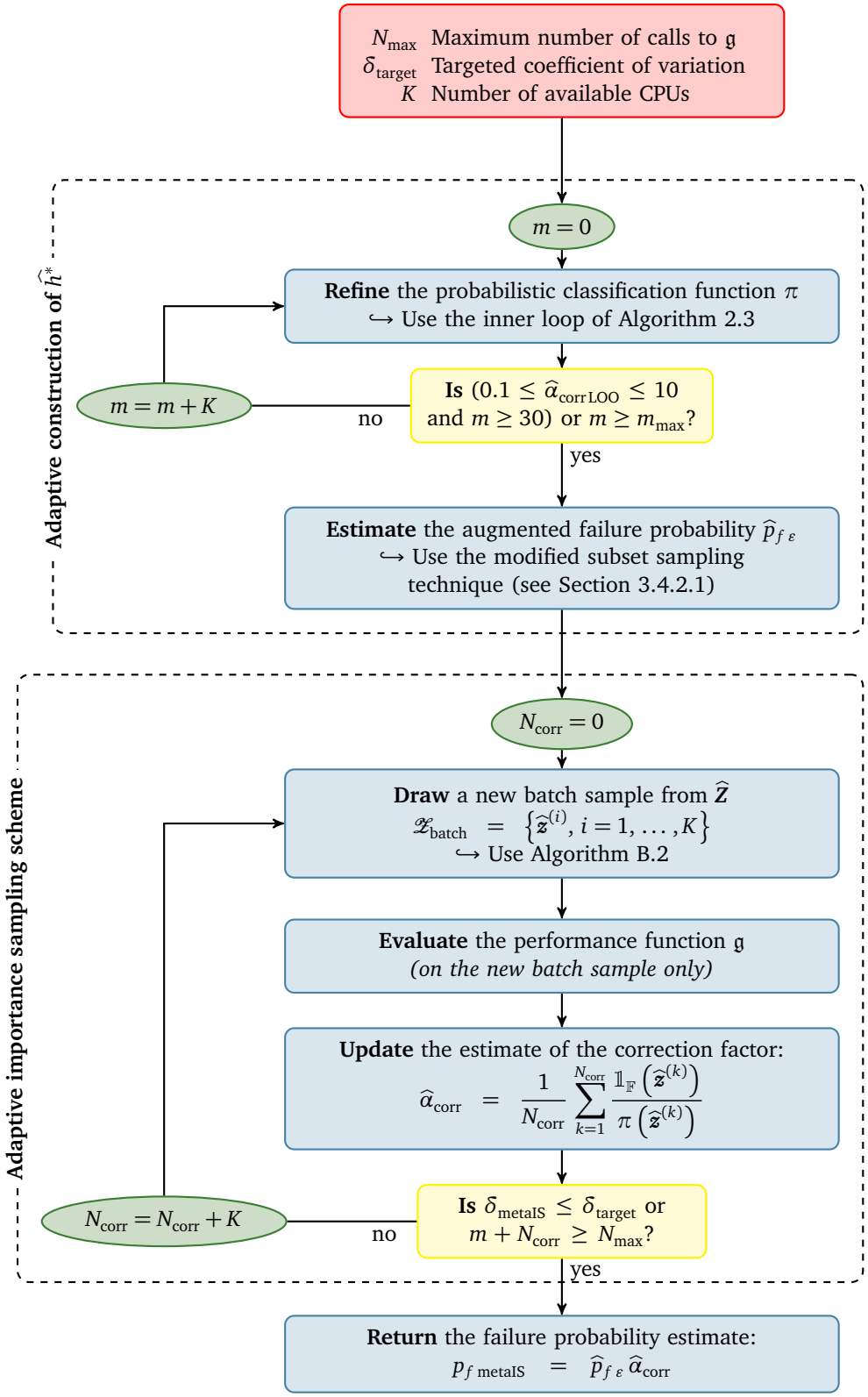
The proposed meta-model-based importance sampling scheme is sketched in Algorithm 3.2. The algorithm only requires the choice of a maximum number of performance function evaluations $N_{\max}$, a targeted coefficient of variation δ_{target} for the final estimate $\widehat{p}_{f\,\text{metaIS}}$ and the parameters of the sampling-based adaptive refinement strategy in Algorithm 2.3. It is essentially divided in two main parts.

The first part concerns the adaptive contruction of the instrumental PDF $\widehat{h}^*$ to make it tend toward the optimal one h^*. This is achieved by applying the sampling-based refinement algorithm introduced in Section 2.4 of Chapter 2 as long as the leave-one-out estimate of the correction factor $\alpha_{\text{corr LOO}}$ is not sufficiently close to one. Once the probabilistic classification function is found sufficiently accurate, the definition of the instrumental PDF $\widehat{h}^*$ is completed by computing its normalizing constant $p_{f\,\varepsilon}$. This resorts to the modified subset sampling technique introduced in Section 3.4.2.1. The sample size per step N_ε shall be chosen as large as reasonably possible (say $N_\varepsilon = 10^4 - 10^5$) so as to ensure a negligible

coefficient of variation δ_ε to minimize its contribution on the final coefficient of variation δ_{metaIS} (see Eq. (3.172)).

The second part consists in using the previously built instrumental PDF in a classical adaptive importance sampling scheme driven by the target coefficient of variation δ_{target}. This resorts to MCMC sampling of the instrumental PDF $\widehat{h}^*$. The present implementation resorts to the modified Metropolis-Hastings algorithm introduced in Section B.3.2 of Appendix B. It is seeded with the points contained in the last sample $\mathscr{X}^{[s]}$ used to estimate the augmented failure probability. In order to avoid the correlation between $\widehat{p}_{f\,\varepsilon}$ and $\widehat{\alpha}_{\text{corr}}$ though it is proposed to *burn* the first samples (say the 10 first elements of each chain) generated by the modified Metropolis-Hastings algorithm. Moreover, the MCMC sampler may retain only one element of each chain every t increments (say $t = 4$) in order to reduce the γ_{corr} coefficient involved in the variance of estimation of $\widehat{\alpha}_{\text{corr}}$. For a more detailed introduction to these two techniques known as *burn-in* and *thinning* (respectively), the reader is referred to Section B.2.4 of Appendix B. Note that rejecting samples is affordable here because the instrumental PDF $\widehat{h}^*$ depends only on the Kriging predictor as opposed to what happens in subset sampling where each sub-optimal instrumental PDF involves the expensive-to-evaluate performance function $\mathfrak{g}$ (through the intermediate subset indicator functions).

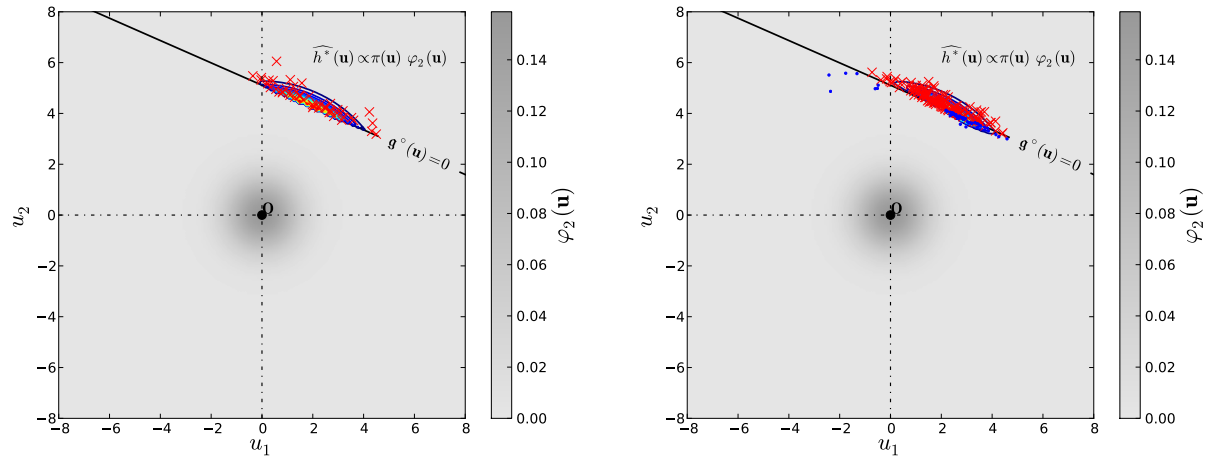
Algorithm 3.2 Metamodel-based importance sampling



3.5.6 Illustration

The meta-model-based importance sampling scheme is applied to the basic structural reliability example defined in Section 3.2.4. All the operations are performed in the standard space spanned by $\boldsymbol{u}$. The first Kriging meta-model is built from a heuristic hypersphere-filling DOE (see Section 2.2.1.2) containing 8 points in the ball of radius $R = 8$. It is then refined by means of the sampling-based refinement technique based on the margin probability criterion weighted by the multivariate standard Gaussian PDF (*i.e.* $w(\boldsymbol{c}) \propto e^{-\|\boldsymbol{c}\|_2^2/2}$). $K = 4$ points are added at each refinement iteration and the refinement is eventually stopped when $\widehat{\alpha}_{\text{corrLOO}}$ reaches a stable value sufficiently close to unity. The size of the DOE at the end of the adaptive refinement procedure is equal to $m = 8 + 9 \times 4 = 44$. The normalizing constant $p_{f\,\varepsilon}$ is estimated by means of the modified subset sampling technique detailed in Section 3.4.2.1 (with $N_\varepsilon = 10^5$ for the sample size per step), and it is found equal to $\widehat{p}_{f\,\varepsilon} \approx 1.53 \times 10^{-6}$ up to a 4% coefficient of variation. Note that it is already in good agreement with the reference solution $p_{f,1} \approx 1.44 \times 10^{-6}$ given by FORM.

The adaptive importance sampling scheme is then triggered. A 10% coefficient of variation is targeted and the batch samples of $\widehat{Z}$ are generated from 50 independent chains. Actually, on this simple example the Kriging predictor is fully accurate after the refinement so that the theoretically impracticable optimal importance sampling scheme is achieved ($\widehat{h^*} = h^*$). Then the first 50-sample allows a sufficient variance reduction with respect to the targeted coefficient of variation. This sample is illustrated in Figure 3.19(a) together with the proposed instrumental PDF $\widehat{h^*}$. It is clear from this illustration that optimality is reached. The correction factor is found equal to one with a zero variance of estimation so that the final coefficient of variation is reduced to that of the augmented failure probability $\delta_{\text{metaIS}} = \delta_\varepsilon \approx 4\%$.



(a) Run #1: Optimal importance sampling

(b) Run #2: Coarser importance sampling

Figure 3.19: Application of the meta-model-based importance sampling scheme to the basic structural reliability example defined in Section 3.2.4.

The proposed hybrid approach is actually designed for more sophisticated reliability examples for which the optimal instrumental PDF can only be approximated (*e.g.* examples

featuring a higher number of random variables and a limit-state surface whose nonlinearity is hard to emulate). To illustrate such a quasi-optimality, the same problem is run again except that the adaptive refinement is arbitrarily stopped at the third iteration. The corresponding instrumental PDF is illustrated in Figure 3.19(b) together with a small sample. In this case the importance sampling scheme yields a correction factor $\widehat{\alpha}_{\text{corr}} \approx 0.98$ (up to a 9% coefficient of variation), an augmented failure probability $\widehat{p}_{f\,\varepsilon} \approx 1.64 \times 10^{-6}$ (up to a 4% coefficient of variation) and eventually the failure probability is $\widehat{p}_{f\,\text{metaIS}} \approx 1.60 \times 10^{-6}$ (up to a 10% coefficient of variation). Finally, this second hybrid reliability analysis used $m + N_{\text{corr}} = (8 + 2 \times 4) + 19 \times 50 = 966$ calls to the performance function $\mathfrak{g}$ which is less optimal than the first run.

3.6 Conclusion

This chapter has reviewed several state-of-the-art reliability methods. First the isoprobabilistic transform was introduced as a mean to work within a convenient *standard space* spanned by independent identically distributed Gaussian random variables. It has been pointed out that the advanced reliability methods resorting to optimization, MCMC sampling or meta-modelling are easier to tune in that standard space than in the original one.

Then a specific emphasis was put on the central role of the importance sampling technique whose concept underlies most state-of-the-art reliability methods (even the first-order reliability method). Subset sampling reveals a particularly robust and efficient technique to estimate low failure probabilities (up to $10^{-6} - 10^{-12}$) so that it should always be preferred to any other technique. Despite it is theoretically biased for every N, it has proven to yield fairly accurate results when N is sufficiently large (say $N > 10^3$).

However, when the performance function is expensive-to-evaluate, even the most efficient variance reduction techniques (including subset sampling) may reveal impracticable. This premise naturally lead to the introduction of the surrogate-based approaches. Such approaches allow a significant reduction of the computational cost induced by the quantification of rare event probabilities. In addition, it has been shown that the substitution error (*i.e.* the error induced by the use of a surrogate instead of the real performance function) can be apprehended when the meta-model provides a local measure of its own accuracy (as Kriging does).

Eventually, a novel hybrid approach has been proposed. It uses the probabilistic classification information provided by a Kriging predictor in order to build a fair approximation of the optimal (although impracticable) instrumental PDF. Using elementary algebra the failure probability is recast as a product of two terms, namely the *augmented failure probability* $p_{f\,\varepsilon}$ which is evaluated by means of the meta-model only, and a correction factor α_{corr} that is computed from evaluations of the original limit-state function. This approach is not significantly more computationally expensive than the usual surrogate-based approaches but it yields an *unbiased* and *convergent* estimator of the quantity of interest. Moreover it has a potentially larger field of application when the Kriging meta-modelling technique loses efficiency (*e.g.* in high dimension).

4

Reliability-based design optimization

Contents

4.1 Introduction

In modern engineering, *design optimization* is the decision-making process that aims at finding the best set of *design variables* which minimizes some *cost* model while satisfying some *performance requirements*. Due to the inconsistency between these two objectives, the optimal solutions often lie on the boundaries of the admissible space. Thus these solutions are rather sensitive to uncertainty either in the parameters (*aleatory*) or in the models themselves (*epistemic*). *Reliability-based design optimization* (RBDO) is a concept that accounts for this uncertainty all along the optimization process.

Basically, the deterministic performance model is transformed into a *probabilistic constraint*. Despite its attractive formulation, the application field of RBDO is still limited to academic examples. This is mostly due to the fact that it is either based on simplifying assumptions (*e.g.* FORM/SORM) that might not hold in practice; or in contrast, it requires computationally intensive stochastic simulations that are not affordable for real industrial problems (involving expensive-to-evaluate computational models).

This chapter proposes a meta-model-based strategy that would *in fine* bring the RBDO application field to more sophisticated examples, closer to real engineering cases. It is essentially based on the use of adaptive surrogate models (see Chapters 1 and 2), variance reduction techniques (see Chapter 3) and reliability sensitivity analysis (introduced in this chapter). The remaining of this section is devoted to the formulation of the RBDO problem. A first introductory analytical example is provided for the sake of illustration. A state-of-the-art is then proposed as a motivation for the RBDO resolution strategy proposed in this manuscript. Section 4.2 recalls a few elements of constrained optimization and mathematical programming before focusing on the presentation of the Polak-He algorithm that was used in this thesis. Section 4.3 is concerned with the computation of the gradient of the failure probability (*i.e.* reliability sensitivity analysis) which is a fundamental ingredient in order to embed a reliability analysis in a design optimization procedure. Section 4.4 eventually presents the proposed surrogate-based RBDO resolution strategy.

4.1.1 Problem formulation

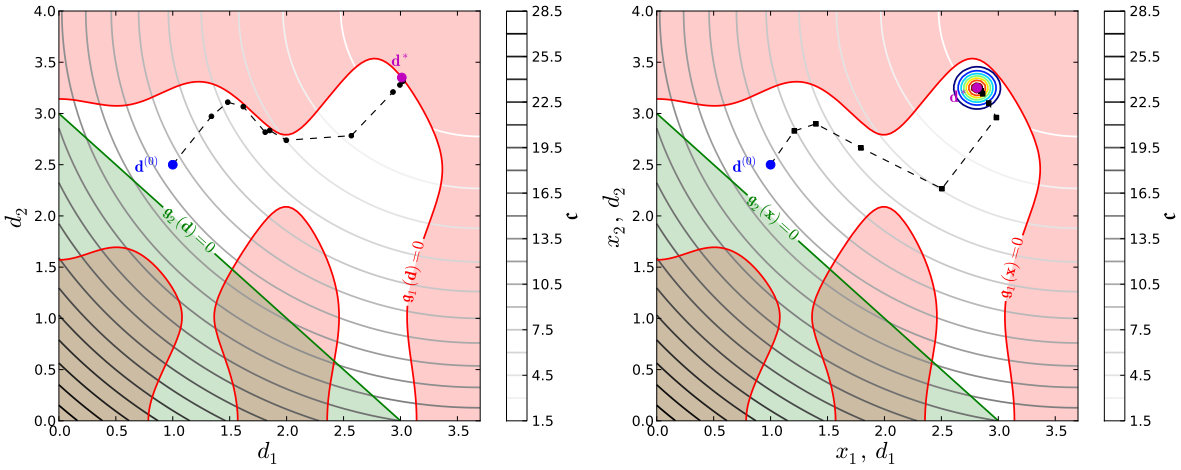
Given a parametric model for the random vector $\boldsymbol{X}$ (as in Section 3.2 of Chapter 3) describing the system to be designed, the most basic formulation for RBDO reads as follows:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}\in\mathbb{D}} \mathfrak{c}(\boldsymbol{d}) \ : \quad \begin{cases} \mathfrak{f}_i(\boldsymbol{d}) \leq 0, \ i = 1, \dots, n_c \\ \mathbb{P}\left[\mathfrak{g}_l(\boldsymbol{X}(\boldsymbol{d})) \leq 0\right] \leq p_{f\,l}^0, \ l = 1, \dots, n_p \end{cases} . \tag{4.1}$$

In this formulation, $\mathfrak{c}$ is the cost function to be minimized with respect to the design variables $\boldsymbol{d} = \{d_i, i = 1, \dots, n_d\} \in \mathbb{D} \subset \mathbb{R}^{n_d}$, while satisfying to n_c deterministic soft constraints $\{\mathfrak{f}_i, \ i = 1, \dots, n_c\}$ bounding the so-called *admissible design space* defined by the analyst. Note that in most applications these soft constraints consists in simple analytical functions that prevent the optimization algorithm from exploring regions of the design space that have no physical meaning (*e.g.* negative or infinite dimensions), so that these constraints

are inexpensive to evaluate. A deterministic design optimization (DDO) would simply require additional performance functions $\{\mathfrak{g}_l,\ l = 1, \ldots, n_p\}$ describing system failures with respect to the specific code of practice. As opposed to the previous soft constraints, these functions often involve the output of an expensive-to-evaluate black-box function $\mathscr{M}$ (*e.g.* a finite element model).

RBDO differs from DDO in the sense that these constraints are transformed into n_p probabilistic constraints $\{\mathbb{P}\left[\mathfrak{g}_l(\boldsymbol{X}) \leq 0\right] \leq p^0_{f\,l},\ l = 1, \ldots, n_p\}$. $p^0_{f\,l}$ is the minimum safety requirement expressed here in the form of an acceptable probability of failure which may be different for each performance function $\mathfrak{g}_l$. DDO and RBDO are graphically opposed in Figure 4.1. Basically, the reliability-based optimal solution steps away from the limit-state surface for the sake of safety with respect to the uncertainty in the random design variables.



(a) Deterministic design optimization

(b) Reliability-based design optimization

Figure 4.1: Comparison of deterministic and reliability-based design optimization.

It should be noticed that the design vector $\boldsymbol{d}$ is a set of parameters defining the random vector $\boldsymbol{X}$ (*e.g.* the mean for the example of Figure 4.1). In other words, in this thesis, design variables are exclusively considered as parameters in the joint probability density function $f_{\boldsymbol{X}}$ of the random vector $\boldsymbol{X}$ because it will later simplify the computation of the derivatives of the probabilistic constraints (*i.e.* reliability sensitivity). There is however no loss of generality since deterministic design variables might possibly be considered as artificially random (*e.g.* Gaussian) with small variance.

It might be argued that this formulation lacks full probabilistic consideration because the cost function is defined in a deterministic manner as it only depends on the parameters $\boldsymbol{d}$ of the random vector $\boldsymbol{X}$. A more comprehensive formulation could account for the randomness of the cost function possibly induced by the one in $\boldsymbol{X}$ but the present formulation is extensively used in the RBDO literature for the sake of simplicity. Note however that thanks to the rather low computational complexity of the usual cost models (at least in the present context of structural mechanics), an accurate sampling-based estimation of a mean cost, say $c(\boldsymbol{d}) = \mathbb{E}\left[c(\boldsymbol{X}(\boldsymbol{d}))\right]$ would not require a large computational effort.

Note that the n_p probabilistic constraints can equivalently be specified in terms of *generalized reliability indices* for the sake of normalization. The so-called *generalized reliability index* was first introduced by Ditlevsen (1979a) (see also Ditlevsen and Madsen, 1996) and it reads as follows:

$$\beta = -\Phi^{-1}\left(p_f\right), \tag{4.2}$$

where the probability of failure p_f can be estimated by means of any technique introduced in Chapter 3. As its name suggests, it is a generalization of the relationship between the Hasofer-Lind reliability index and the first-order approximation of a failure probability as introduced in Section 3.3.4.2 of Chapter 3. Generally speaking, it is easier to work in terms of generalized reliability indices that range from -8 to 8 than to work with probabilities that equivalently range from $1-10^{-15}$ to 10^{-15}.

4.1.2 A first introductory analytical example

In order to provide a better understanding of the previously introduced RBDO formulation, a first fully analytical example is presented and solved in the sequel.

4.1.2.1 Mechanical and probabilistic models

Let us consider a simply supported rectangular and perfectly straight column with cross-section $b \times h$ subjected to a constant service axial load F_{ser}. Provided $h \leq b$, its Euler critical linear elastic buckling load is given by the following formula:

$$F_{\text{cr}} = \frac{\pi^2 E\, b\, h^3}{12\, L^2}. \tag{4.3}$$

This allows one to formulate the performance function which will be involved in the probabilistic constraint as:

$$\mathfrak{g}(\boldsymbol{x}) = \frac{\pi^2 E\, b\, h^3}{12\, L^2} - F_{\text{ser}}. \tag{4.4}$$

For this oversimplified problem, the stochastic model consists of 3 independent lognormal random variables with parameters λ and ζ, namely: $\boldsymbol{X} = (E,\, b,\, h)^{\mathsf{T}}$.

4.1.2.2 Computation of the failure probability

Under these considerations, the limit-state surface $\mathscr{S}_0 = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) = 0\}$ can be transformed into the standard space of independent normal random variables $\boldsymbol{u}$ with zero-mean and unit-variance, and it has an explicit form. Indeed, the isoprobabilistic transform simply reads:

$$x_\bullet = \exp \lambda_\bullet + \zeta_\bullet u_\bullet, \quad \bullet = \{E,\, b,\, h\}\,. \tag{4.5}$$

Hence, recasting the limit-state surface equation as:

$$\mathscr{S}_0 = \left\{\boldsymbol{x} \in \mathbb{X} : \breve{\mathfrak{g}}(\boldsymbol{x}) = \log\left(\frac{\pi^2 E\, b\, h^3}{12\, L^2}\right) - \log\left(F_{\text{ser}}\right) = 0\right\}, \tag{4.6}$$

one gets its expression in the standard normal space which turns out to read:

$$\mathscr{S}_0^\circ = \Bigg\{ \boldsymbol{u} \in \mathbb{R}^3 : \breve{\mathfrak{g}}^\circ(\boldsymbol{u}) = \zeta_E\, u_E + \zeta_b\, u_b + 3\,\zeta_h\, u_h + \left[\log\left(\frac{\pi^2}{12\, F_{\text{ser}}\, L^2} \right) + \lambda_E + \lambda_b + 3\,\lambda_h \right] = 0 \Bigg\}. \tag{4.7}$$

Since it is linear in $\boldsymbol{u}$, the Hasofer-Lind reliability index can be calculated here in a fully closed-form as the distance from the origin to the hyperplane $\mathscr{S}_0^\circ = \{\boldsymbol{u} \in \mathbb{R}^3 : \breve{\mathfrak{g}}^\circ(\boldsymbol{u}) = 0\}$ and reads:

$$\beta_{\text{HL}} = \frac{\log\left(\frac{\pi^2}{12\, F_{\text{ser}}\, L^2} \right) + \lambda_E + \lambda_b + 3\,\lambda_h}{\sqrt{\zeta_E^2 + \zeta_b^2 + 9\zeta_h^2}}. \tag{4.8}$$

Note that the reliability index is also a linear function of λ_b and λ_h and that the associated probability of failure is exactly $p_f = p_{f,1} = \Phi\left(-\beta_{\text{HL}}\right)$ since the limit-state surface is linear in the standard space.

4.1.2.3 Reliability-based design optimization

The purpose of RBDO in this example is to minimize the structural weight which is proportional to the cross section area $\mathfrak{c}(\boldsymbol{d}) = \mu_b\, \mu_h$ while satisfying to a deterministic constraint:

$$f\,(\boldsymbol{d}) = \mu_h - \mu_b \leq 0, \tag{4.9}$$

and the probabilistic constraint (written here in terms of reliability indices):

$$\beta_0 \leq \beta_{\text{HL}}(\boldsymbol{d}), \tag{4.10}$$

where $\beta_0 = 3$ is the minimum target reliability index. The optimization is carried out with a constant coefficient of variation δ for both b and h so that:

$$\zeta_b = \zeta_h = \sqrt{\log\left(1 + \delta^2\right)} \equiv \zeta. \tag{4.11}$$

Noting that minimizing $\mathfrak{c}$ with respect to the means μ_b and μ_h is equivalent to minimizing $\log \mathfrak{c}$ with respect to $\log \mu_b = \lambda_b + \frac{1}{2}\,\zeta^2$ and $\log \mu_h = \lambda_h + \frac{1}{2}\,\zeta^2$ (since ζ is a prescribed constant), this RBDO problem turns out to be a linear optimization problem in $\breve{\boldsymbol{d}} = (\lambda_b,\, \lambda_h)^{\mathsf{T}}$:

$$\breve{\boldsymbol{d}}^* = \arg\min_{(\lambda_b,\, \lambda_h)} \lambda_b + \lambda_h + \zeta^2 \;:\; \begin{cases} \lambda_h - \lambda_b \leq 0 \\ \beta_0 - \dfrac{\log\left(\frac{\pi^2}{12\, F_{\text{ser}}\, L^2} \right) + \lambda_E + \lambda_b + 3\,\lambda_h}{\sqrt{\zeta_E^2 + \zeta_b^2 + 9\zeta_h^2}} \leq 0 \end{cases}, \tag{4.12}$$

whose solution consists in saturating both constraints. In other words, the optimal design is the square section ($\lambda_b = \lambda_h$) with optimal width:

$$\lambda_b^* = \lambda_h^* = \frac{1}{4}\left(\beta_0\, \sqrt{\zeta_E^2 + \zeta_b^2 + 9\zeta_h^2} - \log\left(\frac{\pi^2}{12\, F_{\text{ser}}\, L^2} \right) - \lambda_E \right). \tag{4.13}$$

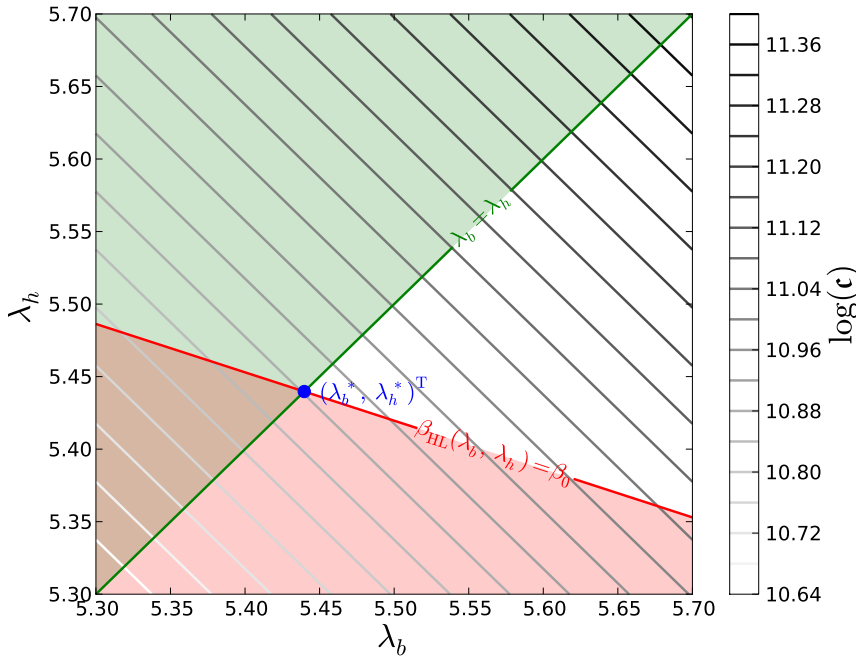


Figure 4.2: Objective and constraint functions for the introductory RBDO analytical example in the (λ_b, λ_h)-space.

4.1.2.4 Illustration

The concrete column is now assumed to have the following properties:

- a lognormal Young's modulus E with mean $\mu_E = 10,000$ MPa and coefficient of variation $\delta_E = 10\%$;
- lognormal width b and height h with constant and equal coefficients of variation $\delta = \delta_b = \delta_h = 5\%$;
- and a deterministic length $L = 3,000$ mm;

It is subjected to a deterministic service load of magnitude:

$$F_{\text{ser}} = \frac{\pi^2 \mu_E I_0}{L^2} \quad \text{with} \quad I_0 = \frac{200 \times 200^3}{12}. \tag{4.14}$$

Under such assumptions, the optimal mean design for a minimum reliability index $\beta_0 = 3$ is $\mu_b^* = \mu_h^* \approx 231$ mm. Note that this is about 15% more than the optimal deterministic design ($\mu_b^{\text{det}} = 200$ mm).

The optimization problem formulated in Eq. (4.13) is illustrated in Figure 4.2. It is clear from this illustration that the optimal solution lies at the intersection of the two linear constraints (the green and red lines) independently of the objective function which is also linear. Note that the linearity is specific to this simple example though.

4.1.3 State-of-the-art and perspectives

The review introduced in this section is mostly based on the book by Tsompanakis et al. (2008) (see the chapter by Chateauneuf and Aoues, 2008, in particular) and the review articles by Eldred et al. (2002), Valdebenito and Schuëller (2010) and Aoues and Chateauneuf (2010).

The most straightforward approach to solve the RBDO problem in Eq. (4.1) consists in nesting a reliability analysis within a nonlinear constrained optimization loop. Such methods are referred to as *double-loop* or *nested* approaches. Despite their conceptual simplicity and their undeniable accuracy these "brute-force" approaches are often argued to lack efficiency since they require too many evaluations of the performance functions that might involve the output of a time-consuming computer code. However, for a broad range of applications where the performance functions are linear or weakly nonlinear – *i.e.* when the first order reliability method is applicable, the nested approach is able to give results within a reasonable number of performance functions evaluations (*e.g.* Enevoldsen and Sørensen (1994) achieve convergence within a few thousands evaluations).

The formulation in Eq. (4.1) is known as the *reliability index approach* (RIA). Note that RIA often refers to nested RBDO algorithm based on FORM despite the fact that the concept can easily be extended to sampling-based reliability methods. Some authors, starting with Tu et al. (1999), proposed an alternative formulation to the RBDO problem known as the *performance measure approach* (PMA). The probabilistic constraints are transformed into quantile constraints that are approximated using the first-order reliability theory. According to Youn and Choi (2004), PMA would be more stable and efficient than RIA because the inner reliability algorithms used to solve the first-order quantile approximation (the so-called *Mean Value* algorithms) are argued to be much more efficient than their probability approximation counterparts (*e.g.* the iHLRF algorithm evoked in Section 3.3.4.1 of Chapter 3).

The nested approach becomes intractable in case of more complex performance functions for which the nested reliability analysis must resort to sampling-based methods. For such cases however, Royset and Polak (2004a,b) proposed the so-called *sample average approximation* method which consists in gradually refining the sampling-based reliability analysis as the optimization algorithm converges towards an optimal design. To do so, they propose an empirical stepwise refinement criteria to define whether the sample size should be raised or not. This can allow a quite significant reduction of the total number of simulation runs but it still requires too many performance functions evaluations to make the approach applicable to real engineering cases. Nevertheless, the application field of this approach is growing with the increasing availability of high performance computational resources (*e.g.* interconnected clusters of PCs).

An alternative to double-loop approach consists in decoupling the optimization loop from the reliability analyses so that both can be sequentially performed in an independent manner. Such approaches are referred to as *sequential approaches* or *decoupled approaches* (Royset et al., 2001; Du and Chen, 2004; Aoues and Chateauneuf, 2010). An advantage of these approaches is that they do not require reliability sensitivity analysis since the op-

timization is performed directly on the performance function. However, the decoupling often relies on the *most probable failure point* (MPFP) assumptions and thus suffers from the possible non-uniqueness of this point and the strong nonlinearities in the performance functions.

Single-loop approaches (Kuschel and Rackwitz, 1997; Kirjner-Neto et al., 1998; Kharmanda et al., 2002; Shan and Wang, 2008) attempt to fully reformulate the original RBDO problem into an equivalent DDO problem that enables a simple and efficient resolution by means of classical optimization algorithms. The approach is mainly based on concepts that are closely related with the notion of partial safety factors. It is certainly the most computationally efficient approach as soon as the assumptions under which the probabilistic-deterministic equivalence is built hold. Once again, most of them are based upon the assumption that the MPFP exists and that it is unique.

Stochastic subset optimization (SSO) is a sampling-based approach recently proposed by Taflanidis (2007); Taflanidis and Beck (2008, 2009a,b) that consists in finding the region of the admissible design space where the failure probability density function is minimal. It is based on conditional sampling in a so-called *augmented reliability space* where the design variables are artificially considered as random with uniform distribution. The range of uncertainty in the design variables is reduced along with the identification of low failure probability density regions. The overall concept is closely related with the subset sampling reliability method proposed by Au and Beck (2001), and further explored by Au (2005) for reliability-based design sensitivity analysis. In the end, the algorithm provides a set of parameters that is likely to contain the optimal solution which can possibly be pinpointed by means of a more refined stochastic search algorithm. However, the problem that SSO attempts to solve is not a full RBDO problem in the sense that it is designed to minimize the failure probability whereas the purpose of RBDO is to minimize a cost function under some maximal failure probability constraint.

Finally, the approach that is investigated here is the *surrogate-based* (or *meta-model-based*) approach. Since the performance function evaluation might involve a time-consuming computational task, the approach consists in replacing this function by a surrogate that is much faster to evaluate (see Section 3.4 of Chapter 3). This approach is argued here to be more flexible than the classical methods based on Taylor series expansions (*i.e.* approaches using the first-order reliability theory) that are extensively used in the RBDO literature. As for surrogate-based reliability analyses the number of different contributions in this field is quite large as it depends on the meta-modelling strategy (type and refinement), the reliability estimation technique, the constrained optimization solver and how these three tools are used.

Eldred et al. (2002) provide a thorough review of the different fundamental principles underlying surrogate-based RBDO (substitution of the performance functions only, substitution of the failure probabilities only, or substitution of both). Papadrakakis and Lagaros (2002) resort to neural network and Monte Carlo sampling. Eldred et al. (2008); Eldred (2009); Coelho et al. (2011) describe the random performance G with a polynomial chaos expansion. Coelho et al. (2011) additionally propose to interpret the coefficients in the expansion as functions of the design parameters and then use Kriging meta-models to sur-

rogate them. The strategy is applied to multi-objective optimization accounting for uncertainty. Basudhar and Missoum (2008, 2009) investigate the use of support vector classifiers in order to surrogate the limit-state surface. The approach is applied to structural reliability examples involving random fields. Lee and Jung (2008); Bichon et al. (2009); Lee et al. (2011a,b) make use of a Kriging-based reliability analysis nested in a constrained optimization solver. The article by Bichon et al. (2009) presents different results depending on the type of optimizer that is used.

The present manuscript focuses on the use of adaptive Kriging predictors for the resolution of the RBDO problem in order to be able to *(i)* apply the strategy to design problems featuring expensive-to-evaluate performance functions and multiple MPFPs, and *(ii)* quantify the substitution error on the reliability of the final design.

4.2 Elements of inequality constrained optimization

The problem in Eq. (4.1) is a *nonlinear inequality constrained optimization* problem of the following form:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}\in\mathbb{D}} \quad \mathfrak{c}(\boldsymbol{d}): \quad \mathfrak{f}_i(\boldsymbol{d}) \leq 0, \quad i = 1, \ldots, n_c, \tag{4.15}$$

where $\mathbb{D}$ is the design space, $\mathfrak{c}$ is the objective function and $\{\mathfrak{f}_i, i = 1, \ldots, n_c\}$ regroups here all the constraints (deterministic and probabilistic) for the sake of clarity. The resolution strategies for such problems fundamentally differ depending on whether the design variables $\boldsymbol{d}$ are *continuous* or *discrete*. In this thesis, only continuous design variables are considered. For such problems, the *nonlinear programming* methods seem to be the most appropriate. This section shortly reviews the basic principle underlying these methods before focusing on the presentation of the Polak-He algorithm that is then used.

4.2.1 The Karush-Kuhn-Tucker optimality conditions

A wide range of constrained optimizers for continuous design variables is based on the so-called *Karush-Kuhn-Tucker* (KKT) *necessary optimality conditions*. To implement these optimality conditions, the objective and constraint functions are assumed to be continuously once-differentiable at any local minimum $\boldsymbol{d}^*$. Hence, it enables the computation of their gradients and the KKT optimality conditions read as follows.

Theorem 4.2.1. ***The Karush-Kuhn-Tucker necessary optimality conditions***
Let $\boldsymbol{d}^* \in \mathbb{D}$ denote a vector such that $\mathfrak{g}_i(\boldsymbol{d}^*) \leq 0$, for $i = 1, \ldots, n_c$. If $\boldsymbol{d}^*$ is a local minimum of the constrained optimization problem formulated in Eq. (4.15), then there exists a vector $\boldsymbol{\lambda}^* \in \mathbb{R}^{+n_c}$ such that:

$$\boldsymbol{\nabla}\mathfrak{c}\left(\boldsymbol{d}^*\right) + \sum_{i=1}^{n_c} \lambda_i^* \boldsymbol{\nabla}\mathfrak{f}_i\left(\boldsymbol{d}^*\right) = \boldsymbol{0}, \tag{4.16}$$

$$\lambda_i^* \mathfrak{f}_i\left(\boldsymbol{d}^*\right) = 0, \quad i = 1, \ldots, n_c. \tag{4.17}$$

where $\boldsymbol{\nabla}$ denotes the gradient with respect to $\boldsymbol{d}$.

Note that the KKT multipliers $\boldsymbol{\lambda} = (\lambda_i, i = 1, \ldots, n_c)^\mathsf{T}$ might be equal to zero. This means that not all constraints need to be *activated* at the local minimum $\boldsymbol{d}^*$ (*i.e.* $\{\mathfrak{f}_i(\boldsymbol{d}^*), i = 1, \ldots, n_c\}$ do not need to be all equal to zero) in order to satisfy Eq. (4.17). This is the main difference with the Lagrange optimality conditions for equality constrained optimization problems. Eq. (4.16) means that the gradient of the cost function is a linear combination of the *activated* constraints at any local optimum (see Figure 4.3).

Example 4.2.1.

For instance, the problem of finding the best linear unbiased predictor in Section 1.4.3 of Chapter 1 (see Proof 1.4.1 in particular) is an equality constrained optimization problem *for which the Lagrange optimality conditions apply. All constraints are activated at the optimal solution.*

Conversely, the problem involved in the determination of the coefficients of a soft *support vector classifier in Section 1.1.2.2 of Chapter 1 (see Eq. (1.6)) is an* inequality constrained optimization problem *for which the abovementioned KKT conditions apply: only a few constraints are activated at the optimal solution (they correspond to the support vectors).*

4.2.2 Dual methods

More specifically, the so-called *dual methods* aim at solving directly the latter optimality conditions. Indeed, despite Theorem 4.2.1 introduces an additional set of unknown variables (the n_c Karush-Kuhn-Tucker multipliers $\boldsymbol{\lambda}^*$), it also yields a set of $n_d + n_c$ equations that enable the resolution of the system formed by Eqs. (4.16) and (4.17). However, these equations might eventually be nonlinear and/or non-explicit so that the system formed by the KKT conditions cannot be solved analytically. Hence, their resolution is approximated numerically by finding the global optimum of an augmented optimization problem named the *dual problem*.

Let us define the Lagrangian associated with the optimization problem in Eq. (4.15):

$$\mathscr{L}(\boldsymbol{d}, \boldsymbol{\lambda}) = \mathfrak{c}(\boldsymbol{d}) + \sum_{i=1}^{n_c} \lambda_i \mathfrak{f}_i(\boldsymbol{d}). \tag{4.18}$$

The dual problem consists in finding the global minimizer of the Lagrangian:

$$(\boldsymbol{d}^*, \boldsymbol{\lambda}^*) = \arg \min_{(\boldsymbol{d}, \boldsymbol{\lambda}) \in \mathbb{D} \times \mathbb{R}^{+n_c}} \mathscr{L}(\boldsymbol{d}, \boldsymbol{\lambda}), \tag{4.19}$$

which is an unconstrained global optimization problem. The relation with the primal problem is clear from the first-order necessary optimality conditions obtained by cancelling the partial derivatives of the Lagrangian:

$$\boldsymbol{\nabla}\mathfrak{c}(\boldsymbol{d}^*) + \sum_{i=1}^{n_c} \lambda_i^* \boldsymbol{\nabla}\mathfrak{f}_i(\boldsymbol{d}^*) = \boldsymbol{0}, \tag{4.20}$$

$$\mathfrak{f}_i(\boldsymbol{d}^*) = 0, \quad i = 1, \ldots, n_c. \tag{4.21}$$

However, a fundamental drawback underlying the resolution of the dual problem instead of the primal one is that it assumes that all the constraints are activated (see Eq. (4.21)) at the optimum which is non-necessarily the case for inequality constrained optimization. As a matter of fact, most implementations of the resolution of this dual problem proceeds by testing all possible combinations of activated constraints (*i.e.* the constraints are sequentially removed and added). Hence, the combinatorial complexity of this procedure rapidly increases with the number of constraints n_c.

4.2.3 The Polak-He algorithm

The Polak-He algorithm (Polak, 1997, Section 2.6) attempts to circumvent the aforementioned combinatorial complexity by penalizing the most violated constraint. Indeed, Polak (1997) showed under reasonable assumptions that solving the problem in Eq. (4.15) is equivalent (in terms of Theorem 4.2.1) to solving the following *min-max* problem:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}\in\mathbb{D}} \quad \mathfrak{c}(\boldsymbol{d}) : \quad \mathfrak{f}(\boldsymbol{d})_+ \leq 0, \tag{4.22}$$

where $\mathfrak{f}(\boldsymbol{d})_+ = \max(0;\ \{\mathfrak{f}_i(\boldsymbol{d}),\ i = 1, \ldots, n_c\})$ denotes the most violated constraint at point $\boldsymbol{d}$.

Polak (1997) then came up with the resolution algorithm introduced in this section. This algorithm proceeds iteratively on j starting from an initial point $\boldsymbol{d}^{(0)}$. As many other nonlinear programming algorithms, the improvement rule that defines the next point at iteration j reads as follows:

$$\boldsymbol{d}^{(j+1)} = \boldsymbol{d}^{(j)} + s^{(j)}\,\boldsymbol{h}^{(j)}, \tag{4.23}$$

where $\boldsymbol{h}^{(j)}$ is the *descent direction*, and $s^{(j)}$ is the *step size* along this direction.

4.2.3.1 Descent direction

The j-th descent direction $\boldsymbol{h}^{(j)}$ is computed as follows:

$$\boldsymbol{h}^{(j)} = -\frac{1}{\delta}\left(\mu_0^*\,\boldsymbol{\nabla}\mathfrak{c}(\boldsymbol{d}^{(j)}) + \sum_{i=1}^{n_c}\mu_i^*\,\boldsymbol{\nabla}\mathfrak{f}_i(\boldsymbol{d}^{(j)})\right), \tag{4.24}$$

where $(\mu_0^*, \boldsymbol{\mu}^{*\mathsf{T}})^\mathsf{T}$ is any solution of the following *linearly constrained quadratic programming* subproblem:

$$\left(\mu_0^*, \boldsymbol{\mu}^{*\mathsf{T}}\right)^\mathsf{T} = \arg\min_{(\mu_0,\boldsymbol{\mu}^\mathsf{T})^\mathsf{T}\in\mathbb{R}^{n_c+1}} \quad \theta^{(j)}\left(\mu_0, \boldsymbol{\mu}\right) : \quad \begin{cases} 0 \leq \mu_i \leq 1, \quad i = 0, \ldots, n_c \\ \sum_{i=0}^{n_c}\mu_i = 1 \end{cases}, \tag{4.25}$$

whose (quadratic) objective function reads:

$$\theta^{(j)}\left(\mu_0, \boldsymbol{\mu}\right) = \mu_0\,\gamma\,\mathfrak{f}(\boldsymbol{d})_+ + \sum_{i=1}^{n_c}\mu_i\,\mathfrak{f}_i(\boldsymbol{d}) + \frac{1}{2\,\delta}\left\|\mu_0^*\,\boldsymbol{\nabla}\mathfrak{c}(\boldsymbol{d}^{(j)}) + \sum_{i=1}^{n_c}\mu_i^*\,\boldsymbol{\nabla}\mathfrak{f}_i(\boldsymbol{d}^{(j)})\right\|^2. \tag{4.26}$$

In the latter expression, $\gamma > 0$ and $\delta > 0$ are two of the four tuning parameters of the Polak-He algorithm (the two others being used for the step size calculation). Polak (1997) guarantees convergence for any admissible values of these two parameters but warns about their influence on the convergence rate.

4.2.3.2 Step size

The step size is determined by means of the Goldstein-Armijo approximate line search rule. Given $0 < \alpha < 1$ and $0 < \beta < 1$, the rule consists in finding $s^{(j)}$ such that:

$$s^{(j)} = \max_{k \in \mathbb{N}} \left\{ \beta^k : M\left(\boldsymbol{d}^{(j)}, \boldsymbol{d}^{(j)} + \beta^k \boldsymbol{h}^{(j)}\right) \leq -\beta^k \alpha \theta^{(j)}\left(\mu_0^*, \boldsymbol{\mu}^*\right) \right\}, \tag{4.27}$$

where the merit function M measures the improvement brought by the step size and is defined here as:

$$M\left(\boldsymbol{d}^{(j)}, \boldsymbol{d}^{(j)} + \beta^k \boldsymbol{h}^{(j)}\right) = \max \begin{pmatrix} \mathfrak{c}\left(\boldsymbol{d}^{(j)} + \beta^k \boldsymbol{h}^{(j)}\right) - \mathfrak{c}\left(\boldsymbol{d}^{(j)}\right) - \gamma \mathfrak{f}\left(\boldsymbol{d}^{(j)}\right)_+ \\ \left\{ \mathfrak{f}_i\left(\boldsymbol{d}^{(j)}\right) - \mathfrak{f}\left(\boldsymbol{d}^{(j)}\right)_+, i = 1, \ldots, n_c \right\} \end{pmatrix}. \tag{4.28}$$

Again, Polak (1997) points out that the two tuning parameters α and β may alter the convergence rate of the algorithm.

4.2.3.3 Implementation

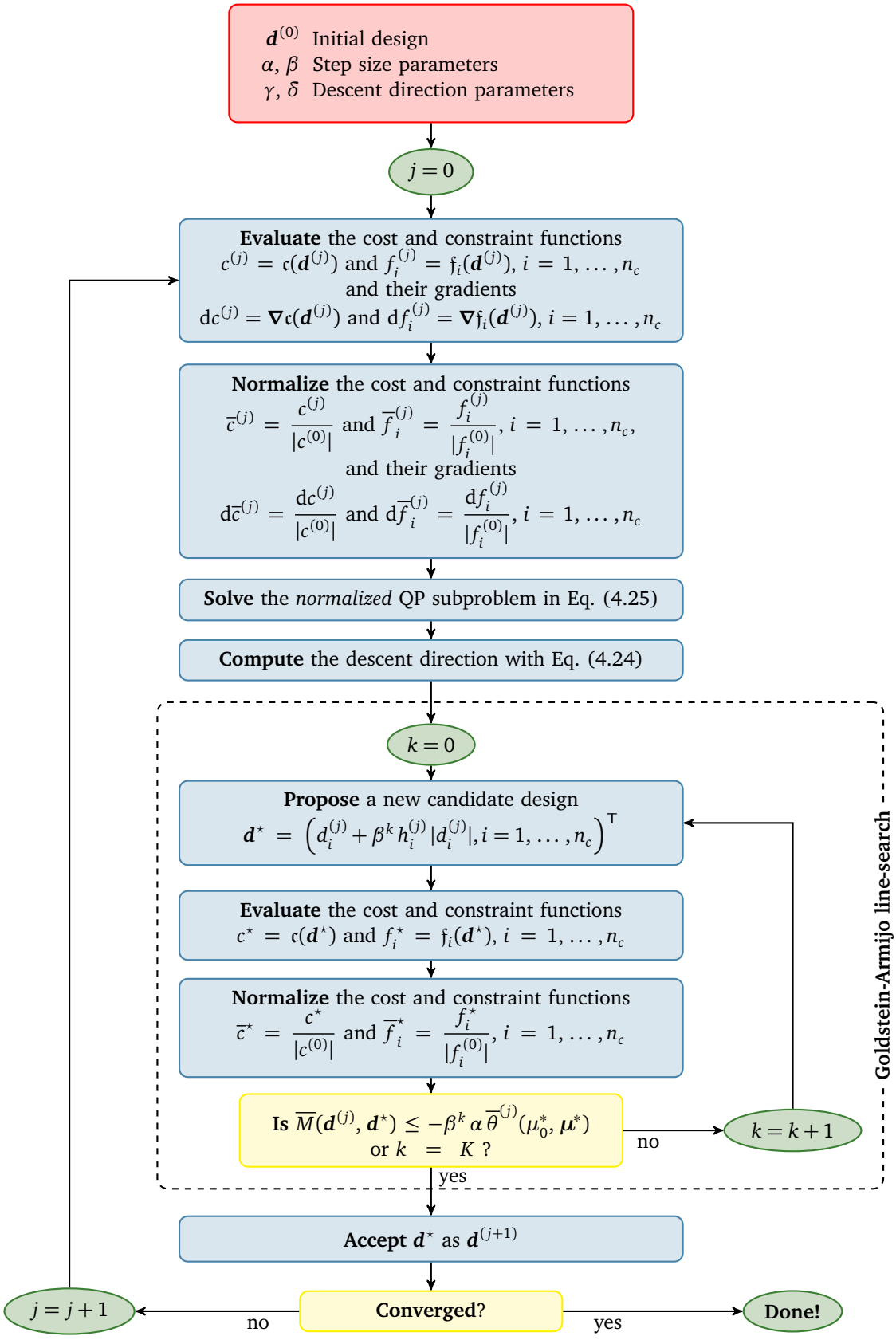
Algorithm 4.1 is provided so as to summarize this section. The algorithm requires an initial point $\boldsymbol{d}^{(0)}$ and the values of the four tuning parameters of the Polak-He algorithm. The present implementation involves the normalization of the design variables $\boldsymbol{d}$, the cost function $\mathfrak{c}$, the constraint functions $\{\mathfrak{f}_i, i = 1, \ldots, n_c\}$ and their respective gradients. More specifically, the cost and constraint functions and their gradients are normalized with respect to their initial value (provided it is not zero), and the design variables are normalized with respect to their value at the previous iteration (provided it is not zero) in order to make the step along the improvement direction significant. Thanks to this normalization, the following set of parameters:

$$\alpha = 0.5, \quad \beta = 0.6, \quad \gamma = 2, \quad \delta = 1 \tag{4.29}$$

yields satisfactory convergence rates independently of the order of magnitudes of the different variables and functions involved in the problem at hand. This set of parameters was initially recommended by Royset and Polak (2004a,b), without pointing out the convergence issues that may occur if the aforementioned normalizations are not applied though.

The algorithm iterates on j starting from the provided initial design $\boldsymbol{d}^{(0)}$. First, the cost function, the constraint functions and their respective gradients are evaluated at the current design. The function values are normalized with respect to their absolute initial values $c^{(0)}$ and $\{f_i^{(0)}, i = 1, \ldots, n_c\}$ (provided they are not equal to zero). The gradients are also normalized with respect to these initial values.

Algorithm 4.1 The Polak-He algorithm



Then the j-th descent direction $\boldsymbol{h}^{(j)}$ in Eq. (4.24) is computed by resolving the *normalized* QP subproblem. This problem reads as in Eq. (4.25) except that the cost and constraint functions values as well as their gradients in Eq. (4.26) are replaced with their normalized counterparts. In the present implementation, the QP subproblem is solved by means of the `quadprog` function available in Matlab's Optimization toolbox. It also proceeds iteratively from a starting vector $(\mu_0, \boldsymbol{\mu}^{\mathsf{T}})^{\mathsf{T}}$ which is presently initialized at random in the admissible space defined in Eq. (4.25).

The second step of the j-th iteration consists in finding the power of β that will determine the j-th step size $s^{(j)}$. The power is initialized at $k = 0$ and a first point is proposed by means of the improvement rule in Eq. (4.23). Note that a slight variant of this improvement rule is proposed in Algorithm 4.1 in order to account for the different orders of magnitude of each design variable in $\boldsymbol{d}$. This normalization greatly improves the versatility of the algorithm in terms of convergence rates. The merit function in Eq. (4.28) also makes use of normalized values for the cost and constraint functions. Polak (1997) ensures that there exists a $k \in \mathbb{N}$ such that the merit function ends up being satisfied. In practice though, the approximate line-search is aborted after K increments (say $K = 10$ so that $s^{(j)} \leq \beta^K = 6 \times 10^{-3}$ in the present implementation).

The algorithm then repeats the steps detailed in the previous paragraphs until it converges. Convergence is detected by means of the usual criteria, namely:

(i) if the latest improvement brought to the objective function is insignificant with respect to the starting value:

$$\left| \frac{\mathfrak{c}\left(\boldsymbol{d}^{(j)}\right) - \mathfrak{c}\left(\boldsymbol{d}^{(j+1)}\right)}{\mathfrak{c}\left(\boldsymbol{d}^{(0)}\right)} \right| \leq \varepsilon_c, \tag{4.30}$$

where ε_c controls the quality of the optimal solution;

(ii) *and* if all the feasibility constraints are fulfilled:

$$\mathfrak{f}_i(\boldsymbol{d}^{(j+1)}) \leq 0, \quad i = 1, \ldots, n_c. \tag{4.31}$$

4.2.4 Illustration

The Polak-He algorithm is applied to a (deterministic) inequality constrained optimization problem inspired from the article by Lee and Jung (2008, Example 1). The problem consists in finding the optimal design $\boldsymbol{d}^*$ defined as follows:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d} \in \mathbb{D}} \quad \mathfrak{c}(\boldsymbol{d}) : \quad \begin{cases} \mathfrak{f}_1(\boldsymbol{d}) \leq 0 \\ \mathfrak{f}_2(\boldsymbol{d}) \leq 0 \end{cases}, \tag{4.32}$$

where:

$\mathbb{D}$ is the admissible domain that is set here equal to:

$$\mathbb{D} = [0;\ 3.7] \times [0;\ 4]\,; \tag{4.33}$$

$\mathfrak{c}$ is the objective function defined as:

$$\mathfrak{c}(\boldsymbol{d}) = (d_1 - 3.7)^2 + (d_2 - 4)^2; \tag{4.34}$$

$\mathfrak{f}$ are the constraint functions defined as:

$$\mathfrak{f}_1(\boldsymbol{d}) = d_1 \sin(4\,d_1) + 1.1\,d_2 \sin(2\,d_2), \tag{4.35}$$
$$\mathfrak{f}_2(\boldsymbol{d}) = 3 - d_1 - d_2. \tag{4.36}$$

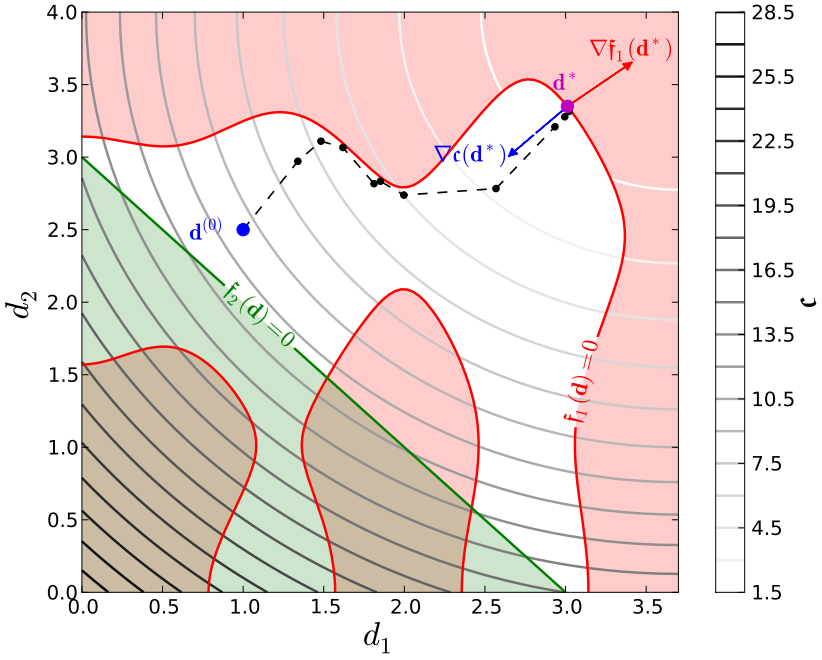


Figure 4.3: Convergence of the Polak-He algorithm for the two-dimensional constrained optimization problem defined in Eq. (4.32).

This problem is illustrated in Figure 4.3. The contours of the objective function are represented in gray. It is easy to show that the quadratic objective function is minimum for $d_1 = 3.7$ and $d_2 = 4$ (*i.e.* in the upper right corner of Figure 4.3). However, this optimal solution is not admissible with respect to the first constraint $\mathfrak{f}_1$ (in red) as illustrated in Figure 4.3. The optimization problem involves 6 constraints: the two constraints explicitly formulated in Eqs (4.35) and (4.36) and the 4 constraints used for limiting the search to the rectangular domain $\mathbb{D}$. The initial design is set equal to $\boldsymbol{d}^{(0)} = (1, 2.5)^\mathsf{T}$ and the algorithm converges at $\boldsymbol{d}^* = (3.01, 3.35)^\mathsf{T}$ for $\epsilon_c = 10^{-4}$ after 14 iterations. The latter point is clearly the best trade-off between the minimization of the objective function and the fulfillment of the constraints. It should also be noticed that the first constraint is activated at the optimum so that this optimal solution is rather sensitive to the possible uncertainty in $\boldsymbol{d}$ or in the constraint $\mathfrak{f}_1(\boldsymbol{d})$ as it lies on the boundary of the admissible space.

In addition, it can be checked that $\boldsymbol{d}^*$ satisfies the KKT first-order necessary optimality conditions. Since $\mathfrak{f}_1$ is the only constraint activated at the optimum, the first KKT multiplier λ_1 is the only nonzero multiplier. Hence, $\boldsymbol{\nabla}\mathfrak{c}(\boldsymbol{d}^*)$ and $\boldsymbol{\nabla}\mathfrak{f}_1(\boldsymbol{d}^*)$ should satisfy the first

condition of Theorem 4.2.1 which reduces here to:

$$\boldsymbol{\nabla} \mathfrak{c}(\boldsymbol{d}^*) + \lambda_1 \boldsymbol{\nabla} \mathfrak{f}_1(\boldsymbol{d}^*) = \mathbf{0}. \tag{4.37}$$

The latter relation is almost checked for $\lambda_1 \approx 0.15$. The two gradients are indeed not fully collinear due to the numerical resolution of the problem. The second condition of Theorem 4.2.1 is also verified because $\mathfrak{f}_1(\boldsymbol{d}^*) \approx 0$ (again the approximation is due to the numerical resolution of the problem).

4.2.5 Conclusion

This section has reviewed the basic principles underlying the resolution of nonlinear inequality constrained optimization problem. Thanks to its min-max-based formulation, the Polak-He algorithm is able to deal with heavily constrained optimization problems such as the RBDO problem of interest in this chapter. Nonetheless, the use of such an algorithm requires the knowledge of the gradients of the objective and constraint functions, including that of the probabilistic constraint. This remark naturally motivates the next section.

4.3 Reliability sensitivity analysis

Sensitivity of the failure probability p_f to the parameters in the joint *probability density function* (PDF) $f_{\boldsymbol{X}}$ is of major importance in reliability-based engineering as it gives crucial information on the decisions to be made. In essence, it answers the fundamental question that follows a reliability analysis:

"How should I change the design variables in the probabilistic model in order to achieve a chosen reliability level?".

In addition, it is a key ingredient for the implementation of a gradient-based optimization algorithm (such as Algorithm 4.1) for the resolution of the RBDO problem formulated in Eq. (4.1).

4.3.1 Problem statement

4.3.1.1 Definition

First, recall from Chapter 3 that the failure probability p_f is defined as follows:

$$p_f(\boldsymbol{d}) = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d}) \, \mathrm{d}\boldsymbol{x} \equiv \mathbb{E}\left[\mathbb{1}_{\mathbb{F}}(\boldsymbol{X})\right]. \tag{4.38}$$

where $\mathbb{1}_{\mathbb{F}}$ is the failure indicator function indicating whether a realization $\boldsymbol{x}$ lies in the failure domain $\mathbb{F}$ or not, and $f_{\boldsymbol{X}}$ is the joint PDF of $\boldsymbol{X}$ parametrized by $\boldsymbol{d}$. According to this

setup, the gradient of the failure probability p_f with respect to the vector of parameters $\boldsymbol{d}$ in the probability density function $f_{\boldsymbol{X}}$ is defined as follows:

$$\boldsymbol{\nabla} p_f(\boldsymbol{d}) = \left(\frac{\partial p_f}{\partial d_k}, k = 1, \dots, n_d \right)^{\mathsf{T}}, \tag{4.39}$$

where the partial derivative with respect to the k-th design variable reads:

$$\frac{\partial p_f(\boldsymbol{d})}{\partial d_k} = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{p_f(\boldsymbol{d} + \varepsilon \boldsymbol{e}_k) - p_f(\boldsymbol{d})}{\varepsilon}. \tag{4.40}$$

where $\boldsymbol{e}_k$ is the k-th canonical basis vector of $\mathbb{R}^{n_d}$ that has all its components equal to zero except the k-th one which is equal to one.

4.3.1.2 Common random numbers

The computational procedure that consists in a direct use of Eq. (4.40) for a sufficiently small ε is known as the *finite forward differences* technique and its use is widespread in numerical optimization. In the present uncertain context though, this technique has two fundamental drawbacks. First, it can be seen that it requires n_d+1 estimations of the failure probability for a vector $\boldsymbol{d}$ of size n_d. Second, if each term in the numerator of the partial derivative in Eq. (4.39) is estimated by means of Monte Carlo sampling then its variance of estimation is proportional to:

$$\begin{aligned} \operatorname{Var}\left[\widehat{p}_{f\,N_2}(\boldsymbol{d} + \varepsilon \boldsymbol{e}_k) - \widehat{p}_{f\,N_1}(\boldsymbol{d}) \right] = \operatorname{Var}\left[\widehat{p}_{f\,N_2}(\boldsymbol{d} + \varepsilon \boldsymbol{e}_k) \right] + \operatorname{Var}\left[\widehat{p}_{f\,N_1}(\boldsymbol{d}) \right] \\ - 2 \operatorname{Cov}\left[\widehat{p}_{f\,N_2}(\boldsymbol{d} + \varepsilon \boldsymbol{e}_k), \widehat{p}_{f\,N_1}(\boldsymbol{d}) \right], \end{aligned} \tag{4.41}$$

where $\widehat{p}_{f\,N_1}$ (resp. $\widehat{p}_{f\,N_2}$) denotes an unbiased Monte Carlo estimator computed from an N_1-sample $\mathscr{X}^{[1]}$ (resp. N_2-sample $\mathscr{X}^{[2]}$) of independent copies of the random vector $\boldsymbol{X}$. The variances of estimation simply sum each others in the most basic case where the two terms are estimated independently.

Hence for the sake of variance reduction, Royset and Polak (2004a,b) and Taflanidis (2007), amongst others, propose to use *common random numbers* (CRN). This procedure essentially consists in using the same seed for generating both samples $\mathscr{X}^{[1]}$ and $\mathscr{X}^{[2]}$. Indeed, CRN introduces dependence in the estimation of the two probabilities (provided ε is small) so that the third covariance term in Eq. (4.41) reduces the variance of estimation of the quantity of interest. However, since CRN is still computationally demanding (it requires $n_d + 1$ reliability analyses for a single evaluation of the gradient $\boldsymbol{\nabla} p_f$) and since it is not applicable to more advanced Monte Carlo sampling techniques such as subset sampling, it was not further considered in this thesis.

4.3.1.3 Other approaches

Reliability sensitivity analysis has been extensively used in the specific context of first- and second-order reliability methods. For instance, Bjerager and Krenk (1989) computed the

gradient of the Hasofer-Lind reliability index β_{HL} (see also Ditlevsen and Madsen, 1996, Chapter 8) which in turns gives the gradient of the first-order approximation of the failure probability. However, the approach is limited to structural reliability problems for which the assumption that the MPFP is unique holds.

More recently, Garnier et al. (2009) provided asymptotic formulæ for the estimation of the first- and second-order derivatives of a probability estimated by means of Monte Carlo sampling or quadrature. However, their investigations are limited to the multivariate Gaussian distribution under the assumption of small variance. The sensitivities are then applied to *chance-constrained optimization* which is another name for the problem formulated in Eq. (4.1).

4.3.2 The score function approach

The approach that is used in this thesis was originally proposed by Rubinstein (1976). It was later named the *score function* (SF) approach in Rubinstein (1986) (see also Rubinstein and Kroese, 2008, Chapter 7) and recently brought to the structural reliability community by Wu (1994).

First, the partial derivative of the failure probability in Eq. (4.38) with respect to the k-th design variable reads:

$$\frac{\partial p_f(\boldsymbol{d})}{\partial d_k} = \frac{\partial}{\partial d_k} \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})\, \mathrm{d}\boldsymbol{x}. \tag{4.42}$$

Assuming that *(i)* the joint PDF $f_{\boldsymbol{X}}$ is continuously differentiable with respect to d_k and that *(ii)* the integration range $\mathbb{X}$ does not depend on d_k, the partial derivative of the failure probability recasts as follows:

$$\frac{\partial p_f(\boldsymbol{d})}{\partial d_k} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) \frac{\partial f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{\partial d_k}\, \mathrm{d}\boldsymbol{x}. \tag{4.43}$$

Note that the latter equation does not hold for the uniform distribution as the bounds of its integration range depends on its statistical moments (Lee et al., 2011a,b). This remark unfortunately holds for most truncated distributions.

Then, in order to compute this integral as an expectation with respect to the same probability measure as the one used for estimating the failure probability itself, it is genuinely proposed to use an *importance sampling trick* (see Section 3.3.3 of Chapter 3). It proceeds as follows:

$$\begin{aligned}
\frac{\partial p_f(\boldsymbol{d})}{\partial d_k} &= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) \frac{\partial f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d}) / \partial d_k}{f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})} f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})\, \mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\boldsymbol{x}) \frac{\partial \log f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{\partial d_k} f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})\, \mathrm{d}\boldsymbol{x} \\
&= \mathbb{E}\left[\mathbb{1}_{\mathbb{F}}(\boldsymbol{X})\, \kappa_k(\boldsymbol{X}, \boldsymbol{d})\right],
\end{aligned} \tag{4.44}$$

where the so-called *score function*:

$$\kappa_k(\boldsymbol{X}, \boldsymbol{d}) = \frac{\partial \log f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{\partial d_k} \tag{4.45}$$

has been introduced.

Thus, given a sample $\mathscr{X} = \{\boldsymbol{X}^{(i)}, i = 1, \ldots, N\}$ of N independent copies of the random vector $\boldsymbol{X}$, the following estimator:

$$\widehat{\frac{\partial p_f(\boldsymbol{d})}{\partial d_k}}_{\text{MCS}} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\mathbb{F}}\left(\boldsymbol{X}^{(i)}\right) \kappa_k\left(\boldsymbol{X}^{(i)}, \boldsymbol{d}\right) \tag{4.46}$$

is unbiased and asymptotically convergent according to the central limit theorem. In addition to these first convenient properties, it can also be seen that the estimation of the failure probability and that of its gradient can be done with the same N-sample. Hence, reliability sensitivity analysis through the score function approach is a *simple post-processing* of a Monte-Carlo-sampling-based reliability analysis. It does not require any additional simulation of the indicator function $\mathbb{1}_{\mathbb{F}}$.

It should also be noticed that the approach extends to both higher order derivatives, provided the joint PDF is sufficiently differentiable (Rubinstein and Kroese, 2008; Millwater, 2009), and statistical moments of any order and of any variable (*i.e.* not only indicator functions).

4.3.3 Estimation using subset sampling

Another interesting feature of this *importance sampling trick* is that it applies to other estimation techniques such as subset sampling. The following presentation closely follows the one given in the article by Song et al. (2009).

First, recall from Section 3.3.5 of Chapter 3 that provided a strictly decreasing sequence of positive reals $q_1 > \ldots > q_s = 0$, the splitted form of the failure probability is as follows:

$$p_f(\boldsymbol{d}) = p_1(\boldsymbol{d}) \prod_{i=2}^{s} p_{i|i-1}(\boldsymbol{d}), \tag{4.47}$$

where the intermediate probabilities are:

$$p_1(\boldsymbol{d}) = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_1}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d}) \, \mathrm{d}\boldsymbol{x}, \tag{4.48}$$

$$p_{i|i-1}(\boldsymbol{d}) = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x}) h_i^*(\boldsymbol{x}, \boldsymbol{d}) \, \mathrm{d}\boldsymbol{x}, \quad i = 2, \ldots, s. \tag{4.49}$$

In the latter expression, it is also recalled that the i-th instrumental PDF reads:

$$h_i^*(\boldsymbol{x}, \boldsymbol{d}) = \frac{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{p_{i-1}(\boldsymbol{d})}, \tag{4.50}$$

with its normalizing constant equal to:

$$p_{i-1}(\boldsymbol{d}) = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})\, \mathrm{d}\boldsymbol{x}. \tag{4.51}$$

Song et al. (2009) showed that the partial derivative of the failure probability in Eq. (4.47) with respect to the k-th design variable reads:

$$\frac{\partial p_f(\boldsymbol{d})}{\partial d_k} = \frac{p_f(\boldsymbol{d})}{p_1(\boldsymbol{d})} \frac{\partial p_1(\boldsymbol{d})}{\partial d_k} + \sum_{i=2}^{s} \frac{p_f(\boldsymbol{d})}{p_{i|i-1}(\boldsymbol{d})} \frac{\partial p_{i|i-1}(\boldsymbol{d})}{\partial d_k}, \tag{4.52}$$

where the partial derivatives of the intermediate probabilities are defined by recurrence:

$$\frac{\partial p_1(\boldsymbol{d})}{\partial d_k} = \mathbb{E}\left[\mathbb{1}_{\mathbb{F}_1}(\boldsymbol{X})\, \kappa_k(\boldsymbol{X}, \boldsymbol{d})\right], \tag{4.53}$$

$$\begin{aligned} \frac{\partial p_{i|i-1}(\boldsymbol{d})}{\partial d_k} = \mathbb{E}_{\boldsymbol{X}^{[i]}}\Bigg[\mathbb{1}_{\mathbb{F}_i}(\boldsymbol{X}^{[i]}) \Bigg(& \kappa_k(\boldsymbol{X}^{[i]}, \boldsymbol{d}) - \frac{1}{p_1(\boldsymbol{d})} \frac{\partial p_1(\boldsymbol{d})}{\partial d_k} \\ & - \sum_{j=2}^{i-1} \frac{1}{p_{j|j-1}(\boldsymbol{d})} \frac{\partial p_{j|j-1}(\boldsymbol{d})}{\partial d_k} \Bigg) \Bigg], \quad i = 2, \ldots, s, \end{aligned} \tag{4.54}$$

where κ_k denotes the score function defined in Eq. (4.45) and $\boldsymbol{X}^{[i]}$ is the random vector that is distributed according to the i-th suboptimal instrumental PDF h_i^*.

Proof 4.3.1. *The proof proceeds in three steps:*

(i) *The expression in Eq. (4.52) can easily be derived from Eq. (4.47) as the derivative of a product of s terms.*

(ii) *The expression for the partial derivative of the first probability p_1 can easily be computed as in Eq. (4.44).*

(iii) *The computation of the partial derivatives of the other $s-1$ intermediate probabilities is somewhat more involved and is as follows. Let us substitute the i-th instrumental PDF h_i^* for its expression in Eq. (4.49):*

$$p_{i|i-1}(\boldsymbol{d}) = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x}) \frac{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{p_{i-1}(\boldsymbol{d})}\, \mathrm{d}\boldsymbol{x}. \tag{4.55}$$

Assuming the two conditions formulated above Eq. (4.43) apply, the partial derivative of the latter integral with respect to the k-th design variable rewrites as follows:

$$\begin{aligned} \frac{\partial p_{i|i-1}(\boldsymbol{d})}{\partial d_k} &= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x}) \frac{\partial}{\partial d_k} \left(\frac{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{p_{i-1}(\boldsymbol{d})} \right) \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x})\, \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) \left(\frac{\frac{\partial f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{\partial d_k} p_{i-1} - f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d}) \frac{\partial p_{i-1}(\boldsymbol{d})}{\partial d_k}}{p_{i-1}(\boldsymbol{d})^2} \right) \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{4.56}$$

Using the same importance sampling trick as in Eq. (4.44), it comes that:

$$\frac{\partial p_{i|i-1}(\boldsymbol{d})}{\partial d_k} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x})\, \mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x}) \left(\frac{\frac{\partial f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})}{\partial d_k}\, p_{i-1} - f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})\, \frac{\partial p_{i-1}(\boldsymbol{d})}{\partial d_k}}{p_{i-1}(\boldsymbol{d})^2} \right) \frac{h_i^*(\boldsymbol{x},\boldsymbol{d})}{h_i^*(\boldsymbol{x},\boldsymbol{d})}\, \mathrm{d}\boldsymbol{x}. \quad (4.57)$$

This expression can be further elaborated by substituting again the i-th instrumental PDF h_i^* for its expression in the denominator on the right:

$$\begin{aligned} \frac{\partial p_{i|i-1}(\boldsymbol{d})}{\partial d_k} &= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x})\, \frac{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})\, p_{i-1}(\boldsymbol{d})}{\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})} \left(\frac{\frac{\partial f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})}{\partial d_k}\, p_{i-1} - f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})\, \frac{\partial p_{i-1}(\boldsymbol{d})}{\partial d_k}}{p_{i-1}(\boldsymbol{d})^2} \right) h_i^*(\boldsymbol{x},\boldsymbol{d})\, \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x})\, \frac{1}{f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})} \left(\frac{\frac{\partial f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})}{\partial d_k}\, p_{i-1} - f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{d})\, \frac{\partial p_{i-1}(\boldsymbol{d})}{\partial d_k}}{p_{i-1}(\boldsymbol{d})} \right) h_i^*(\boldsymbol{x},\boldsymbol{d})\, \mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}_i}(\boldsymbol{x}) \left(\kappa_k(\boldsymbol{x},\boldsymbol{d}) - \frac{1}{p_{i-1}(\boldsymbol{d})} \frac{\partial p_{i-1}(\boldsymbol{d})}{\partial d_k} \right) h_i^*(\boldsymbol{x},\boldsymbol{d})\, \mathrm{d}\boldsymbol{x}, \end{aligned} \quad (4.58)$$

where $\kappa_k(\boldsymbol{x},\boldsymbol{d})$ is the score function defined in Eq. (4.45). Since the normalizing constant of the i-th instrumental PDF is equal to:

$$p_{i-1}(\boldsymbol{d}) = p_1(\boldsymbol{d}) \prod_{j=2}^{i-1} p_{j|j-1}(\boldsymbol{d}), \quad (4.59)$$

its partial derivative with respect to the k-th design variable comes again as the derivative of a product of i terms and it reads as follows:

$$\frac{\partial p_{i-1}(\boldsymbol{d})}{\partial d_k} = \frac{p_{i-1}(\boldsymbol{d})}{p_1(\boldsymbol{d})} \frac{\partial p_1(\boldsymbol{d})}{\partial d_k} + \sum_{j=2}^{i-1} \frac{p_{i-1}(\boldsymbol{d})}{p_{j|j-1}(\boldsymbol{d})} \frac{\partial p_{j|j-1}(\boldsymbol{d})}{\partial d_k}. \quad (4.60)$$

Hence, the partial derivative of the i-th intermediate probability with respect to the k-th design variable finally recasts as the following expectation:

$$\begin{aligned} \frac{\partial p_{i|i-1}(\boldsymbol{d})}{\partial d_k} = \mathbb{E}_{\boldsymbol{X}^{[i]}} \Bigg[\mathbb{1}_{\mathbb{F}_i}(\boldsymbol{X}^{[i]}) \Bigg(\kappa_k(\boldsymbol{X}^{[i]},\boldsymbol{d}) - \frac{1}{p_1(\boldsymbol{d})} \frac{\partial p_1(\boldsymbol{d})}{\partial d_k} \\ - \sum_{j=2}^{i-1} \frac{1}{p_{j|j-1}(\boldsymbol{d})} \frac{\partial p_{j|j-1}(\boldsymbol{d})}{\partial d_k} \Bigg) \Bigg] \end{aligned} \quad (4.61)$$

with respect to $\boldsymbol{X}^{[i]}$ distributed according to the i-th suboptimal instrumental PDF h_i^*.

Hence, the gradient of the failure probability can be estimated along with the failure probability itself in a subset sampling scheme using the same intermediate samples as the ones used to compute $\widehat{p}_1$ and $\{\widehat{p}_{i|i-1},\, i = 2, \ldots, s\}$.

4.3.4 Estimation using meta-model-based importance sampling

The score function approach can also be applied to the original meta-model-based importance sampling estimator proposed in Section 3.5 of Chapter 3.

First, recall that the failure probability was recast as the product of two terms:

$$p_f(\boldsymbol{d}) = \alpha_{\text{corr}}(\boldsymbol{d})\, p_{f\,\varepsilon}(\boldsymbol{d}), \tag{4.62}$$

where the correction factor (resp. the augmented failure probability) reads:

$$\alpha_{\text{corr}}(\boldsymbol{d}) = \int_{\mathbb{X}} \frac{\mathbb{1}_F(\boldsymbol{x})}{\pi(\boldsymbol{x})}\, \widehat{h}^*(\boldsymbol{x}, \boldsymbol{d})\, \mathrm{d}\boldsymbol{x}, \tag{4.63}$$

$$p_{f\,\varepsilon}(\boldsymbol{d}) = \int_{\mathbb{X}} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})\, \mathrm{d}\boldsymbol{x}. \tag{4.64}$$

In these expressions, π denotes the probabilistic classification function that does not depend on $\boldsymbol{d}$ and $\widehat{h}^*$ denotes the proposed meta-model-based instrumental PDF which reads:

$$\widehat{h}^*(\boldsymbol{x}, \boldsymbol{d}) = \frac{\pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{p_{f\,\varepsilon}(\boldsymbol{d})}. \tag{4.65}$$

Then, the partial derivative of the failure probability in Eq. (4.62) with respect to the k-th design variable can easily be derived as that of a product of two terms:

$$\frac{\partial\, p_f(\boldsymbol{d})}{\partial\, d_k} = \frac{\partial\, \alpha_{\text{corr}}(\boldsymbol{d})}{\partial\, d_k}\, p_{f\,\varepsilon}(\boldsymbol{d}) + \alpha_{\text{corr}}(\boldsymbol{d})\, \frac{\partial\, p_{f\,\varepsilon}(\boldsymbol{d})}{\partial\, d_k}, \tag{4.66}$$

where the partial derivative of the augmented failure probability (resp. that of the correction factor) reads as follows:

$$\frac{\partial\, p_{f\,\varepsilon}(\boldsymbol{d})}{\partial\, d_k} = \mathbb{E}\left[\pi(\boldsymbol{X})\, \kappa_k(\boldsymbol{X}, \boldsymbol{d})\right], \tag{4.67}$$

$$\frac{\partial\, \alpha_{\text{corr}}(\boldsymbol{d})}{\partial\, d_k} = \mathbb{E}_{\boldsymbol{Z}}\left[\frac{\mathbb{1}_{\mathbb{F}}(\boldsymbol{Z})}{\pi(\boldsymbol{Z})}\left(\kappa_k(\boldsymbol{Z}, \boldsymbol{d}) - \frac{1}{p_{f\,\varepsilon}(\boldsymbol{d})}\, \frac{\partial\, p_{f\,\varepsilon}(\boldsymbol{d})}{\partial\, d_k}\right)\right], \tag{4.68}$$

where κ_k denotes the score function defined in Eq. (4.45) and $\boldsymbol{Z}$ is the random vector that is distributed according to $\widehat{h}^*$.

Proof 4.3.2. *The proof for the expression of the partial derivative in Eq. (4.66) closely resembles Proof 4.3.1 for subset sampling except that it reduces here to the differentiation of a product of two terms only.*

Again, the gradient of the failure probability can be estimated along with the failure probability itself in a meta-model-based importance sampling scheme using the same samples as the ones used to compute $\widehat{\alpha}_{\text{corr}}$ and $\widehat{p}_{f\,\varepsilon}$.

4.3.5 Evaluation of the score function

The analytical computation of the score function in Eq. (4.45) is available for most parametric univariate distributions as well as in the multivariate case provided the joint distribution is defined using the copula formalism as introduced in Section 3.2 of Chapter 3.

Indeed, recall that if the probabilistic model $\boldsymbol{X}$ is specified in terms of parametric marginal PDFs $\{f_{X_i},\, i = 1, \ldots, n\}$ and a copula density function c, then according to Theorem 3.2.1 of Chapter 3, its joint PDF reads as follows:

$$f_{\boldsymbol{X}}\left(x_1, \ldots, x_n;\, \boldsymbol{d}\right) = c\left(F_{X_1}(x_1, \boldsymbol{d}), \ldots, F_{X_n}(x_n, \boldsymbol{d})\right) \prod_{i=1}^{n} f_{X_i}(x_i, \boldsymbol{d}). \tag{4.69}$$

Note that in the latter equation, it was assumed that the vector of design variables is exclusively involved in the marginal distributions and not in the copula c. Indeed, in this manuscript as well as in most applications of these sensitivities to reliability-based design optimization, the design variables $\boldsymbol{d}$ are means of the marginal distributions (see *e.g.* Dubourg et al., 2011d; Lee et al., 2011a,b).

Assuming that the k-th design variable in vector $\boldsymbol{d}$ is involved in the k-th marginal distribution *only*, the k-th score function reads:

$$\begin{aligned}
\kappa_k(\boldsymbol{x}, \boldsymbol{d}) &= \frac{\partial \log f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{d})}{\partial d_k} \\
&= \frac{\partial \log c\left(F_{X_1}(x_1, \boldsymbol{d}), \ldots, F_{X_k}(x_k, \boldsymbol{d}), \ldots, F_{X_n}(x_n, \boldsymbol{d})\right)}{\partial d_k} + \frac{\partial \log f_{X_k}(x_k, \boldsymbol{d})}{\partial d_k} \\
&= \frac{\partial \log c(u_1, \ldots, u_k, \ldots, u_n)}{\partial u_k} \frac{\partial F_{X_k}(x_k, \boldsymbol{d})}{\partial d_k} + \frac{\partial \log f_{X_k}(x_k, \boldsymbol{d})}{\partial d_k}.
\end{aligned} \tag{4.70}$$

Further analytical developments obviously depend on the copula and the marginal distributions.

Lee et al. (2011a,b) provides the expression for the log-derivatives of most parametric copulas in two dimensions. In this thesis, only the normal copula has been considered. The partial derivative of the logarithm of the normal copula with respect to the k-th uniform margin easily comes from Eq. (3.13), and it reads as follows:

$$\frac{\partial \log c_{\Phi}(u_1, \ldots, u_k, \ldots, u_n;\, \mathbf{R}_0)}{\partial u_k} = -\mathrm{d}\boldsymbol{z}(\boldsymbol{u})^{\mathsf{T}} \left(\mathbf{R}_0^{-1} - \mathbf{I}\right) \boldsymbol{z}(\boldsymbol{u}), \tag{4.71}$$

where $\boldsymbol{z}(\boldsymbol{u}) = (\Phi^{-1}(u_i),\, i = 1, \ldots, n)^{\mathsf{T}}$, $\mathrm{d}\boldsymbol{z}(\boldsymbol{u})$ has its n components equal to zero except the k-th one equal to $1/\varphi(\Phi^{-1}(u_k))$, and $\mathbf{R}_0$ is the positive definite matrix of parameters.

Note that if the probabilistic model assumes independence of the marginal distributions (or equivalently if $\mathbf{R}_0 = \mathbf{I}$ in the previous equation), then it easily comes that the score function reduces to that of the k-th marginal distribution:

$$\kappa_k(\boldsymbol{x}, \boldsymbol{d}) = \frac{\partial \log f_{X_k}(x_k, \boldsymbol{d})}{\partial d_k}. \tag{4.72}$$

The expressions for the PDFs and the CDFs of the normal, lognormal, and Gumbel distributions that are used in this thesis are given in Table 4.1. Table 4.2 gives the expressions for the partial derivative of their CDFs with respect to their mean μ, and the corresponding marginal score functions.

Distribution	Parameters	PDF $f_X(x,\mu)$	CDF $F_X(x,\mu)$
Normal	$\mu,\ \sigma$	$\frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$	$\Phi\left(\frac{x-\mu}{\sigma}\right)$
Lognormal	$\zeta=\sqrt{\log\left(1+\frac{\sigma^2}{\mu^2}\right)}$ $\lambda=\log\mu-\frac{1}{2}\zeta^2$	$\frac{1}{\zeta}\varphi\left(\frac{\log x-\lambda}{\zeta}\right)$	$\Phi\left(\frac{\log x-\lambda}{\zeta}\right)$
Gumbel	$\alpha=\frac{\pi}{\sigma\sqrt{6}}$ $\beta=\mu-\frac{\gamma\sigma\sqrt{6}}{\pi}$	$\alpha\, e^{-\alpha(x-\beta)-e^{-\alpha(x-\beta)}}$	$e^{-e^{-\alpha(x-\beta)}}$

Table 4.1: Probability density functions (PDF) and cumulative distribution functions (CDF) for the normal, lognormal and Gumbel distributions. $\gamma\approx 0.577$ is the Euler constant.

4.3.6 Conclusion

Thanks to the score function approach, the gradient of the failure probability with respect to the parameters in the PDF of the input random variables is now available. It is worth retaining that this approach is applicable to a wide range of sampling-based estimation techniques and that it consists in a simple post-processing of a reliability analysis. Lastly, this section completes the required technical details for implementing a double-loop RBDO algorithm where the optimization problem is solved by means of the Polak-He algorithm based on accurate sampling-based reliability and reliability sensitivity analyses.

4.4 Metamodel-based RBDO

In order to reduce the computational cost induced by the use of a double-loop approach for the resolution of the RBDO problem in Eq. (4.1), it is proposed to use the Gaussian process meta-modelling technique (see Chapter 1). More specifically, it is proposed to substitute a Kriging surrogate $\{\widehat{Y}_l,\ l=1,\dots,n_p\}$ for each performance function $\{\mathfrak{g}_l,\ l=1,\dots,n_p\}$. The work introduced in this section has been published in Dubourg et al. (2011d) based on similar ideas developped in Dubourg et al. (2010a,b).

Distribution	Parameters derivatives	CDF derivatives $\frac{\partial F_X(x,\mu)}{\partial \mu}$	Score functions $\frac{\partial \log f_X(x,\mu)}{\partial \mu}$
Normal	$\frac{\partial \mu}{\partial \mu}=1,\ \frac{\partial \sigma}{\partial \mu}=0$	$-\frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$	$\frac{x-\mu}{\sigma^2}$
Lognormal	$\frac{\partial \zeta}{\partial \mu}=\frac{-\sigma^2}{\zeta\,\mu\,(\mu^2+\sigma^2)}$ $\frac{\partial \lambda}{\partial \mu}=\frac{\mu^2+2\sigma^2}{\mu\,\left(\mu^2+\sigma^2\right)}$	$-\frac{1}{\zeta}\left[\frac{\partial \lambda}{\partial \mu}+\frac{\log x-\lambda}{\zeta}\frac{\partial \zeta}{\partial \mu}\right]\varphi\left(\frac{\log x-\lambda}{\zeta}\right)$	$-\frac{1}{\zeta}\frac{\partial \zeta}{\partial \mu}+\frac{\log x-\lambda}{\zeta^2}\left[\frac{\partial \lambda}{\partial \mu}+\frac{\log x-\lambda}{\zeta}\frac{\partial \zeta}{\partial \mu}\right]$
Gumbel	$\frac{\partial \alpha}{\partial \mu}=0$ $\frac{\partial \beta}{\partial \mu}=1$	$-\alpha\, e^{-\alpha(x-\beta)-e^{-\alpha(x-\beta)}}$	$\alpha^2\left(1-e^{-\alpha(x-\beta)}\right)e^{-\alpha(x-\beta)-e^{-\alpha(x-\beta)}}$

Table 4.2: Partial derivative of the CDFs with respect to their mean μ and corresponding score functions for the normal, lognormal and Gumbel distributions.

4.4.1 The augmented reliability space

In this section, we describe the space where the Kriging surrogates are built and refined. Indeed, observing that building the Kriging surrogates from scratch for each nested reliability analysis (*e.g.* in the standard space) would be particularly inefficient, it is proposed to build and refine one unique *global* Kriging surrogate for all the nested reliability analyses (and for each performance function $\{\mathfrak{g}_l,\, l = 1, \ldots, n_p\}$).

4.4.1.1 Definition

Such a global approach can be achieved by working in a so-called *augmented reliability space*. This term has been used in different meaning in the literature. Thus a clarification is required.

Kharmanda et al. (2002) defined the augmented reliability space as the tensor product between the space of the standardized normal random variables and the design space $\mathbb{U}\times\mathbb{D}$. However, the dimension of this space $(n + n_d)$ is needlessly increased by the number of design variables n_d. It is also argued here that using this space may cause some loss of information since the performance functions are not in bijection with that augmented space. Indeed, a point $\boldsymbol{x} \in \mathbb{X}$ can usually be described by an infinity of couples $(\boldsymbol{u}, \boldsymbol{d}) \in \mathbb{U} \times \mathbb{D}$.

In contrast, in the work by Taflanidis (2007); Taflanidis and Beck (2008, 2009a,b) and in the present approach, the dimension of the augmented reliability space is kept equal to n by considering that the design vector $\boldsymbol{d}$ simply augments the uncertainty in the random vector $\boldsymbol{X}$. Indeed, the augmented random vector $\boldsymbol{V} = \boldsymbol{X}(\boldsymbol{D})$ has a PDF h which accounts for both an *instrumental uncertainty* in the design choices $\boldsymbol{D}$ and the aleatory uncertainty in the random vector $\boldsymbol{X}$. Under such considerations, the augmented PDF h reads as follows:

$$h(\boldsymbol{x}) = \int_{\mathbb{D}} f_{\boldsymbol{X}}(\boldsymbol{x}\,|\boldsymbol{d}\,)\;\pi(\boldsymbol{d})\;\mathrm{d}\boldsymbol{d}, \tag{4.73}$$

where $f_{\boldsymbol{X}}$ is the PDF of $\boldsymbol{X}$ given the parameters $\boldsymbol{d}$ and π is a pseudo PDF for $\boldsymbol{D}$ which is naturally assumed uniform over the design space $\mathbb{D}$. An illustration of the definition of this augmented PDF is provided in Figure 4.4 for the case of a normal distribution with mean value uniformly distributed over $\mathbb{D}$. The augmented reliability space is spanned by the axis $\mathbb{V}$ on the left in this simple case.

4.4.1.2 Confidence region

The DOEs from which the Kriging surrogates are built and refined should uniformly cover a sufficiently large *confidence region* of the augmented PDF h in order to make the surrogate limit-state surfaces accurate wherever they can potentially be evaluated along the optimization process. More precisely, the surrogates should be accurate for extreme design choices $\boldsymbol{d}$ (*i.e.* located onto the boundaries of the design space $\mathbb{D}$) and extreme values of the random vector $\boldsymbol{X}$ (to be able to compute reliability indices as large as *e.g.* $\beta_0 = 8$).

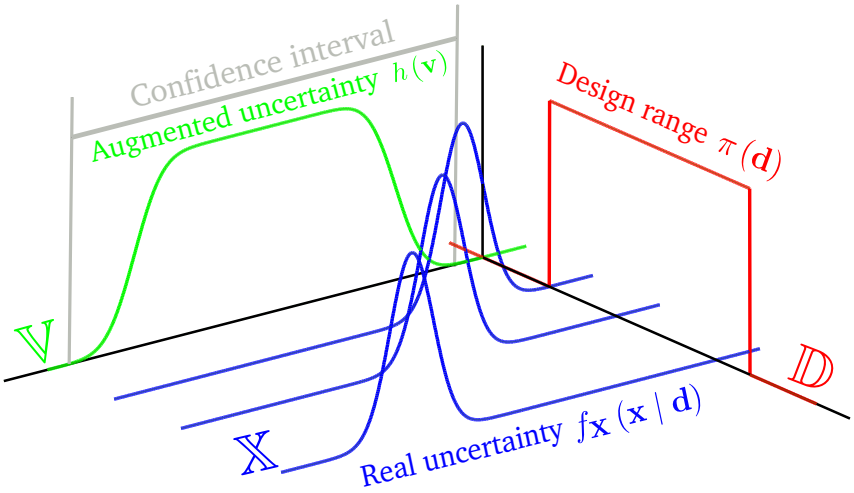


Figure 4.4: The augmented probability density function of a Gaussian random variate with uniform mean.

Such a *confidence region* is essentially the multivariate extension of the univariate concept of *confidence interval*. More specifically, a confidence region $\mathbb{X}_{f_0}$ for a probability density f is delimited by a chosen contour f_0 and it is defined as the following *slice*:

$$\mathbb{X}_{f_0} = \{\boldsymbol{x} \in \mathbb{X} : f(\boldsymbol{x}) > f_0\}. \tag{4.74}$$

However, under the previous general assumptions, it is hard to derive an analytical equation for this confidence region so that it is proposed to use an approximation instead.

Before proceeding to the presentation of this approximate confidence region, we should distinguish two types of random variables in the RBDO problem:

- the *design random variables* $\boldsymbol{X}_d$ whose distributions are defined by the design variables $\boldsymbol{d}$;
- the *basic random variables* $\boldsymbol{X}_p$ whose distributions are independent of the design variables.

Hence, the original complete random vector is $\boldsymbol{X} = (\boldsymbol{X}_d, \boldsymbol{X}_p)$ and the corresponding augmented vector is $\boldsymbol{V} = (\boldsymbol{V}_d, \boldsymbol{V}_p)$. Despite the potential dependence between these two components (as dictated by the copula), it is proposed to approximate the confidence region $\mathbb{V}_{\beta_0}$ of the complete augmented vector $\boldsymbol{V}$ as a tensor product between the marginal confidence regions $\mathbb{V}_{\beta_0,d}$ for $\boldsymbol{V}_d$ and $\mathbb{V}_{\beta_0,p}$ for $\boldsymbol{V}_p$.

Hyperrectangular approximate confidence region for the design random variables

The confidence region for the augmented design random variables $\boldsymbol{V}_d$ is approximated by an hyperrectangular region $\widehat{\mathbb{V}}_{\beta_0,d}$. Such an hyperrectangular region is simply defined as the tensor product of the confidence intervals on the augmented margins

$\{V_{d\,i},\, i = 1, \ldots, n_d\}$ bounded below by a collection of *lower quantiles* $\{q^-_{V_{d\,i}},\, i = 1, \ldots, n_d\}$ and above by a collection of *upper quantiles* $\{q^+_{V_{d\,i}},\, i = 1, \ldots, n_d\}$.

For each margin, the lower quantiles at probability level $1 - \Phi\left(\beta_0\right) = \Phi\left(-\beta_0\right)$ (resp. the upper quantiles at probability level $\Phi\left(+\beta_0\right)$) are solution of the following optimization problems:

$$q^-_{V_{d\,i}} = \min_{\boldsymbol{d}\in\mathbb{D}} F^{-1}_{X_{d\,i}}\left(\Phi(-\beta_0) \,|\, \boldsymbol{d}\right), \quad i = 1, \ldots, n_d, \tag{4.75}$$

$$q^+_{V_{d\,i}} = \max_{\boldsymbol{d}\in\mathbb{D}} F^{-1}_{X_{d\,i}}\left(\Phi(+\beta_0) \,|\, \boldsymbol{d}\right), \quad i = 1, \ldots, n_d, \tag{4.76}$$

where $\{F^{-1}_{X_{d\,i}},\, i = 1, \ldots, n_d\}$ are the quantile functions of the margins in $\boldsymbol{X}_d$. If the domain $\mathbb{D}$ is rectangular and if one is able to derive analytical expressions for these quantile functions and their derivatives with respect to the parameters, then these optimization problems might be solved analytically. However assuming a more general setup where the quantile functions can only be evaluated numerically, these problems can be solved numerically by means of a simple gradient-based algorithm. Finally, the sought hyperrectangle can be easily defined as the following tensor product:

$$\widehat{\mathbb{V}}_{\beta_0,\,d} = \prod_{i=1}^{n_d} \left[q^-_{V_{d\,i}};\, q^+_{V_{d\,i}}\right]. \tag{4.77}$$

Example 4.4.1.

Consider a Gaussian random variable $X_d \sim \mathcal{N}_1(\mu, \sigma^2)$ *whose mean is comprised in* $\left[\mu^-;\, \mu^+\right]$. *In this simple case, the lower (resp. upper) quantile is given by:*

$$q^-_{V_d} = \mu^- - \sigma\,\beta_0, \tag{4.78}$$

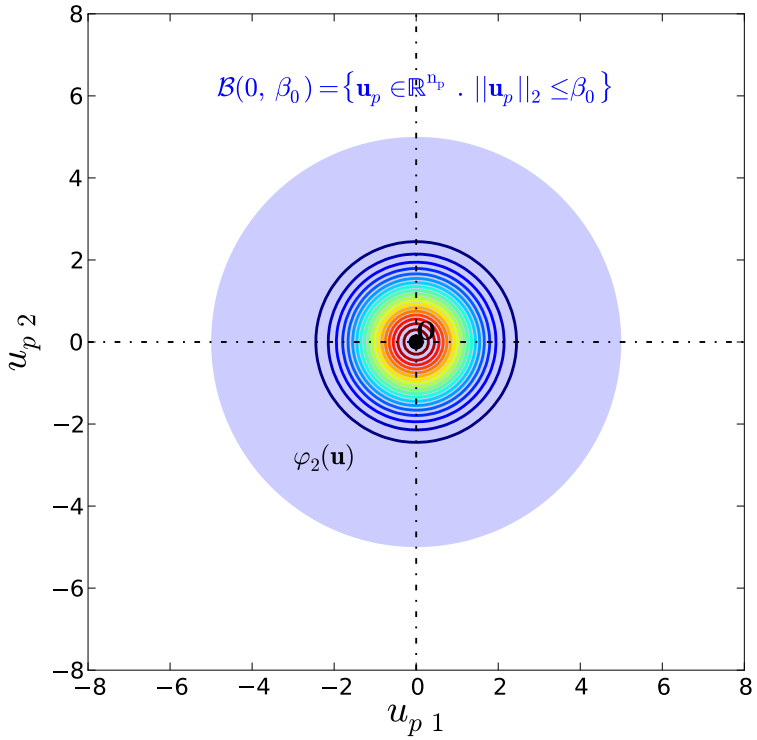
$$q^+_{V_d} = \mu^+ + \sigma\,\beta_0. \tag{4.79}$$

Actual confidence region for the basic random variables

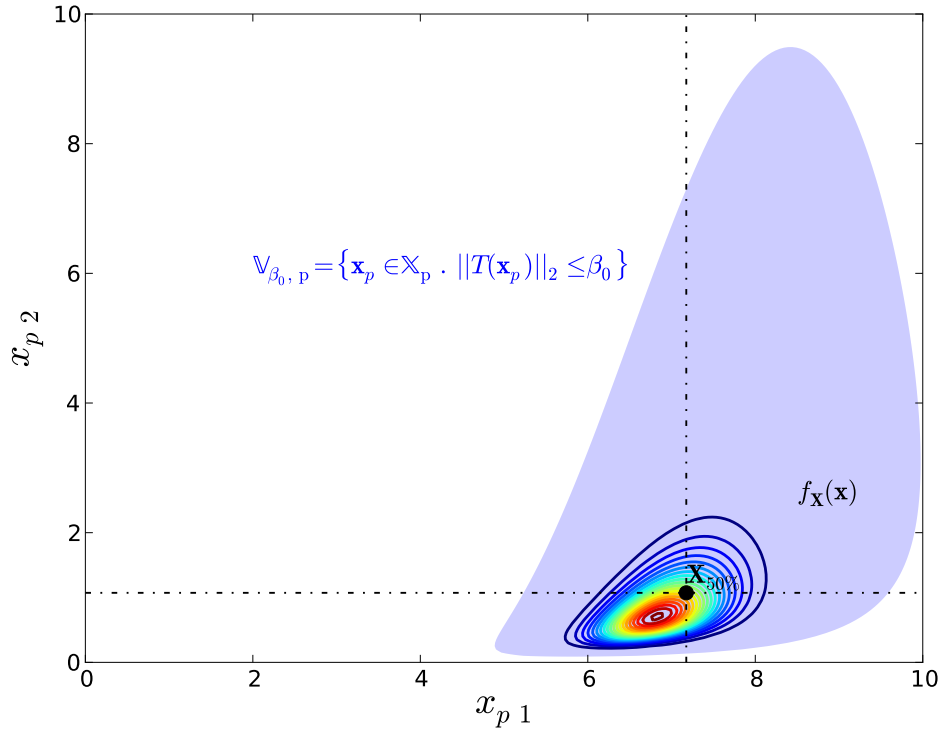
The actual confidence region for the basic random variables is easier to derive thanks to the isoprobabilistic transform T. Indeed, recall from Section 3.2.3 of Chapter 3 that the isoprobabilistic transform T transforms the original random vector $\boldsymbol{X}_p$ into a standard Gaussian random vector $\boldsymbol{U}_p$ with independent components. In this standard space, the contours of the multivariate standard Gaussian PDF φ_{n_p} are spherical. Hence, in order to define the confidence region of $\boldsymbol{X}_p$, it is proposed to first map the realizations $\boldsymbol{x}_p$ to the standard space $\boldsymbol{u}_p$ and then check for the distance of this vector to the origin. As a result, the sought confidence region for a given radius β_0 reads as follows:

$$\mathbb{V}_{\beta_0,\,p} = \left\{\boldsymbol{x}_p \in \mathbb{X}_p : \left\|T(\boldsymbol{x}_p)\right\|_2 \leq \beta_0\right\}. \tag{4.80}$$

Such a confidence region is illustrated in Figure 4.5 for the basic example defined in Section 3.2.4 of Chapter 3.



(a) Confidence region in the standard space



(b) Confidence region in the physical space

Figure 4.5: Confidence regions for the basic random vector defined in Section 3.2.4 of Chapter 3. Panel (a) illustrates the confidence region in the standard space (an hypersphere of radius $\beta_0 = 5$) whereas Panel (b) illustrates the confidence region in the original space spanned by $\boldsymbol{X}_p$. The latter is an image of the former through the inverse isoprobabilistic transform T^{-1}.

4.4.1.3 Weighting density for the refinement criterion

At last, the proposed approximate confidence region for the augmented random vector $\boldsymbol{V}$ reads as follows:

$$\widehat{\mathbb{V}}_{\beta_0} = \widehat{\mathbb{V}}_{\beta_0,d} \times \mathbb{V}_{\beta_0,p}. \tag{4.81}$$

Consequently, in order to implement the sampling-based refinement strategy exposed in Section 2.4 of Chapter 2, it is proposed to resort to the following indicator function:

$$\mathbb{1}_{\widehat{\mathbb{V}}_{\beta_0}}(\boldsymbol{v}) = \begin{cases} 1 & \text{if } \boldsymbol{v} \in \widehat{\mathbb{V}}_{\beta_0} \\ 0 & \text{otherwise} \end{cases}, \tag{4.82}$$

as a weighting pseudo-density for bounding the sampling of the refinement criterion (*e.g.* the margin probability function) to the region of interest.

4.4.2 Implementation

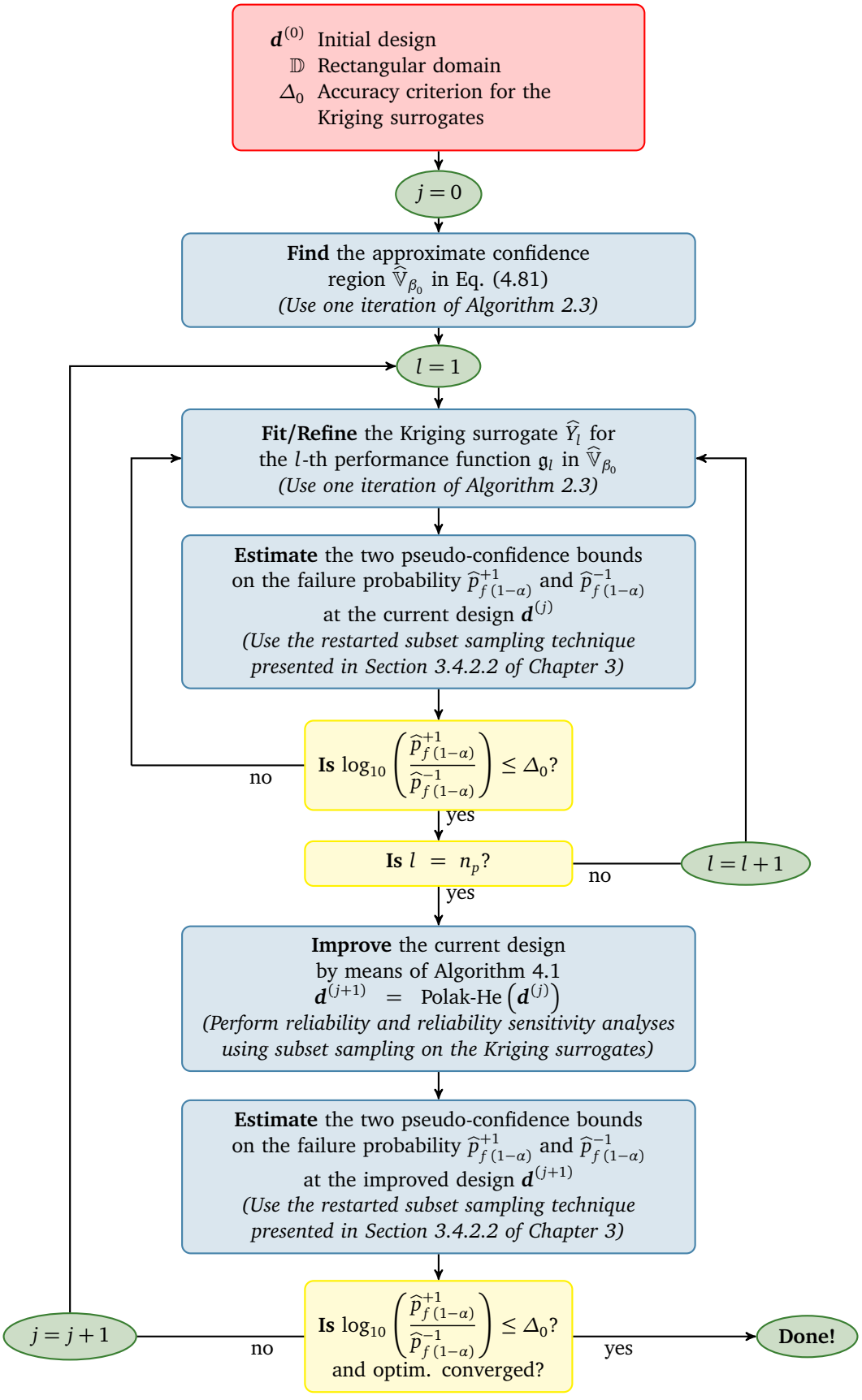
The proposed surrogate-based RBDO strategy is sketched in Algorithm 4.2. The algorithm requires an initial design $\boldsymbol{d}^{(0)}$, a rectangular admissible design space $\mathbb{D}$ and an accuracy criterion Δ_0 for the Kriging surrogates. The initial design $\boldsymbol{d}^{(0)}$ is required by the Polak-He optimization algorithm which proceeds iteratively by a sequence of improvements. The rectangular admissible space is used in order to define the approximate confidence region where the Kriging surrogates are built and refined. The accuracy criterion is required by the refinement procedure in order to decide whether a Kriging surrogate is accurate enough for reliability analysis or not.

The preliminary step of this algorithm consists in finding the bounded approximate confidence region $\widehat{\mathbb{V}}_{\beta_0}$ and the associated indicator function in Eq. (4.82). Provided this region is determined, the algorithm then iterates on j.

First, an initial Kriging surrogate $\{\widehat{Y}_l, l = 1, \ldots, n_p\}$ is built for each performance function $\{\mathfrak{g}_l, l = 1, \ldots, n_p\}$ using a space-filling DOE in $\widehat{\mathbb{V}}_{\beta_0}$. This initial DOE contains K_0 points (say $50 \leq K_0 \leq 100$ depending on the dimension of the space $\widehat{\mathbb{V}}_{\beta_0}$). If the l-th Kriging surrogate already exists (from a preliminary analysis or if $j > 0$) then it is refined by means of the sampling-based refinement strategy introduced in Section 2.4 of Chapter 2. More specifically, Algorithm 2.3 is tuned as follows:

- the refinement criterion $\mathscr{C}$ is the margin probability function;
- the weighting density function w is set here to the approximate confidence region indicator function in Eq. (4.82);
- the number of points K added at each refinement iteration is set equal to the number of available CPUs to evaluate the l-th performance function $\{\mathfrak{g}_l, l = 1, \ldots, n_p\}$.

Algorithm 4.2 Proposed surrogate-based RBDO strategy



The two extreme approximate failure probabilities $p_{f\,(1-\alpha)}^{+1}$ and $p_{f\,(1-\alpha)}^{-1}$ are computed at each refinement iteration. These two approximate probabilities are estimated on the upper (resp. lower) bound of the 95% confidence interval on the surrogate limit-state surfaces. The reader is referred back to Section 3.4.2.2 of Chapter 3 for the definition of these probabilities and for the presentation of the restarted subset sampling technique which is used here to estimate them. Based on these two estimates, the corresponding convergence metric is computed and the adaptive refinement is stopped if:

$$\log_{10}\left(\frac{\widehat{p}_{f\,(1-\alpha)}^{+1}}{\widehat{p}_{f\,(1-\alpha)}^{-1}}\right) \leq \Delta_0. \tag{4.83}$$

Once all the Kriging surrogates are fitted with a sufficient level of accuracy for the current design $\boldsymbol{d}^{(j)}$, the algorithm resorts to one iteration of the Polak-He optimization algorithm (see Algorithm 4.1) in order to improve the current design. The required reliability and reliability sensitivity analyses are performed onto the surrogates using the subset sampling technique exposed in Section 3.3.5 of Chapter 3. Reliability sensitivity analysis resorts to the score function approach as previously detailed in Section 4.3.

The accuracy of the Kriging surrogates is estimated at the new design in order to check if the new design can be accepted. If the required accuracy is reached and if the optimization has converged (see Section 4.2 for the convergence criteria), then the algorithm returns the sought optimal design $\boldsymbol{d}^* = \boldsymbol{d}^{(j+1)}$. Otherwise, the algorithm loops again starting with the adaptive refinement of the Kriging surrogates.

4.4.3 Tricks of the trade for an efficient coupling

4.4.3.1 Nested reliability and reliability sensitivity analyses

During the optimization, the Polak-He algorithm may explore extremely safe regions of the design space so that the failure probabilities may become too low to be accurately estimated by subset sampling. Pointing out that most RBDO applications will target reliability indices between 2 (*i.e.* $p_f \approx 10^{-2}$) and 6 (*i.e.* $p_f \approx 10^{-9}$), the accurate estimation of reliability indices greater than 8 (*i.e.* $p_f \approx 10^{-15}$) is utterly useless. On the other hand, it should be remarked that the subset sampling technique provides a strictly decreasing sequence of intermediate probability estimates that converges toward the actual failure probability. In other words, this means that all the intermediate failure probabilities estimated during the adaptive splitting is a consistent upper bound for the actual failure probability. Hence, it is proposed to abort the adaptive splitting algorithm when the intermediate probability estimate at some step becomes less than 10^{-15} (*i.e.* $\beta > 8$). If this lower bound is reached when the Polak-He algorithm requires the gradient of the failure probability, then the latter is replaced by that of the upper bound as well.

4.4.3.2 Refining the accuracy criteria

For the sake of efficiency and accuracy, Algorithm 4.2 is run several times with decreasing accuracy criteria for the Kriging surrogates. The algorithm is first used to perform a coarse search with $\Delta_0 = 1$. Then, the algorithm is run again from the first coarse optimal design but the accuracy criterion is divided by two in order to refine the former optimal design. The initial design space $\mathbb{D}$ specified by the user is also reduced on purpose to make this refinement more efficient (and the DOE is recycled). More specifically, the new design space is re-centred onto the first coarse optimal design and its spread is divided by two along each dimension of the design space. The new reduced design space is also required to remain in the initial user-specified design space. This refinement procedure is repeated until the optimal solution $\boldsymbol{d}^*$ stabilizes between two successive runs of the algorithm. Practice shows that the optimal solution is not much sensitive to the accuracy criterion so that the algorithm is run two or three times only.

4.4.4 Illustration

The proposed surrogate-based RBDO strategy is applied to the introductory analytical example in order to *(i)* validate its ability to converge to the reference analytical solution provided in Section 4.1.2 and *(ii)* introduce the convergence diagnostics (plots) that are extensively used in the other forthcoming examples.

4.4.4.1 Problem statement

Recall that the performance function involved in the probabilistic constraint is defined as follows:

$$\mathfrak{g}(\boldsymbol{x}) = \frac{\pi^2 E\, b\, h^3}{12\, L^2} - F_{\text{ser}}. \tag{4.84}$$

where $\boldsymbol{X} = (E,\, b,\, h)^{\mathsf{T}}$ is modelled as a random vector. The corresponding probabilistic model consists of the 3 independent random variables defined in Table 4.3. It is also recalled that the service load F_{ser} was determined so that the initial deterministic design $\mu_b^{(0)} = \mu_h^{(0)} = 200$ mm satisfies the limit-state equation $\mathfrak{g}(\boldsymbol{x}) = 0$, namely:

$$F_{\text{ser}} = \frac{\pi^2\, \mu_E\, \mu_b^{(0)}\, \mu_h^{(0)3}}{12\, L^2}. \tag{4.85}$$

The reliability-based design problem consists in finding the optimal mean values μ_b and μ_h of the random width b and height h. The optimal design is the one that minimizes the average cross-section area which is approximated as follows:

$$\mathfrak{c}(\boldsymbol{d}) = \mu_b\, \mu_h. \tag{4.86}$$

It should also satisfy the following deterministic constraint:

$$\mathfrak{f}(\boldsymbol{d}) = \mu_h - \mu_b \leq 0, \tag{4.87}$$

Variable	Distribution	Mean	C.o.V.
E (MPa)	Lognormal	$10,000$	15%
b (mm)	Lognormal	μ_b	5%
h (mm)	Lognormal	μ_h	5%
L (mm)	Deterministic	$3,000$	–

Table 4.3: Probabilistic model for the Euler buckling of a straight column.

in order to ensure that the Euler formula is applicable, as well as the following probabilistic constraint:

$$\mathbb{P}\left[\mathfrak{g}\left(\boldsymbol{X}\right) \leq 0\right] \leq \Phi(-\beta_0), \tag{4.88}$$

where $\beta_0 = 3$ is the minimum generalized reliability index (note that $\Phi(-3) \approx 10^{-3}$).

4.4.4.2 Stepwise run of the proposed algorithm

The proposed strategy is applied in order to solve the RBDO problem numerically. The refinement procedure of the Kriging limit-state surface is initialized with an initial space-filling DOE of $K_0 = 10$ points and $K = 10$ points are sequentially added to the DOE if it is not accurate enough for reliability estimation. The optimization is performed on the location parameter of the lognormal distributions λ_b and λ_h, and the scale parameters ζ are kept constant so that their initial coefficient of variation $\delta = 5\%$ also remains constant along the optimization (as for the analytical solution in Section 4.1.2).

The convergence of the algorithm is depicted in Figures 4.6 and 4.7 for two runs starting from different initial designs. The first run is initiated with the optimal deterministic design $\mu_b = \mu_h = b^* = h^* = 200$ mm whereas the second run is initiated with a slight overdesign $\mu_b = 200$ mm and $\mu_h = 300$ mm. Convergence is achieved in both cases as all the constraints (deterministic and probabilistic) are satisfied and both the cost and design variables have reached a stable value. The algorithm converges to the exact solution computed in Section 4.1.2 which is the squared section with width $\mu_b^* \approx 231$ mm. It can be seen that the probabilistic constraint is activated at the optimal solution. This is generally the case in the RBDO examples that are being dealt with in the literature and it justifies the probabilistic reformulation of the underlying deterministic design rule.

Figures 4.6 and 4.7 features four diagnostic plots that are used throughout this manuscript in order to assess convergence. The upper left panel illustrates the convergence of the design variables. The lower left one illustrates that of the objective function. The lower right panel depicts the convergence of the three generalized reliability indices:

$$\beta_{1-\alpha}^{\,i} = -\Phi^{-1}\left(p_{f\,(1-\alpha)}^{\,i}\right),\ i = -1,\ 0,\ +1, \tag{4.89}$$

where the corresponding failure probabilities are estimated on the three approximate failure subsets $\{\mathbb{F}_{1-\alpha}^{\,i},\ i = -1,\ 0,\ +1\}$ defined from the Kriging surrogates (see Section 3.4.2.2 of Chapter 3). The spread between the two extreme generalized reliability indices indicates how accurate the reliability estimates are.

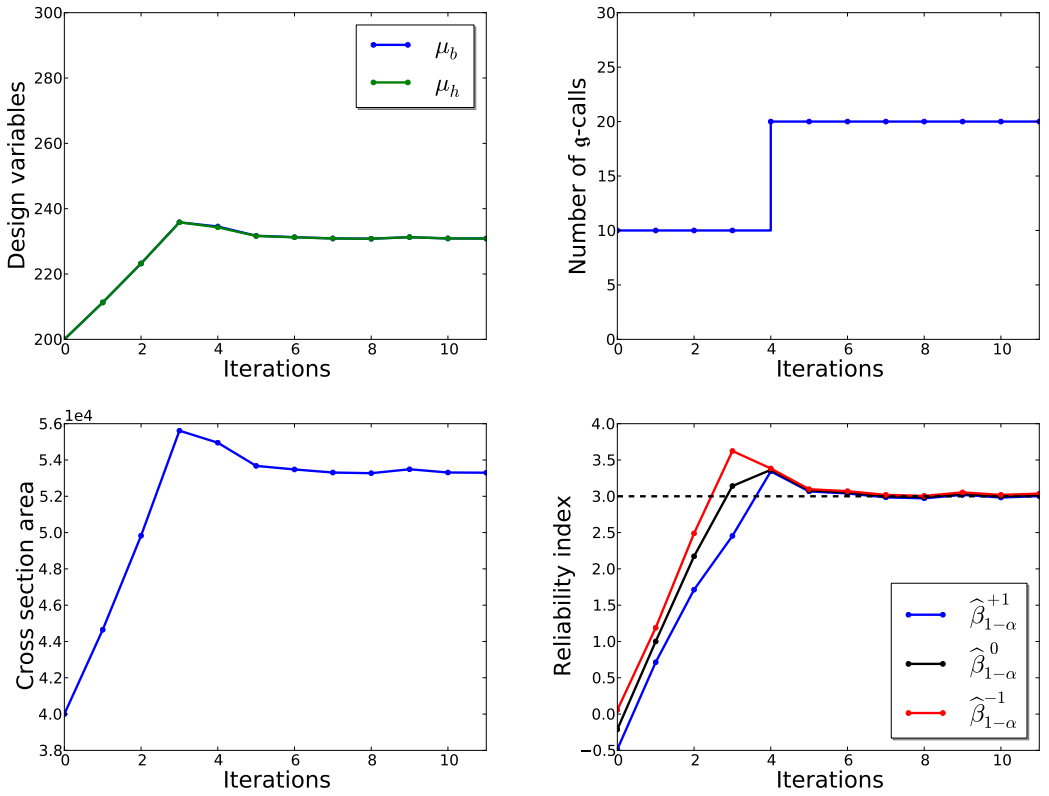


Figure 4.6: Elastic buckling of a straight column. Convergence for run #1, starting from the optimal deterministic design.

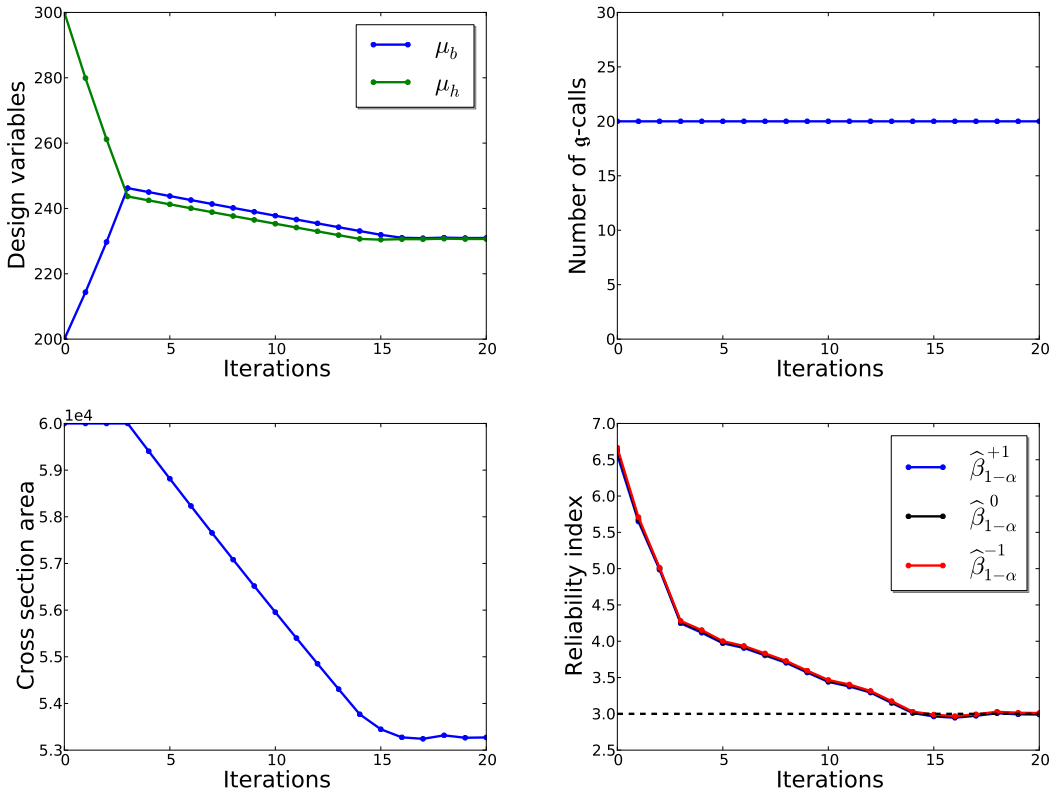


Figure 4.7: Elastic buckling of a straight column. Convergence for run #2, starting from a conservative design.

The reduction of the latter spread should be opposed to the upper right panel which illustrates the cumulative number of evaluations of the performance function that are used in order to refine the Kriging surrogates. It can be seen that for both runs, the optimal design is reached using only 20 evaluations of the performance function. Indeed, the DOE used for building the Kriging predictor is enriched only once and it is then accurate enough for all the remaining design configurations including the optimal design. In order to shed some light on the computational gain brought by the proposed surrogate-based strategy, the latter number of evaluations should be compared with that required by a nested RBDO algorithm resorting to subset sampling on the original performance function. Such a brute-force strategy indeed requires about 4×10^6 evaluations of the performance function for the same number of iterations of the optimizer and converges to the same optimal design. This brute-force strategy will be used in the sequel in order to obtain a reference solution when there does not exist any analytical reference.

4.5 Conclusion

This chapter has presented a surrogate-based approach to reliability-based design optimization. State-of-the-art techniques for resolving this "chance-constrained" optimization problem were first shortly reviewed as a motivation for the use of adaptive surrogate models in the usual nested formulation. A specific emphasis was put on the predominance of most-probable-failure-point-based approaches in the literature and their associated dangers when the MPFP uniqueness assumption does not hold. The reason for this predominance is that sampling-based approaches are clearly not affordable when the performance models are expensive to evaluate. Hence, the use of analytical surrogates appears unavoidable at some point.

Rather than using a collection of linear approximations of the limit-state surfaces in the standard spaces associated to each design (*i.e.* using nested FORM analyses), it is proposed to use one unique global surrogate for each performance function on a sufficiently large domain spanned by their input variables. Thanks to these global surrogates, the reliability analysis is *decoupled* from the optimization loop and it thus enables a significant reduction of the overall computational cost induced by the resolution of the RBDO problem.

However, the Kriging representation that is used to surrogate the performance functions does not yield analytical estimates of the failure probability as FORM does. So, it is proposed to resort to subset sampling in order to compute the failure probability and its partial derivatives with respect to the design variables. The partial derivatives are computed by means of the so-called score function approach. Thanks to this smart importance-sampling-like technique, the partial derivatives of interest can be obtained with no additional calls to the performance functions. Actually, their estimation resorts to the same sample than the one used to estimate the failure probability itself. This approach extends to most sampling-based reliability techniques such as subset sampling or meta-model-based importance sampling.

5

Academic validation examples

Contents

5.1 Meta-model-based importance sampling

5.1.1 Introduction

This section presents and comments a selected set of examples that have been investigated in order to validate the *meta-model-based importance sampling* technique proposed in Section 3.5 of Chapter 3. The purpose is to estimate the failure probability p_f of a physical system given its mathematical performance model $\mathfrak{g}$ and a probabilistic model $f_{\boldsymbol{X}}$ for its random parameters $\boldsymbol{X}$. It is recalled that the failure domain is canonically defined as:

$$\mathbb{F} = \{\boldsymbol{x} \in \mathbb{X} : \mathfrak{g}(\boldsymbol{x}) \leq 0\}, \tag{5.1}$$

and that the corresponding failure probability reads as follows:

$$p_f = \mathbb{P}[\boldsymbol{X} \in \mathbb{F}]. \tag{5.2}$$

The Kriging surrogates are initially built from an hypersphere-filling DOE of size K_0 using the procedure detailed in Section 2.2.1.2 of Chapter 2. The radius of the hypersphere is set equal to 8. They are then refined by means of the sampling-based refinement strategy introduced in Section 2.4 of Chapter 2. K cluster centres are added per refinement iteration. Both K_0 and K are given in the sequel depending on the example. The Kriging surrogates are always built using a *constant regression model* $\boldsymbol{f} = \{1\}$ and a *squared exponential covariance model* (see Eq. (1.77) in Chapter 1).

5.1.2 A simple concave failure domain

This mathematical example was originally proposed in Rackwitz (2001). It is presently used for illustrating that the proposed importance sampling estimator remains unbiased and convergent when the dimension n of the input $\boldsymbol{x}$ gets large.

5.1.2.1 Problem definition

The problem involves n independent lognormal random variates with mean value $\mu = 1$ and standard deviation $\sigma = 0.2$. The performance function reads as follows:

$$\mathfrak{g}(\boldsymbol{x}) = \left(n + a\,\sigma\,\sqrt{n}\right) - \sum_{i=1}^{n} x_i, \tag{5.3}$$

where a is set equal to 3 for the present application. The dimension of the problem is successively set equal to $n = \{2, 50, 100\}$ in order to assess the influence of the dimension on the proposed estimator. The problem is represented in Figure 5.1. It can be seen that the limit-state surface features a single MPFP and a quite significant curvature that makes the FORM approximation non-conservative (*i.e.* $p_{f\,1} < p_f$). It should also be noticed that this example was first built to show that even the second-order asymptotic approximations (SORM) are mislead by the strong curvature (see Rackwitz, 2001, for the details).

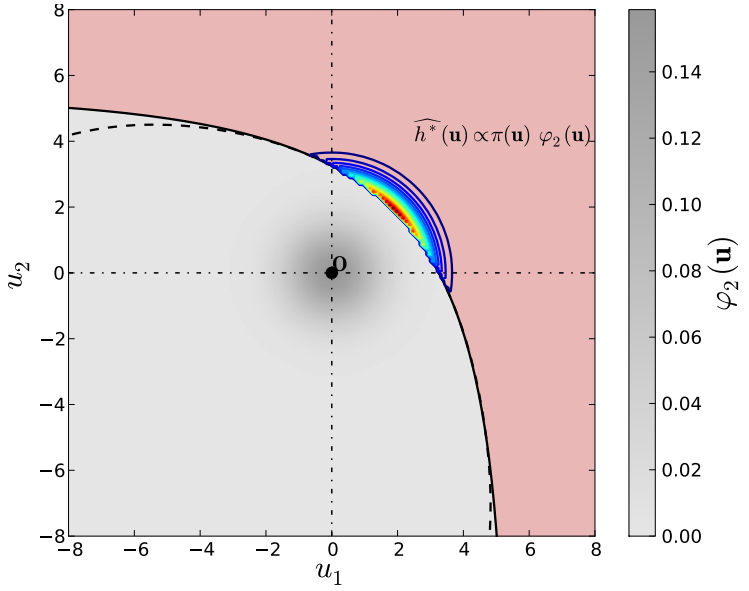


Figure 5.1: Rackwitz' limit-state surface in the standard space (black solid line). The contours represent the proposed instrumental PDF. It is optimal here for $n = 2$. Note that the surrogate limit-state surface (dashed black line) is also relatively accurate.

5.1.2.2 Results

The results are given in Table 5.1. The first block provides the reference results obtained by crude Monte Carlo sampling whereas the second block provides the results yielded by the proposed importance sampling scheme. For both methods a 2% coefficient of variation on the failure probability estimates is targeted. First, it can be observed that the proposed estimator is in good agreement with the reference Monte Carlo estimates. The estimates of both the augmented failure probability $\widehat{p}_{f\,\varepsilon}$ and the correction factor $\widehat{\alpha}_{\text{corr}}$ are also provided. Note that the correction factor decreases with the dimension. In low dimension ($n = 2$) the Kriging limit-state surface almost exactly fits its exact counterpart (see Figure 5.1) hence $\widehat{\alpha}_{\text{corr}} = 1$ and $\widehat{p}_{f\,\text{metaIS}} = \widehat{p}_{f\,\varepsilon}$ (no misclassification). In larger dimensions the surrogate loses accuracy, and the correction factor becomes more influent.

The size of the DOE from which the Kriging predictor is built is given as the product between the number of refinement iterations and the number K of cluster centres added per iteration. K is chosen equal to the number of input random variables ($K = n = \{2, 50, 100\}$) so that the cost for the definition of the instrumental PDF is equivalent to the cost induced by a gradient-based design point search algorithm (when the gradient is approximated using a *finite forward differentiation* technique). In all cases 6 iterations are required (see Table 5.1). Note that the total number of calls to the real performance function $\mathfrak{g}$ for the proposed importance sampling scheme ($m + N_{\text{corr}}$) is much less than the number of calls induced by Monte Carlo simulation (N) for the same targeted coefficient of variation.

n	2	50	100
Crude Monte Carlo sampling (ref.)			
$\widehat{p}_{f\,\mathrm{MC}}$	4.78×10^{-3}	1.91×10^{-3}	1.73×10^{-3}
δ_{MC}	$\leq 2\%$	$\leq 2\%$	$\leq 2\%$
N	522,000	1,100,000	1,450,000
Metamodel-based importance sampling			
m	6×2	6×50	6×100
$\widehat{p}_{f\,\varepsilon}$	4.85×10^{-3}	1.95×10^{-3}	1.83×10^{-3}
δ_{ε}	$\leq 1.41\%$	$\leq 1.41\%$	$\leq 1.41\%$
N_{corr}	100	1,500	2,100
$\widehat{\alpha}_{\mathrm{corr}}$	1.00	0.99	0.93
δ_{corr}	0%	$\leq 1.41\%$	$\leq 1.41\%$
$m + N_{\mathrm{corr}}$	**112**	**1,800**	**2,700**
$\widehat{p}_{f\,\mathrm{metaIS}}$	4.85×10^{-3}	1.93×10^{-3}	1.70×10^{-3}
δ_{metaIS}	$\leq 1.41\%$	$\leq 2\%$	$\leq 2\%$

Table 5.1: Results for the reliability example from Rackwitz (2001).

5.1.3 A two-dimensional four-branch serial system

This mathematical example was originally proposed by Waarts (2000, AE12). It is used here for demonstrating the ability of the proposed instrumental PDF to fit to multiple MPFPs. In addition, thanks to the low computational complexity of the performance function, the sampling-based reliability analyses are repeated 30 times on purpose to check the required properties of the estimators, their unbiasedness and their respective variance of estimation.

5.1.3.1 Problem definition

The performance function consists in a set of 4 functions whose minimum defines the failure domain $\mathbb{F}$. It reads as follows:

$$\mathfrak{g}(\boldsymbol{x}) = \min \begin{pmatrix} 3 + (x_1 - x_2)^2/10 - (x_1 + x_2)/\sqrt{2} \\ 3 + (x_1 - x_2)^2/10 + (x_1 + x_2)/\sqrt{2} \\ x_1 - x_2 + 7/\sqrt{2} \\ x_2 - x_1 + 7/\sqrt{2} \end{pmatrix}, \quad \boldsymbol{x} \in \mathbb{X}. \tag{5.4}$$

The random vector $\boldsymbol{X}$ is composed with two independent standard normal random variates. The problem is illustrated in Figure 5.2, where the failure domain corresponds to the outer region of the star-shaped limit-state (shaded in red).

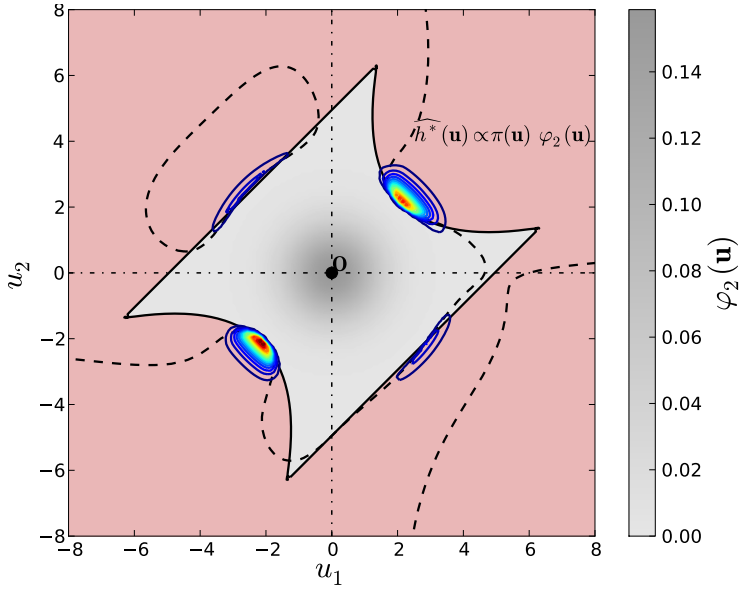


Figure 5.2: Waarts' limit-state surface in the standard space (black solid line). The contours represent the proposed instrumental PDF. It can be seen that despite the surrogate limit-state surface (dashed black line) is not fully accurate, the corresponding instrumental PDF is close to be optimal and covers the four MPFPs.

5.1.3.2 Results

The sampling-based adaptive refinement procedure is tuned as follows. The initial DOE fills the hypersphere of radius $\beta_0 = 8$ with $K_0 = 8$ points and $K = 4$ points are added at each refinement iteration. The procedure is stopped according to the leave-one-out estimate of the correction factor as detailed in Section 3.5.4 of Chapter 3. The adaptive importance sampling scheme targets a 5% coefficient of variation, and the sampling resorts to 100 Markov chains incremented in parallel.

For the sake of validation, the results yielded by the proposed approach are compared with the ones obtained by means of crude Monte Carlo sampling and subset sampling. Crude Monte Carlo sampling is ran by batches (of size $1,000$) until the coefficient of variation is less than 5%. Subset sampling uses $100,000$ Markov chain increments per step.

	Monte Carlo	**Subset Sampling**	**Meta-IS**[a]
N	172,000	284,195	**40 + 200**
$\widehat{p}_f$	2.26×10^{-3}	2.28×10^{-3}	2.38×10^{-3}
C.o.V.	<5%	<3%	<5%
Averaged $\widehat{p}_f$ (×30)	2.24×10^{-3}	2.23×10^{-3}	2.25×10^{-3}
Empirical C.o.V.	<5%	<3%	<5%

[a] $N = m + N_{\text{corr}}$.

Table 5.2: Results for the reliability example from Waarts (2000).

The results are given in Table 5.2. The first block of rows gives the number of function evaluations, the value of the failure probability and that of the coefficient of variation estimated by means of a single run of each sampling algorithm. The second block validates the respective estimators of the failure probability and of the variance of estimation by empirically averaging the results over a set of 30 independent runs of each algorithm. More specifically, the probability values (resp. that of their coefficients of variation) are the means (resp. the coefficients of variation) of 30 independent failure probability estimates. It can be seen that the empirical estimates are in reasonable agreement with the results yielded by a single run.

To conclude, this example confirmed two aspects of the proposed strategy. First, it demonstrates its ability to deal with multiple most probable failure points (note that there are four of them here). Second, it empirically validates the properties of the associated estimator enunciated in Section 3.5 of Chapter 3. It is *unbiased* and the formula for computing its coefficient of variation is in agreement with the empirical estimator.

5.1.4 A two-degree-of-freedom damped oscillator

This structural reliability example was first proposed in the report by Der Kiureghian and De Stefano (1990). It was then used for benchmark purposes in the recent article by Bourinet et al. (2011). It is used here for demonstrating the robustness of the proposed approach for estimating low failure probabilities and their gradients.

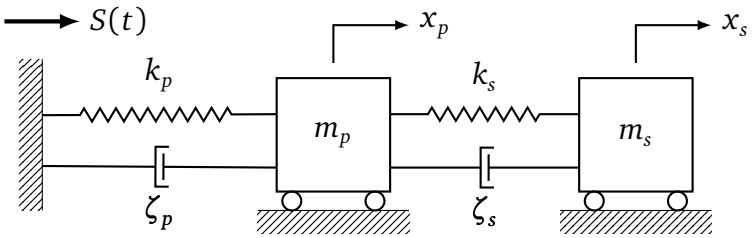


Figure 5.3: A two-d.o.f. damped oscillator under a white-noise base acceleration.

5.1.4.1 Problem definition

Let us consider the two-degree-of-freedom primary-secondary system sketched in Figure 5.3. This system is characterized by the masses m_p and m_s, spring stiffnesses k_p and k_s, natural frequencies $\omega_p = (k_p/m_p)^{1/2}$ and $\omega_s = (k_s/m_s)^{1/2}$, and damping ratios ζ_p and ζ_s, where the subscript p and s respectively refer to the primary and secondary oscillator. It has been shown by Igusa and Der Kiureghian (1985) that the mean-square relative displacement of the secondary spring under a white noise base acceleration S is given by:

$$\mathbb{E}_S\left[x_s^2\right] = \pi \frac{S_0}{4\,\zeta_s\,\omega_s^3} \frac{\zeta_a\,\zeta_s}{\zeta_p\,\zeta_s\,(4\,\zeta_a^2+\theta^2)+\gamma\,\zeta_a^2} \frac{(\zeta_p\,\omega_p^3+\zeta_s\,\omega_s^3)\,\omega_p}{4\,\zeta_a\,\omega_a^4}, \tag{5.5}$$

where S_0 is the intensity of the white noise, $\gamma = m_s/m_p$, $\omega_a = (\omega_p + \omega_s)/2$, $\zeta_a = (\zeta_p + \zeta_s)/2$ and $\theta = (\omega_p - \omega_s)/\omega_a$ is the tuning parameter.

Let F_s denote the force capacity of the secondary spring. The reliability of the system can be reduced to that of the secondary spring in the form of the following performance function:

$$\mathfrak{g}(\boldsymbol{x}) = F_s - k_s \max_{t\in[0;\,\tau]} |x_s(t)|, \tag{5.6}$$

where τ denotes the duration of the excitation. Der Kiureghian and De Stefano (1990) then remarked that the uncertainty in the peak response is usually small so that the maximum can be replaced by its mean amplitude. Hence, the performance function recasts as follows:

$$\mathfrak{g}(\boldsymbol{x}) = F_s - p\, k_s \sqrt{\mathbb{E}_S\left[x_s^2\right]}, \tag{5.7}$$

where p denotes the so-called peak factor set here equal to 3 as in the original report.

The probabilistic model for the random vector $\boldsymbol{X}$ is given in Table 5.3. All random variables are independent. The mean of the secondary spring force capacity μ_{F_s} is successively set equal to 15, 21.5 and 27.5 in order to decrease the failure probability.

Variable	Distribution	Mean	C.o.V.
m_p	Lognormal	1.5	10%
m_s	Lognormal	0.01	10%
k_p	Lognormal	1	20%
k_s	Lognormal	0.01	20%
ζ_p	Lognormal	0.05	40%
ζ_s	Lognormal	0.02	50%
F_S	Lognormal	{15, 21.5, 27.5}	10%
S_0	Lognormal	100	10%

Table 5.3: Probabilistic model for the two-d.o.f. damped oscillator from Der Kiureghian and De Stefano (1990).

5.1.4.2 Results

The adaptive refinement of the Kriging surrogate is initialized with $K_0 = 32$ points and $K = 16$ new points are sequentially added until the leave-one-out estimate of the correction factor reaches a stable value between 0.1 and 10. The adaptive importance sampling scheme uses 100 Markov chains incremented in parallel and a 5% coefficient of variation is targeted.

Table 5.4 presents the results obtained on this example for the three values of μ_{F_s}. For the sake of validation, the results yielded by the proposed meta-model-based importance sampling technique are compared with those obtained by subset sampling (using a sample size per step of 10^5) which is considered here as reference. The FORM results were obtained

by Der Kiureghian and De Stefano (1990) and confirmed by Bourinet et al. (2011) using the iHLRF algorithm (Zhang and Der Kiureghian, 1995) with a small and constant step size of 0.025. It can be seen that the bias yielded by the FORM approximation is rather significant.

The results yielded by the proposed strategy are finally compared to those obtained by Bourinet et al.'s surrogate-based approach. To make it short here, this approach consists in using a subset sampling scheme where each subset is (adaptively) approximated by a support vector classifier in order to limit the total number of calls to the performance function $\mathfrak{g}$. The reader is referred to the original article or the Ph. D. thesis by Deheeger (2008) for more details. From the computational cost point of view, it can be seen that:

(i) the total number of calls to the original performance function for the proposed strategy is comparable with that required by their substitution approach,

(ii) both approaches yield unbiased estimates of the quantity of interest.

μ_{F_s} values		FORM[a]	Subset Sampling[b]	Meta-IS[c]	SVM + Subset[a]
15	N	1,179	300,000	**464 + 200**	1,719
	p_f	2.19×10^{-2}	4.63×10^{-3}	4.80×10^{-3}	4.78×10^{-3}
	C.o.V.	–	<3%	<5%	<4%
21.5	N	2,520	500,000	**336 + 400**	2,865
	p_f	3.50×10^{-4}	4.75×10^{-5}	4.46×10^{-5}	4.42×10^{-5}
	C.o.V.	–	<4%	<5%	<7%
27.5	N	2,727	700,000	**480 + 200**	4,011
	p_f	3.91×10^{-6}	3.47×10^{-7}	3.76×10^{-7}	3.66×10^{-7}
	C.o.V.	–	<5%	<5%	<10%

[a] As computed by Bourinet et al. (2011).
[b] Reference results.
[c] $N = m + N_{\text{corr}}$.

Table 5.4: Results for the two-d.o.f. damped oscillator from Der Kiureghian and De Stefano (1990).

The topology of the failure domain and the instrumental PDF for the case where $\mu_{F_s} = 27.5$ are illustrated in Figures 5.4 and 5.5. The figures are composed with the 4 most significant cuts centred at the MPFP in the standard space. It can be seen that the failure region has the shape of a needle pointing towards the centre of the standard space. This shape confirms that the SORM results provided in the article by Bourinet et al. (2011) are reasonably accurate. As a final remark, it should be noticed that the proposed instrumental PDF is close to be optimal so that it justifies the variance reduction yielded by the associated importance sampling scheme. Actually for the three values of $\mu_{F_s} = \{15, 21.5, 27.5\}$, the correction factor estimates are respectively equal to $\widehat{\alpha}_{\text{corr}} = \{0.94, 0.92, 0.93\}$.

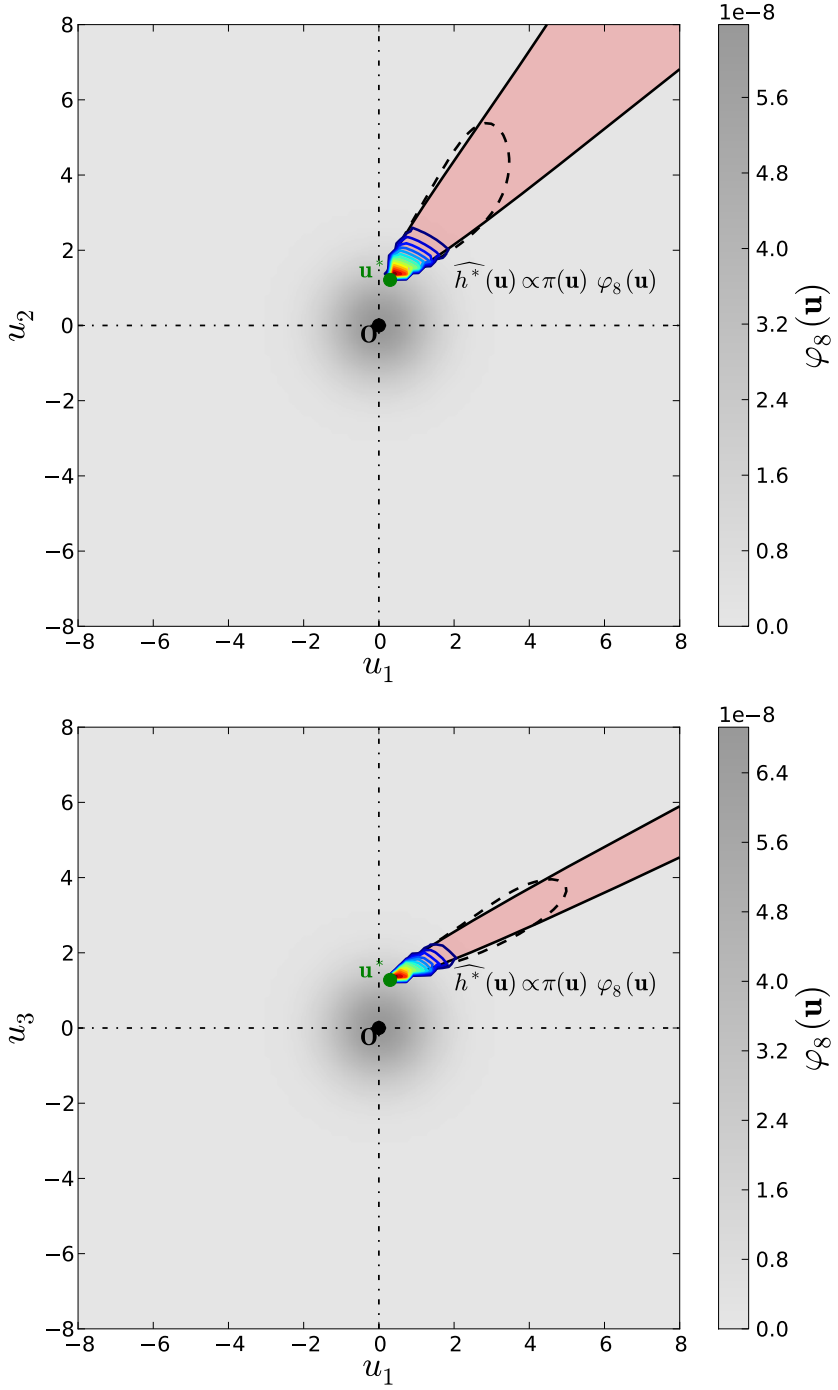


Figure 5.4: Topology of the failure domain in the standard space using cuts centred at the MPFP (green dot). The original (resp. surrogate) limit-state is represented as the black solid (resp. dashed) line. The contours represent the proposed instrumental PDF which is again close to be optimal.

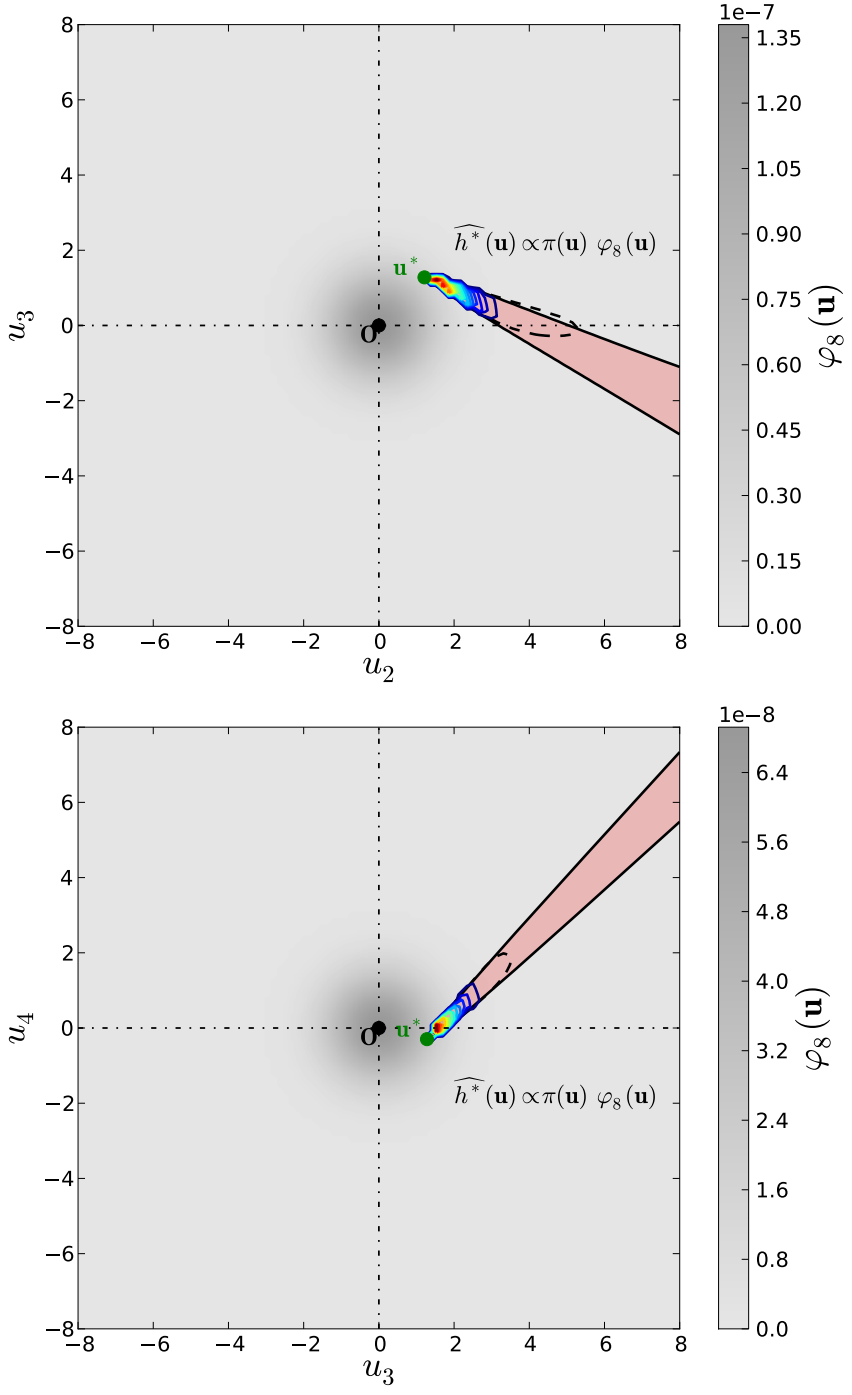
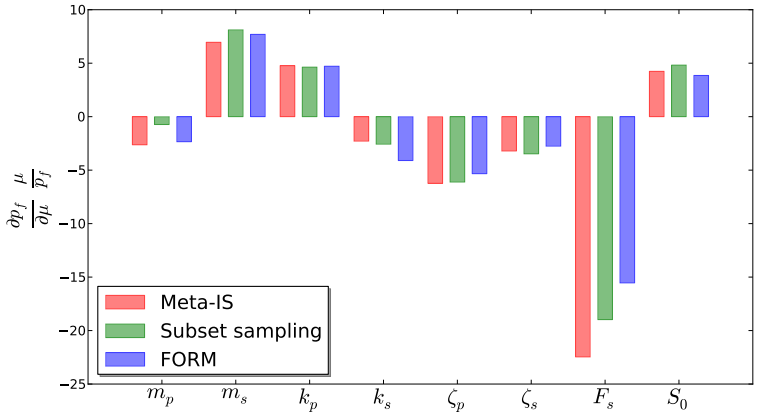
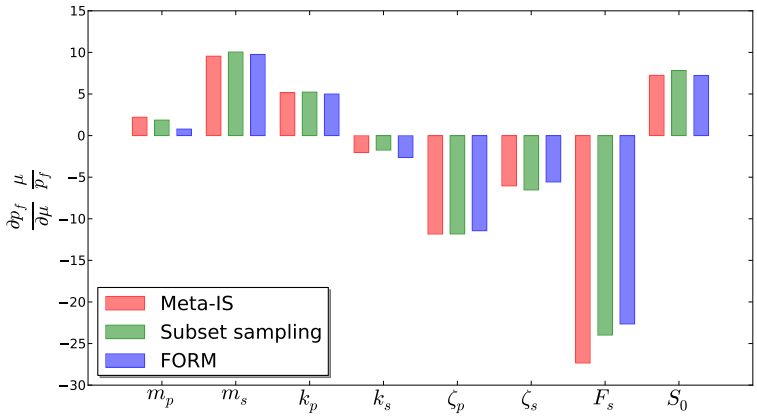


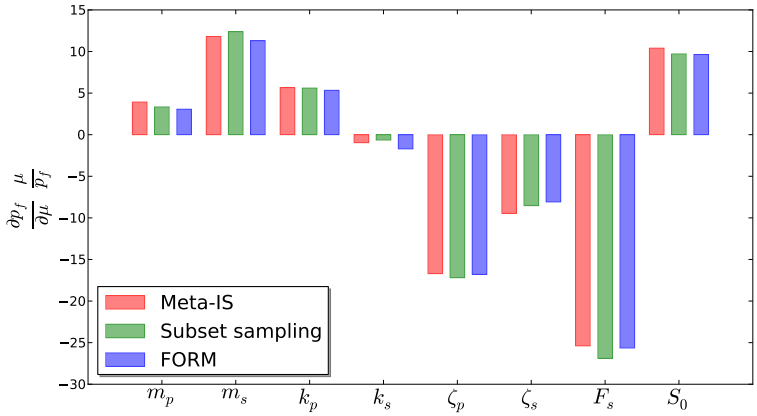
Figure 5.5: *(continued from Figure 5.4)* Topology of the failure domain in the standard space using cuts centred at the MPFP (green dot). The original (resp. surrogate) limit-state is represented as the black solid (resp. dashed) line. The contours represent the proposed instrumental PDF which is again close to be optimal.



(a) Mean force capacity: $\mu_{F_s} = 15$



(b) Mean force capacity: $\mu_{F_s} = 21.5$



(c) Mean force capacity: $\mu_{F_s} = 27.5$

Figure 5.6: Elasticity of the failure probability with respect to the mean of each variable $e = \frac{\partial p_f}{\partial \mu} \frac{\mu}{p_f}$.

At last, Figure 5.6 illustrates the results of the reliability sensitivity analysis carried out on this example for each value of μ_{F_s}. The bars represent the elasticities of the failure probability with respect to the mean of each variable. Such "elasticities" are defined in the book by Lemaire (2009) as the partial derivatives of the failure probability with respect to the mean of each variable multiplied by the means and divided by the failure probability estimate (of each reliablity method):

$$e = \frac{\partial p_f}{\partial \mu} \frac{\mu}{p_f}. \tag{5.8}$$

Thanks to the latter normalization, it clearly appears that the FORM approximations of the gradient of the failure probability (see Bjerager and Krenk, 1989, for the expression) are consistent with those obtained by the sampling-based approaches even if FORM is inaccurate here. It should be noticed that failure is most sensitive to the mean force capacity of the secondary spring μ_{F_s}. Hence, it justifies the decrease of the failure probability in Table 5.4 when μ_{F_s} augments. These latest results validate the score function approach for reliability sensitivity analysis (see Section 4.3 of Chapter 4).

5.1.5 An 8-hole plate under tension

This structural reliability example was first proposed in the Ph. D. thesis of Deheeger (2008). It was also used in Dubourg et al. (2011c) as an application example. It is a preliminary step for studying the applicability of the proposed meta-model-based importance sampling technique to structural reliability examples involving spatially varying material properties.

5.1.5.1 Problem definition

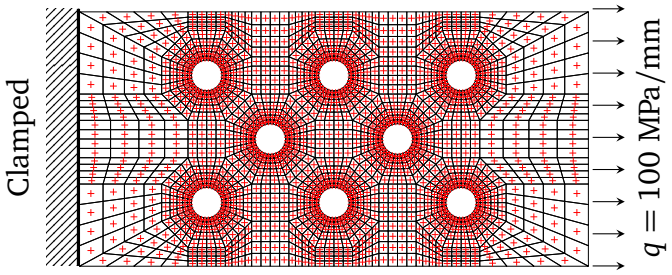
Mechanical model

The structure under consideration is depicted in Figure 5.7(a). It is a rectangular plate of size 200×100 mm with 8 10-mm-diameter-holes. The left end of the plate is clamped (horizontally and vertically) and a distributed line load with magnitude $q = 100$ MPa/mm is applied onto its right end. Plain stress, linear elasticity and a nonzero Poisson's ratio $\nu = 0.3$ are assumed. The probabilistic model for the spatially varying Young's modulus E is described in the sequel. The Von Mises equivalent stress field on the plate $\mathfrak{P}$ is computed with Code_Aster v10[1] for each realization of the Young's modulus. The limit-state surface is defined with respect to the following performance function:

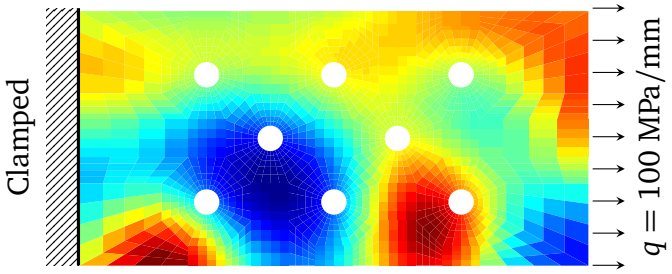
$$\mathfrak{g}(\boldsymbol{x}) = \sigma_y - \max_{\boldsymbol{p} \in \mathfrak{P}} \left(\sigma_{\text{eq, Von Mises}}(\boldsymbol{p}) \right), \tag{5.9}$$

where $\sigma_y = 450$ MPa is the yield strength of the plate's constitutive material and $\boldsymbol{p}$ denotes the position in the plate $\mathfrak{P}$.

[1] EDF finite element code, see http://www.codeaster.org.



(a) Mesh, boundary conditions and loads.



(b) One realization of the Young's modulus random field.

Figure 5.7: A clamped 8-hole plate under tension assuming plain stress linear elasticity, a nonzero Poisson's ratio and a randomly and spatially varying Young's modulus.

Probabilistic model

The Young's modulus $\{E(\boldsymbol{p}), \boldsymbol{p} \in \mathfrak{P}\}$ is modelled by a lognormal random field whose mean (resp. coefficient of variation) is set equal to $\mu_E = 200,000$ MPa (resp. $\delta_E = 25\%$). The realizations of this random field are represented by a *translated* Karhunen-Loève expansion. More specifically, the representation resorts to the simulation of a Gaussian random field $\{U(\boldsymbol{p}), \boldsymbol{p} \in \mathfrak{P}\}$ with zero mean and unit variance and the realizations of the corresponding lognormal random field are obtained by means of the following so-called *translation* (see Grigoriu, 1998):

$$E(\boldsymbol{p}) = \exp\left(\lambda_E + \zeta_E U(\boldsymbol{p})\right), \quad \boldsymbol{p} \in \mathfrak{P}, \tag{5.10}$$

where ζ_E and λ_E are the parameters of the lognormal distribution respectively defined as follows:

$$\zeta_E = \sqrt{\log\left(1 + \delta_E^2\right)}, \tag{5.11}$$

$$\lambda_E = \log(\mu_E) - \frac{1}{2}\zeta_E^2. \tag{5.12}$$

The definition of the Gaussian random field is completed with that of its autocorrelation function which reads as follows:

$$R(\boldsymbol{p}, \boldsymbol{p}') = \exp\left(\frac{\left\|\boldsymbol{p} - \boldsymbol{p}'\right\|_2^2}{\ell^2}\right), \quad (\boldsymbol{p}, \boldsymbol{p}') \in \mathfrak{P} \times \mathfrak{P}, \tag{5.13}$$

where $\ell = \sqrt{20}$ mm. The standard Gaussian random field is represented by means of its truncated Karhunen-Loève expansion which reads as follows:

$$U(\boldsymbol{p}) = \sum_{k=1}^{M} \Xi_k \sqrt{\lambda_k}\, \varphi_k(\boldsymbol{p}), \quad \boldsymbol{p} \in \mathfrak{P}, \tag{5.14}$$

where $\boldsymbol{\Xi} = (\Xi_k,\, k = 1, \ldots, M)$ is a vector of independent standard normal random variates. It is recalled that the M couples $\{(\lambda_k, \varphi_k), k = 1, \ldots, M\}$ are the elements of the spectral decomposition of the correlation function corresponding to the greatest eigenvalues λ_k (see Eq. (1.69) in Chapter 1). The latter spectral decomposition is performed here numerically using the wavelet-Galerkin scheme proposed by Phoon et al. (2002). The expansion in Eq. (5.14) involves here $M = 20$ terms.

A realization of the Young's modulus random field described above is illustrated in Figure 5.7(b). The Von Mises equivalent stress field and the performance function $\mathfrak{g}$ are then computed for each realization of the random field $\{E(\boldsymbol{p}), \boldsymbol{p} \in \mathfrak{P}\}$ using Code_Aster on a cluster of PCs with 56 cores in a distributed manner (*i.e.* one core per realization). The realizations of the Young's modulus random field are evaluated at the centroid of the elements (the red "+" in Figure 5.7(a)). Each Von Mises equivalent stress field is extrapolated at the nodes of the finite element mesh illustrated in Figure 5.7(a) before its maximum value is picked in order to compute the associated limit-state function value.

5.1.5.2 Results

The Kriging surrogates are built from an initial DOE of size $K_0 = 100$ and sequentially refined by adding $K = 50$ cluster centres at each iteration. The procedure is stopped here when the DOE size reaches $m = 1,000$.

The results for this structural reliability example are given in Table 5.5. The meta-model-based importance sampling failure probability estimate is confirmed by subset sampling at a lesser computational expense though. The multiple FORM analysis resorted here to the restarted iHLRF algorithm proposed by Der Kiureghian and Dakessian (1998), but it failed here in finding all the MPFPs. As a result, the multi-FORM approximation is inaccurate.

	Multi-FORM	**Subset Sampling**	**Meta-IS**[a]
N	1,168	25,000	1,000 + 250
$\widehat{p}_f$	6.50×10^{-6}	1.70×10^{-5}	1.41×10^{-5}
C.o.V.	–	<15%	<10%

[a] $N = m + N_{\text{corr}}$.

Table 5.5: Results of the reliability analyses for the 8-hole plate under tension.

5.1.6 Conclusion

The collection of examples presented above demonstrates the applicability of *meta-model-based importance sampling* to a wide variety of structural reliability examples featuring

(i) a reasonably large number of random variates (up to 100), *(ii)* reasonably nonlinear performance functions, *(iii)* low failure probabilities, and *(iv)* multiple design points.

The properties of the estimator were empirically validated. It is *unbiased* and the expression for its *variance of estimation* is confirmed empirically by simulation. Since it takes advantage of the information provided by the Kriging surrogate, it offers a significant variance reduction that makes it competitive with respect to the other reliability methods (including the surrogate-based approaches). However, it is worth remembering that the adaptive refinement step is of utmost importance for ensuring a high fidelity of the Kriging surrogate with respect to the actual performance function. *The finer the surrogate, the greater the variance reduction.*

Note that in the light of the previous selected set of examples (amongst others not exposed here), the following rule of thumb can be proposed for tuning the number of points added per refinement iteration:

$$K = \min\{2\,n;\ 50\} \quad \text{and} \quad K_0 = 2K. \tag{5.15}$$

Eventually, the score function approach was successfully applied to the computation of the gradient of the failure probability with respect to the means of the random variables in the probabilistic model. In addition to the fact that it is computationally efficient (it does not require any additional evaluation of the performance function) it was shown to yield meaningful unbiased estimates of the sensitivities of interest.

5.2 Meta-model-based RBDO

5.2.1 Introduction

The purpose of this section is to validate the *meta-model-based RBDO strategy* exposed in Section 4.4 of Chapter 4. All the problems that are being dealt with in the sequel consists in finding the following probabilistically constrained optimizer:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}\in\mathbb{D}} \mathfrak{c}(\boldsymbol{d}) \ : \quad \begin{cases} \mathfrak{f}_i(\boldsymbol{d}) \leq 0,\ i = 1, \ldots, n_c \\ \mathbb{P}\left[\mathfrak{g}_l(\boldsymbol{X}(\boldsymbol{d})) \leq 0\right] \leq p^0_{f\,l},\ l = 1, \ldots, n_p \end{cases}, \tag{5.16}$$

where $\boldsymbol{d}$ are mean values of some components of the random vector $\boldsymbol{X}$.

For the sake of benchmarking, the results are compared with the ones obtained in the original papers from which the examples are inspired. However, the reliability indices obtained at the optimum that are provided in the original papers are systematically checked using subset sampling. The validation of the proposed approach is performed with respect to a so-called *brute force technique* that consists in nesting a subset sampling scheme (using $N = 10^4$ samples per step) in the optimization loop without substituting the performance functions. For obvious reasons, the latter validation can only be performed when the performance functions are inexpensive to evaluate.

The Kriging surrogates assume a *constant regression model* $\boldsymbol{f} = \{1\}$ and a *squared exponential covariance model* (see Eq. (1.77) in Chapter 1). The K_0 points in the initial DOE are obtained by means of the sampling-clustering approach so that they uniformly span the approximated confidence region of the augmented reliability space. The sampling-based adaptive refinement strategy is then applied using the margin probability criterion bounded to the approximated confidence region as introduced in Section 4.4.1 of Chapter 4. Both K_0 and K are given in the sequel depending on the example. The surrogate-based reliability analyses are performed by means of the *restarted subset sampling algorithm* in order to compute the three probabilities $\{p^{\,i}_{f\,(1-\alpha)}, i = -1,\ 0, +1\}$ and the corresponding generalized reliability indices. This makes use of 10^5 Markov chain increments per subset step.

5.2.2 A highly nonlinear limit-state surface

This two-dimensional mathematical example was first proposed in the article by Lee and Jung (2008). It is used here for illustration purposes and it can be reused for validating implementations of the proposed strategy.

5.2.2.1 Problem definition

In essence this problem is the probabilistic counterpart of the deterministic design optimization problem already exposed in Section 4.2 of Chapter 4. The two performance functions read as follows:

$$\mathfrak{g}_1(\boldsymbol{x}) = -x_1 \sin(4\,x_1) - 1.1\,x_2 \sin(2\,x_2), \tag{5.17}$$

$$\mathfrak{g}_2(\boldsymbol{x}) = x_1 + x_2 - 3, \tag{5.18}$$

where $\boldsymbol{x}$ is the realization of the random vector $\boldsymbol{X}$ composed with two independent normal variables with the same standard deviation $\sigma = 0.1$. The optimization is performed with respect to the mean of the latter random variables in order to minimize the following quadratic objective function:

$$\mathfrak{c}(\boldsymbol{d}) = (\mu_1 - 3.7)^2 + (\mu_2 - 4)^2. \tag{5.19}$$

The search is restrained to the following hyperrectangular design space:

$$\mathbb{D} = [0;\ 3.7] \times [0;\ 4]\,. \tag{5.20}$$

The optimum must satisfy the following two probabilistic constraints:

$$\mathbb{P}\left[\mathfrak{g}_l\left(\boldsymbol{X}(\boldsymbol{d})\right) \leq 0\right] \leq \Phi(-\beta), \quad l = 1,\, 2, \tag{5.21}$$

where the generalized reliability index β is set here equal to 2 for both constraints.

5.2.2.2 Results

All results are presented in Table 5.6. The first line gives the results obtained when the problem is solved in a deterministic fashion. Note that the generalized reliability index of the first active constraint is close to zero, meaning that the corresponding failure probability is close to $\Phi(0) = 50\%$. The second constraint is far from being activated as illustrated in Figure 5.8. The computation of the true reliability index β_2 was actually aborted because it is greater than 8 ($p_{f\,2} < \Phi(-8) \approx 10^{-15}$).

Method	μ_1	μ_2	Cost	${\beta_1}^a$	${\beta_2}^a$	$\mathfrak{g}_1/\mathfrak{g}_2$-calls	Opt. iter.
DDO	2.97	3.41	0.89	-0.10	$> 8.00^b$	70/70	10
RBDO ($\beta_l \geq 2, \quad l = 1, 2$)							
Brute force	2.84	3.23	1.33	2.00	$> 8.00^b$	$\approx 10^7/10^7$	10
PMA[c]	2.82	3.30	1.26	1.67	$> 8.00^b$	296	7
PMA w/ kriging[c]	2.82	3.30	1.26	1.67	$> 8.00^b$	90	7
Meta-RBDO	2.81	3.25	1.35	2.00	$> 8.00^b$	80/10	10

[a] Computed using subset sampling on the original performance function $\mathfrak{g}$.
[b] Subset sampling is aborted if the failure probability is less than $\Phi(-8)$.
[c] Optimal solution computed by Lee and Jung (2008).

Table 5.6: Comparative results for the example by Lee and Jung (2008).

The other lines compare the RBDO results obtained with different approaches. The first line corresponds to the reference solution computed with the brute force technique (a nested formulation resorting to subset sampling on the original performance functions). The lines involving "PMA" refer to the results of Lee and Jung (2008) except that the reliability indices were checked using subset sampling at the optimal solution provided in the original paper. It can be seen that the PMA (FORM-based) approximation overestimated the actual reliability index. This is due to the concave shape of the failure domain in the vicinity of the optimal solution that makes the FORM approach non-conservative. The "PMA with Kriging" approach of Lee and Jung (2008) is mistaken as well as it makes use of a FORM approximation on the Kriging surrogates.

At last, the adaptive-Kriging-based RBDO strategy is applied using $K_0 = 10$ points for the initial design and $K = 10$ points for the subsequent refinement iterations if they are required. The surrogate of the first (strongly nonlinear) limit-state surface is refined 7 times while the other linear limit-state surface is accurately estimated from the very first DOE. It can be seen from Table 5.6 that the proposed strategy yields accurate results with respect to the reference brute force solution.

Convergence is illustrated here in the two-dimensional space spanned by the random design variables in Figure 5.8. It can be seen that the optimal probabilistically constrained design has been found. Both the surrogate-based and the brute force RBDO strategy stepped away from the limit-state surface in order to guarantee the required level of safety ($\beta = 2$).

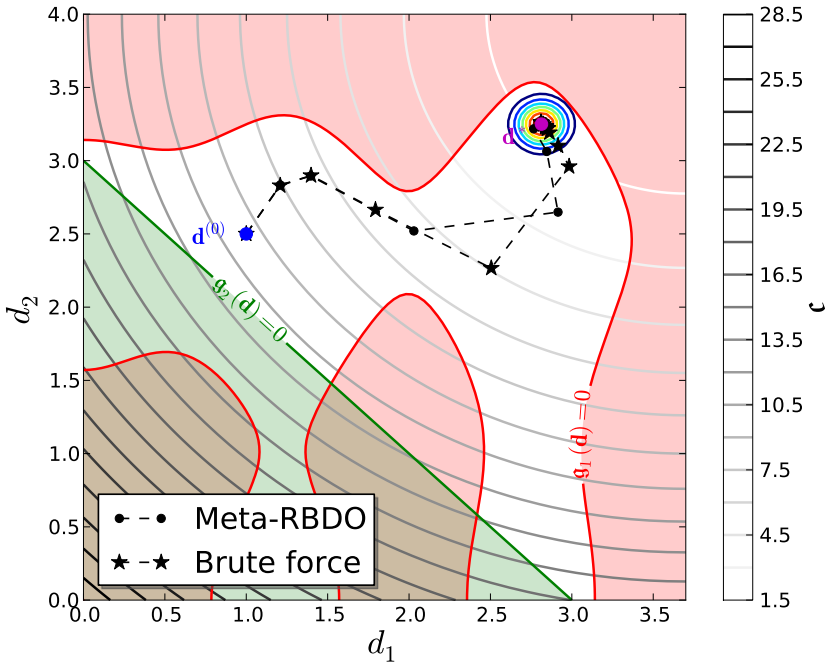


Figure 5.8: Convergence of the meta-model-based RBDO algorithm for the mathematical example by Lee and Jung (2008) in the two-dimensional random design space.

5.2.3 Three nonlinear limit-states

This mathematical example is taken from the paper by Shan and Wang (2008, Example 1). It was also dealt with by Du and Chen (2004).

5.2.3.1 Problem definition

The problem involves the three following two-dimensional analytical performance functions:

$$\mathfrak{g}_1(\boldsymbol{x}) = x_1^2 \frac{x_2}{20} - 1, \tag{5.22}$$

$$\mathfrak{g}_2(\boldsymbol{x}) = \frac{(x_1 + x_2 - 5)^2}{30} + \frac{(x_1 - x_2 - 12)^2}{120} - 1, \tag{5.23}$$

$$\mathfrak{g}_3^{\text{ini}}(\boldsymbol{x}) = \frac{80}{x_1^2 + 8\,x_2 + 5} - 1, \tag{5.24}$$

where x_1 and x_2 are the realizations of two independent Normal random variates X_1 and X_2 with mean value d_1 and d_2 and with the same constant standard deviation $\sigma_1 = \sigma_2 = \sigma = 0.3$.

Noting that the third performance function is incompatible with the probabilistic model (*i.e.* the two normal random variates in its denominator makes it undefined for $x_1^2 + 8\,x_2 +$

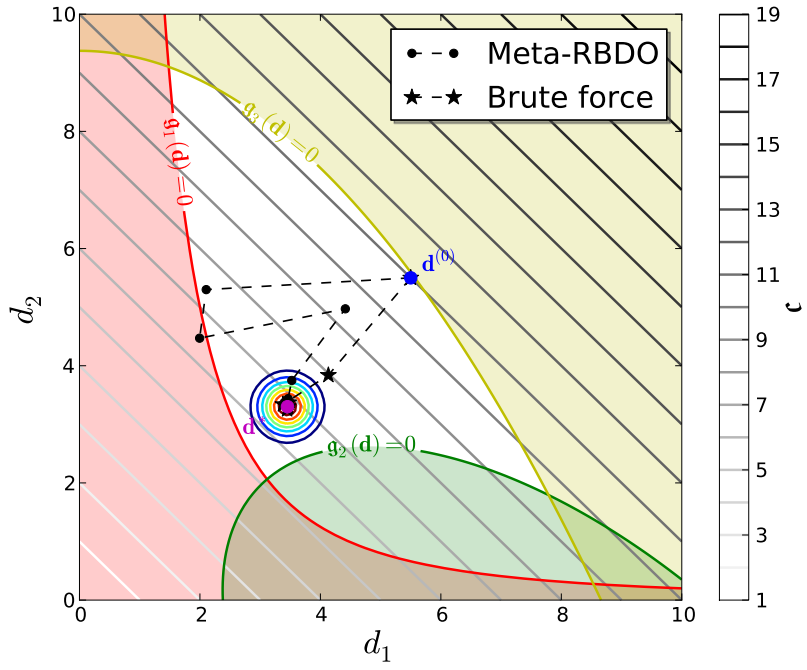


Figure 5.9: Convergence of the meta-model-based RBDO algorithm for the mathematical example by Shan and Wang (2008).

$5 = 0$), it is slightly modified so that $G_3 = \mathfrak{g}_3(\boldsymbol{X})$ is a second-order random variate. Indeed, the failure domain of interest is equivalently defined as follows:

$$\mathbb{F}_3 = \left\{ \boldsymbol{x} \in \mathbb{X} : \mathfrak{g}_3(\boldsymbol{x}) = 80 - \left(x_1^2 + 8\,x_2 + 5 \right) \leq 0 \right\}. \tag{5.25}$$

The RBDO problem consists in finding the optimal mean vector $\boldsymbol{d}^*$ which reads as follows:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d} \in \mathbb{D}} \quad d_1 + d_2 \ : \quad \mathbb{P}\left[\mathfrak{g}_l\left(\boldsymbol{X}(\boldsymbol{d}) \right) \leq 0 \right] \leq \Phi\left(-\beta_0 \right),\ l = 1, \ldots, 3, \tag{5.26}$$

where $\beta_0 = 3$ is the minimum generalized reliability index and $\mathbb{D} = [0;\ 10]^2$ is the admissible design space. The low dimension of the problem enables a graphical illustration given in Figure 5.9.

5.2.3.2 Results

Table 5.7 shows the benchmark results obtained on this mathematical example. The reference results were taken from Shan and Wang (2008). The original article does not indicate whether the number of function calls is the sum of the evaluations on the three performance functions, or the evaluations per performance function. For the proposed algorithm the number of function calls are provided separately. Deterministic design optimization (DDO) leads to an unreliable design onto the limit-state surfaces (*i.e.* $\mathfrak{g}_1(\boldsymbol{x}) = \mathfrak{g}_2(\boldsymbol{x}) = 0$) as it does not account for the uncertainty in $\boldsymbol{X}$. Note that the third constraint is not activated at the optimum as illustrated in Figure 5.9. The equivalent reliability indices obtained by subset sampling are given here to check if the first order reliability theory assumptions

hold in this case since all the RBDO strategies but the proposed one are FORM-based. All the underlying failure probabilities were accurately estimated with a coefficient of variation less than 5%. For this simple example, the curvature of the limit-state surfaces are low enough to be neglected and the more flexible Kriging surrogate seems useless with respect to the first-order approximation, but it can be seen that it is at least as efficient as the best FORM-based method.

Method	d_1	d_2	Cost	β_1	β_2	β_3	$\mathfrak{g}$-calls	Opt. iter.
DDO[a]	3.11	2.06	5.18	-0.04	0.08	> 8.00[b]	27	6
RBDO ($\beta_i \geq 3, \quad i = 1, 2, 3$)								
Brute force	3.45	3.27	6.72	2.98	3.07	> 8.00[b]	$\approx 10^6/10^6/10^6$	4
PMA[a]	3.44	3.29	6.73	2.96	3.06	> 8.00[b]	1,566	4
SORA[c]	3.44	3.29	6.73	2.97	3.06	> 8.00[b]	151	4
SLA[a]	3.44	3.29	6.73	2.96	3.05	> 8.00[b]	19	4
RDS[a]	3.44	3.28	6.72	2.96	3.02	> 8.00[b]	27	6
Meta-RBDO	3.46	3.27	6.74	3.04	2.98	> 8.00[b]	20/10/10	11

[a] As computed by Shan and Wang (2008).
[b] Subset sampling is aborted if the failure probability is less than $\Phi(-8)$.
[c] As computed by Du and Chen (2004).

Table 5.7: Comparative results for the mathematical example by Shan and Wang (2008).

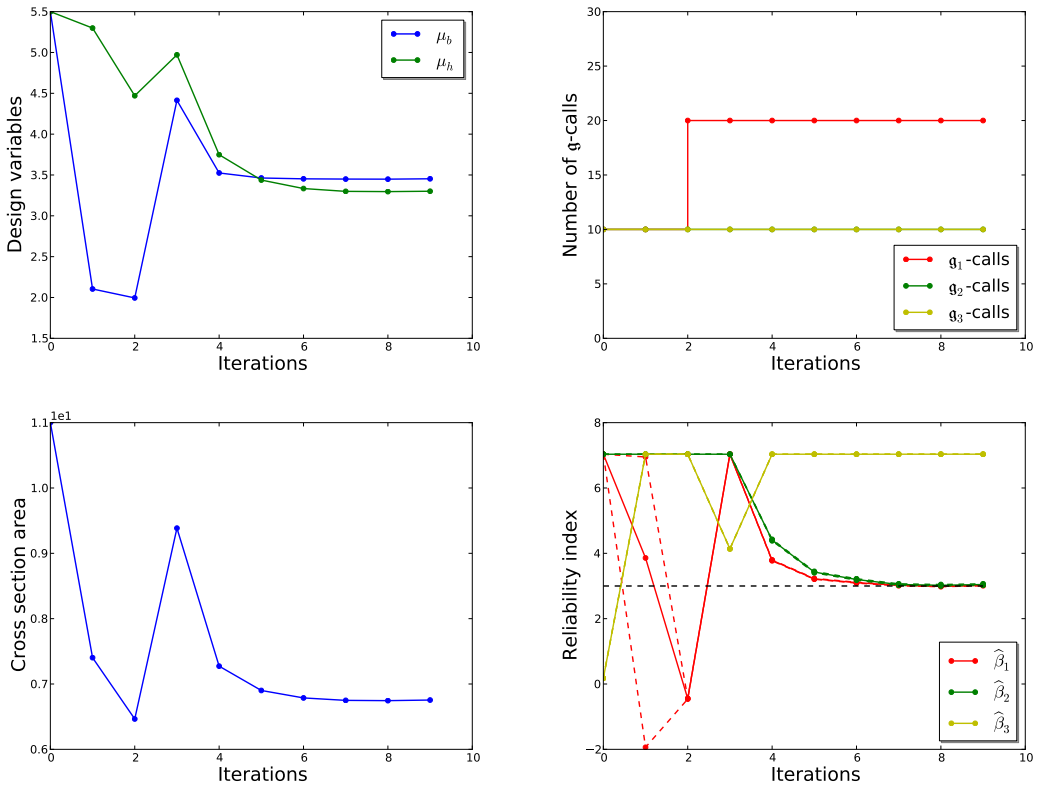


Figure 5.10: Convergence of the meta-model-based RBDO algorithm for the mathematical example by Shan and Wang (2008).

The proposed meta-model-based RBDO algorithm is initialized with $K_0 = 10$ simulation runs for each limit-state surface and the first surrogate required one update with $K = 10$ new evaluations at initialization. The surrogates are then accurate enough for all the subsequent iterations of the optimization loop.

The convergence of the proposed algorithm is depicted in Figure 5.10. In the reliability indices panel, the solid lines represent the evolution of the equivalent reliability indices estimated by subset sampling onto the kriging surrogates, and the dashed lines represent their pseudo-confidence interval estimated by means of the restarted subset sampling approach detailed in Section 3.4.2.2 of Chapter 3. It is hard to distinguish the dashed lines from the solid lines which shows how narrow these confidence intervals are in this case.

5.2.4 A short column under oblique bending

This structural reliability example is extensively used in the RBDO literature. It is inspired by the article by Royset et al. (2001).

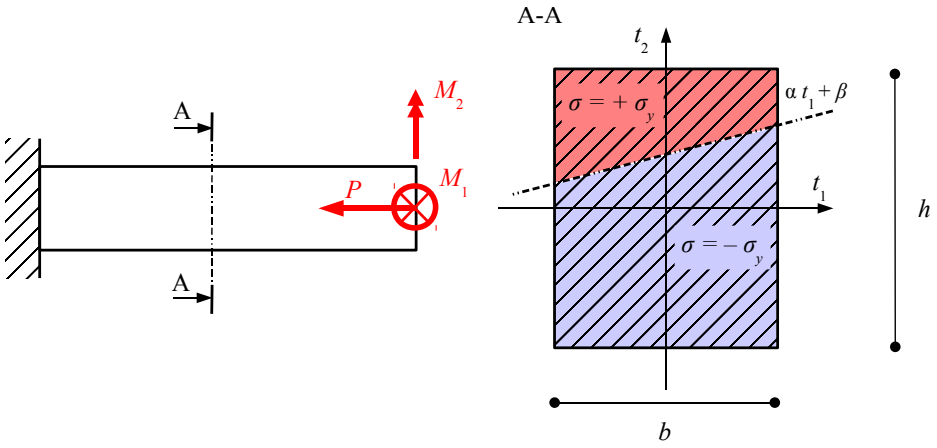


Figure 5.11: A short column under oblique bending.

5.2.4.1 Problem definition

Mechanical model

The structure under consideration is a short column with rectangular cross-section $b \times h$. It is subjected to an axial load F and two bending moments M_1 and M_2 whose axes are defined with respect to the two principal axes of inertia of the cross-section. Such a load is referred to as *oblique bending* due to the rotation of the neutral axis as illustrated in Figure 5.11. According to the original reference, the performance function describing the ultimate serviceability of the column with respect to its yield stress σ_y reads as follows:

$$\mathfrak{g}(\boldsymbol{x}) = 1 - \frac{4\,M_1}{b\,h^2\,\sigma_y} - \frac{4\,M_2}{b^2\,h\,\sigma_y} - \left(\frac{F}{b\,h\,\sigma_y}\right)^2. \tag{5.27}$$

It is obtained by assuming that the constitutive material has an elastic perfectly plastic behaviour, and it results from the superposition of the 3 ultimate forces in the cross-section under oblique bending which are defined as:

$$N_u = \int_{-h/2}^{h/2}\int_{-b/2}^{b/2} \sigma_u\left(t_1, t_2\right)\, \mathrm{d}t_1\, \mathrm{d}t_2, \tag{5.28}$$

$$M_{1u} = \int_{-h/2}^{h/2}\int_{-b/2}^{b/2} \sigma_u\left(t_1, t_2\right)\, t_1\, \mathrm{d}t_1\, \mathrm{d}t_2, \tag{5.29}$$

$$M_{2u} = \int_{-h/2}^{h/2}\int_{-b/2}^{b/2} \sigma_u\left(t_1, t_2\right)\, t_2\, \mathrm{d}t_1\, \mathrm{d}t_2, \tag{5.30}$$

where $\sigma_u\left(t_1, t_2\right)$ is the *ultimate stress distribution* under oblique bending defined as:

$$\sigma_u\left(t_1, t_2\right) = \begin{cases} +\sigma_y & \text{if} \quad t_2 > \alpha\, t_1 + \beta \\ -\sigma_y & \text{otherwise} \end{cases}, \tag{5.31}$$

and illustrated in Figure 5.11.

Probabilistic model and RBDO problem

The stochastic model originally involves three independent random variables whose distributions are given in Table 5.8. Note that in the original paper, the design variables b and h are considered as deterministic. Since the present approach only deals with design variables that define the joint PDF of the random vector $\boldsymbol{X}$, they are considered here as normally distributed with a small coefficient of variation (1%), and the optimization is performed with respect to their mean μ_b and μ_h.

The objective function is formulated as follows:

$$\begin{aligned} \mathfrak{c}(\boldsymbol{d}) &= \mathfrak{c}_0(\boldsymbol{d}) + p_f(\boldsymbol{d})\, \mathfrak{c}_f(\boldsymbol{d}) \\ &= \mathfrak{c}_0(\boldsymbol{d})\, \left(1 + 100\, p_f(\boldsymbol{d})\right) \\ &= \mu_b\, \mu_h\, \left(1 + 100\, p_f(\boldsymbol{d})\right), \end{aligned} \tag{5.32}$$

where $p_f\, c_f$ is the expected failure cost which is chosen here to be proportional to the construction cost c_0. The search for the optimal design is limited to the designs that satisfy the following geometrical constraints: $1/2 \leq \mu_b/\mu_h \leq 2$ with $100 \leq \mu_b, \mu_h \leq 1,000$, and the minimum reliability index is set equal to $\beta_0 = 3$.

5.2.4.2 Results

The results are given in Table 5.9. In this table, β_{HL} denotes the Hasofer-Lind reliability index (FORM-based), and β_{SS} denotes the generalized reliability index estimated by *subset*

Variable		Distribution	Mean	C.o.V.
M_1	(N.mm)	Lognormal	250×10^6	30%
M_2	(N.mm)	Lognormal	125×10^6	30%
P	(N)	Lognormal	2.5×10^6	20%
σ_y	(MPa)	Lognormal	40	10%
b	(mm)	Normal	μ_b	1%[a]
h	(mm)	Normal	μ_h	1%[a]

[a]Artificial uncertainty introduced for the reliability sensitivity analyses.

Table 5.8: Probabilistic model for the short column under oblique bending.

sampling (the coefficient of variation is less than 5%). The deterministic design optimization (DDO) was performed using the mean values of all the variables in Table 5.8 without considering any uncertainty and it thus yields a 50% probability of failure. Note that the optimal deterministic cost does not account for the expected failure cost in this case (*i.e.* it only accounts for the construction cost $\mathfrak{c}_0$).

The other lines of Table 5.9 shows the results of the RBDO problem. The first row gives the reference results obtained by Royset et al. (2001). The number of calls to the performance functions was not given in the original paper. It has thus been estimated given the methodology the authors used and assuming they targeted a 5% coefficient of variation on the failure probability that they appeared to estimate by crude Monte Carlo sampling. The second row provides the results from a FORM-based nested RBDO algorithm. It can be seen that the first-order reliability theory assumptions are not conservative in this case since $\beta_{\text{SS}} = 3.19 < \beta_{\text{HL}} = 3.38$. The third row gives the results obtained by the same nested RBDO algorithm, using however the subset sampling technique instead of FORM. Finally, the last row gives the results obtained when Kriging is used to surrogate the limit-state surface. The Kriging surrogate was initialized with a 50-point-DOE and sequentially refined with $K = 10$ points per refinement iteration.

Method	Opt. design (mm)	Cost (mm^2)	$\mathfrak{g}$-calls	Reliability
DDO	$b = 258$ $h = 500$	1.29×10^5	50	$\beta_{\text{SS}} \approx 0.01$
RBDO ($\beta \geq 3$)				
DSA	$b = 372$ $h = 559$	2.15×10^5	$\approx 4 \times 10^6$	$\beta_{\text{SS}} \approx 3.38$
Nested FORM	$b = 399$ $h = 513$	2.12×10^5	9,472	$\beta_{\text{HL}} = 3.38$
Check of FORM assumptions		2.20×10^5	4×10^5	$\beta_{\text{SS}} \approx 3.19$
Brute force	$b = 369$ $h = 560$	2.16×10^5	19×10^6	$\beta_{\text{SS}} \approx 3.35$
Meta-RBDO	$b = 358$ $h = 580$	2.15×10^5	70	$\beta_{\text{SS}} \approx 3.32$

Table 5.9: Comparative results for the short column under oblique bending.

Another interesting fact about this example is that the reliability constraint is not activated at the optimum. Indeed, the algorithm converges at a higher reliability level as illustrated in Figure 5.12. This is due to the specific formulation of the cost function in

Eq. (5.32) that accounts for a failure cost that is depending on the failure probability. Actually, the cost function behaves itself as a constraint and the optimal reliability level is formulated in terms of an acceptable risk p_f ¢ $_f$ instead of an acceptable reliability index β_0.

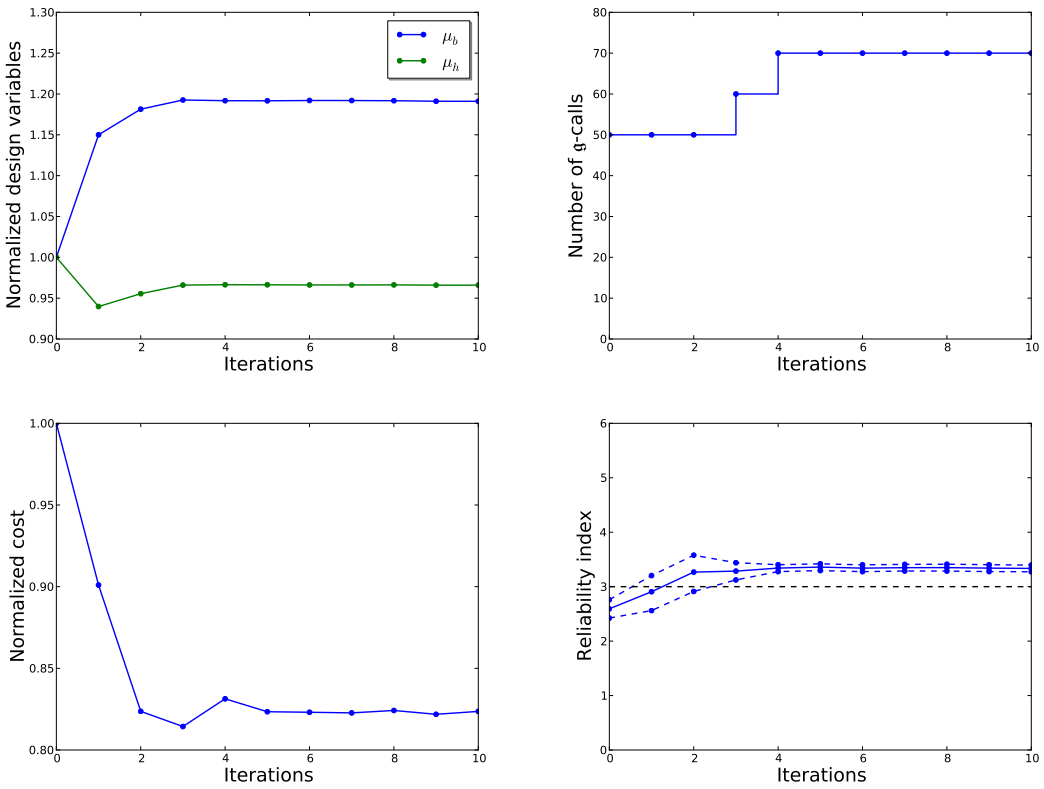


Figure 5.12: Convergence of the proposed meta-model-based approach for the short column under oblique bending. The design variables have been normalized with respect to the initial design ($\mu_b^{(0)} = 300$ mm and $\mu_h^{(0)} = 600$ mm) in order to emphasize the improvement brought by the algorithm.

In the end, it is worth noting that the proposed strategy allowed to save a significant number of calls to the performance function with respect to both the FORM-based and the brute force approaches.

5.2.5 A bracket structure

This section deals with the bracket structure considered by Chateauneuf and Aoues in the book by Tsompanakis et al. (2008).

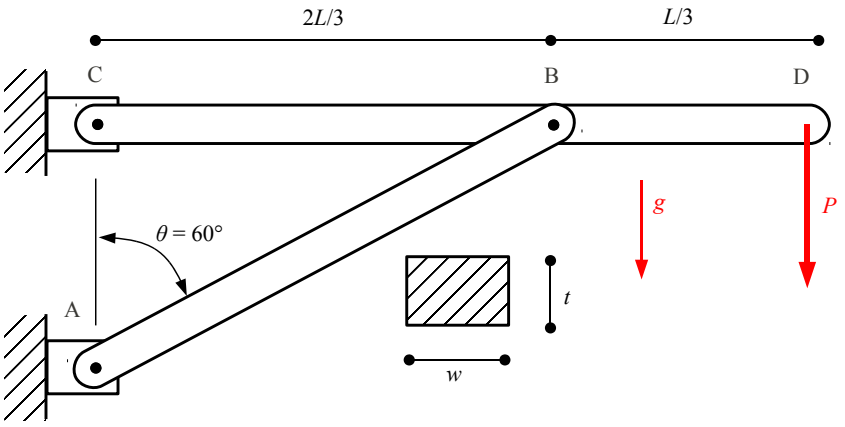


Figure 5.13: The bracket structure proposed by Chateauneuf and Aoues (2008).

5.2.5.1 Problem definition

Mechanical model

Let us consider the bracket structure sketched in Figure 5.13. It is loaded by its own dead load and by a vertical load applied at its right tip. The two failure modes under consideration are as follows.

$\mathbb{F}_1$: The maximum bending stress in the horizontal beam (CD, at point B) should not exceed the yield strength of the constitutive material, so that the first performance function reads as follows:

$$\mathfrak{g}_1(\boldsymbol{x}) = \sigma_y - \sigma_{\mathrm{B}}(\boldsymbol{x}), \tag{5.33}$$

where the maximum bending stress reads:

$$\sigma_{\mathrm{B}}(\boldsymbol{x}) = \frac{6\,M_{\mathrm{B}}}{w_{\mathrm{CD}}\,t^2} \quad \text{with:} \quad M_{\mathrm{B}} = \frac{P\,L}{3} + \frac{\rho\,g\,w_{\mathrm{CD}}\,t\,L^2}{18}. \tag{5.34}$$

$\mathbb{F}_2$: The maximum axial load in the inclined member (AB) should not exceed the Euler critical buckling load (neglecting its own weight), so that the second performance function reads as follows:

$$\mathfrak{g}_2(\boldsymbol{x}) = F_{\mathrm{buckling}}(\boldsymbol{x}) - F_{\mathrm{AB}}(\boldsymbol{x}), \tag{5.35}$$

where the critical Euler buckling load F_{buckling} is defined as:

$$F_{\mathrm{buckling}}(\boldsymbol{x}) = \frac{\pi^2\,E\,I}{L_{\mathrm{AB}}^2} = \frac{\pi^2\,E\,t\,w_{\mathrm{AB}}^3\,9\,\sin^2\theta}{48\,L^2}, \tag{5.36}$$

and the resultant of axial forces in member AB reads (neglecting its own weight):

$$F_{\mathrm{AB}}(\boldsymbol{x}) = \frac{1}{\cos\theta}\left(\frac{3\,P}{2} + \frac{3\,\rho\,g\,w_{\mathrm{CD}}\,t\,L}{4}\right). \tag{5.37}$$

Both members have the same thickness t but different widths w_{AB} and w_{CD}.

Probabilistic model and RBDO problem

The probabilistic model for this example is the collection of independent random variables given in Table 5.10. Note that the coefficients of variation of the random design variables are kept constant along the optimization as in the original paper.

Variable		Distribution	Mean	C.o.V.
P	(kN)	Gumbel	100	15%
E	(GPa)	Gumbel	200	8%
σ_y	(MPa)	Lognormal	225	8%
ρ	(kg/m^3)	Weibull	7,860	10%
L	(m)	Normal	5	5%
w_{AB}	(mm)	Normal	$\mu_{w_{\mathrm{AB}}}$	5%
w_{CD}	(mm)	Normal	$\mu_{w_{\mathrm{CD}}}$	5%
t	(mm)	Normal	μ_t	5%

Table 5.10: Probabilistic model for the bracket structure proposed by Chateauneuf and Aoues (2008).

The RBDO problem consists in finding the rectangular cross-sections of the two structural members that minimize the expected overall structural weight which is computed as follows:

$$\mathfrak{c}\left(\mu_{w_{\mathrm{AB}}}, \mu_{w_{\mathrm{CD}}}, \mu_t\right) = \mu_\rho\, \mu_t\, \mu_L \left(\frac{4\sqrt{3}}{9}\mu_{w_{\mathrm{AB}}} + \mu_{w_{\mathrm{CD}}}\right). \tag{5.38}$$

The optimal design should satisfy a minimum reliability requirement equal to $\beta_0 = 2$ with respect to the two performance functions defined in Eq. (5.33) and Eq. (5.35). The two failure modes are considered independently (*i.e.* the reliability of the series system made of the two failure modes is not considered). The search for the optimal design is bounded to a reasonable hyperrectangle: $50 \leq \mu_{w_{\mathrm{AB}}}, \mu_{w_{\mathrm{CD}}}, \mu_t \leq 300$ (mm).

5.2.5.2 Results

The comparative results for this example are provided in Table 5.11 considering the ones obtained by the brute force approach as a reference. The first row gives the deterministic optimal design that was obtained by Chateauneuf and Aoues (2008) using *partial safety factors* (PSF). It can be seen from the reliability indices that these PSF provide a significant safety level. However, one could potentially argue that the design is over-reliable. Hence, RBDO is used here in order to find an even lighter design allowing for a lower safety level $\beta_0 = 2$. All the RBDO designs presented in Table 5.11 were obtained by taking the PSF-based design as the initial design.

Chateauneuf and Aoues (2008) resorted to the *sequential optimization and reliability approach* (SORA, Du and Chen, 2004) which is a decoupled FORM-based approach. The reliability indices at the optimal design were checked by means of the subset sampling

technique (the coefficient of variation is less than 5%) and revealed that the first-order approximation is fairly accurate in this case. The nested FORM approach provides the same solution but it is less computationally efficient. The brute force approach (nested subset sampling) converges to the same results but clearly proves that crude-sampling-based approaches are not tractable for RBDO. At last, the meta-model-based approach converged to the same design using only a few dozens of runs. The Kriging surrogates were initialized with $K_0 = 50$ points and $K = 10$ points are added at each refinement iteration if required. The convergence of the algorithm is depicted in Figure 5.14.

Method	**Opt. design** (mm)	**Cost** (kg)	**g-calls**	**Reliability**
DDO w/ PSF[a]	$w_{AB} = 61$ $w_{CD} = 202$ $t = 269$	2,632	40	$\beta_{SS1} \approx 4.83$ $\beta_{SS2} \approx 2.83$
RBDO ($\beta_l \geq 2$, $l = 1, 2$)				
Brute force	$w_{AB} = 58$ $w_{CD} = 119$ $t = 241$	1,550	5×10^6 5×10^6	$\beta_{SS1} \approx 1.99$ $\beta_{SS2} \approx 2.00$
SORA[a]	$w_{AB} = 61$ $w_{CD} = 157$ $t = 209$	1,675	1,340	$\beta_{SS1} \approx 1.96$ $\beta_{SS2} \approx 1.98$
Nested FORM[a]	$w_{AB} = 61$ $w_{CD} = 157$ $t = 209$	1,675	2,340	$\beta_{SS1} \approx 1.96$ $\beta_{SS2} \approx 1.98$
Meta-RBDO	$w_{AB} = 58$ $w_{CD} = 128$ $t = 233$	1,584	160 90	$\beta_{SS1} \approx 1.98$ $\beta_{SS2} \approx 1.94$

[a] As computed by Chateauneuf and Aoues (2008).

Table 5.11: Comparative results for the bracket structure proposed by Chateauneuf and Aoues (2008).

Again, the proposed strategy saved a significant number of calls to the performance function and it revealed even more computationally efficient than its FORM-based counterparts.

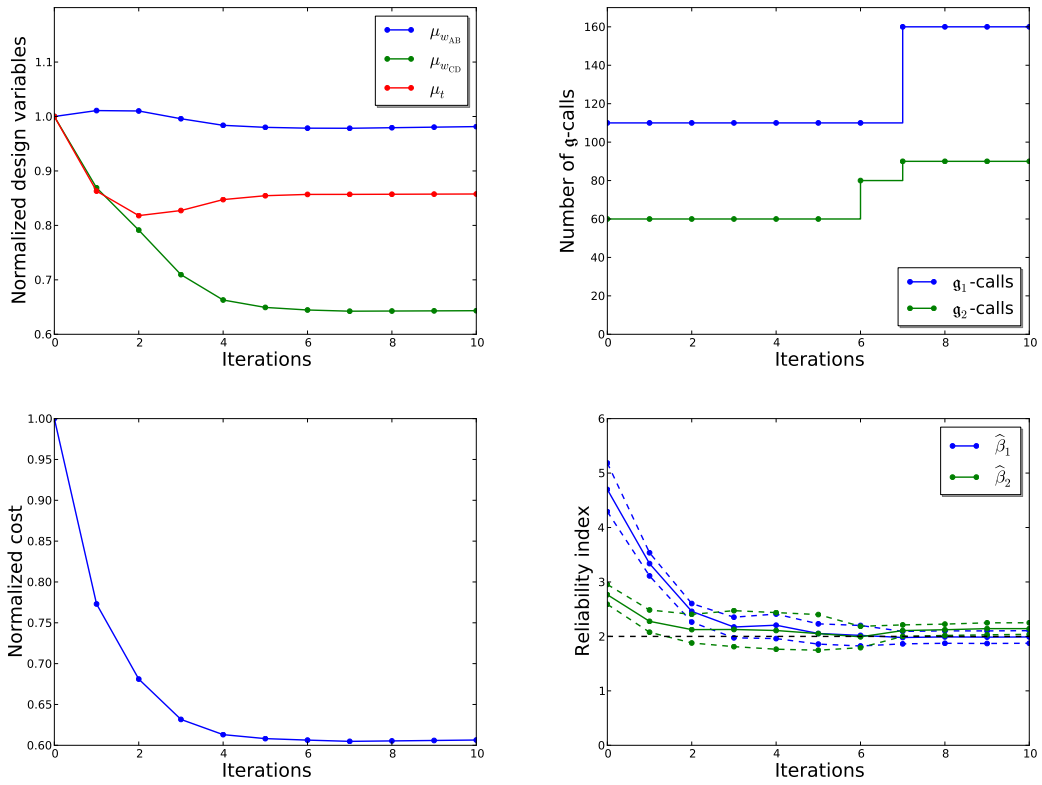


Figure 5.14: Convergence of the meta-model-based RBDO strategy for the bracket structure proposed by Chateauneuf and Aoues (2008). The design variables have been normalized with respect to the initial design chosen here as the PSF-based design of Table 5.11 (first row).

5.2.6 A 23-member plane truss bridge

The mechanical and probabilistic models for this example come from the article by Blatman and Sudret (2008b) and the RBDO problem was originally formulated in Dubourg et al. (2011e).

5.2.6.1 Problem definition

Mechanical model

This example addresses the design of the truss structure illustrated in Figure 5.15. It is composed of 11 horizontal bars and 12 inclined bars. The nodes in the upper horizontal bars are submitted to a vertical load. The maximum vertical displacement of the structure (denoted by V_1 in Figure 5.15) is provided by a two-dimensional bar finite element model. The model comprises 23 bar elements and 13 nodes with 2 degrees of freedom before applying the boundary conditions. The lower left node is clamped horizontally and

vertically while the lower right node is simply clamped vertically in order to get a statically determined structure.

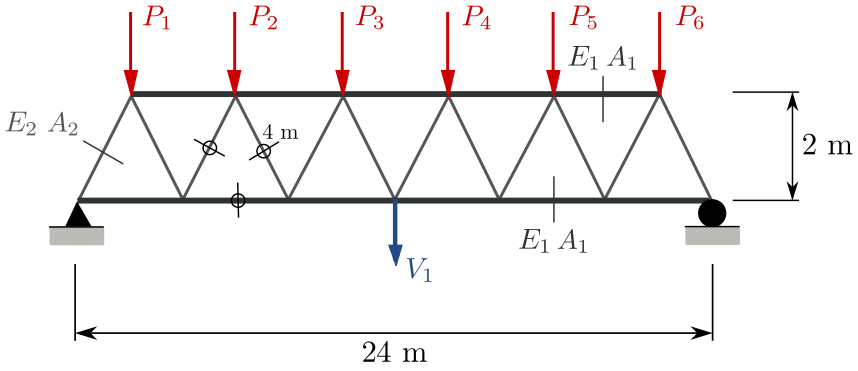


Figure 5.15: A 23-member plane truss bridge.

The maximum vertical displacement of the structure is equal to 7.78 cm when all the variables in the probabilistic model of Table 5.12 are set to their mean values. The design rule is such that the maximum vertical displacement of the structure should not exceed $V_{\max} = 10$ cm, so that the performance function reads as follows:

$$\mathfrak{g}(\boldsymbol{x}) = V_{\max} - \left|V_1(\boldsymbol{x})\right|. \tag{5.39}$$

Note that it is assumed here that the maximum displacement always occurs at the middle node of the lower part of the bridge, although it may not hold for all the realizations of the stochastic model.

Probabilistic model and RBDO problem

Variable	Distribution	Mean	Standard deviation
E_1, E_2 (Pa)	Lognormal	2.10×10^{11}	2.10×10^{10}
A_1 (m^2)	Lognormal	μ_{A_1}	2×10^{-4}
A_2 (m^2)	Lognormal	μ_{A_2}	1×10^{-4}
$P_1, \ldots, P_6$ (N)	Gumbel	5×10^4	7.5×10^3

Table 5.12: Probabilistic model for the 23-member plane truss bridge.

The initial probabilistic model used by Blatman and Sudret (2008b) is given in Table 5.12. It features 10 independent random variables modelling the Young's modulus and the cross-sections of the two groups of bars (resp. horizontal and inclined), as well as the 6 vertical loads applied to the nodes of the upper horizontal bars. The initial mean cross-sections are respectively set to $\mu_{A_1} = 2 \times 10^{-3}$ m^2 and $\mu_{A_2} = 1 \times 10^{-3}$ m^2.

The design optimization problem consists in finding the two cross-sections that minimize the volume of the structure while ensuring a minimal reliability index β_0 set equal to 3. This reads as follows:

$$\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}\in\mathbb{D}} \quad L_1\,\mu_{A_1} + L_2\,\mu_{A_2} \;: \quad \mathbb{P}\left[\mathfrak{g}(\boldsymbol{X}(\boldsymbol{d})) \leq 0\right] \leq \Phi(-\beta_0), \tag{5.40}$$

where L_1 (resp. L_2) is the cumulative length of all horizontal members (resp. inclined members). The search for the optimal solution is bounded to the following hyperrectangle $\mathbb{D} = \left[6\times10^{-4};\, 6\times10^{-3}\right] \times \left[3\times10^{-4};\, 3\times10^{-3}\right]$.

5.2.6.2 Results

The results are presented in Table 5.13. For the sake of validation the results yielded by the proposed meta-model-based RBDO strategy are opposed to that obtained by two nested approaches. The first one is the brute force strategy resorting to subset sampling and the other one resorts to FORM. The provided reliability indices were computed by means of the subset sampling technique in order to check the FORM assumptions (the coefficient of variation is less than 5%). It should be noticed that the results of the nested FORM approach are a bit more optimal because the reliability index was slightly overestimated by the first order approximation. In addition, it can be seen that the use of the Kriging surrogate enables a significant reduction of the computational cost without altering the quality of the reliability index estimation. For this example, the Kriging surrogates were initialized with $K_0 = 100$ points and $K = 50$ points were sequentially added if required. The convergence of the algorithm is depicted in Figure 5.16.

Method	**Opt. design** (m^2)	**Cost** (m^3)	**$\mathfrak{g}$-calls**	**Reliability**
DDO	$A_1 = 1.59\times10^{-3}$ $A_2 = 6.95\times10^{-4}$	0.0939	30	-0.08
RBDO ($\beta \geq 3$)				
Brute force	$\mu_{A_1} = 2.53\times10^{-3}$ $\mu_{A_2} = 7.98\times10^{-4}$	0.1383	3×10^5	3.05
Nested FORM	$\mu_{A_1} = 2.43\times10^{-3}$ $\mu_{A_2} = 7.99\times10^{-4}$	0.1341	1,621	2.74
Meta-RBDO	$\mu_{A_1} = 2.53\times10^{-3}$ $\mu_{A_2} = 8.13\times10^{-4}$	0.1388	350	3.05

Table 5.13: Results for the 23-members plane truss bridge.

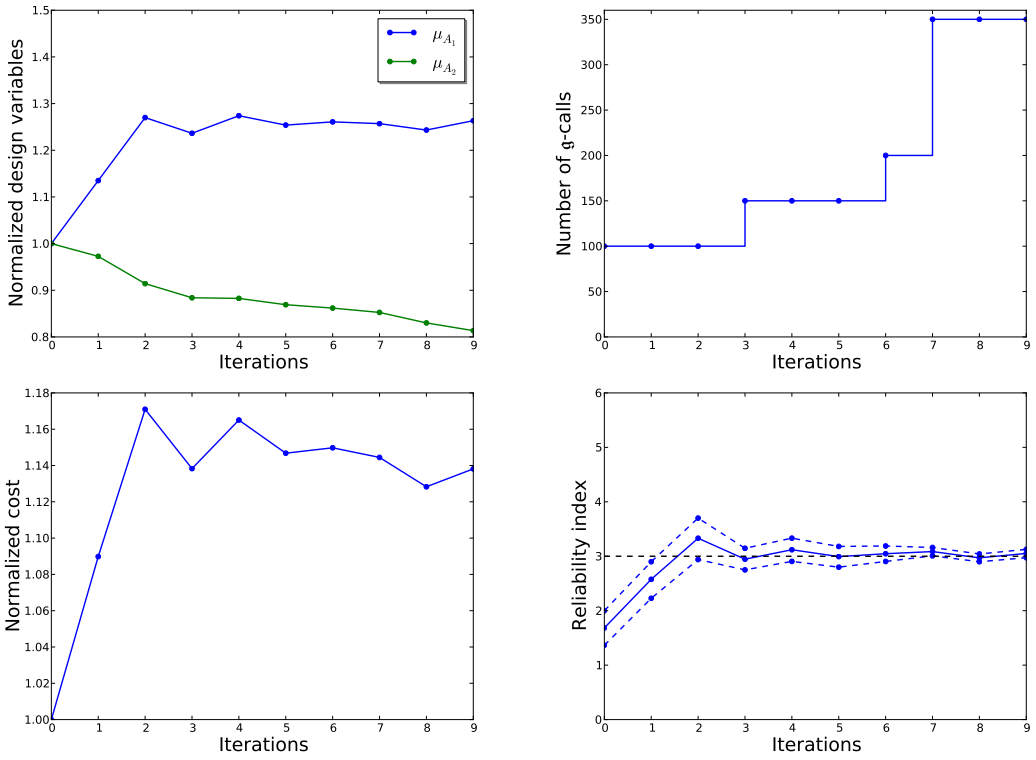


Figure 5.16: Convergence of the meta-model-based RBDO strategy for the 23-member plane truss bridge.

5.2.7 Conclusion

This section validated the proposed global surrogate-based RBDO strategy on a set of structural reliability examples featuring: *(i)* a reasonable dimension (up to 10 variables), *(ii)* reasonably nonlinear performance functions, *(iii)* starting from empirically over- (resp. under-) reliable initial designs.

It has been shown that the use of global Kriging surrogates enables a significant reduction of the computational cost while conserving the good properties of sampling-based estimators for both the reliability indices and their gradients with respect to the design variables. The confidence-bounds-based refinement rule was also proved to be objective as the final surrogate-based reliability indices were all found to be reasonably close to the actual reliability indices estimated on the original performance functions.

Besides the use of Kriging surrogates, it is worth remembering that the subset sampling procedure is of utmost importance in an RBDO context. Indeed, the use of crude Monte Carlo sampling for the nested reliability and reliability sensitivity analyses is not conceivable even on the surrogates due to the possibly low failure probabilities encountered during the optimization. It should also be noted that the score function approach turned out to be a very efficient tool for estimating the gradients of the reliability indices without any additional call to the Kriging predictors.

Finally, it has to be admitted that the use of Kriging predictors as simple plug-in estimators for the performance functions may fail for structural reliability examples featuring a much larger number of variables (say $n > 20$). Hence, for such cases, it is of the author's belief that hybrid approaches such as *meta-model-based importance sampling* could be used for estimating the failure probability in the very last iterations of the optimization (*i.e.* once the location of the optimal solution has been spotted). Actually, in higher dimension the margin of uncertainty of the Kriging predictors will not shrink so that the proposed confidence bound metric will stall for a large number of refinement iterations. Thus, the meta-model-based importance sampling scheme could be triggered at this point for checking the reliability indices and for refining the optimal design based on the sensitivities that are also available for this estimator.

6

Application to the buckling of imperfect shells

Contents

6.1 Introduction

Shell structures occupy a predominant part of our landscape (see the article by Ramm and Wall, 2004, for a review of their applications). They owe this predominance to their curvature which allow them to withstand large transverse loading by a membrane-dominated stress state. As a result, they can be used to form large-span shelters such as roofs, fuselages or boat and submarine hulls without requiring too many intermediate supports such as stiffening beams or rims. Nonetheless, as many optimized structures, thin shells also exhibit a significant sensitivity with respect to their geometrical, material and other environmental conditions which are typically unknown to some extent.

Early work on the stability analysis of slender structures (such as beams or shells) is often attributed to Euler (1744), although most of the theoretical concepts in practice today are due to Lagrange (1788), Timoshenko (1936) and Von Mises (1914). Timoshenko (resp. Von Mises) is known for having derived an analytical expression for the critical buckling load of a cylindrical shell under axial load (resp. external pressure) amongst other related work in structural mechanics. However, parallel experimental studies revealed embarrassing discrepancies between their theory and practice. Koiter (1945) was certainly the first researcher to point out that these discrepancies are mostly explained by the imperfect geometry and boundary conditions of the experimental specimens. This premise is now fully acknowledged by the whole community of engineers and scientists in structural mechanics in the light of other studies by Arbocz and Babcock (1969); Singer et al. (1971); Singer and Abramovich (1995) amongst others.

A key aspect of these imperfections though is that they are extremely varying in terms of shape and amplitude. Hence, for the sake of designing safe structures, designers have to account for extreme (although likely) imperfections in their calculations and then resort to advanced numerical schemes in order to justify their design since the aforementioned analytical formulæ do not hold for imperfect structures. However, this approach, referred to as the *worst case approach* in the sequel, introduces an unknown degree of conservatism which may not suit the safety requirements fixed by the stakeholders.

In this thesis as in previous work by Faulkner and Das (1990); Groen and Kaminski (1996); Bourinet et al. (2000); Schenk and Schuëller (2003, 2007); Noirfalise (2009), it is argued that a more comprehensive approach to the design of imperfect shells necessarily falls under a *probabilistic formulation*. A specific emphasis is put on the probabilistic description of shape and material imperfections. This work does not consider the uncertainty in the boundary conditions (see the work by Schenk and Schuëller, 2003, 2007).

This chapter is organized as follows. Section 6.2 first introduces the fundamental concepts of nonlinear stability analysis in solid mechanics. It also defines the assumptions that are accounted for in the subsequent probabilistic buckling analyses. The remaining sections are devoted to the analysis and the design of two thin-shell structures. Section 6.3 is concerned with the computation of the buckling probability of an imperfect cylindrical shell roof featuring spatially varying random properties. In Section 6.4, the reliability-based design optimization methodology is applied to an imperfect submarine pressure hull. This probabilistic design philosophy is opposed to the state-of-the-art worst case approach.

6.2 Elements of shell nonlinear stability analysis

Buckling is a structural instability problem triggered by some excessive load that needs to be identified. This load will be referred to as the critical buckling (or collapse) load in the sequel. In practice, it is determined by applying a so-called *load proportionality factor* (LPF) which is initialized to zero and then incrementally increased until collapse is observed.

This section shortly formalizes the mechanical problem at hand and presents an original perturbation method proposed by Damil and Potier-Ferry (1990) to solve it in an efficient manner: the *asymptotic numerical method*. The interested reader is referred to the Ph. D. theses by Zahrouni (1998), Baguet (2001) and Noirfalise (2009) and the articles by Cochelin (1994); Cochelin et al. (1994) for a deeper introduction to these advanced concepts of structural mechanics.

6.2.1 Equilibrium path

In continuum mechanics, the equilibrium state of a conservative mechanical system is basically characterized in terms of a zero elementary variation of its total energy denoted by E. This fundamental principle leads to the establishment of the following so-called *variational formulation* of equilibrium states:

$$\delta E = E_{,\boldsymbol{u}}(\lambda, \boldsymbol{u})\,\delta\boldsymbol{u} = 0, \tag{6.1}$$

where $\boldsymbol{u}$ denotes any admissible displacement of the structure and $E_{,\boldsymbol{u}}(\lambda, \boldsymbol{u})$ is the first-order functional derivative of the potential energy which depends on the LPF λ.

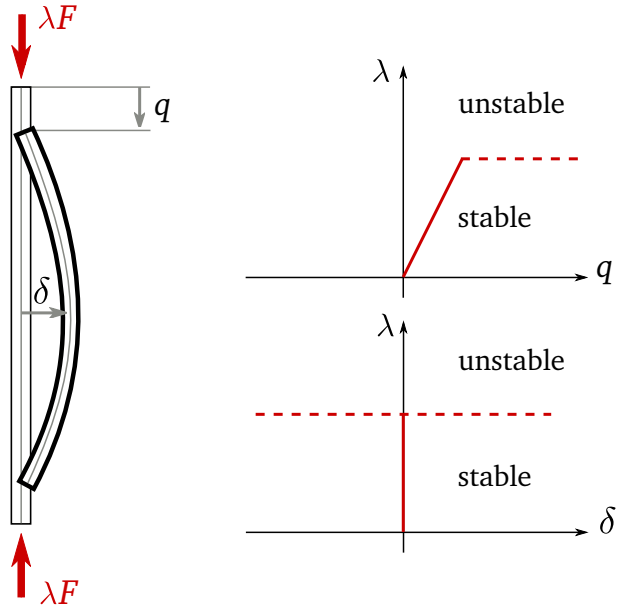


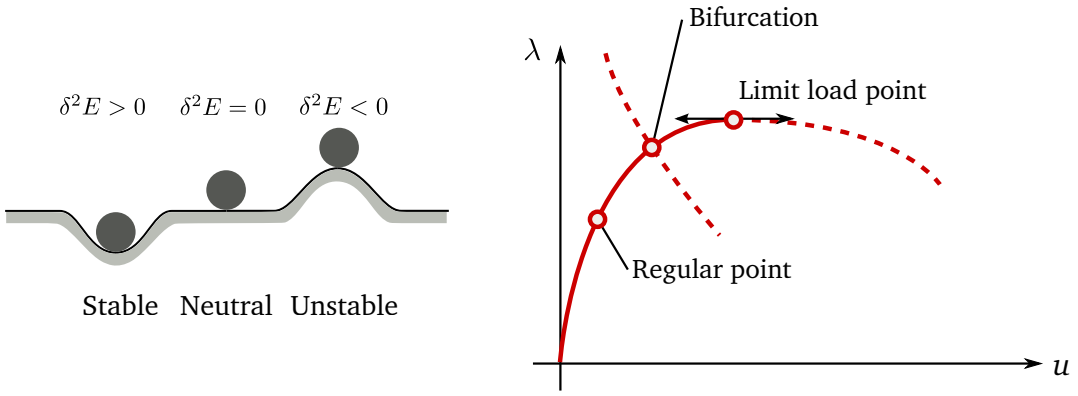
Figure 6.1: Elastic stability of a straight column.

Hence, the purpose of nonlinear mechanics is to solve the latter equation for any value of λ in order to characterize all the equilibrium states of the structure. This infinite set of

equilibrium states is usually referred to as the *equilibrium path* of the structure. It is often represented as a graph opposing a local displacement in some direction to the values of the LPF (see *e.g.* Figure 6.1). For stable structures, the equilibrium state associated with the value $\lambda = 1$ is the only state of interest, although it is obtained incrementally by increasing λ from zero (reference state of the free structure) or from any other *known* equilibrium state $(\lambda_0, \boldsymbol{u}_0)$.

6.2.2 Equilibrium stability

Unstable (resp. stable) equilibrium states of a mechanical system are characterized by a negative (resp. strictly positive) second-order functional derivative of the total energy $E_{,\boldsymbol{u}\,\boldsymbol{u}}$, meaning that they correspond to local maxima (resp. minima) of the the total energy as illustrated in Figure 6.2(a).



(a) Characterization of stability

(b) Limit and bifurcation points

Figure 6.2: Equilibrium stability.

There are basically two kinds of instabilities in nonlinear shell analyses, both potentially leading to buckling and/or plastic collapse: bifurcation points and limit points, see Figure 6.2(b) for illustration. Regarding bifurcation points, the structure may lose its stability along the equilibrium path, resulting in sudden and large displacements which often lead to collapse. Regarding limit points, this occurs when the structure is no longer able to withstand loads, due to nonlinear geometrical and/or material effects. For shell structures, these two kinds of points interact in a rather joint manner, one triggering the other and conversely.

Practical detection of these instability is a rather non-trivial task and it involves the resolution of a *perturbed* equilibrium problem along with the resolution of Eq. (6.1). The present study focuses on imperfect structures which are commonly assumed to fail at their limit load. Indeed, Koiter (1945) showed that the presence of initial imperfections in the structure[1] smooths the equilibrium paths. As a result, imperfect structures regularly feature

[1]The initial study was carried out on cylinders under a compressive axial load with modal imperfections, but this generalizes to other imperfect structures.

smooth limit load points rather than sharp bifurcation points which are only observed on perfect structures (see Figure 6.3). Therefore, in the sequel the detection of singular points along the equilibrium path will be restricted to limit points characterized by *horizontal tangents*.

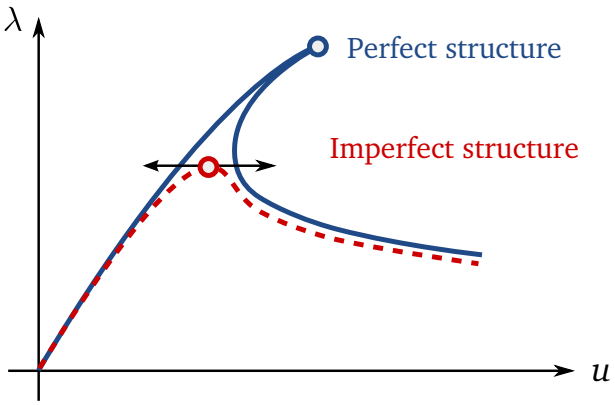


Figure 6.3: Influence of initial imperfections on the equilibrium path.

6.2.3 General formulation of the static equilibrium equations

In static analysis, the total energy of a structure of volume $\mathcal{V}$ is given by:

$$E(\lambda, \boldsymbol{u}) = \int_{\mathcal{V}} W_{\text{int}}(\boldsymbol{\varepsilon})\,\mathrm{d}v - \lambda\, W_{\text{ext}}(\boldsymbol{u}), \tag{6.2}$$

where W_{int} is the strain energy density in the structure and W_{ext} is the virtual work of external forces. $\boldsymbol{\varepsilon}$ denotes the Green-Lagrange strain tensor which is defined as:

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \underbrace{\frac{1}{2}\left(\boldsymbol{\nabla u} + \boldsymbol{\nabla}^{\mathsf{T}}\boldsymbol{u}\right)}_{\boldsymbol{\varepsilon}_l(\boldsymbol{u})} + \underbrace{\frac{1}{2}\left(\boldsymbol{\nabla u}\,\boldsymbol{\nabla}^{\mathsf{T}}\boldsymbol{u}\right)}_{\boldsymbol{\varepsilon}_{nl}(\boldsymbol{u},\,\boldsymbol{u})}, \tag{6.3}$$

where $\boldsymbol{\varepsilon}_l(\boldsymbol{u})$ (resp. $\boldsymbol{\varepsilon}_{nl}(\boldsymbol{u}, \boldsymbol{u})$) denotes the linear (resp. symmetric quadratic bilinear) term of $\boldsymbol{\varepsilon}$. Assuming elastic small displacements, the strain energy density W_{int} reduces to the following quadratic form:

$$W_{\text{int}}(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbf{D} : \boldsymbol{\varepsilon}, \tag{6.4}$$

where $\mathbf{D}$ is the tensor of elasticity modelling the material behaviour.

Differentiating the latter relation with respect to the strain yields the second Piola-Kirchhoff stress tensor:

$$\boldsymbol{\Pi} = W_{\text{int},\boldsymbol{\varepsilon}} = \mathbf{D} : \boldsymbol{\varepsilon}, \tag{6.5}$$

and the equilibrium equation (see Eq. (6.1)) rewrites as the following set of equations:

$$\begin{cases} \delta E &= \displaystyle\int_{\mathcal{V}} \boldsymbol{\Pi} : \delta\boldsymbol{\varepsilon}\,\mathrm{d}v - \lambda\, W_{\text{ext}}(\boldsymbol{u}) = 0, & \text{(equilibrium)} \\ \boldsymbol{\Pi} &= \mathbf{D} : \boldsymbol{\varepsilon} & \text{(stress-strain relationship)} \end{cases}, \tag{6.6}$$

where:

$$\delta\boldsymbol{\varepsilon}(\boldsymbol{u}) = \boldsymbol{\varepsilon}_l(\delta\boldsymbol{u}) + \frac{1}{2}\boldsymbol{\varepsilon}_{nl}(\boldsymbol{u}, \delta\boldsymbol{u}). \tag{6.7}$$

6.2.4 Sources of nonlinearity

6.2.4.1 Large displacements

Accounting for large displacement in buckling analysis is of utmost importance for studying the influence of small imperfections on the critical load. Indeed, the linear buckling eigenproblem which is obtained by approximating the second quadratic term in the Green-Lagrange strain tensor $\boldsymbol{\varepsilon}_{nl}$ reveals a poor approximation for the modelling of imperfect (actual) structures. Therefore, the present imperfection-based analysis accounts for large displacements.

6.2.4.2 Large rotations and shear strain of shells

In order to account for the large rotations and shear strains (*i.e.* a virtual deformation through the thickness) of thin-walled shell structures, the present analyses resort to a three-dimensional seven-parameter shell formulation proposed by Büchter et al. (1994). This formulation, based on the *enhanced assumed strain* (EAS) concept, disables the usual locking problems featured by shell elements. However, it introduces another set of nonlinear *compatibility* equations (see *e.g.* Baguet, 2001, pp. 43–48).

6.2.4.3 Follower forces

The concept of follower forces refers to the property of loads to follow the deformed shape of the structure. A typical example of a follower force field is the hydrostatic pressure exerted on a submarine hull. Indeed, such a pressure is exerted along the normal to the boundary surface of the structure whatever its shape is. From a computational viewpoint, this additional assumption introduces a dependence of the virtual work of external forces on the LPF (see *e.g.* Noirfalise, 2009, pp. 81–86).

6.2.4.4 Material properties

All the structures considered in the sequel are assumed to be made of the same steel which is supposed to behave according to a nonlinear elastic Ramberg-Osgood constitutive law (also known as the *power law*). It is characterized by the following nonlinear stress-strain relationship:

$$\varepsilon = \frac{\sigma}{E} + \alpha\frac{\sigma}{E}\left(\frac{|\sigma|}{\sigma_y}\right)^{n-1}, \tag{6.8}$$

where ε is the strain, σ is the stress, α and n are the two Ramberg-Osgood parameters, E is the Young's modulus and σ_y is the yield strength.

Note that plasticity cannot be taken into account within the FE code that is used here. Even though, it is argued that the two structures under concern in this chapter do not present any significant local unloading until the collapse load of interest is reached. In this case, nonlinear elasticity is a fairly accurate model for the constitutive laws.

6.2.5 The asymptotic numerical method

The problem in Eq. (6.6) is usually solved by means of the so-called incremental iterative methods such as the Newton-Raphson algorithm. The present work is based on an original alternative known as the *asymptotic numerical method* (ANM) proposed by Damil and Potier-Ferry (1990) and Cochelin (1994).

6.2.5.1 The idea

It is first proposed to rewrite the problem in Eq. (6.6) into the following convenient quadratic form:

$$\boldsymbol{R}(\boldsymbol{Y}, \lambda) = L(\boldsymbol{Y}) + Q(\boldsymbol{Y}, \boldsymbol{Y}) - \lambda \boldsymbol{F} = 0, \tag{6.9}$$

where $\boldsymbol{R}$ is a vector of residuals, L is a linear operator, Q is a bilinear quadratic operator, $\boldsymbol{F}$ is a constant vector, and $\boldsymbol{Y} = (\boldsymbol{u}^\mathsf{T}, \boldsymbol{S}^\mathsf{T})^\mathsf{T}$ regroups the unknowns of the problem.

A key idea of the ANM then consists in expanding the unknowns $\boldsymbol{Y}$ and λ over a unique parameter noted a in the form of the following polynomial series expansions:

$$\left\{ \begin{array}{rcl} \boldsymbol{Y}(a) & = & \boldsymbol{Y}_0 + a\,\boldsymbol{Y}_1 + a^2\,\boldsymbol{Y}_2 + \ldots + a^N\,\boldsymbol{Y}_N \\ \lambda(a) & = & \lambda_0 + a\,\lambda_1 + a^2\,\lambda_2 + \ldots + a^N\,\lambda_N \end{array} \right. , \tag{6.10}$$

where $(\lambda_0, \boldsymbol{Y}_0)$ describes the initial state of the system, supposedly known. In this study, the polynomial expansions are truncated after $N = 30$ terms.

Introducing these expansions into Eq. (6.9) and grouping the terms with the same power of a then yield the following succession of linear systems for orders $p = 1, \ldots, N$:

$$\left\{ \begin{array}{rcl} L_t(\boldsymbol{Y}_1) & = & \lambda_1 \boldsymbol{F} \\ L_t(\boldsymbol{Y}_2) & = & \lambda_2 \boldsymbol{F} - Q(\boldsymbol{Y}_1, \boldsymbol{Y}_1) \\ & \vdots & \\ L_t(\boldsymbol{Y}_N) & = & \lambda_N \boldsymbol{F} - \displaystyle\sum_{p=1}^{N-1} Q(\boldsymbol{Y}_p, \boldsymbol{Y}_{N-p}) \end{array} \right. , \tag{6.11}$$

where $L_t(\bullet) = L(\bullet) + 2Q(\boldsymbol{Y}_0, \bullet)$ is the *tangent operator*, which is the same at all orders.

At this stage, the problem involves one too many unknown, namely the parameter a. As in the classical incremental iterative methods, the ANM uses a pseudo arc-length technique by setting:

$$a = (\boldsymbol{Y} - \boldsymbol{Y}_0)\,\boldsymbol{Y}_1 + (\lambda - \lambda_0)\,\lambda_1, \tag{6.12}$$

which completes the system of equations in Eq. (6.11).

Hence, it can be seen that the initial nonlinear problem in Eq. (6.9) has been genuinely transformed into a set of N linear systems by rejecting all nonlinearities to the right-hand side of Eq. (6.11). In addition, the N linear systems composing Eq. (6.11) feature a single linear operator L_t which is the same at all order. When switching to the discrete form of the problem (by means of a classical finite element method), the resolution of the N linear systems require only one decomposition of the tangent stiffness matrix $\mathbf{K}_t$ which is the discrete counterpart of L_t. This latter remark makes the ANM very efficient as the tangent stiffness matrix $\mathbf{K}_t$ may be large in practice.

Eventually, the ANM provides a continuous representation of the equilibrium path for any value of λ thanks to the series expansion in a. This is an interesting property with respect to the incremental iterative methods that need to solve the problem for each value of λ.

6.2.5.2 Validity of the expansion

However, due to the use of finite expansions in Eq. (6.10), the solution becomes invalid for large values of a. Thus, it is proposed to truncate the solution below a maximum value of a denoted by $a_{\max}$. This maximum value is based on a study of the residual's norm $\|\boldsymbol{R}\|$.

Cochelin (1994) proved that it is reasonable to approximate the norm of the residual by that of the first omitted term in the expansion, so that:

$$\|\boldsymbol{R}(a)\| \approx \left\| a^{N+1} \boldsymbol{R}_{N+1} \right\|. \tag{6.13}$$

Based on this approximation, Cochelin then came up with the following expression for $a_{\max}$:

$$a_{\max} = \left(\epsilon \frac{\|\boldsymbol{F}\|}{\left\|\boldsymbol{R}_{N+1}\right\|} \right)^{\frac{1}{N+1}}, \tag{6.14}$$

where ϵ is the maximum tolerance on the norm of the residual. This tolerance is usually set equal to a small value (here 10^{-8}) thanks to the normalization of the residual with respect to the right hand-side $\|\boldsymbol{F}\|$ of Eq. (6.9).

The description of the whole equilibrium path is therefore made piecewise by repeating the procedure incrementally (*i.e.* by resetting the initial state of the system $(\lambda_0, \boldsymbol{Y}_0)$ to $(\lambda(a_{\max}), \boldsymbol{Y}(a_{\max}))$. Even though, the ANM remains more computationally efficient than its incremental iterative counterparts because it is incremental only. Indeed, incremental iterative methods need to iterate within the increments in order to remain on the equilibrium path, each iteration involving an expensive decomposition of the tangent stiffness matrix.

6.2.5.3 Determination of the limit-load carrying capacity

The determination of the limit-load carrying capacity exploits the parametric approximation of the load proportionality factor. Indeed, limit points are characterized by horizontal

tangents on the equilibrium path thus meaning that the derivative of the load proportionality factor with respect to a equals zero at the critical limit-load. This naturally lead to the following definition:

$$\lambda_{\text{limit}} = \lambda(a_{\text{limit}}), \quad \text{with} \quad a_{\text{limit}} = \left\{ a \in [0;\, a_{\max}] : \frac{\mathrm{d}\lambda}{\mathrm{d}a} = 0 \right\}. \tag{6.15}$$

Thanks to the chosen polynomial series expansion for the LPF, finding the the limit load simply consists in finding the roots of a polynomial of order $N-1$ and retaining the lowest positive root that is less than $a_{\max}$, provided it exists.

6.2.6 The EVE finite element code

The asymptotic numerical method has been coupled to a finite element model within the EVE code whose development was initiated by Cochelin (1994) and Baguet (2001). The present studies resorts to a more recent version developed by Noirfalise (2009) who implemented the follower forces and some other functionalities that are used in the sequel (such as local boundary conditions for the Büchter-Ramm shell element).

6.3 Reliability analysis of the Scordelis-Lo shell roof

This section is concerned with the reliability analysis of the cylindrical shell roof illustrated in Figure 6.4. This mechanical example is inspired from the article by Scordelis and Lo (1961) (see also Ramm and Wall, 2004, pp. 405–406), but the load case and the material properties have been modified as in Dubourg et al. (2009a, 2011b).

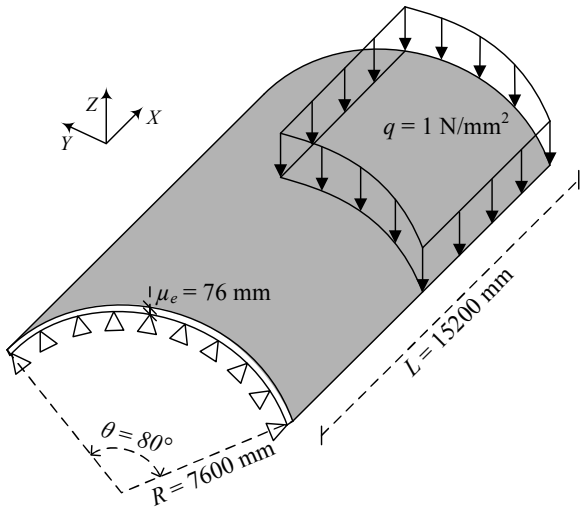
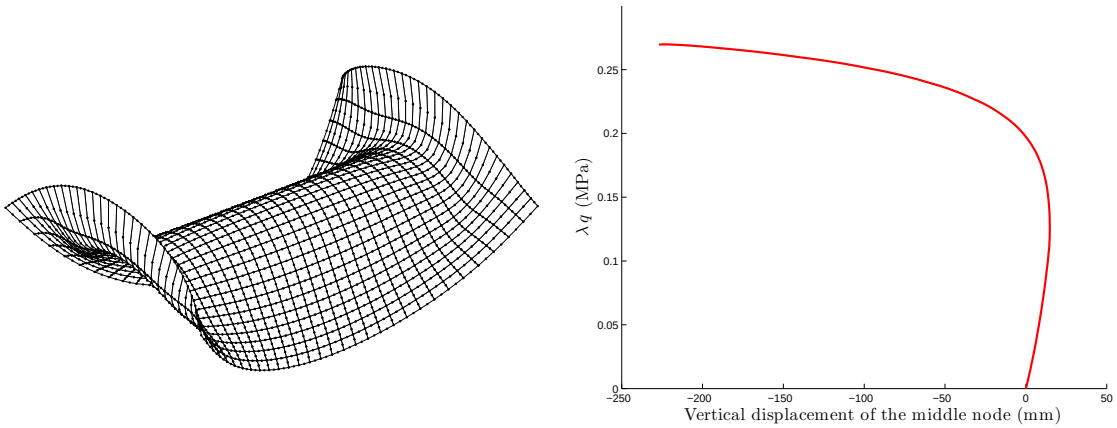


Figure 6.4: A shell roof under a uniformly distributed vertical load (Scordelis and Lo, 1961).

6.3.1 The Scordelis-Lo shell roof

The dimensions of the cylindrical shell roof under consideration are given in Figure 6.4. Its longitudinal edges are free while its circumferential edges are simply supported by rigid diaphragms, meaning that the radial displacement is fixed equal to zero. It is subjected to a distributed (non-follower) vertical load $q = 1$ MPa (over the whole roof upper face). This load emulates the dead load of the structure plus an additional load due to a snow layer so that the force field is exerted vertically whatever the shape of the deformed structure.

The constitutive material is assumed to follow a nonlinear elastic Ramberg-Osgood law as described in Section 6.2.4.4. The problem is discretized with a mesh containing 30×30 8-node Büchter-Ramm shell elements, thus featuring 2,821 nodes and 16,926 degrees of freedom before applying the boundary conditions. The limit load is computed by differentiating the piecewise polynomial series expansion of the equilibrium path as explained in Section 6.2.5.3. Figure 6.5(a) illustrates the amplified deformed shape of the perfect structure (*i.e.* the one corresponding to the mean of the forthcoming probabilistic model) at its limit load capacity which is found equal to $q_{\text{limit}} = 0.2676$ MPa. Figure 6.5(b) represents the equilibrium path for the vertical displacement of the node located at midspan of the roof.



(a) Deformed shape at limit load

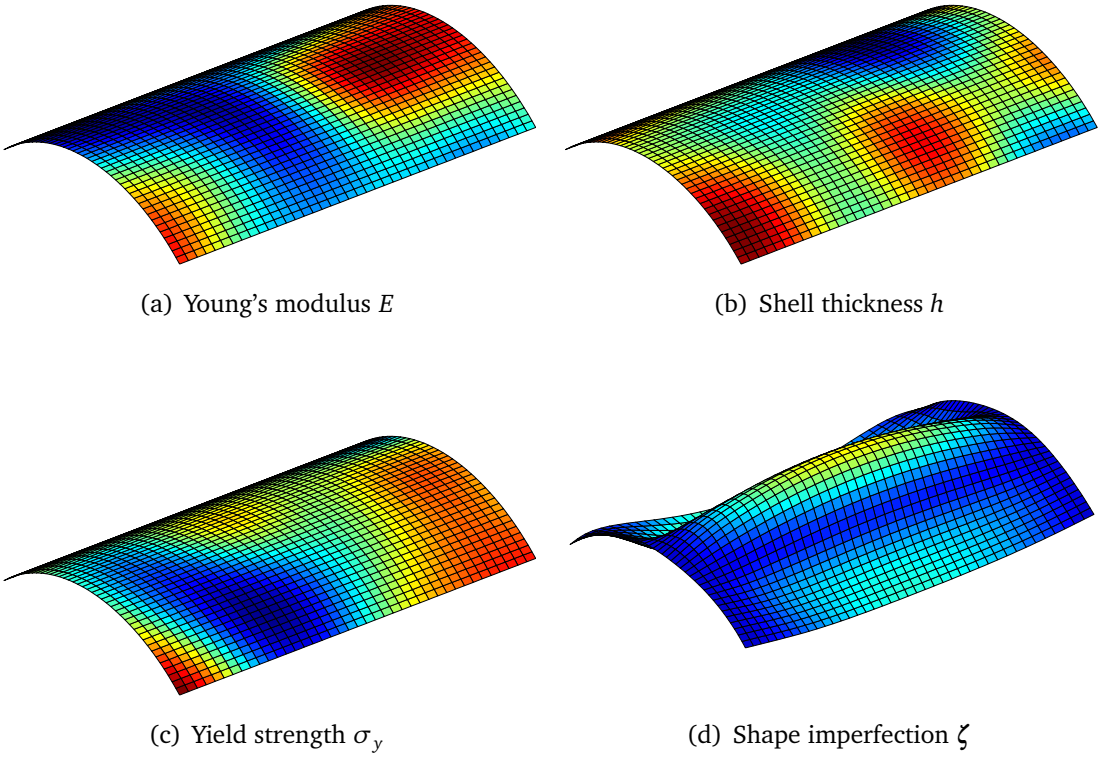
(b) Equilibrium path

Figure 6.5: Nonlinear stability analysis of the perfect shell roof.

6.3.2 Spatially varying random properties

The probabilistic model for this structure resorts to random fields for modelling the spatial variability of the shell properties. More specifically, the material properties (the Young's modulus E and the yield strength σ_y) and the shell thickness h are modelled by independent lognormal random fields represented by translated Karhunen-Loève expansions, and the shape imperfection is represented by a random linear combination of the three most critical Euler buckling modes of the shell roof.

The probabilistic model described in the sequel features 93 independent Gaussian random variables grouped in the vector $\boldsymbol{\Xi} = (\boldsymbol{\Xi}_E, \boldsymbol{\Xi}_{\sigma_y}, \boldsymbol{\Xi}_h, \boldsymbol{\Xi}_\zeta)$. A random realization of the four random fields is illustrated in Figure 6.6.



(a) Young's modulus E (b) Shell thickness h

(c) Yield strength σ_y (d) Shape imperfection ζ

Figure 6.6: A random realization of the four random fields.

6.3.2.1 Representation of the lognormal random fields

The three lognormal random fields E, σ_y and h are obtained by transforming the realizations of three independent standard Gaussian random fields U_E, U_{σ_y} and U_h through the following translation:

$$F(x, \theta) = \exp\left(\lambda_F + \zeta_F U_F(\boldsymbol{p})\right), \quad F = E, \sigma_y, h, \tag{6.16}$$

where λ_F and ζ_F are the location and scale parameters of the lognormal distributions whose mean and coefficient of variation are given in Table 6.1.

Variable		Distribution	Mean	C.o.V.
E	(MPa)	Lognormal	200,000	3%
σ_y	(MPa)	Lognormal	390	7%
h	(mm)	Lognormal	76	5%

Table 6.1: Parameters of the lognormal random fields for the material properties of the shell roof.

The three underlying Gaussian random fields are assumed to have a zero mean, a unit variance and the following isotropic squared exponential covariance function:

$$C\left(x-x',\,\theta-\theta'\right)=\exp\left[\frac{(x-x')^2}{\ell^2}+\frac{R^2\,(\theta-\theta')^2}{\ell^2}\right], \tag{6.17}$$

where the correlation length ℓ is set equal to 3 500 mm. These random fields are represented by means of their truncated Karhunen-Loève expansion:

$$U_F(x,\theta)=\sum_{k=1}^{M}\Xi_{F\,k}\,\sqrt{\lambda_k}\,\varphi_k(x,\,\theta),\quad F=E,\,\sigma_y,\,h, \tag{6.18}$$

where $\boldsymbol{\Xi}_F=(\Xi_{F\,k},\,k=1,\ldots,M)$ is a vector of independent standard normal random variates. It is recalled that the M couples $\left\{\left(\lambda_k,\,\varphi_k\right),k=1,\ldots,M\right\}$ are the elements of the spectral decomposition of the correlation function corresponding to the greatest eigenvalues λ_k (see Eq. (1.69) in Chapter 1). The latter spectral decomposition is performed here numerically using the wavelet-Galerkin scheme proposed by Phoon et al. (2002). The size M of the expansion is the same for the three fields and it is equal to $M=30$ so that the relative mean squared error with respect to the non-truncated expansion is less than 3.70%. The Haar wavelet family used for discretizing the Fredholm integral equation is truncated at level 7.

6.3.2.2 Representation of the shape imperfection

The modelling of the shape imperfection is based on mechanical considerations. It is proposed to model the shape imperfection as a random combination of its most critical buckling patterns. Such patterns are obtained by means of a linear buckling analysis of the perfect structure. They are illustrated in Figure 6.7. The shape imperfection being modelled as a linear combination of these three modes, it will conserve their symmetry properties with respect to both the longitudinal and transversal axes.

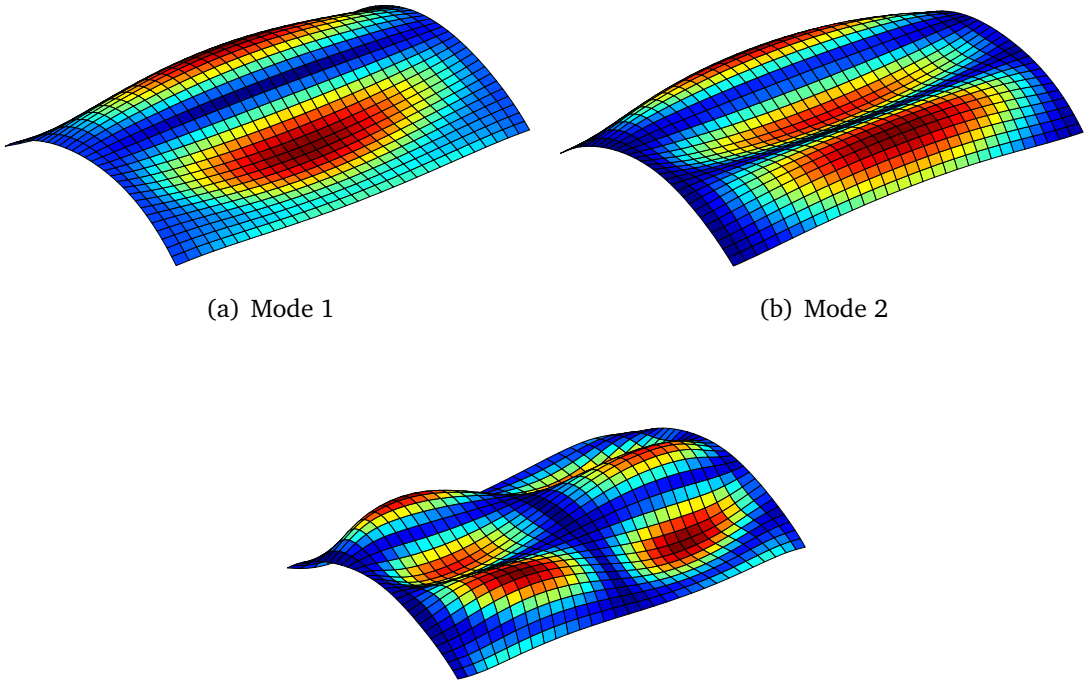
Hence, the realizations of the shape imperfection vector field are given by:

$$\boldsymbol{\zeta}\,(x,\,\theta)=\sum_{k=1}^{3}\Xi_{\zeta\,k}\,\boldsymbol{U}_k\,(x,\,\theta), \tag{6.19}$$

where $\left\{\boldsymbol{U}_k,\,k=1,\,2,\,3\right\}$ are the displacement vector fields between the three first buckling modes and the perfect reference structure. The three Gaussian variables $(\Xi_{\zeta\,k},k=1,\,2,\,3)^{\mathsf{T}}$ are independent, zero-mean and their standard deviation is set equal to $\sigma_\zeta=9.5$ mm. This empirical moments were determined with the two following objectives.

(i) The mean shape should match the perfect structure, hence $\mu_\zeta=0$.

(ii) The maximum amplitude of the shape imperfection should not exceed a fraction of the mean thickness for the sake of validity of the chosen shell formulation. Hence the standard deviation has been adjusted by Monte Carlo sampling so that the maximum amplitude at ±2 standard deviations matches the half of the roof mean thickness $\mu_h/2=38$ mm. Such a criterion is met for $\sigma_\zeta=9.5$ mm.



(a) Mode 1 (b) Mode 2

(c) Mode 3

Figure 6.7: The three most critical buckling patterns of the shell roof.

6.3.3 Reliability analysis

The present reliability analysis follows the classical formulation used in the probabilistic buckling literature. It consists in estimating the following failure probability:

$$\mathbb{P}\left[\boldsymbol{\Xi} \in \mathbb{F}\right] = \mathbb{P}\left[q_{\text{limit}}(\boldsymbol{\Xi}) \leq q_{\text{ser}}\right], \tag{6.20}$$

where the service load is arbitrarily set equal to $q_{\text{ser}} = 0.18$ MPa in order to make the failure probability sufficiently low. Recall that $\boldsymbol{\Xi}$ denotes here the 93-component random vector which is used to sample the four random fields.

Three different reliability methods have been used on this example on purpose to investigate the features of the failure domain $\mathbb{F}$. First, the subset sampling approach is used in order to provide a sound reference solution. Then, the meta-model-based importance sampling strategy is applied so as to prove its scalability to advanced structural reliability problems featuring a reasonably large stochastic model and a nonlinear performance model. Eventually, the restarted iHLRF algorithm (Der Kiureghian and Dakessian, 1998) is used in an attempt to shed some light on the most probable configurations of the probabilistic model that lead to failure.

The three failure probability estimates are given in Table 6.2. Note that the meta-model-based importance sampling required a quite significant number of finite element runs although it is still less computationally expensive than subset sampling. For this example,

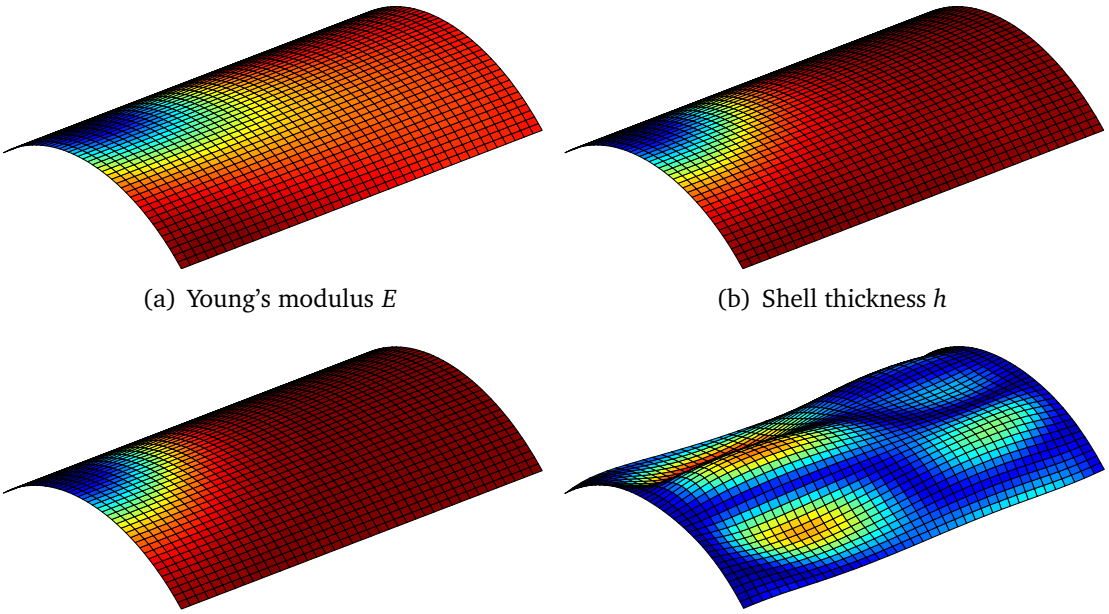
the correction factor equals $\alpha_{\text{corr}} \approx 0.641$. This is significantly far from unity with respect to the other examples introduced in Chapter 5. This is because the Kriging surrogate does not accurately fit the limit-state surface $\{\boldsymbol{\xi} \in \mathbb{R}^{93} : g_{\text{limit}}(\boldsymbol{\xi}) = g_{\text{ser}}\}$ in this case due to its significant nonlinearity and the important dimension of the input space. Even though, this example truly reveals the interest of the proposed meta-model-based importance sampling scheme as it yields unbiased estimates of the probability of interest.

Method	Subset sampling	Multi-FORM	Meta-IS
# of FE-runs	20,000	$\approx 10,000$	$6 \times 93 + 9\,500$
Probability	1.27×10^{-4}	1.22×10^{-4}	1.32×10^{-4}
C.o.V.	<13%	–	<14%

Table 6.2: Results of the reliability analysis of the shell roof.

The multiple FORM approximation was computed by means of the Φ_n approach on the series system composed with the Taylor series expansion of the limit-state surface in the vicinity of the four most probable failure points (see Section 3.3.4.2 of Chapter 3). These four points were found by means of the restarted algorithm iHLRF algorithm tuned as in the original paper by Der Kiureghian and Dakessian (1998).

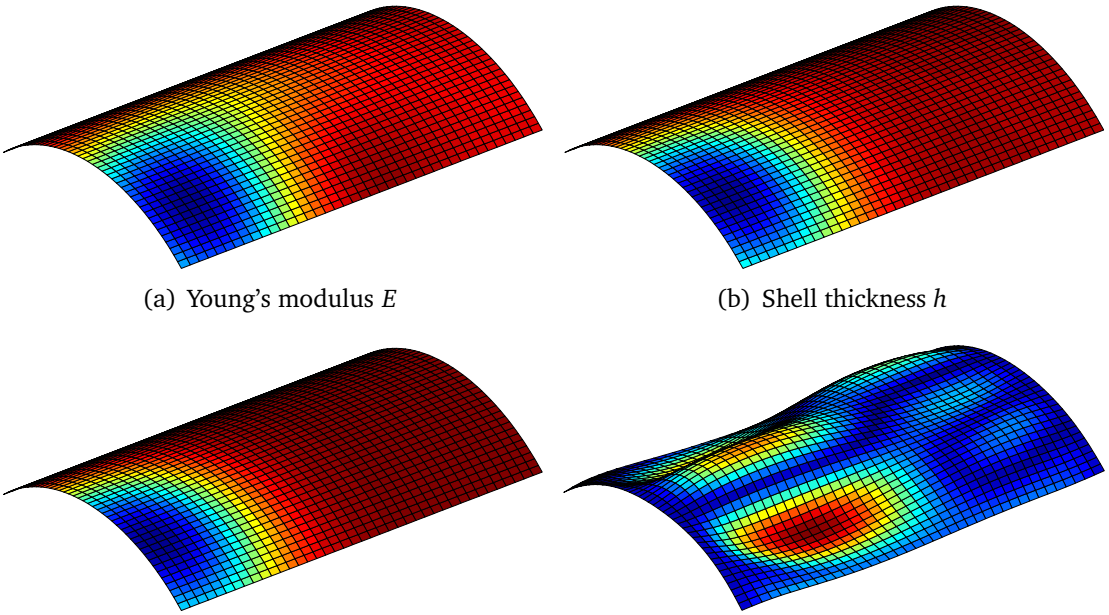
Besides the fact that the Multi-FORM approximation of the failure probability is in reasonable agreement with the other sampling-based approach, the coordinates of the four most probable failure configurations are proved to be physically meaningful as illustrated in Figures 6.8–6.11. Actually, these four configurations feature locally low capacities (*i.e.* low values of the Young's modulus, the shell thickness and the yield strength) and high demand (*i.e.* a high shape imperfection). Due to the symmetry of the three buckling patterns used to build the random shape imperfection, there exists four most probable failure configurations located at the four corners of the roof. Note that the consideration of the four failure modes in a series system is of utmost importance for the FORM approximation because each component has a failure probability of 3×10^{-5} only.



(a) Young's modulus E (b) Shell thickness h

(c) Yield strength σ_y (d) Shape imperfection ζ

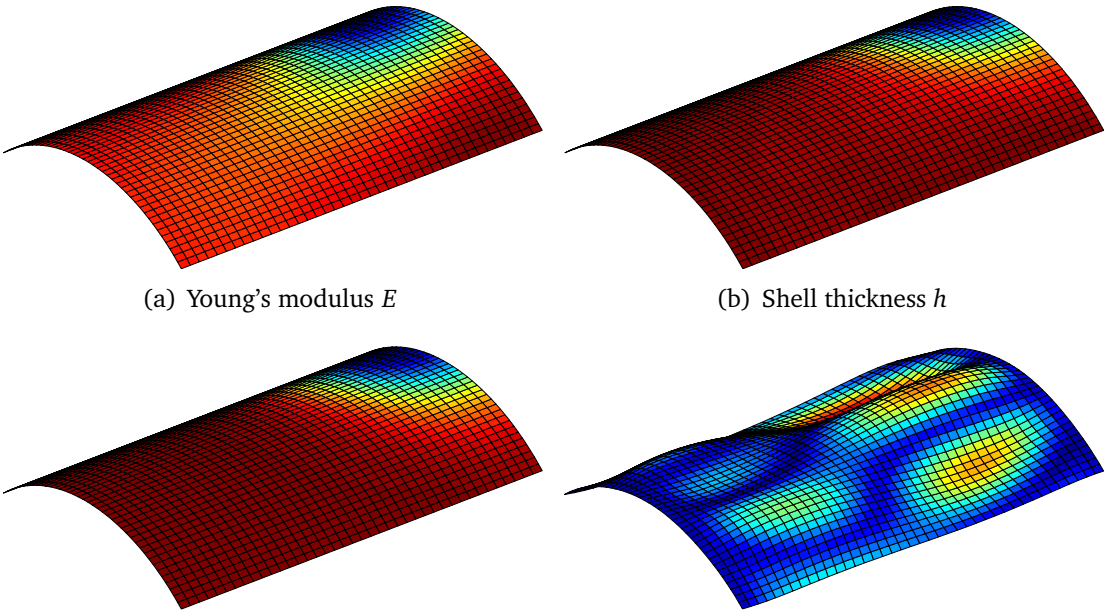
Figure 6.8: Most probable failure configuration #1.



(a) Young's modulus E (b) Shell thickness h

(c) Yield strength σ_y (d) Shape imperfection ζ

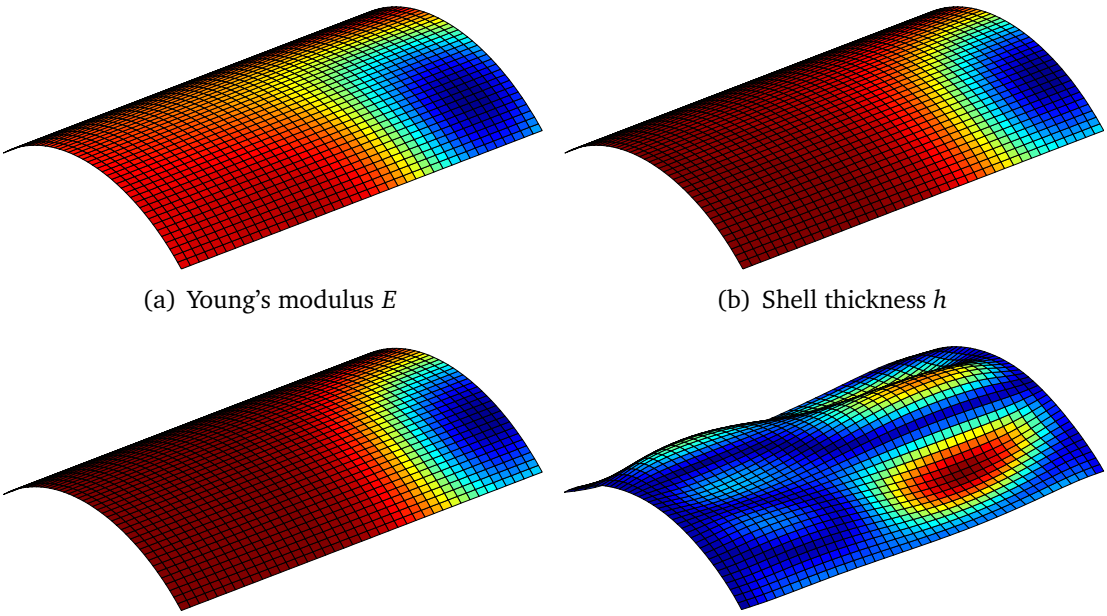
Figure 6.9: Most probable failure configuration #2.



(a) Young's modulus E (b) Shell thickness h

(c) Yield strength σ_y (d) Shape imperfection ζ

Figure 6.10: Most probable failure configuration #3.



(a) Young's modulus E (b) Shell thickness h

(c) Yield strength σ_y (d) Shape imperfection ζ

Figure 6.11: Most probable failure configuration #4.

Eventually, it is worth mentioning that subset sampling interestingly provides the full CDF of the critical buckling load as illustrated in Figure 6.12. This curve is interesting in probabilistic buckling analysis as it gives a relationship between the serviceability threshold q_{ser} and the associated failure probability.

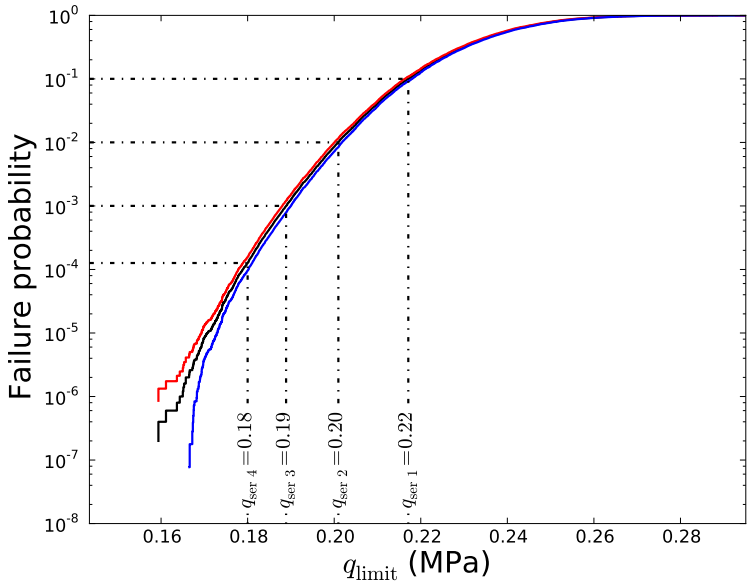


Figure 6.12: Cumulative distribution function of the critical buckling load estimated by subset sampling and its $\pm 2\,\sigma$ confidence interval.

6.3.4 Conclusion

The buckling collapse of the Scordelis-Lo shell roof has been investigated by means of a probabilistic approach. This comprehensive modelling has accounted for spatially and randomly varying shell properties described by random fields. The Gaussian shape imperfection modelling has been based on a combination of the most critical linear elastic buckling modes as typically considered in engineering practice. The other three lognormal random fields were represented with translated Karhunen-Loève expansions. The resulting probabilistic model is driven by 93 random variables.

Three different reliability methods have been applied. The subset sampling technique was first used to provide a sound reference solution. It also enables the reconstitution of the CDF of the critical buckling load which is an interesting by-product for probabilistic buckling analyses. The restarted iHLRF of Der Kiureghian and Dakessian (1998) was then applied in order to perform an exhaustive search of the most probable failure configurations. Fortunately, the original tuning of the algorithm allowed to find 4 configurations whose interpretation reveals in good agreement with the physical intuition.

At last, despite the strong nonlinearity and the high dimensionality of the problem, the meta-model-based importance sampling scheme provided an unbiased estimate of the failure probability. This example confirms the interest one should have in using such an hybrid approach as it enables the use of a coarse meta-model for accurately estimating failure probabilities in a reasonably high dimension.

6.4 Reliability-based design of a submarine pressure hull

Submarine pressure hulls such as the one illustrated in Figure 6.13 are mainly composed of a combination of ring-stiffened cylinders, cones, elliptical or spherical ends, internal diaphragms, bulkheads and deep frames. When immersed, these structures are subjected to an external hydrostatic pressure that is proportional to the diving depth I:

$$p = \rho_{\text{water}}\, g\, I, \tag{6.21}$$

where ρ_{water} is the sea water density (set here equal to $1{,}000$ kg/m^3), and $g \approx 10$ m/s^2 is the gravitational constant. Such a load induces a compression stress state that is mostly membrane dominated. Design against buckling therefore constitutes a key point for submarines.

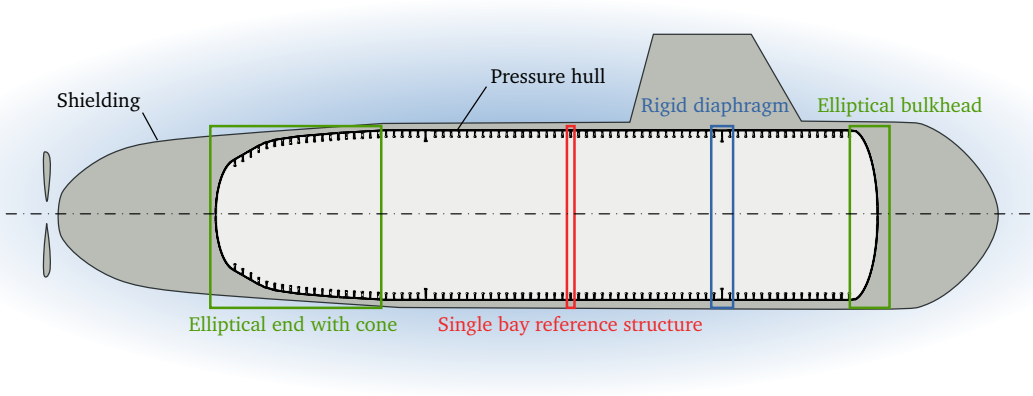


Figure 6.13: Composition of a typical submarine pressure hull.

The design practice is usually based on specific standards and design codes such as the *British standard 5500* (BS5500) or the more recent *Eurocode 3*, possibly along with finite-element-based simulations. It makes often use of long-term-experience-based safety factors at various design stages, which eventually implies an *unknown* degree of conservatism. Hence, structural reliability methods reveal a promising tool for investigating the safety margins attached to the current submarine design practices (see *e.g.* Faulkner and Das, 1990; Groen and Kaminski, 1996; Bourinet et al., 2000).

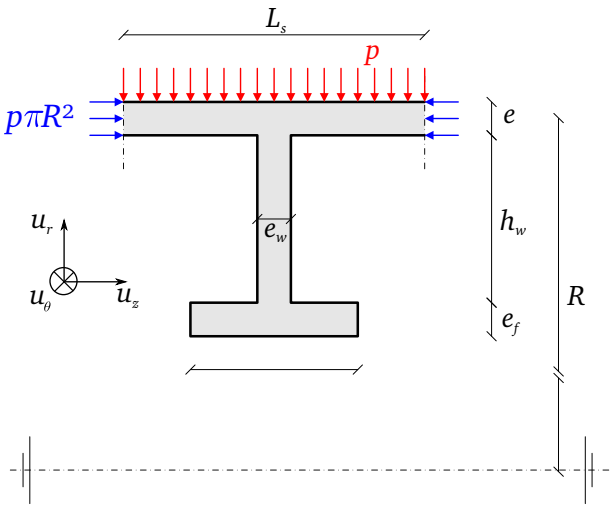
Another major challenge for the designer consists in finding an optimal ratio between the volume of the resistant structure and the boarding capacity of the submersible. The latter point falls under the *reliability-based design optimization* (RBDO) formulation. The work presented in the sequel is based on preliminary studies published in Dubourg et al. (2008, 2011a).

Note that for the sake of confidentiality, the submarine design study exposed in the sequel does not make use of DCNS design criteria nor data, and the final design does not actually meet the requirements imposed by the architect. This work is based on elements that are publicly available in the literature (Bourinet et al., 2000; Gayton et al., 2003; Dubourg et al., 2008).

6.4.1 The single bay reference structure

6.4.1.1 A ring-stiffened shell cylinder

The present study does not consider the submarine hull as a whole. It focuses instead on a single bay reference structure which consists in a ring-stiffened shell cylinder. The dimensions of this elementary structure are shown in Figure 6.14. The outer cylinder, the ring and the inner cylinder are respectively referred to as the shell plating, the web and the flange. This simplified modelling makes sense for the bays that are located far from the bulkheads and the other singularities exhibited by actual submarine hulls.



(a) Ring-stiffened shell cylinder

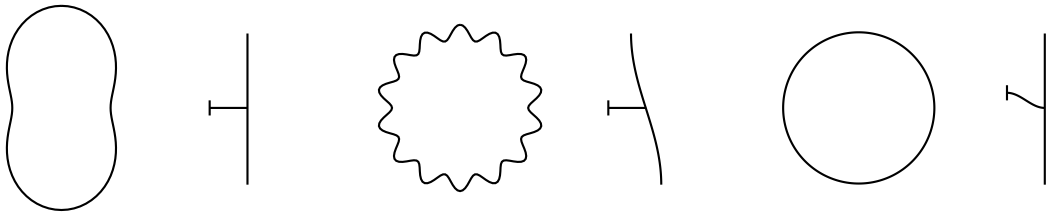
Variable		Value
R	(mm)	2,488
L_s	(mm)	600
e	(mm)	24
h_w	(mm)	156
e_w	(mm)	10
w_f	(mm)	120
e_f	(mm)	24

(b) Initial design

Figure 6.14: Single bay reference structure and initial design.

The linear elastic stability analysis of this ring-stiffened shell exhibits some typical buckling patterns. The three most critical kinds of buckling patterns are known as overall buckling, interframe buckling and frame tripping and they are illustrated in Figure 6.15. Actual structures exhibit some unavoidable shape imperfections due to the manufacturing process (mostly cold-bending- and welding-based) and heavy loads connected to the hull (*e.g.* the nuclear reactor). These initial imperfections may trigger buckling or premature plastic collapse at pressure far below those corresponding to elastic buckling, even if these imperfections are of moderate amplitude due to the stringent tolerances used in fabrication.

Predicting the collapse pressure for any given imperfect geometry is not straightforward though because the structure may feature a considerable degree of interaction between the aforementioned buckling modes. For solving the buckling problem at hand, the designer may resort to closed-form solutions or other semi-numerical methods available in the codes of practice (*e.g.* the BS5500). Another alternative that is investigated here consists in using an appropriate finite-element-based simulation.



(a) Overall mode (b) Interframe mode (c) Frame tripping mode

Figure 6.15: Most critical buckling patterns of a ring-stiffened shell cylinder.

6.4.1.2 Modelling of the shape imperfections

Note that the present analysis is restricted to the effects of overall and interframe shape imperfections. The frame tripping mode is avoided here by checking for conservative safety constraints (from the BS5500) regarding the proportions of the stiffener's web and flange during the optimization. The overall (resp. interframe) radial imperfection is given by:

$$\zeta_n(z,\theta) = A_n \cos(n\theta), \tag{6.22}$$

$$\zeta_m(z,\theta) = A_m \cos\left(\frac{\pi}{L_s} z\right) \cos(m\theta), \tag{6.23}$$

where n (resp. m) is the number of circumferential waves that typically ranges from 2 to 6 (resp. 10 to 20), and A_n (resp. A_m) denotes the amplitude of the radial imperfection. The origin of the z axis is located at the left end of the bay in its vertical position. In the present application, only two modes are considered: $n = 2$ and $m = 14$. The two modes correspond to the most critical buckling patterns of the initial design. A finer study would consist in considering a larger spectrum of imperfections depending on the design.

6.4.1.3 Nonlinear finite element model

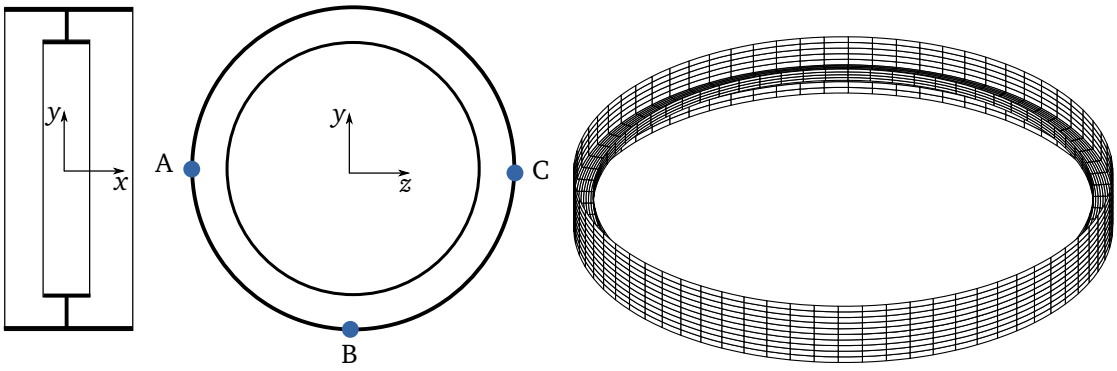
It is proposed to compute the collapse pressure by means of the asymptotic numerial method. Material and geometric nonlinearities are taken into account. The steel that constitutes the hull is assumed to follow a nonlinear elastic Ramberg-Osgood constitutive law as described in Section 6.2.4.4. Plasticity is neglected here as it can be reasonably assumed that there is no significant unloading anywhere in the structure until collapse is reached. Follower forces are taken into account for the hydrostatic pressure field p so that it is always exerted normally with respect to the deformed structure.

Rigid body modes are eliminated in three points as illustrated in Figure 6.16(a) in such a way that it does not modify the shape of the deformed structure, namely:

- in A, the three degrees of translation are set equal to zero;
- in B, the degree of translation along the x-axis is set equal to zero;
- in C, the degrees of translation along the x- and y-axes are set equal to zero;

The orthoradial rotations of the two circumferential ends are set equal to zero and the longitudinal displacements of each end are assigned equal to the longitudinal displacement of one node (for each end) in order to account for the presence of repeated adjacent bays. As a result, each circumferential end remain in a (moving) plane and the deformed plating is tangent to the normal of this plane. In addition, a membrane compressive stress of amplitude $p\,\pi R^2$ is exerted as indicated in Figure 6.14. This additional load emulates the background effect (*i.e.* the hydrostatic pressure exerted on the two ends of the cylindrical hull).

The structure is meshed with 1,540 Büchter-Ramm elements featuring about 40,000 degrees of freedom. The hull is meshed with 70×10 elements, 70×8 elements for the stiffener's web and 70×4 elements for the stiffener's flange. The collapse pressure was shown to stabilize for a coarser mesh featuring 15,000 degrees of freedom although it has been raised here in order to accurately represent the highest modal imperfection featuring 14 waves along the circumference (one wave being represented here with $70/14 = 5$ elements). Since the Büchter-Ramm element is quadratic, this yields a fair number of 11 nodes to represent one wave. Amplified superpositions of the two imperfections considered here are illustrated in Figure 6.17.

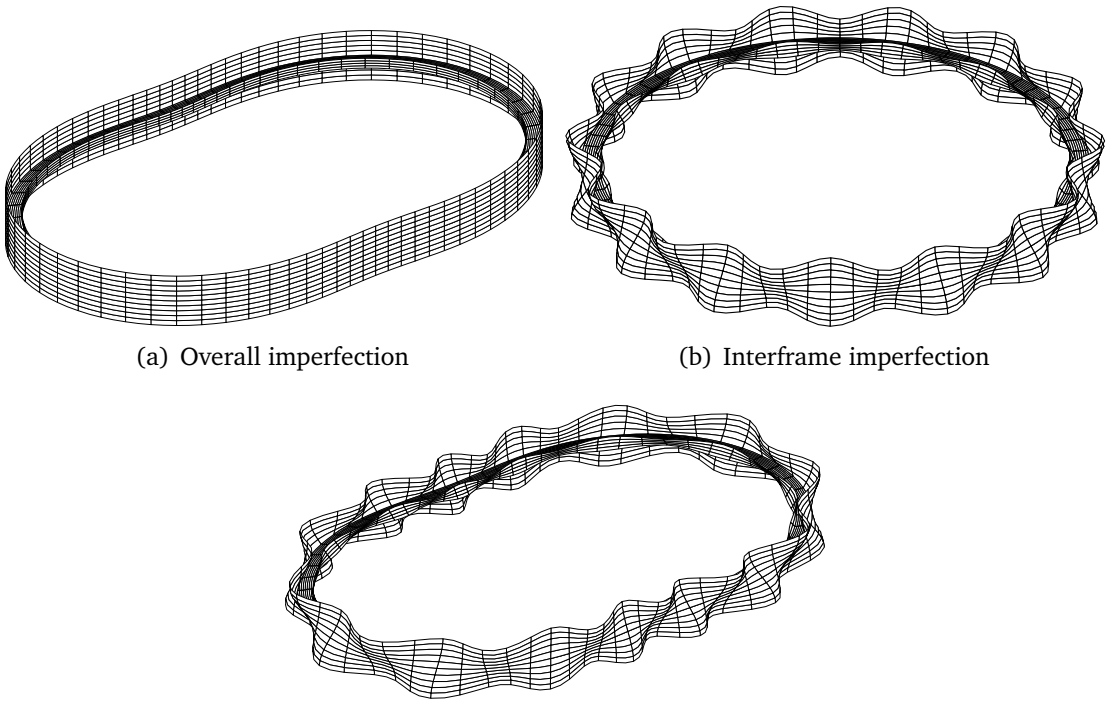


(a) Nodes for the elimination of rigid body moves (b) Perfect mesh

Figure 6.16: Finite element modelling of the ring-stiffened shell cylinder.

6.4.1.4 Semi-numerical model

In the sequel the designs obtained with the finite element model are compared with the ones based on approximate semi-numerical solutions available in the shell design codes of practice (see Dubourg et al., 2008, for a review). These approximations are not able to account for the possible interactions between the buckling modes. Indeed, the model for predicting the overall (resp. interframe) plastic collapse pressure $p_{n\,\mathrm{pl}}$ (resp. $p_{m\,\mathrm{pl}}$) depends on the amplitude of an overall (resp. interframe) imperfection A_n (resp. A_m) only. $p_{n\,\mathrm{pl}}$ is determined here according to the Bryant formula embedded in the BS5500, and $p_{m\,\mathrm{pl}}$ resorts to an interpolated table of finite element solutions derived by the *Krylov shipbuilding research institute* (KSRI).



(a) Overall imperfection (b) Interframe imperfection

(c) Both imperfections

Figure 6.17: Meshed ring-stiffened shell cylinder with amplified imperfections.

The final semi-numerical model yielding the plastic collapse pressure for the two modes is approximated as follows:

$$p_{\text{critical}}\left(A_n, A_m\right) = \min\left(p_{n\,\text{pl}}(A_n),\, p_{m\,\text{pl}}(A_m)\right). \tag{6.24}$$

6.4.2 Formulations of the design optimization problem

In this section two design philosophies are opposed. The first one resorts to the so-called *worst case approach* that consists in designing for an extreme configuration specified by the experts. The other one uses a more comprehensive probabilistic model and eventually falls under the RBDO formulation.

6.4.2.1 Objective and constraints

First, the objective of the design optimization is to find the set of parameters ruling the stiffness of the structure $\boldsymbol{d} = (e, h_w, e_w, w_f, e_f)^{\mathsf{T}}$ that minimizes the ratio between the structural weight and the weight of water displaced by the structure. The latter ratio reads as follows:

$$\mathfrak{c}(\boldsymbol{d}) = \frac{\rho_{\text{steel}}\, \mathscr{V}_{\text{steel}}(\boldsymbol{d})}{\rho_{\text{water}}\, \pi\, (R + e/2)^2\, L_s}. \tag{6.25}$$

The admissible design space is bounded with the following constraints.

(i) Since the semi-numerical model lacks consideration of the frame tripping mode, it is proposed to resort to the following conservative safety criteria prescribed in the BS5500:

$$h_w \leq 1.1\sqrt{\frac{E}{\sigma_y}}\,e_w, \tag{6.26}$$

$$w_f \leq \sqrt{\frac{E}{\sigma_y}}\,e_f. \tag{6.27}$$

These two constraints actually bound the slenderness ratios of the stiffener.

(ii) The stiffener's flange should not be too large with respect to the interframe distance in order to be able to place sufficiently large openings through the hull:

$$445 \text{ mm} \leq L_s - w_f. \tag{6.28}$$

(iii) The design space is bounded with the following reasonable boundaries:

$$\frac{p\,R}{\sigma_y} \leq e \leq 50 \text{ mm}, \tag{6.29}$$

$$w_f \leq h_w \leq 2\,w_f, \tag{6.30}$$

$$5 \text{ mm} \leq e_w \leq 25 \text{ mm}, \tag{6.31}$$

$$70 \text{ mm} \leq w_f \leq 150 \text{ mm}, \tag{6.32}$$

$$15 \text{ mm} \leq e_f \leq 50 \text{ mm}. \tag{6.33}$$

The first lower constraint on the hull thickness e means that the circumferential stress in the equivalent non-stiffened cylinder should not exceed the yield strength σ_y.

At last, the predictive models for the collapse pressure (namely the finite element model and the semi-numerical solutions) are used for guaranteeing that collapse does not occur at some prescribed diving depth I_{acc}. Therefore, it leads to the establishment of the following last constraint:

$$I_{\text{acc}}\,\rho_{\text{water}}\,g \leq p_{\text{critical}}(\boldsymbol{d}). \tag{6.34}$$

It is assumed that the present submarine is designed for an accidental diving depth I_{acc} of 250 m.

6.4.2.2 The worst case approach

The worst case approach basically consists in setting all the demand (resp. capacity) variables to their highest (resp. lowest) possible values and to find the optimal design for this

worst scenario. In the present context of shell design, this resorts to *(i)* prescribed maximum imperfection amplitudes and *(ii)* a destruction diving depth I_{des} that is significantly larger than the accidental diving depth I_{acc}.

Here, the maximum imperfection amplitudes are taken from the BS5500 recommendations and are set equal to $A_{2\,\max} = 5R/1,000$ and $A_{14\,\max} = L_s/100$. The destruction diving depth is arbitrarily fixed to 340 m.

6.4.2.3 The probabilistic approach

Arguing that the previous worst case approach introduces an unknown degree of conservatism, it is proposed to resort to a more comprehensive probabilistic model for describing the possible configurations of the hull. This probabilistic model is specified in Table 6.3.

Variable		Distribution	Mean	C.o.V.
E	(MPa)	Lognormal	200,000	5%
σ_y	(MPa)	Lognormal	390	5%
σ_u	(MPa)	Lognormal	570	3%
e	(mm)	Lognormal	μ_e	3%
h_w	(mm)	Lognormal	μ_{h_w}	3%
e_w	(mm)	Lognormal	μ_{e_w}	3%
w_f	(mm)	Lognormal	μ_{w_f}	3%
e_f	(mm)	Lognormal	μ_{e_f}	3%
A_2	(mm)	Lognormal	$\frac{1}{3}\frac{5R}{1,000}$	50%
A_{14}	(mm)	Lognormal	$\frac{1}{3}\frac{L_s}{100}$	50%

Table 6.3: Probabilistic model for the ring-stiffened shell cylinder.

Since no data is available, the probabilistic model for the material properties is built from the recommendations available in the JCSS probabilistic modelling code[2] (Vrouwenvelder, 1997). This code also prescribes a linear correlation between the yield strength σ_y and the ultimate stress σ_u in the form of a Pearson correlation coefficient $\rho = 0.75$, which is taken into account in the present analysis. The right-skewed probabilistic model for the amplitudes of the imperfections was built with an empirical coefficient of variation of 50% and the mean is such that the previous worst imperfections $A_{2\,\max}$ and $A_{14\,\max}$ matches the 99.5%-quantile of the present probabilistic model. This thus leads approximately to set the mean value equal to one third of the latter worst imperfection amplitudes as indicated in Table 6.3.

Given this probabilistic model, the original deterministic design optimization problem is transformed into a reliability-based design problem where safety is measured by means

[2] See http://www.jcss.ethz.ch/publications/publications_pmc.html.

of the following probability of failure:

$$p_f(\boldsymbol{d}) = \mathbb{P}\left[p_{\text{critical}}(\boldsymbol{d}, \boldsymbol{X}) \leq I_{\text{acc}}\, \rho_{\text{water}}\, g\right]. \tag{6.35}$$

where $\boldsymbol{X}$ is the random vector that collects all the random variables of the probabilistic model. The optimization is performed with respect to the means of the random design variables e, h_w, e_w, w_f and e_f. The probabilistic constraint reads as follows:

$$p_f(\boldsymbol{d}) \leq \Phi(-\beta_0), \tag{6.36}$$

where $\beta_0 = 6$ in the present application (*i.e.* $p_{f\,0} \leq 10^{-9}$).

6.4.2.4 Resolution strategies

The deterministic design optimization problem underlying the worst case approach is solved here by means of the Polak-He gradient-based optimizer (see Algorithm 4.1 in Chapter 4). It uses the two mechanical models for the buckling strength of the structure, namely the semi-numerical (SN) and the finite element (FE) models.

The reliability-based design optimization problem underlying the probabilistic approach is solved with the meta-model-based RBDO strategy. Again, two designs are computed with either of the mechanical models.

Once the four optimal designs are found, a reliability analysis is performed in order to compute the safety level of the optimally designed structures at both the accidental and the destruction diving depth using the probabilistic model of Table 6.3. Since the finite element model is expensive to evaluate, this resorts to the proposed meta-model-based importance sampling technique with a 5% target coefficient of variation on the failure probability. For the less expensive semi-numerical model, it is proposed to resort to direct subset sampling in order to compute the whole CDF of the critical pressure which yields a relationship between the failure probability and the diving depth in a single run for each design.

6.4.3 Results

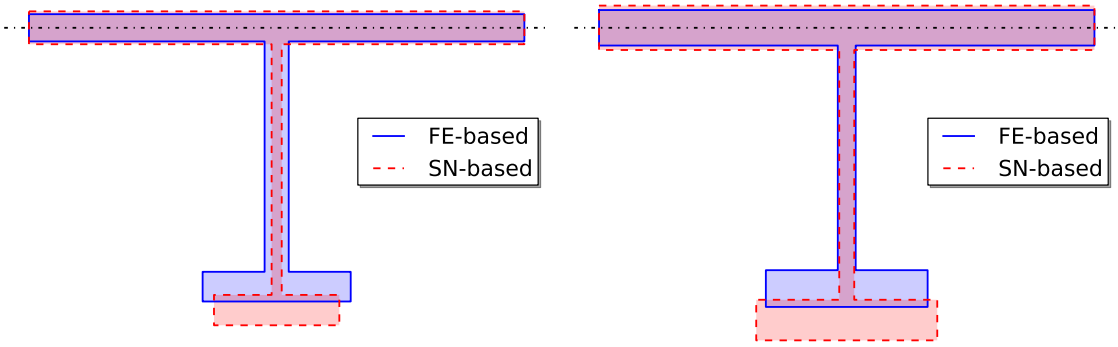
The results are given in Table 6.4 and the corresponding designs are illustrated in Figure 6.18. First, it should be noticed that the FE-based (finite element) design is always more cost-optimal than its SN (semi-numerical) counterpart. Actually, this confirm the initial intuition as the semi-numerical solutions involve a set of built-in safety factors that eventually lead to an important (although *unknown*) degree of conservatism. In the worst case approach, the gain in using a finite element model is only 4%, whereas it reaches 17% in the RBDO approach.

It should also be noticed that the SN-based design always features a more slender stiffener web than the FE-based designs. This is because the SN-solution lacks an explicit consideration of the frame tripping buckling mode. This lack is such that in the deterministic worst case approach, the BS5500 safety constraint regarding this mode is active at the optimal design (see Eq. (6.26)). Indeed, in this case, the stiffener web is clearly too slender as illustrated in Figure 6.18(a).

Method		Worst case approach		RBDO ($\beta = 6$)	
		FE-based	SN-based	FE-based	SN-based
e	(mm)	21.99	26.56	28.65	35.85
h_w	(mm)	186.01	[a]202.38	181.37	201.66
e_w	(mm)	19.47	[a]8.14	14.44	12.11
w_f	(mm)	119.57	101.22	130.62	146.18
e_f	(mm)	23.97	24.53	29.68	32.77
Cost	(%)	19.60	20.04	23.56	28.47
$\beta(I_{\text{acc}})$		4.99	3.81	6.06	6.11
$\beta(I_{\text{des}})$		1.40	2.00	4.42	4.99

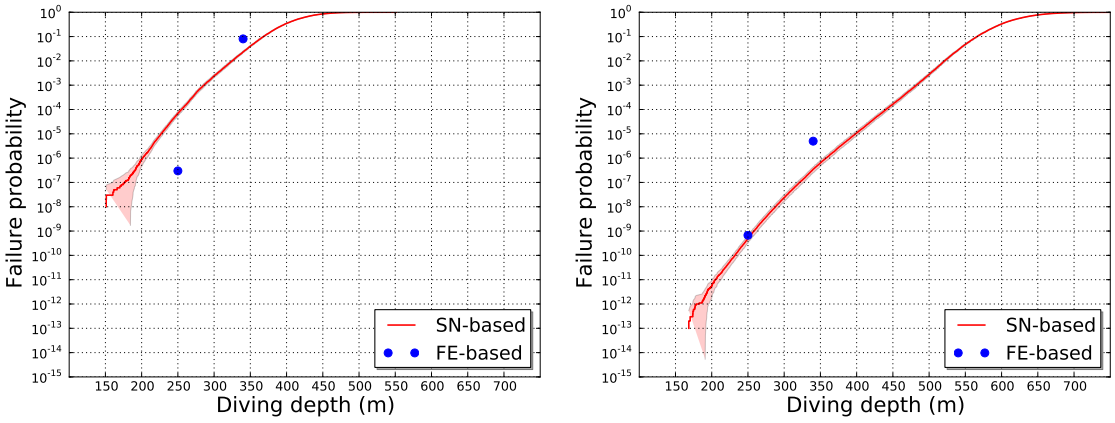
[a]The frame tripping safety constraint is active.

Table 6.4: Results for the design optimization of the imperfect ring-stiffened shell cylinder.



(a) Worst case optimal designs (b) RBDO ($\beta = 6$) optimal designs

Figure 6.18: Comparison of the optimal designs for the imperfect ring-stiffened shell cylinder.



(a) Worst case optimal designs (b) RBDO ($\beta = 6$) optimal designs

Figure 6.19: Relation between the diving depth and the failure probability for the imperfect ring-stiffened shell cylinder.

As expected, the worst case approach offers a significant degree of safety at the accidental diving depth and it even remains a little margin at the destruction diving depth although the failure probability is much greater there ($p_f = \Phi(-\beta) \approx 10^{-2}$). The probabilistic approach enables an explicit control of the safety level at the accidental diving depth. Due to the important targeted level of safety ($p_f < 10^{-9}$), the reliability-based optimal designs are of course less optimal than their worst case counterparts.

The relationship between the diving depth and the failure probability is illustrated in Figure 6.19. The subset sampling technique applied with the semi-numerical model enables a reconstruction of the full CDF. The meta-model-based importance sampling applied on the expensive-to-evaluate finite-element model only yields the failure probability estimates at the two diving depths of interest. It can be seen from Figure 6.19(b) that the failure probability matches the maximum tolerance set here equal to $p_f = \Phi(-6) < 10^{-9}$.

6.4.3.1 Computational details

Convergence of the meta-model-based RBDO strategy is obtained within 850 calls to the buckling strength models. Note that it is of utmost importance for the FE-based application due to the important numerical effort required by a single finite element analysis (about 10 minutes of CPU time). For the sake of completeness, the convergence of the algorithm is depicted in Figure 6.20 and Figure 6.21 for each buckling strength model.

The complementary meta-model-based importance sampling analyses on the FE-based designs revealed that the Kriging surrogates accurately fit the limit-state surfaces as the correction factor is always close to unity. Hence, the required coefficient of variation of 5% is obtained within a few hundred calls to the finite element model.

6.4.4 Conclusion

This section applied the reliability-based design philosophy to the design of an imperfect submarine pressure hull prone to buckling.

First, the safety margin associated with the current worst case design methodology has been quantified in the form of a failure probability. It reveals that this common practice yields a significant level of safety although it is not truly mastered.

Second, in order to address this latter remark it is proposed to explicitly account for the uncertainties in the optimization problem. This eventually falls under the so-called RBDO formulation which is commonly identified to be too computationally demanding for being applied to industrial problems. In this context, the proposed meta-model-based RBDO strategy truly reveals interesting to come up with a solution within less than a thousand runs of the finite element code.

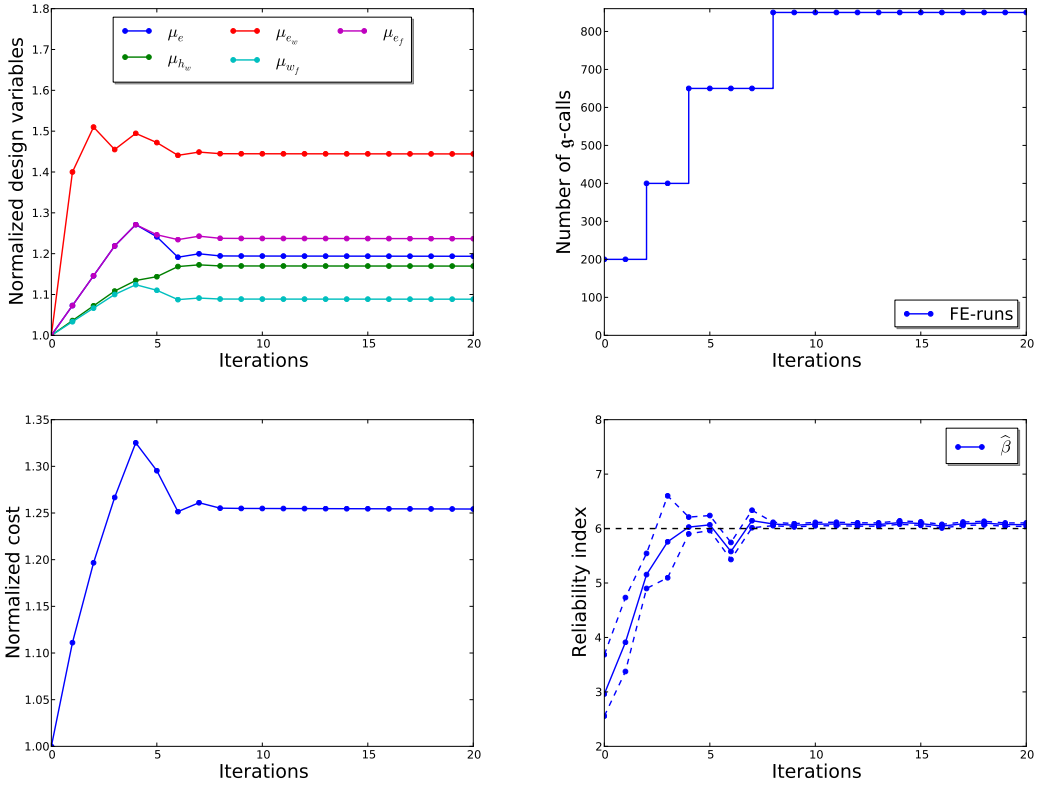


Figure 6.20: Convergence of the meta-model-based RBDO algorithm for the FE-based analysis.

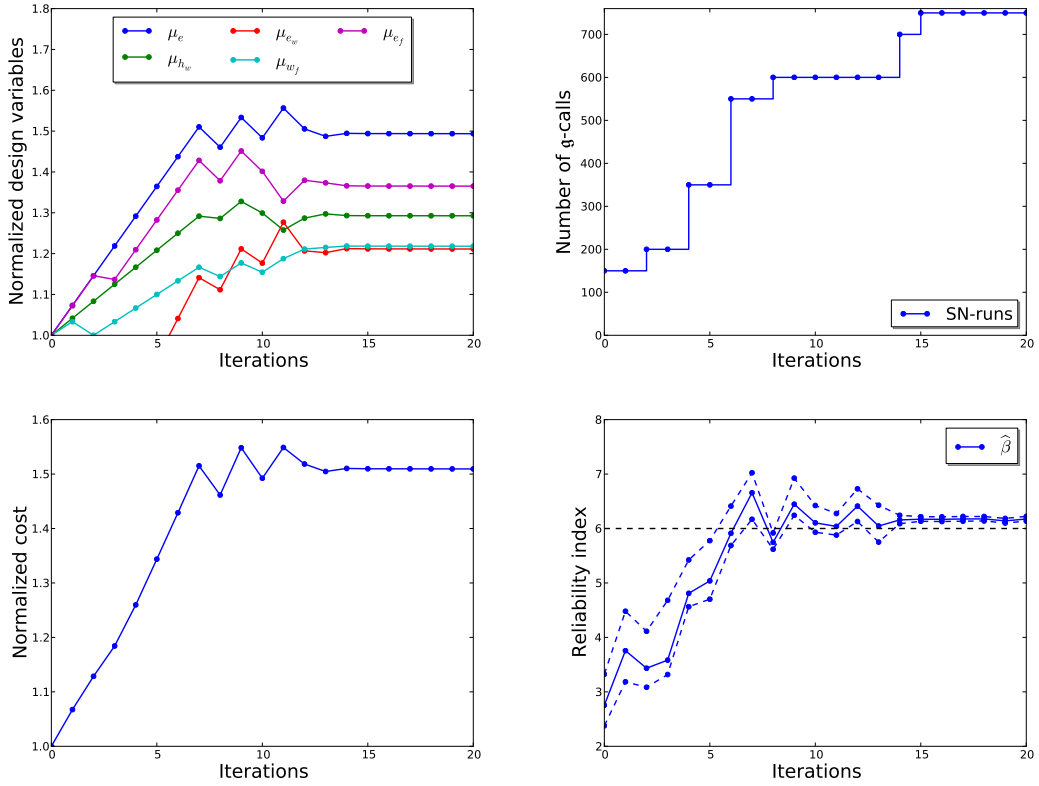


Figure 6.21: Convergence of the meta-model-based RBDO algorithm for the SN-based analysis.

Conclusion

Summary and main contributions

The work presented in this thesis was aimed at solving the reliability-based design problem when the limit-state function involves an expensive-to-evaluate numerical model. This seems to be achieved in the light of the examples presented in Chapters 5 and 6.

Probabilistic surrogate modelling

To achieve this objective it has been decided to resort to *adaptive surrogates* for the limit-state function. In Chapter 1, the Gaussian process (or Kriging) prediction methodology has revealed a useful tool for performing such a substitution. Indeed, it provides a local measure of its own accuracy in the convenient form of a probabilistic measure that allows multiple post-processing operations such as:

(i) the formulation of confidence intervals on the prediction;

(ii) probabilistic classification with respect to a specific contour (*e.g.* a limit-state surface);

(iii) the adaptive refinement of an initial coarse meta-model for different purposes such as global optimization or contour approximation.

Sampling-clustering-based adaptive designs of experiments

Building on a quite exhaustive review of the existing techniques for building adaptive designs of experiments with a Kriging meta-model, an alternative strategy is proposed in Chapter 2. It focuses on the use of the so-called refinement criteria rather than on the development of new ones as typically addressed in the literature. More specifically, it is proposed to interpret the refinement criteria as probability density functions for the improvement points. This vision naturally leads to a sampling-based approach. Sampling of this criteria resorts to Markov chain Monte Carlo samplers such as the slice sampler which revealed robust with respect to the high multi-modality of the contour refinement criteria.

As expected, the sampled population mainly occupies the vicinity of the region of interest (*i.e.* the limit-state surface). Clustering is then employed on purpose to find a subset of points that best represents the latter population. This sampling-clustering approach solves the problem of multiple best improvement points and enables the use of a distributed computing platform.

Meta-model-based importance sampling for reliability analysis

The quantification of failure probabilities is a key step of the reliability-based design methodology. Hence, Chapter 3 has reviewed a subset of probability estimation techniques available in the literature along with comments on their pros and cons. It was concluded that the subset sampling approach offers the best trade-off between versatility and computational effort (with respect to the order of magnitude of the failure probability in particular). Nonetheless, it remains inapplicable when the limit-state function involves an expensive-to-evaluate numerical model.

For this reason, a whole section of Chapter 3 has been devoted to the presentation of the so-called *surrogate-based reliability analyses* that became quite popular in the last decade. They basically consists in using a meta-model instead of the original model. However, it was pointed out that this substitution unavoidably introduces a bias in the failure probability estimate. Thus, in order to *apprehend* this bias, two different metrics have been presented. The *augmented failure probability* is defined as the sum of the aleatory uncertainty in the probabilistic model and the epistemic uncertainty induced by the Kriging prediction of the failure domain. The second metric is a *pseudo-confidence interval* on the failure probability whose lower (resp. upper) bound is computed on the lower (resp. upper) confidence bound of the Kriging-predicted limit-state surface. None of these two metrics constitutes a real measure of the bias though. They only provide an intuition.

Hence, an hybrid strategy is proposed in an attempt to quantify the actual bias. This strategy uses the probabilistic classification function from a Gaussian process predictor for approximating the optimal instrumental PDF in an importance sampling scheme. After basic algebra, it turned out that the failure probability recasts as the product of the augmented failure probability and a correction factor. This *correction factor* measures the bias of interest. The estimation of this correction factor ineluctably resorts to cross-sampling between the Kriging meta-model and the original model. However, it is worth remembering that this sampling is optimal when the Kriging meta-model is accurate, so that it is not necessarily less efficient than the usual plug-in estimators (see the examples in Chapter 5).

Reliability-based design optimization using Kriging surrogates and subset sampling

Chapter 4 eventually tackled the main topic of this research work which is concerned with the efficient resolution of the reliability-based design optimization problem. It is presented as a classical design optimization problem for which the hard inequality constraints are

simply replaced by their probabilistic counterparts. Hence, a first part of Chapter 4 concentrates on a few elements of inequality constrained optimization and eventually come up with the need for the gradient of the failure probability with respect to the design variables.

The estimation of this quantity is commonly identified as a reliability sensitivity analysis in the associated literature. The strategy that is used here takes advantage of the specific formulation of the RBDO problem considered in this thesis where the design variables are exclusively involved in the definition of the probabilistic model (*i.e.* they are means of random variables). Indeed, this enables the use of the so-called *score function approach* which is shown to yield the gradient of interest as a by-product of a reliability analysis. In addition, it can be applied to crude Monte Carlo, subset and importance sampling.

Then, in order to make the nested surrogate-based reliability analyses more efficient, it is proposed to construct the Kriging predictors of the limit-state surfaces in a global space. This space is referred to as the *augmented reliability space* and it is obtained by considering that the design parameters are uniformly distributed over the admissible design space. As a result, this augments the uncertainty in the probabilistic model and the range on which the Kriging meta-model should be accurate.

The final RBDO algorithm resorts to subset sampling on the Kriging surrogates for the nested reliability and reliability sensitivity analyses. It may ask for a refinement of the Kriging surrogates based on the spread of the pseudo-confidence interval on the failure probability previously introduced in Chapter 3. This somewhat empirical metric revealed appropriate for the RBDO examples that have been dealt with in Chapters 5 and 6.

Probabilistic design of imperfect shells prone to buckling

The reliability analysis and reliability-based design of imperfect shells prone to buckling is then provided so as to illustrate the applicability of the proposed approaches. First, it has been shown that the prediction of the critical buckling load involves a significant number of fine assumptions that makes the resulting numerical model expensive-to-evaluate. Besides, the probabilistic representation of the spatially varying random shell properties requires a fairly large number of random variables. In this context, the application of meta-model-based strategies is quite challenging.

The meta-model-based importance sampling technique has been able to yield an unbiased estimate of the failure probability even though the underlying Kriging predictor does not accurately fit the actual limit-state surface. This example truly reveals the interest one should have in such hybrid approaches when using meta-models for such challenging reliability analyses.

The meta-model-based RBDO strategy has been applied to the design of an imperfect submarine pressure hull and it was able to provide an optimal design within less than a thousand runs of the finite element model. A complementary meta-model-based importance sampling analysis at the optimal design showed that the Kriging surrogates accurately fit the actual limit-state surface.

Future work

The use of adaptive surrogate models is inescapable at some point when decision-making analyses such as risk analysis or design optimization are based on expensive-to-evaluate numerical models. Even though this thesis contributed to the resolution of the specific problem of reliability-based design optimization, some elementary tools and concepts might be reused to address a broader range of problems.

RBDO using meta-model-based importance sampling

First, it has been admitted that the proposed surrogate-based RBDO strategy may lose efficiency when both the number of random variables is greater than 20 and the limit-state function features a high degree of nonlinearity. For such cases, hybrid approaches such as *meta-model-based importance sampling* could be used for estimating the failure probability and its gradient within the optimization loop. Complementary studies are required though in order to choose *(i)* when to use meta-model-based importance sampling (*i.e.* for computing the direction and (or not) during the approximate line-search) and *(ii)* how to recycle the design of experiments.

Application to a broader range of RBDO problems

Second, the design optimization problems considered in this thesis exclusively aimed at tuning the means of the random design variables. Further investigations are required in order to apply the proposed methodology to the optimal tolerancing of manufactured products. This problem essentially consists in balancing products quality with cost-effective production capabilities by specifying adequate conformity tolerances (*i.e.* by tuning the variances of the random design variables). Such an application obviously requires the development of real cost models (functions of the so-called capability coefficients) as well as an extended test of the proposed surrogate-based strategy.

Sampling-clustering-based efficient global optimization

The sampling-clustering approach to the construction of adaptive designs of experiments has only been applied to the contour refinement problem. It is the author's belief that it can easily be extended to the resolution of both constrained and unconstrained global optimization problems. Indeed, the positive-valued expected improvement criterion featured in the efficient global optimization algorithm can also be interpreted as a probability density function for the location of the minimizer of interest. Hence, it can potentially be sampled from.

This approach would additionally allow one to tackle the convergence criterion problem for the EGO algorithm. Indeed, scrutinizing the maximum value of the expected improvement as originally proposed by Jones et al. (1998) does not reveal particularly appropriate

as the search for this maximum might be mislead due to the significant number of modes featured by this criterion. It could instead be proposed to look at the spread of the candidate population at the sampling stage to compute confidence intervals on the location of the minimizer. Convergence would then be based on a comparison of the latter spread to the width of the initial admissible space.

In addition, such a sampling-based approach enables the use of informative prior densities for the location of the minimizer as weighting densities of the expected improvement. This would ultimately turn the optimization problem into a fully Bayesian inverse problem.

Bayesian updating featuring expensive-to-evaluate likelihood functions

Besides decision-making problems, inverse parametric identification has also gained a large interest in the past decade. The purpose of such problems is to find the environmental conditions in which a system has evolved based on indirect measurements of its performance (*e.g.* find the initial geometry of a structure given its deformed shape and an uncertain description of its load). This typically resorts to a numerical model that is able to predict the measured data and which is a function of the sought environmental conditions. Then the so-called Bayesian updating formalism consists in inverting the model on purpose to find the distribution of its input parameters given that of its output. The output distribution is based on the data measurements and is referred to as the *likelihood function*.

Such an inversion typically resorts to Markov chain Monte Carlo samplers which are known to be rather computationally demanding due to their possibly long tuning period and high rejection rates (Perrin, 2008; Berveiller et al., 2011). Then the idea already followed by Taflanidis and Cheung (2011) would consist in replacing the expensive-to-evaluate likelihood function by a meta-model. This research work obviously requires the definition of new metrics for assessing and reducing the loss of accuracy induced by this substitution.

Bibliography

Abdo, T. and R. Rackwitz (1990). A new β-point algorithm for large time invariant and time-variant reliability problems. In Der Kiureghian, A. and P. Thoft-Christensen (Eds.), *Proc. 3rd WG 7.5 IFIP Conference on Reliability and Optimization of Structural Systems, Berkeley, California, USA – Lecture notes in Engineering, Vol. 61*. Springer.

Abramowitz, M. and I. Stegun (1970). *Handbook of mathematical functions*. Dover Publications, Inc.

Aizerman, M., E. Braverman, and L. Rozonoer (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25, 821–837.

Allen, D. (1971). The prediction sum of squares as a criterion for selecting prediction variables. Technical Report 23, Dept. of Statistics, University of Kentucky.

Andrieu, C., N. De Freitas, A. Doucet, and M. Jordan (2003). An introduction to MCMC for machine learning. *Machine learning*, 50(1–2), 5–43.

Aoues, Y. and A. Chateauneuf (2010). Benchmark study of numerical methods for reliability-based design optimization. *Struct. Multidisc. Optim.*, 41(2), 277–294.

Arbocz, J. and C. Babcock (1969). The effect of general imperfections on the buckling of cylindrical shells. *J. Applied Mech.*, 36, 28–38.

Arthur, D. and S. Vassilvitskii (2007). K-means++: the advantages of careful seeding. In *Proc. 18th annual ACM-SIAM symposium on discrete algorithms*, pp. 1027–1035.

Au, S.-K. (2005). Reliability-based design sensitivity by efficient simulation. *Computers & Structures*, 83(14), 1048–1061.

Au, S.-K. and J. Beck (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Prob. Eng. Mech.*, 16(4), 263–277.

Au, S.-K., J. Ching, and J. Beck (2007). Application of subset simulation methods to reliability benchmark problems. *Structural Safety*, 29(3), 183–193.

Auder, B. (2011). *Classification et modélisation de sorties fonctionnelles de codes de calcul*. Ph. D. thesis, Université Pierre et Marie Curie - Paris 6.

Baguet, S. (2001). *Stabilité des structures minces et sensibilité aux imperfections par la Méthode Asymptotique Numérique*. Ph. D. thesis, Université d'Aix-Marseille II.

Barbillon, P. (2010). *Méthodes d'interpolation à noyaux pour l'approximation de fonctions type boîte noire coûteuses*. Ph. D. thesis, Université Paris-Sud 11.

Basudhar, A. and S. Missoum (2008). Adaptive explicit decision functions for probabilistic design and optimization using support vector machines. *Computers & Structures*, 86(19-20), 1904–1917.

Basudhar, A. and S. Missoum (2009). A sampling-based approach for probabilistic design with random fields. *Comput. Methods Appl. Mech. Engrg.*, 198(47-48), 3647–3655.

Basudhar, A. and S. Missoum (2010). An improved adaptive sampling scheme for the construction of explicit boundaries. *Struct. Multidisc. Optim.*, 42, 517–529.

Bect, J., D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez (2011). Sequential design of computer experiments for the estimation of a probability of failure. *Stat. Comput.*, In press.

Berveiller, M., Y. Le Pape, B. Sudret, and F. Perrin (2011). Updating the long-term creep strains in concrete containment vessels by using MCMC simulation and polynomial chaos expansions. *Struct. Infrastruct. Eng*, 7. In press.

Berveiller, M., B. Sudret, and M. Lemaire (2006). Stochastic finite elements: a non intrusive approach by regression. *Eur. J. Comput. Mech.*, 15(1-3), 81–92.

Bichon, B. (2010). *Efficient surrogate modeling for reliability analysis and design*. Ph. D. thesis, Faculty of the Graduate School of Vanderbilt University., Nashville, Tennessee.

Bichon, B., M. Eldred, L. Swiler, S. Mahadevan, and J. McFarland (2008). Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA Journal*, 46(10), 2459–2468.

Bichon, B., S. Mahadevan, and M. Eldred (2009). Reliability-based design optimization using efficient global reliability analysis. In *50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, volume 47.

Bjerager, P. (1988). Probability integration by directional simulation. *J. Eng. Mech.*, 114(8), 1285–1302.

Bjerager, P. and S. Krenk (1989). Parametric sensitivity in first order reliability theory. *J. Eng. Mech.*, 115(7), 1577–1582.

Blatman, G. (2009). *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*. Ph. D. thesis, Université Blaise Pascal - Clermont II.

Blatman, G. and B. Sudret (2008a). Adaptive sparse polynomial chaos expansions - application to structural reliability. In *Proc. 4th Int. ASRANet Colloquium, Athens, Greece*.

Blatman, G. and B. Sudret (2008b). Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus Mécanique*, 336(6), 518–523.

Blatman, G. and B. Sudret (2010a). An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Prob. Eng. Mech.*, 25(2), 183–197.

Blatman, G. and B. Sudret (2010b). Reliability analysis of a pressurized water reactor vessel using sparse polynomial chaos expansions. In Straub, D., L. Esteva, and M. Faber (Eds.), *Proc. 15th IFIP WG7.5 Conference on Reliability and Optimization of Structural Systems, Munich, Germany*, pp. 9–16. CRC Press.

Blatman, G., B. Sudret, and M. Berveiller (2007). Quasi-random numbers in stochastic finite element analysis. *Mécanique & Industries*, 8, 289–297.

Bourinet, J.-M., F. Deheeger, and M. Lemaire (2011). Assessing small failure probabilities by combined subset simulation and support vector machines. *Structural Safety*, In press.

Bourinet, J.-M., N. Gayton, M. Lemaire, and A. Combescure (2000). Reliability analysis of the stability of shells based on combined finite element and response surface methods. In *Proc. 4th Int. Coll. on Comput. of Shell & Spat. Struct. (IASS-IACM 2000)*.

Bourinet, J.-M. and M. Lemaire (2008). FORM sensitivities to correlation, application to fatigue crack propagation based on Virkler data. In *Proc. 4th Int. ASRANet Colloquium, Athens, Greece*.

Bourinet, J.-M., C. Mattrand, and V. Dubourg (2009). A review of recent features and improvements added to FERUM software. In *Proc. 10th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2009), Osaka, Japan*.

Breitung, K. (1984). Asymptotic approximation for multinormal integrals. *J. Eng. Mech.*, 110(3), 357–366.

BS5500 (1997). Unfired fusion-welded pressure vessels. British Standard Institutions.

Bucher, C. (2009). Asymptotic sampling for high-dimensional reliability analysis. *Prob. Eng. Mech.*, 24(4), 504–510.

Bucher, C. and U. Bourgund (1990). A fast and efficient response surface approach for structural reliability problems. *Structural Safety*, 7(1), 57–66.

Büchter, N., E. Ramm, and D. Roehl (1994). Three-dimensional extension of non-linear shell formulation based on the enhanced assumed strain concept. *Int. J. Num. Meth. Eng.*, 37(15), 2551–2568.

Cambier, S., P. Guihot, and G. Coffignal (2002). Computational methods for accounting of structural uncertainties, applications to dynamic behavior prediction of piping systems. *Structural Safety*, 24(1), 29–50.

Cannamela, C., J. Garnier, and B. Iooss (2008). Controlled stratification for quantile estimation. *Annals of Applied Statistics*, 2(4), 1554–1580.

Cérou, F. and A. Guyader (2007). Adaptive multilevel splitting for rare event analysis. *Stochastic analysis and applications*, 25(2), 417–443.

Charpentier, A. (2006). *Dependence structures and limiting results, with applications in finance and insurance*. Ph. D. thesis, Katolieke Universiteit Leuven.

Chateauneuf, A. and Y. Aoues (2008). *Structural design optimization considering uncertainties*, chapter 9, pp. 217–246. Taylor & Francis.

Chilès, J.-P. and P. Delfiner (1999). *Geostatistics: modeling spatial uncertainty*. Wiley, New York.

Ching, J., S.-K. Au, and J. Beck (2005a). Reliability estimation for dynamical systems subject to stochastic excitation using subset simulation with splitting. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16), 1557–1579.

Ching, J., J. Beck, and S.-K. Au (2005b). Hybrid subset simulation method for reliability estimation of dynamical systems subject to stochastic excitation. *Prob. Eng. Mech.*, 20(3), 199–214.

Cochelin, B. (1994). A path-following technique via an asymptotic-numerical method. *Computers & Structures*, 53, 1181–1192.

Cochelin, B., N. Damil, and M. Potier-Ferry (1994). The asymptotic-numerical method: an efficient perturbation technique for non-linear structural mechanics. *Rev. Eur. Elts Finis*, 3(2), 281–297.

Coelho, R., J. Lebon, and P. Bouillard (2011). Hierarchical stochastic metamodels based on moving least squares and polynomial chaos expansion – application to the multiobjective reliability-based optimization of space truss structures. *Struct. Multidisc. Optim.*, 43, 707–729.

Cressie, N. (1993). *Statistics for spatial data, revised edition*. John Wiley & Sons Inc.

Cumbus, C., P. Damien, and S. Walker (1996). Uniform sampling in the hypersphere via latent variables and the Gibbs sampler. Technical report, University of Michigan Business School.

Damil, N. and M. Potier-Ferry (1990). A new method to compute perturbed bifurcations: application to the buckling of imperfect elastic structures. *Int. J. Eng. Sci.*, 28, 943–957.

Das, P.-K. and Y. Zheng (2000). Cumulative formation of response surface and its use in reliability analysis. *Prob. Eng. Mech.*, 15(4), 309–315.

Deheeger, F. (2008). *Couplage mécano-fiabiliste, ^{2}SMART méthodologie d'apprentissage stochastique en fiabilité*. Ph. D. thesis, Université Blaise Pascal - Clermont II.

Deheeger, F. and M. Lemaire (2007). Support vector machine for efficient subset simulations: ^{2}SMART method. In *Proc. 10th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP10), Tokyo, Japan*.

Der Kiureghian, A. and T. Dakessian (1998). Multiple design points in first and second-order reliability. *Structural Safety*, 20(1), 37–49.

Der Kiureghian, A. and M. De Stefano (1990). An efficient algorithm for second-order reliability analysis. Technical Report UCB/SEMM-90/20, Dept of Civil and Environmental Engineering, University of California, Berkeley.

Der Kiureghian, A. and O. Ditlevsen (2009). Aleatory or epistemic? does it matter? *Structural Safety*, 31(2), 105–112.

Der Kiureghian, A., H. Lin, and S. Hwang (1987). Second order reliability approximations. *J. Eng. Mech.*, 113(8), 1208–1225.

Desceliers, C., C. Soize, and R. Ghanem (2007). Identification of chaos representations of elastic properties of random media using experimental vibration tests. *Comput. Mech.*, 39(6), 831–838.

Ditlevsen, O. (1979a). Generalized second moment reliability index. *J. Struct. Mech.*, 7(4), 435–451.

Ditlevsen, O. (1979b). Narrow reliability bounds for structural systems. *J. Struct. Mech.*, 7(4), 453–472.

Ditlevsen, O. and H. Madsen (1996). *Structural reliability methods*. J. Wiley and Sons, Chichester.

Du, X. and W. Chen (2004). Sequential optimization and reliability assessment method for efficient probabilistic design. *J. Mech. Des.*, 126, 225–233.

Dubourg, V., J.-M. Bourinet, and B. Sudret (2009a). Analyse fiabiliste du flambage des coques avec prise en compte du caractère aléatoire et de la variabilité spatiale des défauts de forme et d'épaisseur, et des propriétés matériaux. In *Proc. 19^e Congrès Français de Mécanique (CFM19), Marseille*.

Dubourg, V., J.-M. Bourinet, and B. Sudret (2010a). A hierarchical surrogate-based strategy for reliability-based design optimization. In Straub, D., L. Esteva, and M. Faber (Eds.), *Proc. 15th IFIP WG7.5 Conference on Reliability and Optimization of Structural Systems, Munich, Germany*, pp. 53–60. CRC Press.

Dubourg, V., J.-M. Bourinet, and B. Sudret (2010b). Reliability based design optimization using hierarchical Gaussian processes surrogates. In *Proc. 4th European Congress on Computational Mechanics (ECCM4), Paris, France*.

Dubourg, V., J.-M. Bourinet, B. Sudret, and M. Cazuguel (2011a). Reliability-based design optimization of an imperfect submarine pressure hull. In Faber, M., J. Köhler, and K. Nishijima (Eds.), *Proc. 11th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP11), Zurich, Switzerland*.

Dubourg, V. and F. Deheeger (2010). Une alternative à la substitution pour les métamodèles en analyse de fiabilité. In *Proc. 6^e Journées Fiabilité des Matériaux et des Structures, Toulouse*.

Dubourg, V., F. Deheeger, and B. Sudret (2011b). Metamodel-based importance sampling for structural reliability analysis. *Submitted to Prob. Eng. Mech.*

Dubourg, V., F. Deheeger, and B. Sudret (2011c). Metamodel-based importance sampling for the simulation of rare events. In Faber, M., J. Köhler, and K. Nishijima (Eds.), *Proc. 11th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP11), Zurich, Switzerland*.

Dubourg, V., C. Noirfalise, and J.-M. Bourinet (2008). Reliability-based design optimization: An application to the buckling of imperfect shells. In *Proc. 4th Int. ASRANet Colloquium, Athens, Greece*.

Dubourg, V., C. Noirfalise, J.-M. Bourinet, and M. Fogli (2009b). FE-based reliability analysis of the buckling of shells with random shape, material and thickness imperfections. In Furuta, H., D. Frangopol, and M. Shinozuka (Eds.), *Proc. 10th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2009), Osaka, Japan*.

Dubourg, V., B. Sudret, and J.-M. Bourinet (2011d). Reliability-based design optimization using kriging surrogates and subset simulation. *Struct. Multidisc. Optim.*, 44(5), 673–690.

Dubourg, V., B. Sudret, J.-M. Bourinet, and M. Cazuguel (2011e). Optimisation sous contrainte de fiabilité d'une structure en treillis. In *Proc. 10^e Colloque National en Calcul des Structures, Giens*.

Dubourg, V., B. Sudret, and M. Cazuguel (2011f). Modélisation probabiliste de champs d'imperfections géométriques de coques résistantes de sous-marins. In *Proc. 10^e Colloque National en Calcul des Structures, Giens*.

Dubrule, O. (1983). Cross validation of kriging in a unique neighborhood. *Mathematical Geology*, 15, 687–699.

Echard, B., N. Gayton, and M. Lemaire (2011). AK-MCS: an active learning reliability method combining kriging and Monte Carlo simulation. *Structural Safety*, 33(2), 145–154.

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2004). Least Angle Regression. *Annals of Statistics*, 32, 407–499.

Eldred, M. (2009). Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design. In *Proc. 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*.

Eldred, M., C. Webster, and P. Constantine (2008). Design under uncertainty employing stochastic expansion methods. In *Proc. 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. Victoria, British Columbia, Canada.

Eldred, M. S., A. A. Giunta, S. F. Wojtkiewicz Jr, and T. G. Trucano (2002). Formulations for surrogate-based optimization under uncertainty. In *Proc. 9th AIAA/ISSMO Symposium and Exhibit on Multidisciplinary Analysis and Optimization*, Atlanta (GA), United States.

Enevoldsen, I. and J. D. Sørensen (1994). Reliability-based optimization in structural engineering. *Structural Safety*, 15(3), 169–196.

Euler, L. (1744). *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate*, chapter Appendix: De Curvis Elasticis. Bousquet, M.M., Lausanne et Genève.

Faulkner, D. and P. Das (1990). Application of reliability theory to structural design and assessment of submarines and other externally pressurized cylindrical structures. In *Proc. 4th Int. Symp. on Integrity of Offshore Structures, Glasgow, UK*, pp. 199–230.

Franco, J. (2008). *Planification d'expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*. Ph. D. thesis, École Nationale Supérieure des Mines de Saint-Étienne.

Garnier, J. and P. Del Moral (2006). Simulations of rare events in fiber optics by interacting particle systems. *Opt. Commun.*, 267(1), 205–214.

Garnier, J., A. Omrane, and Y. Rouchdy (2009). Asymptotic formulas for the derivatives of probability functions and their Monte Carlo estimations. *Eur. J. Oper. Res.*, 198, 848–858.

Gayton, N., J.-M. Bourinet, and M. Lemaire (2003). CQ2RS, a new statistical approach to the response surface method for reliability analysis. *Structural Safety*, 25, 99–121.

Genz, A. (1992). Numerical computation of the multivariate normal probabilities. *J. Comp. Graph. Stat.*, 1, 141–150.

Genz, A. (1993). Comparison of methods for the computation of multivariate normal probabilities. *Computing Science and Statistics*, 25, 400–405.

Ghanem, R. and A. Doostan (2006). On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data. *J. Comput. Phys.*, 217, 63–81.

Ghanem, R. and P. Spanos (2003). *Stochastic Finite Elements : A Spectral Approach*. Courier Dover Publications.

Ginsbourger, D. (2009). *Multiples métamodèles pour l'approximation et l'optimisation de fonctions numériques multivariables*. Ph. D. thesis, École Nationale Supérieure des Mines de Saint-Étienne.

Ginsbourger, D., R. Le Riche, and L. Carraro (2010). Kriging is well-suited to parallelize optimization. In Hiot, L., Y. Ong, Y. Tenne, and C.-K. Goh (Eds.), *Computational Intelligence in Expensive Optimization Problems*, volume 2 of *Adaptation, Learning, and Optimization*, pp. 131–162. Springer Berlin Heidelberg.

Glasserman, P., P. Heidelberger, P. Shahabuddin, and T. Zajic (1999). Multilevel splitting for estimating rare event probabilities. *Operations Research*, 47(4), 585–600.

Grigoriu, M. (1998). Simulation of non-Gaussian translation processes. *J. Eng. Mech.*, 124(2), 121–126.

Groen, H. and M. Kaminski (1996). Optimisation of pressure vessels under reliability constraints. In *Proc. 15th Int. Conf. Offshore Mech. & Arctic Eng. (OMAE'96)*, pp. 177–185.

Guilleminot, J., C. Soize, D. Kondo, and C. Binetruy (2008). Theoretical framework and experimental procedure for modelling mesoscopic volume fraction stochastic fluctuations in fiber reinforced composites. *Int. J. Solids Struct.*, 45(21), 5567–5583.

Gunn, S. (1998). Support vector machines for classification and regression. Technical report, University of Southampton.

Handcock, M. and M. Stein (1993). A Bayesian analysis of Kriging. *Technometrics*, 35(4), 403–410.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their application. *Biometrika*, 57(1), 97–109.

Hohenbichler, M., S. Gollwitzer, W. Kruse, and R. Rackwitz (1987). New light on first- and second order reliability methods. *Structural Safety*, 4, 267–284.

Hohenbichler, M. and R. Rackwitz (1988). Improvement of second-order reliability estimates by importance sampling. *J. Eng. Mech.*, 114(12), 2195–2199.

Hurtado, J. (2004a). An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. *Structural Safety*, 26, 271–293.

Hurtado, J. (2004b). *Structural reliability – Statistical learning perspectives*, volume 17 of *Lecture notes in applied and computational mechanics*. Springer.

Hurtado, J. (2007). Filtered importance sampling with support vector margin: a powerful method for structural reliability analysis. *Structural Safety*, 29(1), 2–15.

Hurtado, J.-E. and D.-A. Alvarez (2000). Reliability assessment of structural systems using neural networks. In *Proc. European Congress on Computational Methods in Applied Sciences and Engineering, ECCOMAS 2000 (Barcelona, 11-14 Sept. 2000)*. Paper #290.

Igusa, T. and A. Der Kiureghian (1985). Dynamic characterization of two degree-of-freedom equipment-structure systems. *J. Eng. Mech.*, 111(1), 1–19.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. In *Proc. R. Soc. Lond. A, Math. Phys. Sci.*, volume 186, pp. 453–461.

Johnson, M., L. Moore, and D. Ylvisaker (1990). Minimax and maximin distance design. *J.Stat. Plan. Inf.*, 26, 131–148.

Jones, D., M. Schonlau, and W. Welch (1998). Efficient global optimization of expensive black-box functions. *J. Global Optim.*, 13(4), 455–492.

Katafygiotis, L. and S. Cheung (2005). A two-stage subset simulation-based approach for calculating the reliability of inelastic structural systems subjected to Gaussian random excitations. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16), 1581–1595.

Katafygiotis, L. and S. Cheung (2007). Application of spherical subset simulation method and auxiliary domain method on a benchmark reliability study. *Structural Safety*, 29(3), 194–207.

Kaymaz, I. (2005). Application of kriging method to structural reliability problems. *Structural Safety*, 27(2), 133–151.

Kharmanda, G., A. Mohamed, and M. Lemaire (2002). Efficient reliability-based design optimization using a hybrid space with application to finite element analysis. *Struct. Multidisc. Optim.*, 24(3), 233–245.

Kim, S.-H. and S.-W. Na (1997). Response surface method using vector projected sampling points. *Structural Safety*, 19(1), 3–19.

Kirjner-Neto, C., E. Polak, and A. Der Kiureghian (1998). An outer approximation approach to reliability-based optimal design of structures. *J. Optim. Theory Appl.*, 98, 1–16.

Koiter, W. (1945). *On the stability of elastic equilibrium (in Dutch)*. Ph. D. thesis, Delft University.

Koutsourelakis, P., H. Pradlwarter, and G. Schuëller (2004). Reliability of structures in high dimensions, part I: algorithms and applications. *Prob. Eng. Mech.*, 19, 409–417.

Krige, D. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. of the Chem., Metal. and Mining Soc. of South Africa*, 52(6), 119–139.

KSRI (1998). Etablissement de la méthode de calcul de la stabilité des coques. Technical report, Krylov Shipbuilding Research Institute.

Kuschel, N. and R. Rackwitz (1997). Two basic problems in reliability-based structural optimization. *Math. Meth. Oper. Research*, 46(3), 309–333.

Lagrange, J.-L. (1788). *Mécanique analytique*. Courcier, Paris.

Lebrun, R. and A. Dutfoy (2009a). A generalization of the Nataf transformation to distributions with elliptical copula. *Prob. Eng. Mech.*, 24(2), 172–178.

Lebrun, R. and A. Dutfoy (2009b). Do Rosenblatt and Nataf isoprobabilistic transformations really differ? *Prob. Eng. Mech.*, 24, 577–584.

Lebrun, R. and A. Dutfoy (2009c). An innovating analysis of the Nataf transformation from the copula viewpoint. *Prob. Eng. Mech.*, 24(3), 312–320.

L'Ecuyer, P., F. Le Gland, P. Lezaud, and B. Tuffin (2009). *Rare Event Simulation using Monte Carlo Methods*, chapter Splitting techniques, pp. 39–62. Wiley.

Lee, I., K. Choi, Y. Noh, L. Zhao, and D. Gorsich (2011a). Sampling-based stochastic sensitivity analysis using score functions for RBDO problems with correlated random variables. *J. Mech. Des.*, 133(2), 10.

Lee, I., K. Choi, and L. Zhao (2011b). Sampling-based RBDO using the stochastic sensitivity analysis and Dynamic Kriging method. *Struct. Multidisc. Optim.*, In press.

Lee, T. and J. Jung (2008). A sampling technique enhancing accuracy and efficiency of metamodel-based RBDO: Constraint boundary sampling. *Computers & Structures*, 86(13-14), 1463–1476.

Lemaire, M. (2009). *Structural Reliability*. Wiley.

Li, C.-C. and A. Der Kiureghian (1993). Optimal discretization of random fields. *J. Eng. Mech.*, 119(6), 1136–1154.

Liu, P.-L. and A. Der Kiureghian (1986). Multivariate distribution models with prescribed marginals and covariances. *Prob. Eng. Mech.*, 1(2), 105–112.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Trans. Information Theory*, 28, 129–136.

Loève, M. (1977). *Probability theory*. Springer.

Lophaven, S., H. Nielsen, and J. Søndergaard (2002). *DACE, A Matlab Kriging Toolbox*. Technical University of Denmark.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Le Cam, L.M. & Neyman, J. (Ed.), *Proc. 5^{th} Berkeley Symp. on Math. Stat. & Prob.*, volume 1, pp. 281–297, Berkeley, CA. University of California Press.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11, 431–441.

Marrel, A. (2005). Modélisation des codes de calcul dans le cadre des processus gaussiens stationnaires. Technical report, CEA (Stage INSA Toulouse).

Marrel, A. (2008). *Mise en œuvre et utilisation du métamodèle processus gaussien pour l'analyse de sensibilité de modèles numériques*. Ph. D. thesis, Institut National des Sciences Appliquées de Toulouse.

Marrel, A., B. Iooss, F. Van Dorpe, and E. Volkova (2008). An efficient methodology for modeling complex computer codes with Gaussian processes. *Comput. Statist. Data Anal.*, 52, 4731–4744.

Matheron, G. (1962). *Traité de géostatistique appliquée*. Editions Technip.

McKay, M. D., R. J. Beckman, and W. J. Conover (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 2, 239–245.

Melchers, R. (1989). Importance sampling in structural systems. *Structural Safety*, 6(1), 3–10.

Metropolis, N. (1987). The beginning of the Monte Carlo method. *Los Alamos Science, Special Issue*.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller (1953). Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6), 1087–1092.

Metropolis, N. and S. Ulam (1949). The Monte Carlo method. *J. Amer. Stat. Assoc.*, 44(247), 335–341.

Micol, A. (2007). *Approche probabiliste dans la conception des modules de puissance*. Ph. D. thesis, Université Paul Sabatier - Toulouse III.

Millwater, H. (2009). Universal properties of kernel functions for probabilistic sensitivity analysis. *Prob. Eng. Mech.*, 24(1), 89 – 99.

Mockus, J. (1994). Application of Bayesian approach to numerical methods of global and stochastic optimization. *J. Global Optim.*, 4(4), 347–365.

Montgomery, D. (2004). *Design and analysis of experiments*. Wiley.

Naess, A., B. Leira, and O. Batsevych (2009). System reliability analysis by enhanced Monte Carlo simulation. *Structural Safety*, 31(5), 349 – 355.

Nataf, A. (1962). Détermination des distributions dont les marges sont données. *C. R. Acad. Sci. Paris*, 225, 42–43.

Neal, R. (2003). Slice sampling. *Annals Stat.*, 31, 705–767.

Nelsen, R. (1999). *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer.

Nie, J. and B. Ellingwood (2000). Directional methods for structural reliability analysis. *Structural Safety*, 22(3), 233–249.

Nie, J. and B. Ellingwood (2004). A new directional simulation method for system reliability. Part I: application of deterministic point sets. *Prob. Eng. Mech.*, 19(4), 425–436.

Niederreiter, H. (1992). *Random number generation and quasi-Monte Carlo methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Nishijima, K., J. Qin, and M. Faber (2010). A scalable parametric approximation to multinormal probability integrals. In Straub, D., L. Esteva, and M. Faber (Eds.), *Proc. 15th IFIP WG7.5 Conference on Reliability and Optimization of Structural Systems, Munich, Germany*, pp. 139–146. CRC Press.

Noirfalise, C. (2009). *Analyse fiabiliste de la stabilité des coques minces avec imperfections par la Méthode Asymptotique Numérique*. Ph. D. thesis, Université Blaise Pascal - Clermont II.

Oakley, J. (2004). Estimating percentiles of uncertain computer code outputs. *J. Roy. Statist. Soc. Ser. C*, 53(1), 83–93.

Owen, A. (1992). A central limit theorem for Latin hypercube sampling. *J. Royal Stat. Soc., Series B*, 54, 541–551.

Papadrakakis, M. and N. Lagaros (2002). Reliability-based structural optimization using neural networks and Monte Carlo simulation. *Comput. Methods Appl. Mech. Engrg.*, 191(32), 3491–3507.

Patil, A., D. Huard, and C. Fonnesbeck (2010). PyMC: Bayesian stochastic modelling in Python. *J. Stat. Software*, 35(4), 1–81.

Perrin, F. (2008). *Prise en compte des données expérimentales dans les modèles probabilistes pour la prévision de la durée de vie des structures*. Ph. D. thesis, Université Blaise Pascal - Clermont II.

Phoon, K., S. Huang, and S. Quek (2002). Simulation of second-order processes using Karhunen-Loève expansion. *Computers & Structures*, 80(12), 1049–1060.

Picheny, V. (2009). *Improving accuracy and compensating for uncertainty in surrogate modeling*. Ph. D. thesis, University of Florida.

Picheny, V., D. Ginsbourger, O. Roustant, and R. Haftka (2010a). Adaptive designs of experiments for accurate approximation of a target region. *J. Mech. Des.*, 132(7), 071008.

Picheny, V., N. Kim, and R. Haftka (2010b). Application of bootstrap method in conservative estimation of reliability with limited samples. *Struct. Multidisc. Optim.*, 41(2), 205–217.

Piera-Martìnez, M. (2008). *Modélisation des comportements extrêmes en ingénierie*. Ph. D. thesis, Université Paris XI - UFR scientifique d'Orsay.

Piera-Martìnez, M., E. Vazquez, E. Walter, and G. Fleury (2007). RKHS classification for multivariate extreme-value analysis. In *Proc. 7th IASC - Statistics for Data Mining, Learning and Knowledge Extraction*.

Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in large margin classifiers*, pp. 61–74. MIT Press.

Polak, E. (1997). *Optimization algorithms and consistent approximations*. Springer.

Rackwitz, R. (2001). Reliability analysis – A review and some perspectives. *Structural Safety*, 23, 365–395.

Ramm, E. and W. A. Wall (2004). Shell structures - a sensitive interrelation between physics and numerics. *Int. J. Num. Meth. Eng.*, 60(1), 381–427.

Rasmussen, C. and C. Williams (2006). *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Massachusetts, Internet edition.

Robert, C. and G. Casella (2004). *Monte Carlo statistical methods (2^{nd} Ed.)*. Springer Series in Statistics. Springer Verlag.

Roustant, O., D. Ginsbourger, and Y. Deville (2010). *DiceKriging: Kriging methods for computer experiments*. R package version 1.1.

Royset, J. and E. Polak (2004a). Reliability-based optimal design using sample average approximations. *Prob. Eng. Mech.*, 19, 331–343.

Royset, J. O., A. Der Kiureghian, and E. Polak (2001). Reliability-based optimal structural design by the decoupling approach. *Reliab. Eng. Sys. Safety*, 73(3), 213–221.

Royset, J. O. and E. Polak (2004b). Implementable algorithm for stochastic optimization using sample average approximations. *J. Optim. Theory Appl.*, 122(1), 157–184.

Rubinstein, R. (1976). A Monte Carlo method for estimating the gradient in a stochastic network. Technion, Haifa, Israel.

Rubinstein, R. (1982). Generating random vectors uniformly distributed inside and on the surface of different regions. *Eur. J. Oper. Res.*, 10(2), 205–209.

Rubinstein, R. (1986). The score function approach for sensitivity analysis of computer simulation models. *Mathematics and Computers in Simulation*, 28(5), 351–379.

Rubinstein, R. and D. Kroese (2008). *Simulation and the Monte Carlo method*. Wiley Series in Probability and Statistics. Wiley.

Sacks, J., S. Schiller, and W. Welch (1989a). Designs for computer experiments. *Technometrics*, 31(1), 41–47.

Sacks, J., W. Welch, T. Mitchell, and H. Wynn (1989b). Design and analysis of computer experiments. *Stat. Sci.*, 4(4), 409–423.

Santner, T., B. Williams, and W. Notz (2003). *The design and analysis of computer experiments*. Springer series in Statistics. Springer.

Saporta, G. (2006). *Probabilités, analyse des données et statistique*. Editions Technip, 2nd edition.

Schenk, C. and G. Schuëller (2003). Buckling analysis of cylindrical shells with random geometric imperfections. *Int. J. Nonlinear Mech.*, 38(7), 1119–1132.

Schenk, C. and G. Schuëller (2007). Buckling analysis of cylindrical shells with cutouts including random boundary and geometric imperfections. *Comput. Methods Appl. Mech. Engrg.*, 196(35–36), 3424–3434.

Schölkopf, B. and A. Smola (2003). Kernel methods and Support Vector Machines. Technical report, Max-Planck-Institut fur biologische Kybernetik.

Schonlau, M. (1997). *Computer experiments and global optimization*. Ph. D. thesis, University of Waterloo, Ontario, Canada.

Schuëller, G. and H. Pradlwarter (2007). Benchmark study on reliability estimation in higher dimensions of structural systems – An overview. *Structural Safety*, 29, 167–182.

Schuëller, G., H. Pradlwarter, and P. Koutsourelakis (2004). A critical appraisal of reliability estimation procedures for high dimensions. *Prob. Eng. Mech.*, 19(4), 463–474.

Scordelis, A. C. and K. S. Lo (1961). Computer analysis of cylindrical shells. *ACI Journal*, 61, 539–561.

Sekhon, J. and W. Mebane (2011). Genetic optimization using derivatives: the `rgenoud` package for R. *J. Stat. Software*, 42(11), 1–26.

Shan, S. and G. Wang (2008). Reliable design space and complete single-loop reliability-based design optimization. *Reliab. Eng. Sys. Safety*, 93(8), 1218–1230.

Sichani, M., S. Nielsen, and C. Bucher (2011). Efficient estimation of first passage probability of high dimensional non-linear systems. *Prob. Eng. Mech.*, In press.

Singer, J. and H. Abramovich (1995). The development of shell imperfection measurement techniques. *Thin-Walled Structures*, 23, 379–398.

Singer, J., J. Arbocz, and C. Babcock (1971). Buckling of imperfect stiffened cylindrical shells under axial compression. *AIAA Journal*, 9(1), 68–75.

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.

Smola, A. and B. Schölkopf (2006). A tutorial on support vector regression. *Stat. Comput.*, 14, 199–222.

Song, S., Z. Lu, and H. Qiao (2009). Subset simulation for structural reliability sensitivity analysis. *Reliab. Eng. Sys. Safety*, 94(2), 658–665.

Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, 29, 143–151.

Steinhaus, H. (1956). Sur la division des corps matériels en parties. *Bull. Acad. Polon. Sci.*, 4(12), 801–804.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *J. Royal Stat. Soc., Series B*, 36, 111–147.

Sudret, B. (2007). *Uncertainty propagation and sensitivity analysis in mechanical models – Contributions to structural reliability and stochastic spectral methods*. Habilitation à diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France, 173 pages.

Sudret, B. and A. Der Kiureghian (2002). Comparison of finite element reliability methods. *Prob. Eng. Mech.*, 17, 337–348.

Taflanidis, A. (2007). *Stochastic system design and applications to stochastically robust structural control*. Ph. D. thesis, California Institute of Technology (CalTech).

Taflanidis, A. and J. Beck (2008). Stochastic subset optimization for optimal reliability problems. *Prob. Eng. Mech.*, 23(2-3), 324–338.

Taflanidis, A. and J. Beck (2009a). Life-cycle cost optimal design of passive dissipative devices. *Structural Safety*, 31(6), 508–522.

Taflanidis, A. and J. Beck (2009b). Stochastic subset optimization for reliability optimization and sensitivity analysis in system design. *Computers & Structures*, 87(5-6), 318–331.

Taflanidis, A. and S.-H. Cheung (2011). Stochastic sampling using moving least squares response surface approximations. *Prob. Eng. Mech.*, In press.

Timoshenko, S. (1936). *Theory of elastic stability*. McGraw-Hill.

Tsompanakis, Y., N. Lagaros, and M. Papadrakis (Eds.) (2008). *Structural design optimization considering uncertainties*. Taylor & Francis.

Tu, J., K. Choi, and Y. Park (1999). A new study on reliability-based design optimization. *J. Mech. Des.*, 121, 557–564.

Valdebenito, M. and G. Schuëller (2010). A survey on approaches for reliability-based optimization. *Struct. Multidisc. Optim.*, 42(5), 645–663.

Vapnik, V. (1995). *The nature of statistical learning theory*. Springer.

Vazquez, E. (2005). *Modélisation comportementale de systèmes non-linéaires multivariables par méthodes à noyaux et applications*. Ph. D. thesis, Université Paris XI - UFR scientifique d'Orsay.

Verhoosel, C. and M. Gutiérrez (2009). Reliability analysis of fracture in piezoelectric components with a random microstructure. In *Proc. 10th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2009), Osaka, Japan*.

Villemonteix, J., E. Vazquez, and E. Walter (2009). An informational approach to the global optimization of expensive-to-evaluate functions. *J. Global Optim.*, 44(4), 509–534.

Von Mises, R. (1914). Der kritische Außendruck zylindrischer Rohre. *Zeitschr. Ver. Deut. Ingr.*, 58, 750–755.

Vrouwenvelder, T. (1997). The JCSS probabilistic model code. *Structural Safety*, 19(3), 245–251.

Waarts, P.-H. (2000). *Structural reliability using finite element methods: an appraisal of DARS: Directional Adaptive Response Surface Sampling*. Ph. D. thesis, Technical University of Delft, The Netherlands.

Welch, W., R. Buck, J. Sacks, H. Wynn, T. Mitchell, and M. Morris (1992). Screening, predicting, and computer experiments. *Technometrics*, 34(1), 15–25.

Wu, Y. (1994). Computational methods for efficient structural reliability and reliability sensitivity analysis. *AIAA Journal*, 32(8), 1717–1723.

Youn, B. and K. Choi (2004). Selecting probabilistic approaches for reliability-based design optimization. *AIAA Journal*, 42, 124–131.

Zahrouni, H. (1998). *Méthode Asymptotique Numérique pour les coques en grandes rotations*. Ph. D. thesis, Université Paul Verlaine, Metz, LPMM, Metz.

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. John Wiley.

Zhang, Y. and A. Der Kiureghian (1995). Two improved algorithms for reliability analysis. In Rackwitz, R., G. Augusti, and A. Bori (Eds.), *Proc. 6th IFIP WG7.5 on Reliability and Optimization of Structural systems, Assisi, Italy*. Chapman & Hall, London.

A

Gaussian identities

Contents

A.1 Introduction

The multivariate Gaussian distribution is widely used in probability and statistics due to its convenient properties. Some of them are reviewed in this Appendix because they are used in the text of Chapter 1. The reason why they do not appear in the text is that they might be admitted in a first time. The following definitions and theorems presented here come from the book by Santner et al. (2003).

A.2 The multivariate Gaussian distribution

Definition A.2.1. *An n-component vector $\boldsymbol{X}$ is said to have a* multivariate Gaussian distribution *if it expresses as the following linear form of r independent standard Gaussian random variates $\boldsymbol{\Xi} = (\Xi_i \sim \mathcal{N}(0,1), i = 1, \ldots, r)^\mathsf{T}$:*

$$\boldsymbol{X} = (X_1, \ldots, X_n)^\mathsf{T} = \mathbf{L}\boldsymbol{\Xi} + \boldsymbol{m}, \tag{A.1}$$

where $\mathbf{L}$ is an $n \times r$ matrix and $\boldsymbol{m}$ is an n-component vector, both with real coefficients.

The two first second-order moments of the multivariate distribution come easily from this definition. First, its mean vector reads:

$$\boldsymbol{\mu} \equiv \mathbb{E}[\boldsymbol{X}] = \mathbb{E}[\mathbf{L}\boldsymbol{\Xi} + \boldsymbol{m}] = \boldsymbol{m}, \tag{A.2}$$

and its covariance matrix reads:

$$\boldsymbol{\Sigma} \equiv \mathrm{Cov}[\boldsymbol{X}] = \mathbb{E}\left[(\mathbf{L}\boldsymbol{\Xi} + \boldsymbol{m} - \boldsymbol{\mu})(\mathbf{L}\boldsymbol{\Xi} + \boldsymbol{m} - \boldsymbol{\mu})^\mathsf{T}\right] = \mathbf{L}\mathbf{L}^\mathsf{T}. \tag{A.3}$$

It is clear from this definition that the covariance matrix is always *symmetric* and *positive definite*. The symmetry comes from the fact that it is defined as the product of a real matrix by its transposed matrix; positive definiteness holds because:

$$\boldsymbol{z}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{z} = \boldsymbol{z}^\mathsf{T}\mathbf{L}\mathbf{L}^\mathsf{T}\boldsymbol{z} = \|\mathbf{L}^\mathsf{T}\boldsymbol{z}\|_2^2, \tag{A.4}$$

where $\|\bullet\|_2$ denotes the usual $\mathscr{L}_2$ norm in $\mathbb{R}^n$. Provided $\boldsymbol{z} \neq 0$, one gets:

$$\boldsymbol{z}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{z} \geq 0, \tag{A.5}$$

which is the definition of a (non-strictly) positive definite matrix. A strictly positive definite matrix is obtained if and only if $\mathbf{L}$ is full rank, namely if $r = \mathrm{rank}(\mathbf{L}) = n$, meaning that it cannot be reduced to a lower dimensional matrix due to the presence of lines in $\mathbf{L}$ which would be linear combinations of the others.

A multivariate Gaussian distribution defined with respect to a non full rank matrix $\mathbf{L}$ is said to be *degenerate*, and this can be checked by computing the condition number of $\boldsymbol{\Sigma}$ which is defined as the ratio between its lower and higher singular values from its *singular value decomposition* (SVD). If this condition number is zero, this means $\boldsymbol{X}$ is degenerate.

Theorem A.2.1. ***Multivariate normal probability density function***
Provided $\mathbf{L}$ *has full rank (*i.e. $r = n$*), the multivariate Gaussian distribution has the following* probability density function*:*

$$\varphi_n(\boldsymbol{x}) = \frac{1}{(2\,\pi)^{n/2}\,\left[\det \boldsymbol{\Sigma}\right]^{1/2}}\,\exp\left[-\frac{1}{2}\,(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\,\boldsymbol{\Sigma}^{-1}\,(\boldsymbol{x}-\boldsymbol{\mu})\right], \quad \boldsymbol{x}\in\mathbb{R}^n, \tag{A.6}$$

where $[\det \boldsymbol{\Sigma}]$ *denotes the determinant of the covariance matrix.*

Proof A.2.1. *Let* $\boldsymbol{\Xi}$ *denote a random vector whose n components are independent and identically distributed standard Gaussian random variates. Then, the joint PDF of this vector is simply obtained as the product of the n marginal PDFs:*

$$f_{\boldsymbol{\Xi}}(\boldsymbol{\xi}) = \prod_{i=1}^{n}\frac{1}{(2\,\pi)^{1/2}}\exp\left[-\frac{1}{2}\,\xi_i^2\right] = \frac{1}{(2\,\pi)^{n/2}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\xi_i^2\right] = \frac{1}{(2\,\pi)^{n/2}}\exp\left[-\frac{1}{2}\,\boldsymbol{\xi}^{\mathsf{T}}\,\boldsymbol{\xi}\right]. \tag{A.7}$$

Recall from Definition A.2.1 that a Gaussian random vector is defined as the following linear combination of $\boldsymbol{\Xi}$*:*

$$\boldsymbol{X} = \mathbf{L}\,\boldsymbol{\Xi} + \boldsymbol{\mu}, \tag{A.8}$$

where $\mathbf{L} = \boldsymbol{\Sigma}^{1/2}$ *is additionally assumed to be full rank here. Then since the linear application is monotone, the joint PDF of* $\boldsymbol{X}$ *can be simply obtained by applying the change of variable theorem:*

$$\varphi_n(\boldsymbol{x}) = f_{\boldsymbol{\Xi}}\left(\mathbf{L}^{-1}\,(\boldsymbol{x}-\boldsymbol{\mu})\right)\,[\det \mathbf{J}] \tag{A.9}$$

where $\mathbf{J}$ *is the Jacobian (square) matrix of the one-to-one mapping* $\boldsymbol{\xi}\mapsto\boldsymbol{X}$ *with terms:*

$$J_{i\,j} = \frac{\partial\,\xi_i}{\partial\,x_j}, \quad i,\,j = 1,\,\ldots,\,n. \tag{A.10}$$

The final formula for the joint PDF results from the fact that:

$$\mathbf{J} = \mathbf{L}^{-1} = \boldsymbol{\Sigma}^{-1/2} \quad \Rightarrow \quad [\det \mathbf{J}] = \left[\det \mathbf{L}^{-1}\right] = \left[\det \boldsymbol{\Sigma}^{-1/2}\right] = \frac{1}{[\det \boldsymbol{\Sigma}]^{1/2}}. \tag{A.11}$$

A.3 Linear form of Gaussian random vectors

Theorem A.3.1. ***Linear form of a Gaussian random vector***
Assume $\boldsymbol{X}$ *is distributed according to a multivariate Gaussian distribution* $\boldsymbol{X}\sim\mathcal{N}_n(\boldsymbol{\mu},\,\boldsymbol{\Sigma})$*,* $\mathbf{A}$ *is a real matrix and* $\boldsymbol{b}$ *is a real vector, then:*

$$\mathbf{A}\boldsymbol{X} + \boldsymbol{b} \sim \mathcal{N}_m\left(\mathbf{A}\boldsymbol{\mu} + \boldsymbol{b},\,\mathbf{A}\,\boldsymbol{\Sigma}\,\mathbf{A}^{\mathsf{T}}\right). \tag{A.12}$$

Proof A.3.1. *First, the fact that* $\mathbf{A}\boldsymbol{X}+\boldsymbol{b}$ *is still Gaussian is due to Definition A.2.1. Indeed, considering a random vector* $\boldsymbol{U}$ *whose components are uncorrelated standard Gaussian random variates:*

$$\boldsymbol{X} = \mathbf{L}\,\boldsymbol{U} + \boldsymbol{\mu}, \tag{A.13}$$

where $\mathbf{L}$ *is such that* $\boldsymbol{\Sigma} = \mathbf{L}\,\mathbf{L}^{\mathsf{T}}$*, one may further write:*

$$\mathbf{A}\boldsymbol{X} + \boldsymbol{b} = \mathbf{A}(\mathbf{L}\boldsymbol{U} + \boldsymbol{\mu}) + \boldsymbol{b} = \mathbf{A}\mathbf{L}\boldsymbol{U} + (\mathbf{A}\boldsymbol{\mu} + \boldsymbol{b}), \tag{A.14}$$

which is nothing but another linear form of n uncorrelated standard Gaussian random variates for which Definition A.2.1 still holds. Second-order moments come again easily. Its mean vector reads

$$\mathbb{E}\left[\mathbf{A}\boldsymbol{X}+\boldsymbol{b}\right]=\mathbf{A}\boldsymbol{\mu}+\boldsymbol{b}, \tag{A.15}$$

and its covariance matrix reads:

$$\mathrm{Cov}\left[\mathbf{A}\boldsymbol{X}+\boldsymbol{b}\right]=\mathbb{E}\left[\mathbf{A}\mathbf{L}\boldsymbol{U}\left(\mathbf{A}\mathbf{L}\boldsymbol{U}\right)^{\mathsf{T}}\right]=\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathsf{T}}. \tag{A.16}$$

A.4 Marginal and conditional distributions

In this section, we consider a Gaussian random vector $\boldsymbol{X}$ of size n divided into two sub-vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ of respective sizes n_1 and n_2. It is also assumed that $\boldsymbol{X}$ has its parameters partitioned as follows:

$$\boldsymbol{X}=\begin{bmatrix}\boldsymbol{X}_1\\ \boldsymbol{X}_2\end{bmatrix}\sim\mathcal{N}_n\left(\begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix},\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}\right), \tag{A.17}$$

requiring only that $\boldsymbol{\Sigma}_{12}=\boldsymbol{\Sigma}_{21}^{\mathsf{T}}$ so that the overall covariance matrix remains symmetric.

Theorem A.4.1. ***Marginal distributions of a Gaussian random vector***
The marginal distribution of $\boldsymbol{X}_1$ is also Gaussian with parameters:

$$\boldsymbol{X}_1\sim\mathcal{N}_{n_1}\left(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_{11}\right). \tag{A.18}$$

This can be proved by elaborating the required probability density function from its definition. The marginal distribution of $\boldsymbol{X}_1$ is defined as follows:

$$f_{\boldsymbol{X}_1}(\boldsymbol{x}_1)=\int_{\mathbb{R}^{n_2}} f_{\boldsymbol{X}}(\boldsymbol{x}_1,\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_2, \tag{A.19}$$

The explicit computations involve advanced linear algebra operations among which the computations of the inverse and the determinant of a partitioned symmetric matrix.

A simple interpretation of this theorem is that to marginalize a partitioned Gaussian random vector, one may only retain those terms corresponding to the part of interest and omit the cross-covariance matrices $\boldsymbol{\Sigma}_{12}=\boldsymbol{\Sigma}_{21}^{\mathsf{T}}$.

Theorem A.4.2. ***Conditional distributions of a Gaussian random vector***
The conditional distributions of $\boldsymbol{X}_1\,|\,\boldsymbol{X}_2$ is also Gaussian,

$$\boldsymbol{X}_1\,|\,\boldsymbol{X}_2\sim\mathcal{N}_{n_1}\left(\boldsymbol{\mu}_{1|2},\boldsymbol{\Sigma}_{1|2}\right), \tag{A.20}$$

with mean:

$$\boldsymbol{\mu}_{1|2}=\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}^{\mathsf{T}}\,\boldsymbol{\Sigma}_{22}^{-1}\,(\boldsymbol{x}_2-\boldsymbol{\mu}_2), \tag{A.21}$$

and covariance:

$$\boldsymbol{\Sigma}_{1|2}=\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}^{\mathsf{T}}\,\boldsymbol{\Sigma}_{22}^{-1}\,\boldsymbol{\Sigma}_{21}. \tag{A.22}$$

Again, this can be proved by elaborating the required probability density function from its definition. The conditional distribution of $\boldsymbol{X}_1 \mid \boldsymbol{X}_2$ and $\boldsymbol{X}_2 \mid \boldsymbol{X}_1$ is defined as follows:

$$f_{\boldsymbol{X}_1|\boldsymbol{X}_2}(\boldsymbol{x}_1) = \frac{f_{\boldsymbol{X}}(\boldsymbol{x}_1, \boldsymbol{x}_2)}{f_{\boldsymbol{X}_2}(\boldsymbol{x}_2)}. \tag{A.23}$$

As for the proof of Theorem A.4.1, the explicit computations involve advanced linear algebra operations among which the computations of the inverse and the determinant of a partitioned symmetric matrix. In addition, it requires the prior computations of the marginal distribution $f_{\boldsymbol{X}_2}(\boldsymbol{x}_2)$.

This theorem is used in Section 1.4.6 to derive the posterior distribution of a Bayesian predictor under a Gaussian process prior model.

A.5 Simulation of a Gaussian random vector

A.5.1 The Cholesky decomposition method

The simulation of a non-degenerate Gaussian random vector $\boldsymbol{X}$ of size n given its parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ requires the computation of the Cholesky decomposition of its covariance matrix:

$$\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\mathsf{T}, \tag{A.24}$$

From that decomposition, realizations of $\boldsymbol{X}$ may be simulated from its definition:

$$\boldsymbol{X} = \mathbf{L}\boldsymbol{\Xi} + \boldsymbol{\mu}, \tag{A.25}$$

where $\boldsymbol{\Xi} = (\Xi_i \sim \mathcal{N}(0,1), i = 1, \ldots, m)^\mathsf{T}$. Algorithm A.1 summarizes the simulation procedure.

Algorithm A.1 Simulation of N samples from a non-degenerate Gaussian random vector.

Input: $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$
$\mathbf{L} :=$ Cholesky($\boldsymbol{\Sigma}$) *If it fails use Algorithm A.2 instead.*
$\mathscr{X} = \emptyset$
for $i = 1 \to N$ **do**
  $\boldsymbol{\xi} :=$ StandardGaussianRandomNumbers(m)
  $\boldsymbol{x}^{(i)} := \boldsymbol{\mu} + \mathbf{L}\boldsymbol{\xi}$
  $\mathscr{X} = \left\{\mathscr{X}, \boldsymbol{x}^{(i)}\right\}$
end for
return $\mathscr{X}$

A.5.2 The spectral representation method

This procedure is more robust as it applies to both degenerate and non-degenerate Gaussian random vectors, but it requires the computation of the *eigenvalue decomposition* (EVD) of its covariance matrix instead of its unavailable Cholesky decomposition.

The EVD problem consists in finding the m couples $\left[(\lambda_i, \boldsymbol{\phi}_i), i = 1, \ldots, m\right]$ that are solution of:

$$\boldsymbol{\Sigma}\,\boldsymbol{\phi}_i = \lambda_i\,\boldsymbol{\phi}_i, \quad i = 1, \ldots, m. \tag{A.26}$$

Retaining the r couples associated with the non-zero singular values in the $r \times r$ diagonal matrix Λ and their associated singular vectors grouped in the $m \times r$ matrix $\boldsymbol{\Phi} = \left[\boldsymbol{\phi}_i, i = 1, \ldots, r\right]$, finally gives the following decomposition:

$$\boldsymbol{\Sigma} = \boldsymbol{\Phi}\,\boldsymbol{\Lambda}\,\boldsymbol{\Phi}^{\mathsf{T}} = (\boldsymbol{\Phi}\,\boldsymbol{\Lambda}^{1/2})(\boldsymbol{\Phi}\,\boldsymbol{\Lambda}^{1/2})^{\mathsf{T}}. \tag{A.27}$$

In practice, one neglects small singular values that are negligible in front of the largest singular value, say smaller than $\varepsilon = 10^{-10} \times \max\left\{\lambda_i, i = 1, \ldots, m\right\}$. Hence, another representation of $\boldsymbol{X}$ is proposed:

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{\Lambda}^{1/2}\,\boldsymbol{\Xi}. \tag{A.28}$$

This new representation may then be used in Algorithm A.2 to sample realizations from any Gaussian random vector. This representation only needs the realizations of $r \leq m$ independent random variates $\boldsymbol{\Xi} = \langle \Xi_i \sim \mathcal{N}(0, 1), i = 1, \ldots, r\rangle^{\mathsf{T}}$.

Algorithm A.2 Simulation of N samples from any Gaussian random vector.

Input: $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, ε
$\{(\lambda_i, \boldsymbol{\phi}_i), i = 1, \ldots, m\} :=$ SingularValueDecomposition($\boldsymbol{\Sigma}$)
$\{(\lambda_i, \boldsymbol{\phi}_i), i = 1, \ldots, m\} :=$ SortInDescendingOrder($\{\lambda_i, i = 1, \ldots, m\}$)
$r :=$ Card $\left(\{\lambda_i > \varepsilon, i = 1, \ldots, m\}\right)$
$\boldsymbol{\Lambda}^{1/2} :=$ DiagonalMatrix $\left(\left\{\sqrt{\lambda_i}, i = 1, \ldots, r\right\}\right)$
$\boldsymbol{\Phi} := \left[\boldsymbol{\phi}_i, i = 1, \ldots, r\right]$
$\mathscr{X} = \emptyset$
for $i = 1 \rightarrow N$ **do**
    $\boldsymbol{\xi} :=$ StandardGaussianRandomNumbers(r)
    $\boldsymbol{x}^{(i)} := \boldsymbol{\mu} + \boldsymbol{\Phi}\,\boldsymbol{\Lambda}^{1/2}\,\boldsymbol{\xi}$
    $\mathscr{X} = \left\{\mathscr{X}, \boldsymbol{x}^{(i)}\right\}$
end for
return $\mathscr{X}$

This spectral representation does not introduce any underestimation of the true spread of the degenerate random vector $\boldsymbol{X}$, because the spectral content of the covariance matrix is conserved. However, it is worth pointing out here that Li and Der Kiureghian (1993) extend this technique to the simulation of highly correlated random vectors resulting from the discretization of a stationary Gaussian process. They propose to neglect a larger set of the smallest singular values (possibly significantly positive) in order to minimize the number r of random variables and thus ease structural reliability analyses. This variant does introduce a relative mean-squared error that can be quantified as the amount of neglected spectral content:

$$\mathrm{RMSE}(r) = \frac{\sum_{i=r}^{m} \lambda_i}{\sum_{i=1}^{m} \lambda_i} < 1. \tag{A.29}$$

B

Markov chain Monte Carlo

Contents

B.1 Introduction

Markov chain Monte Carlo (MCMC) techniques arise in many application fields ranging from numerical optimization (*e.g.* simulated annealing) to statistical learning (see *e.g.* the Bayesian predictors in Chapter 1). The reader is referred to the article by Andrieu et al. (2003) which provides a short but quite exhaustive review of these applications.

In this manuscript they are used for two distinct purposes. First, they are used in Section 2.4 of Chapter 2 as a tool to sample from pseudo-PDFs in order to build an adaptive experimental design which concentrates in a specific region of interest. Second, it is used at the end of Chapter 3 for a proper statistical purpose: *i.e.* estimating a mathematical expectation from non-parametric instrumental PDFs. In both cases, the target distribution p is jointly defined up to an unknown but finite normalizing constant. In this setup, MCMC techniques should be seen as the only alternative to sample from p since the more conventional Monte Carlo methods such as the inverse transform technique are inapplicable there. The reader is referred to the reference books by Robert and Casella (2004) and Rubinstein and Kroese (2008) for a thorough review of these modern sampling techniques that appeared in the late 40's (Metropolis and Ulam, 1949; Metropolis, 1987).

B.2 Essentials for MCMC

This section summarizes the basic notions of MCMC sampling algorithms based on the reference book by Robert and Casella (2004) and the article by Andrieu et al. (2003).

B.2.1 Markov chains

First, let us define a stochastic process $\left\{\boldsymbol{X}^{(i)},\, i \in \mathbb{N}\right\}$ indexed by the infinite countable set of positive integers $\mathbb{N}$ that takes its values in an n-dimensional continuous *state space* $\mathbb{X}$. This stochastic process is a first-order Markov chain if and only if its PDF reduces to:

$$p\left(\boldsymbol{x}^{(i+1)} \middle| \boldsymbol{x}^{(i)}, \ldots, \boldsymbol{x}^{(0)}\right) = K\left(\boldsymbol{x}^{(i+1)} \middle| \boldsymbol{x}^{(i)}\right), \quad i \in \mathbb{N}, \tag{B.1}$$

where K is the *transition kernel* of the chain. The latter transition kernel is a (conditional) PDF, therefore it satisfies:

$$\int_{\mathbb{X}} K\left(\boldsymbol{x}^{(i+1)} \middle| \boldsymbol{x}^{(i)}\right) \mathrm{d}\boldsymbol{x}^{(i+1)} = 1, \quad i \in \mathbb{N}. \tag{B.2}$$

Example B.2.1. *The* random walk *process is the most intuitive example of first-order Markov chain. It is defined as follows:*

$$\boldsymbol{X}^{(i+1)} = \boldsymbol{X}^{(i)} + \boldsymbol{\epsilon}, \quad i \in \mathbb{N}, \tag{B.3}$$

provided the first element $\boldsymbol{X}^{(0)}$ is given. This first element is referred to as a seed *in the MCMC technical jargon. $\boldsymbol{\epsilon}$ is a random vector of size n whose probability distribution is independent*

of the previous states of the chains $\left\{\boldsymbol{X}^{(j)}, j = 0, \ldots, i\right\}$. *In addition, if* $\boldsymbol{\epsilon}$ *is symmetric about zero, then the random walk process is a* symmetrical random walk *process.*

B.2.2 Required properties for MCMC

A Markov chain is said to reach *a stationary (or invariant) probability distribution* p if and only if there exists a finite number of increments n_0 such that:

$$\boldsymbol{X}^{(i)} \sim p \quad \Rightarrow \quad \boldsymbol{X}^{(i+1)} \sim p, \quad \forall i > n_0, \quad n_0 \in \mathbb{N} \setminus \infty. \tag{B.4}$$

Such a stationary distribution can be reached if and only if the transition kernel K allows for free moves all over the state space $\mathbb{X}$. The latter property is referred to as the *irreducibility* of the chain. It is a fundamental property required by MCMC algorithms because the chains are initialized from arbitrary seeds in practice. Hence, this arbitrary choice should not affect the ability of the chain to converge. It is also said *aperiodic* if it does not get trapped in cycles. These important properties are hard to check in practice, although some empirical convergence diagnostics are available (Robert and Casella, 2004).

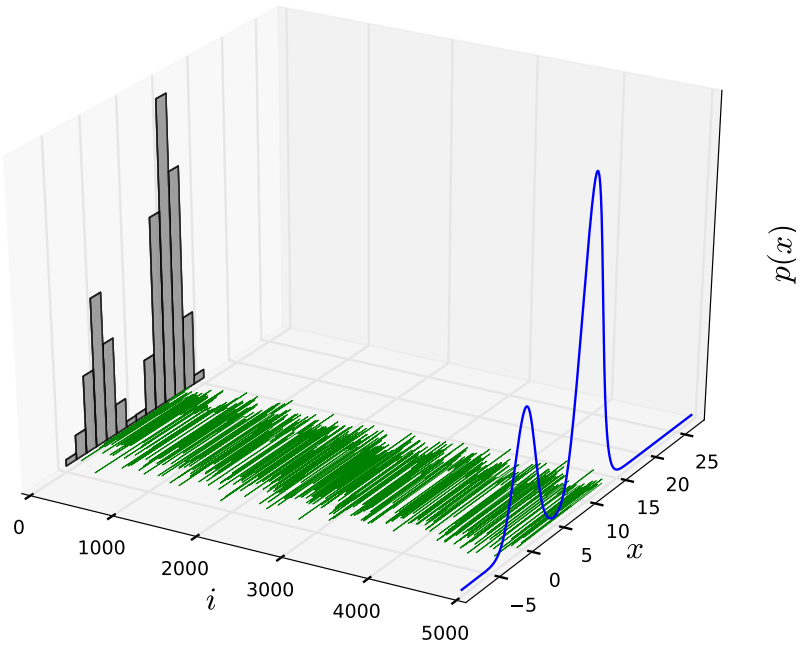


Figure B.1: A Markov chain whose transition kernel K is built so that the chain reaches $p(x) \propto 0.7\exp(-0.2\,x^2) + 0.3\exp(-0.2\,(x-10)^2)$ as stationary distribution. This illustration is inspired from the article by Andrieu et al. (2003).

MCMC algorithms aim at constructing Markov chains from a transition kernel K such that the chain eventually reaches a *targeted* probability distribution p. A non-necessary but sufficient condition on K to insure that a given p is the desired stationary distribution is the following *reversibility* (or *detailed balance*) condition:

$$p\left(\boldsymbol{x}^{(i+1)}\right) K\left(\boldsymbol{x}^{(i)} \middle| \boldsymbol{x}^{(i+1)}\right) = p\left(\boldsymbol{x}^{(i)}\right) K\left(\boldsymbol{x}^{(i+1)} \middle| \boldsymbol{x}^{(i)}\right), \quad \forall i > n_0, \quad n_0 \in \mathbb{N} \setminus \infty. \tag{B.5}$$

Indeed, integrating both sides over $\boldsymbol{x}^{(i)}$ yields:

$$p\left(\boldsymbol{x}^{(i+1)}\right) = \int_{\mathbb{X}} p\left(\boldsymbol{x}^{(i)}\right) K\left(\boldsymbol{x}^{(i+1)} \middle| \boldsymbol{x}^{(i)}\right) \mathrm{d}\boldsymbol{x}^{(i)}, \quad \forall i > n_0, \quad n_0 \in \mathbb{N} \setminus \infty, \tag{B.6}$$

which means that p is the stationary distribution of the chain with kernel K. MCMC algorithms such as the Metropolis-Hastings sampler are specifically designed to satisfy this reversibility condition.

B.2.3 Central limit theorem for stationary Markov chains

B.2.3.1 Statement

It can be proved under some additional conditions that are beyond the author's reach, that a central limit theorem holds for *irreducible*, *aperiodic* and *reversible* Markov chains.

Theorem B.2.1. ***Central limit theorem for stationary Markov chains*** *(see* e.g. *Robert and Casella, 2004, p. 244)*
Let ϕ denote a real-valued function, and let $\{\{\phi(\boldsymbol{X}^{[k](i)}), i = 1, \ldots, N/K\}, k = 1, \ldots, K\}$ be a collection of K irreducible, aperiodic, reversible *Markov chains with p as* invariant distribution. *Then, provided the chains have reached* stationarity, *the following average:*

$$\widehat{\overline{\phi}} = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N/K} \phi\left(\boldsymbol{X}^{[k](i)}\right) \tag{B.7}$$

converges towards the following mathematical expectation:

$$\overline{\phi} = \mathbb{E}\left[\phi\left(\boldsymbol{X}\right)\right] = \int_{\mathbb{X}} \phi\left(\boldsymbol{x}\right) p\left(\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x}, \tag{B.8}$$

and the convergence is as follows:

$$\left[\widehat{\overline{\phi}} - \overline{\phi}\right] \xrightarrow[N\to\infty]{a.s.} \mathcal{N}_1\left(0, \sigma^2_{\overline{\phi}}\right), \tag{B.9}$$

where $\sigma^2_{\overline{\phi}}$ is the variance of estimation.

This theorem is of utmost importance for the methods introduced in Chapter 3. The convergence can be measured in terms of the variance of estimation which is proved to decay in N as in the sequel as for ordinary Monte Carlo methods.

The variance of estimation $\sigma^2_{\overline{\phi}}$ in Theorem B.2.1 is derived as follows:

$$\begin{aligned}
\sigma^2_{\overline{\phi}} &= \mathrm{Var}\left[\widehat{\overline{\phi}}\right] \\
&= \mathbb{E}\left[\left(\widehat{\overline{\phi}} - \overline{\phi}\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{1}{N}\sum_{k=1}^{K}\sum_{i=1}^{N/K} \phi\left(\boldsymbol{X}^{[k](i)}\right) - \overline{\phi}\right)^2\right].
\end{aligned} \tag{B.10}$$

Since the chains are assumed to be generated with the same MCMC mechanism, they are *independent and identically distributed* according to a unique Markov chain denoted by $\{\phi(\boldsymbol{X}^{(i)}), i = 1, \ldots, N\}$, and the variance of estimation further reads:

$$\begin{aligned}\sigma^2_{\overline{\phi}} &= K\,\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N/K}\left(\phi\left(\boldsymbol{X}^{(i)}\right)-\overline{\phi}\right)\right)^2\right] \\ &= \frac{K}{N^2}\,\mathbb{E}\left[\sum_{i=1}^{N/K}\sum_{j=1}^{N/K}\left(\phi\left(\boldsymbol{X}^{(i)}\right)-\overline{\phi}\right)\left(\phi\left(\boldsymbol{X}^{(j)}\right)-\overline{\phi}\right)\right]. \end{aligned} \tag{B.11}$$

Thanks to the linearity of the mathematical expectation, one gets:

$$\begin{aligned}\sigma^2_{\overline{\phi}} &= \frac{K}{N^2}\sum_{i=1}^{N/K}\sum_{j=1}^{N/K}\mathbb{E}\left[\left(\phi\left(\boldsymbol{X}^{(i)}\right)-\overline{\phi}\right)\left(\phi\left(\boldsymbol{X}^{(j)}\right)-\overline{\phi}\right)\right] \\ &= \frac{K}{N^2}\sum_{i=1}^{N/K}\sum_{j=1}^{N/K}\mathrm{Cov}\left[\phi\left(\boldsymbol{X}^{(i)}\right),\phi\left(\boldsymbol{X}^{(j)}\right)\right], \end{aligned} \tag{B.12}$$

by definition of the chains autocovariance. Provided the chains have reached stationarity, their autocovariance (the same for all chains) is a function of the *lag* $\ell = i - j$ and their variance (the same for all chains) is constant, so that it recasts as follows:

$$\mathrm{Cov}\left[\phi\left(\boldsymbol{X}^{(i)}\right),\phi\left(\boldsymbol{X}^{(j)}\right)\right] = \sigma^2_\phi\, R(i-j), \tag{B.13}$$

where the variance $\sigma^2_\phi = \mathrm{Var}\left[\phi(\boldsymbol{X})\right]$ and the stationary autocorrelation R of the chain have been introduced. Substituting the latter simplified expression for the covariance in Eq. (B.12) yields:

$$\sigma^2_{\overline{\phi}} = \frac{K}{N^2}\sum_{i=1}^{N/K}\sum_{j=1}^{N/K}\sigma^2_\phi\, R(i-j). \tag{B.14}$$

Noting that the autocorrelation R is symmetric about zero where it is equal to one by definition, one eventually gets:

$$\begin{aligned}\sigma^2_{\overline{\phi}} &= \frac{K\,\sigma^2_\phi}{N^2}\left[\sum_{i=1}^{N/K}R(0)+\sum_{i=1}^{N/K}\sum_{\substack{j=1\\ j\neq i}}^{N/K}R(i-j)\right] \\ &= \frac{K\,\sigma^2_\phi}{N^2}\left[\frac{N}{K}+2\sum_{i=1}^{N/K}\sum_{\ell=1}^{N/K-i}R(\ell)\right] \\ &= \frac{K\,\sigma^2_\phi}{N^2}\left[\frac{N}{K}+2\sum_{\ell=1}^{N/K-1}\left(\frac{N}{K}-\ell\right)R(\ell)\right] \\ &= \frac{\sigma^2_\phi}{N}\left[1+\gamma\right], \end{aligned} \tag{B.15}$$

where the following coefficient has been introduced:

$$\gamma \equiv 2 \sum_{\ell=1}^{N/K-1} \left(1 - \frac{\ell K}{N}\right) R(\ell). \tag{B.16}$$

B.2.3.2 Numerical approximation of the variance of estimation

In practice, the computation of the variance of estimation in Eq. (B.15) requires the estimation of the chains' variance σ_ϕ^2 and that of the chains' autocorrelation R at lags $\ell = 1, \ldots, N/K - 1$. They can usually be estimated by averaging on both the K Markov chains and their elements with equivalent lags using ergodicity. The variance (resp. the autocorrelation) is computed as follows:

$$\sigma_\phi^2 \approx \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N/K} \phi\left(\boldsymbol{X}^{[k](i)}\right)^2 - \overline{\phi}^2, \tag{B.17}$$

$$R(\ell) \approx \frac{1}{\sigma_\phi^2 \, (N - \ell K)} \sum_{k=1}^{K} \sum_{i=1}^{N/K-\ell} \left(\phi\left(\boldsymbol{X}^{(i)}\right) - \overline{\phi}\right) \left(\phi\left(\boldsymbol{X}^{(i+\ell)}\right) - \overline{\phi}\right). \tag{B.18}$$

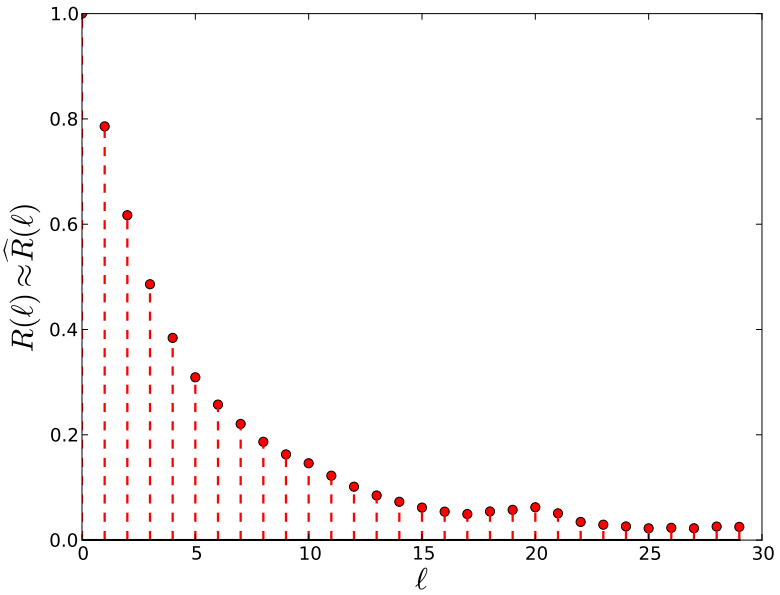


Figure B.2: Autocorrelation of the Markov chain illustrated in Figure B.1.

Example B.2.2. *The autocorrelation of the chain illustrated in Figure B.1 has been estimated from the single chain available (*i.e. $K = 1$*) using Eq. (B.18). Its illustration in Figure B.2 is truncated after the thirtieth lag. In this case, the estimator of the mean $\overline{x} = \mathbb{E}[X]$ is $\widehat{\overline{x}} \approx 6.77$. The correlation coefficient $\gamma \approx 0.99$ and the variance of X $\sigma_X^2 \approx 25.47$ eventually yields a fair approximation of the variance of estimation $\sigma_{\overline{x}} \approx 0.01$. The MCMC-based estimator is approximately $(1 + \gamma) \approx 2$ times less efficient than an ordinary-Monte-Carlo-based estimator.*

*To reduce this inefficiency, a common practice (used for the implementation of the reliability methods at the end of Chapter 3) consists in using several independent Markov chains in parallel (*i.e. $K > 1$*).*

B.2.3.3 Remark for binary-valued functions

Note that in the particular case where ϕ is a binary-valued function (equal to 0 or 1), $\phi(\boldsymbol{X})$ is a Bernoulli variable. Hence, the variance of the chains comes analytically (see *e.g.* Lemaire, 2009, pp. 250–251):

$$\sigma_\phi^2 = \mathrm{Var}\left[\phi(\boldsymbol{X})\right] = p\,(1-p), \tag{B.19}$$

where $p = \mathbb{E}\left[\phi(\boldsymbol{X})\right]$ is the rate parameter of the Bernoulli variable $\phi(\boldsymbol{X})$. Thus, the variance of estimation reduces to:

$$\sigma_{\widehat{p}}^2 = \frac{p\,(1-p)}{N}\,(1+\gamma)\,. \tag{B.20}$$

This eventually proves the following expression for the coefficient of variation:

$$\delta = \frac{\sigma_{\widehat{p}}}{p} = \sqrt{\frac{1-p}{N\,p}\,(1+\gamma)}, \tag{B.21}$$

which is used in Section 3.3.5 of Chapter 3.

B.2.4 Burn-in and thinning

In practice, the chains are initialized with arbitrary seeds $\{\boldsymbol{x}^{[k](0)}, k = 1, \ldots, K\}$, so that they may require a few iterations before they reach stationarity. *Burn-in* is an empirical technique that consists in rejecting (burning) the n_0 first values of the chain in order to remove their non-stationary part. However, finding an objective value for n_0 is rather non-trivial as it depends on both the seed and how the MCMC sampler performs on the target distribution p. In addition, this technique is quite computationally demanding if the target distribution is expensive-to-evaluate. For instance, in Chapter 3, the sub-optimal instrumental PDFs used for subset sampling involve the performance function $\mathfrak{g}$ in their definition (see Section 3.3.5). On the contrary, burn-in is conceivable for estimating the augmented failure probability in Section 3.4.2.1 as it only involves the Kriging predictor.

In the light of Eq. (B.15), it is clear that the variance of estimation of an average estimated from K stationary Markov chains is $(1+\gamma)$ times greater than that of an average that would be estimated by ordinary Monte Carlo sampling due to the non-zero positive γ coefficient. Note that this coefficient depends on the autocorrelation of the chain which might be large when the MCMC sampler does not perform well (*e.g.* on highly multimodal target distributions). For such cases, there exists another simple procedure known as *thinning* which consists in retaining one element of each chain every t states ($t \in \mathbb{N}^*$). Again the parameter t is rather non-trivial to tune as it is case-dependent, and it considerably

augments the total number of increments of the chain. Indeed, if both burn-in and thinning are applied, and if an N-sample is required, then the total number of increments is $N_{\text{tot}} = n_0 + t\,N$.

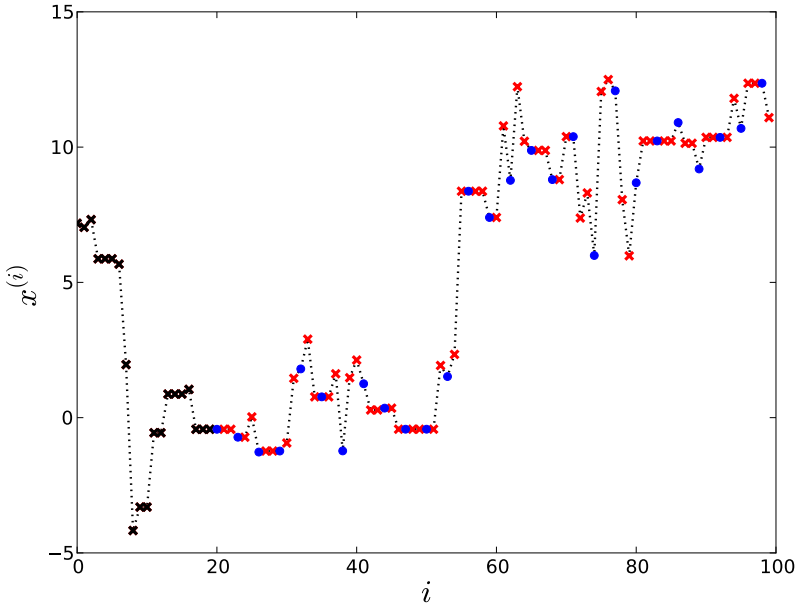


Figure B.3: Burnt and thinned Markov chain. The $n_0 = 20$ first increments are burnt, and then one element is kept every $t = 3$ states.

Finally, it is worth mentioning that the variance of estimation in Eq. (B.15) assumed that the K chains were generated in parallel from different seeds *but* according to a *same* MCMC mechanism. Hence, the burn-in and thinning procedures should be identical for all the chains in an estimation context.

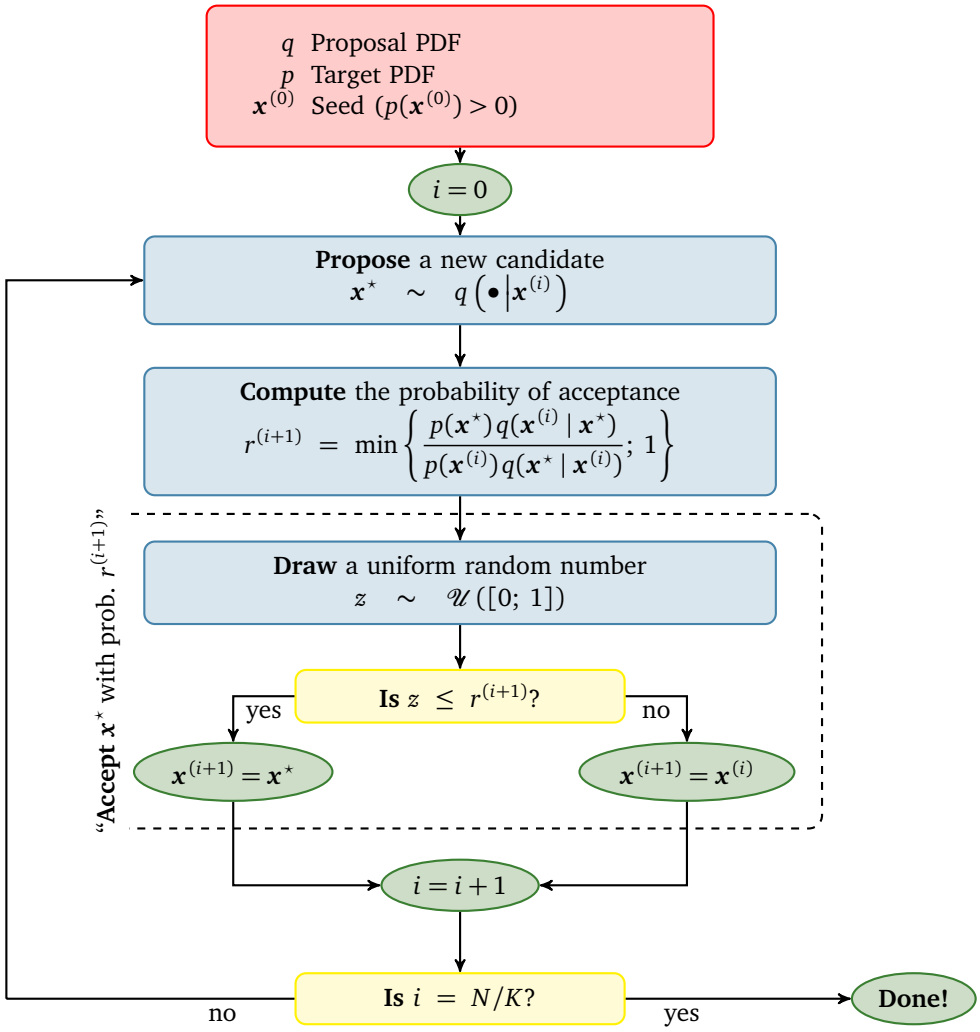
B.3 Markov chain samplers

This section shortly describes some state-of-the-art Markov chain samplers that were used in this thesis. All the procedures are detailed for one chain only for the sake of clarity. They may of course be used from a collection of seeds $\{\boldsymbol{x}^{[k](0)},\, k = 1, \ldots, K\}$ in order to generate a collection of chains $\{\{\boldsymbol{X}^{[k](i)},\, i = 1, \ldots, N/K\},\, k = 1, \ldots, K\}$ in parallel.

B.3.1 The Metropolis-Hastings sampler

The Metropolis-Hastings sampler (Metropolis et al., 1953; Hastings, 1970) is built in order to satisfy the reversibility condition formulated earlier in Eq. (B.5). It proceeds as detailed in Algorithm B.1.

Algorithm B.1 Metropolis-Hastings sampler



It requires a target PDF p, a proposal PDF q and a seed $\boldsymbol{x}^{(0)}$ that is distributed according to p, meaning that $p(\boldsymbol{x}^{(0)}) > 0$. The algorithm then iterates on i until it has performed N/K increments.

At each increment, the algorithm draws a random candidate according to the proposal PDF $q(\bullet \mid \boldsymbol{x}^{(i)})$. Hence, it is important to choose a proposal distribution which is easy to sample from. Then the algorithm computes the acceptance probability which is defined as follows:

$$r^{(i+1)} = \min\left\{\frac{p(\boldsymbol{x}^\star)\,q(\boldsymbol{x}^{(i)} \mid \boldsymbol{x}^\star)}{p(\boldsymbol{x}^{(i)})\,q(\boldsymbol{x}^\star \mid \boldsymbol{x}^{(i)})};\ 1\right\}, \quad i = 0, \ldots, N/K - 1 \tag{B.22}$$

In fact, this transition probability was specifically defined to ensure that the associated kernel satisfies the reversibility condition in Eq. (B.5) (see *e.g.* Andrieu et al., 2003, p. 16–17). The candidate is then accepted with probability $r^{(i+1)}$ as detailed in Algorithm B.1.

One advantage of this algorithm is that it does not require the knowledge of the normalizing constant of the target PDF p. Indeed p only needs to be defined up to an unknown (but finite) normalizing constant as it simplifies in the computation of the acceptance probability in Eq. (B.22). This is indeed a basic property for most MCMC samplers.

There exists many variant of this algorithm (see Robert and Casella, 2004, Chapter 7). The most widely used is the *symmetric Metropolis-Hastings sampler* which resorts to a symmetric proposal PDF q such that:

$$q\left(\boldsymbol{x}^{\star}\left|\boldsymbol{x}^{(i)}\right.\right)=q\left(\boldsymbol{x}^{(i)}\left|\boldsymbol{x}^{\star}\right.\right),\quad i=0,\ldots,N/K-1, \tag{B.23}$$

which in turns simplifies the expression of the acceptance probability in Eq. (B.22):

$$r^{(i+1)}=\min\left\{\frac{p(\boldsymbol{x}^{\star})}{p(\boldsymbol{x}^{(i)})};\,1\right\},\quad i=0,\ldots,N/K-1. \tag{B.24}$$

Another variant is the *independent sampler* whose proposal PDF q does not depend on the current state of the chain.

According to Andrieu et al. (2003), this algorithm is proved to converge to the target probability distribution p provided that the support of q includes the support of p. Even though, Au and Beck (2001) recommend the use of a uniform proposal PDF q to sample from a PDF with infinite support and obtain accurate results for their purpose.

Example B.3.1. *The chain illustrated in Figure B.1 was indeed generated with the Metropolis-Hastings sampler with the target PDF defined in the legend of Figure B.1. The proposal PDF q is a symmetrical Gaussian distribution with variance 100.*

B.3.2 The modified Metropolis-Hastings sampler

The modified Metropolis-Hastings sampler refers to the algorithm proposed by Au and Beck (2001) to sample from the sub-optimal instrumental PDFs in a subset sampling scheme. Recall from Section 3.3.5 of Chapter 3 that these targeted PDFs are defined as follows:

$$h_i(\boldsymbol{x})\propto\mathbb{1}_{\mathbb{F}_{i-1}}(\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x}),\quad i=2,\ldots,s, \tag{B.25}$$

where $f_{\boldsymbol{X}}$ is the original PDF of the random vector $\boldsymbol{X}$ and $\mathbb{1}_{\mathbb{F}_{i-1}}$ is the $(i-1)$-th intermediate subset indicator function which is equal to 1 if $\boldsymbol{x}\in\mathbb{F}_{i-1}$ and 0 otherwise. The algorithm is introduced here for a slightly more general case where the target PDF expresses as follows:

$$p(\boldsymbol{x})\propto\phi(\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x}), \tag{B.26}$$

to match the need of the reliability methods proposed at the end of Chapter 3. Actually, in the setup of Sections 3.4.2.1 and 3.5, ϕ is a *probabilistic classification function* π which takes real values on $[0;\,1]$.

The main motivation underlying the proposition of this modified algorithm is that the usual Metropolis-Hastings algorithm becomes inefficient in large dimension. Indeed, experience proves that the rejection rate dramatically increases with the number of components

n in $\boldsymbol{X}$. As a result, the Markov chains are constituted with many repeated elements and in turn, the γ coefficient significantly increases.

The modified Metropolis-Hastings algorithm takes advantage of the availability of an isoprobabilistic transform T that allows one to map the initial random vector $\boldsymbol{X}$ to a standard Gaussian random vector $\boldsymbol{U}$ with independent components (see Section 3.2.3 of Chapter 3). Indeed, sampling in that standard space is a lot more efficient for the following reason. Since the variables are independent, the target PDF in the standard space recasts as the following product:

$$p(\boldsymbol{u}) \propto \phi(\boldsymbol{u}) \prod_{j=1}^{n} \varphi(u_j). \tag{B.27}$$

Hence, the Metropolis-Hastings algorithm can be run sequentially on each term of this product without altering its convergence properties. In addition, the transformation ease the choice for the proposal PDF q which is itself expressed as a product of independent PDFs $\{q_j,\, j = 1, \ldots, n\}$. Au and Beck (2001) propose to use centred uniform PDFs with width 2 because all standard random variables $\{U_j,\, j = 1, \ldots, n\}$ concentrate in intervals of width ≈ 16 centred about zero. This choice yields accurate results in many examples arising in structural reliability (see the references at the beginning of Section 3.3.5 in Chapter 3).

Algorithm B.2 summarizes this modified sampler. The algorithm requires that the original random vector $\boldsymbol{X}$ can be transformed to a standard Gaussian random vector through an isoprobabilistic transform T (whose inverse T^{-1} is also available). The algorithm then proceeds iteratively on each term of the transformed target PDF in Eq. (B.27) from a given seed in the standard space $\boldsymbol{u}^{(0)}$ (or $\boldsymbol{u}^{(0)} = T(\boldsymbol{x}^{(0)})$). This seed must be distributed according to the transformed target PDF, meaning that $p(\boldsymbol{u}^{(0)})$ is not zero.

First, Algorithm B.1 is applied on each component of the seed. The present implementation from the article by Au and Beck (2001) uses a built-in uniform proposal PDF q centred on the current state with width 2. Noting that this proposal PDF is symmetrical, the first acceptance probability at the i-th increment for the j-th component simplifies as follows:

$$r_{1\,j} = \min\left\{ \frac{\varphi\left(u_j^{\star}\right)}{\varphi\left(u_j^{(i)}\right)};\, 1 \right\} = \min\left\{ \exp\left[\frac{1}{2}\left(u_j^{(i)^2} - u_j^{\star 2} \right) \right];\, 1 \right\}. \tag{B.28}$$

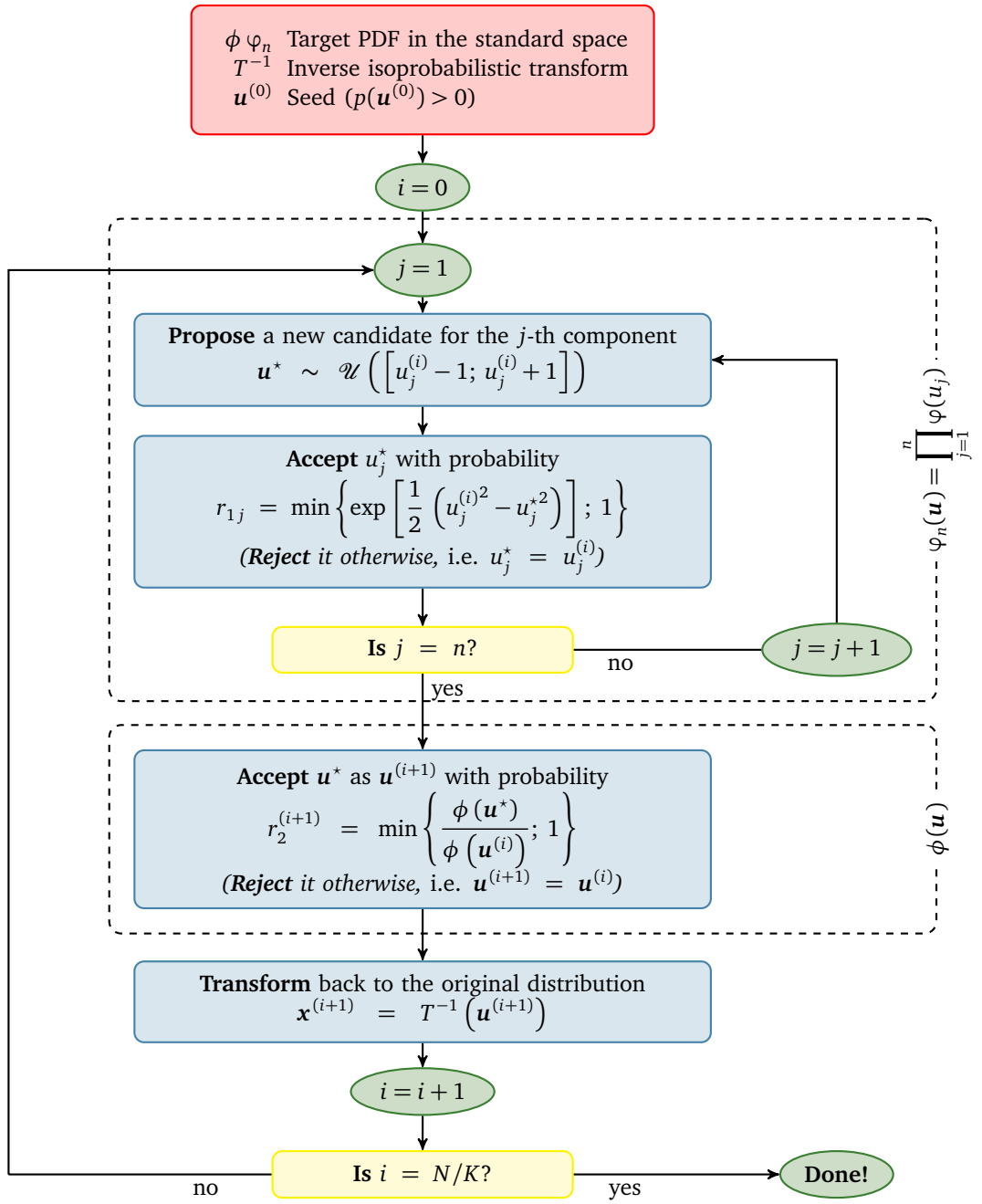
Second, the i-th candidate state $\boldsymbol{u}^{\star}$ proposed by the latter componentwise Metropolis sampler is accepted with a new probability:

$$r_2^{(i+1)} = \min\left\{ \frac{\phi\left(\boldsymbol{u}^{\star}\right)}{\phi\left(\boldsymbol{u}^{(i)}\right)};\, 1 \right\}. \tag{B.29}$$

Note that in the specific case of subset sampling, ϕ is reduced to a binary-valued indicator function. Hence, the second acceptance with probability $r_2^{(i+1)}$ reduces to a deterministic decision depending on the value of ϕ. Even though, Algorithm B.2 is presented here for the more general case where ϕ is real-valued as for the reliability methods introduced at the end of Chapter 3 (see Sections 3.4.2.1 and 3.5).

This algorithm is a lot more efficient than the usual Metropolis-Hastings algorithm applied to the original target PDF p as a block. However, it requires the availability of an isoprobabilistic transform to split the target PDF in a product of independent one-dimensional PDFs times another joint PDF for the conditioning term ϕ so that it recasts as in Eq. (B.27).

Algorithm B.2 Modified Metropolis-Hastings sampler



B.3.3 The slice sampler

B.3.3.1 Motivations

One major difficulty underlying the Metropolis-Hastings sampler lies in the choice of a well-suited proposal PDF q that will ensure *fast convergence*. Indeed, it was previously stated that convergence is ensured if and only if the chain is *irreducible*. This means that any admissible state with respect to the target PDF (any $\boldsymbol{x}$ such that $p(\boldsymbol{x}) > 0$) can be reached from any arbitrary point $\boldsymbol{x}^{(0)}$ (such that $p(\boldsymbol{x}^{(0)}) > 0$) within a finite number of increments. Andrieu et al. (2003) pointed out that the Metropolis-Hastings sampler is proved to converge provided the proposal PDF q dominates the target PDF p. Despite this might theoretically hold, experience shows that a narrow-banded proposal PDF (even with theoretically infinite support) does not yield convergence if the target PDF is highly multi-modal with large zero-probability regions in between the modes such as illustrated in Figure B.4.

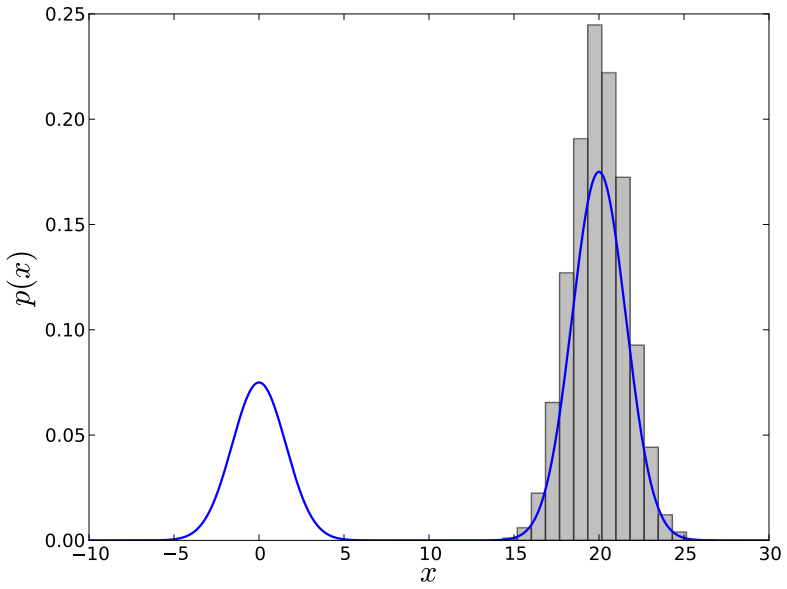


Figure B.4: A bimodal target PDF featuring two distinct modes: $p(x) \propto 0.7\exp(-0.2\,x^2) + 0.3\exp(-0.2\,(x-20)^2)$. This pathological case illustrates the fundamental difficulty underlying the optimal choice of a proposal PDF for the Metropolis-Hastings sampler. Here, the proposal PDF is a standard Gaussian distribution.

A workaround consists in using a larger-banded proposal PDF, but it will results in higher rejection rates (therefore slow convergence). Note also that the probability to jump from one mode to another is rather low, so that *thinning* should definitely be considered to reduce the variance of estimation in Eq. (B.15) if the samples are generated for averaging purposes.

Such pathological cases are not very frequent in structural reliability, so that Au and Beck's modified Metropolis-Hastings sampler performs well for the specific subset sampling application (see Section 3.3.5 of Chapter 3). Another reason for the global performance of the subset sampling algorithm is that it resorts to several Markov chains initialized from all the modes of the target PDF, so that the overall sample constituted by the K Markov chains covers the whole target PDF.

However, it has been pointed out that the refinement criteria introduced in Chapter 2 feature a large number of modes separated by large zero-probability region, so that Metropolis-Hastings samplers yield very poor performance. For such cases, the so-called *slice sampler* proposed by Neal (2003) offers a better performance. Indeed, it adaptively tunes the proposal PDF at each increment of the chain so that it may jump from one mode to the other even when they are separated by large zero-probability regions.

B.3.3.2 Principle

This simulation technique consists in introducing an auxiliary scalar random variate U such that the joint PDF of the augmented vector $(U, \boldsymbol{X})$ reads:

$$J(u, \boldsymbol{x}) = \begin{cases} 1 & \text{if} \quad 0 \leq u \leq p(\boldsymbol{x}) \\ 0 & \text{otherwise} \end{cases} . \tag{B.30}$$

Note that this definition is nothing but a translation of the fundamental theorem of simulation (see Robert and Casella, 2004, Theorem 2.15) whose key idea is to sample the augmented vector "under" the density curve (or surface) $p(\boldsymbol{x})$.

It is easy to show that the *marginalization* of $J(u, \boldsymbol{x})$ *w.r.t.* u results in the target PDF p:

$$\int_{\mathbb{R}} J(u, \boldsymbol{x}) \, \mathrm{d}u = \int_0^{p(\boldsymbol{x})} 1 \, \mathrm{d}u = p(\boldsymbol{x}). \tag{B.31}$$

Hence, sampling $\boldsymbol{X}$ from p is equivalent to sampling $(U, \boldsymbol{X})$ from J and then ignoring U.

The slice sampling procedure relies on the conditional distributions derived from J which are easily shown to read as follows:

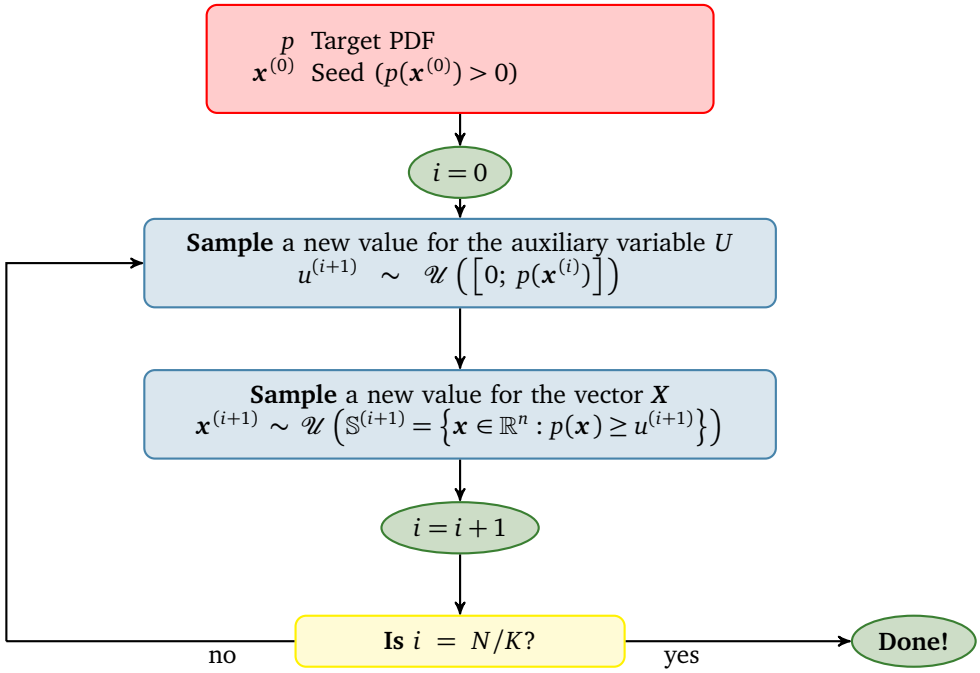
$$U \mid \boldsymbol{X} = \boldsymbol{x} \quad \sim \quad \mathcal{U}\left([0;\, p(\boldsymbol{x})]\right), \tag{B.32}$$

$$\boldsymbol{X} \mid U = u \quad \sim \quad \mathcal{U}\left(\mathbb{S} = \left\{\boldsymbol{x} \in \mathbb{R}^n : p(\boldsymbol{x}) \geq u\right\}\right). \tag{B.33}$$

Algorithm B.3 summarizes the key steps of the slice sampler. As for the other Metropolis-Hastings algorithm the target PDF may only be defined up to an unknown but finite normalizing constant. It does not require any proposal PDF.

The procedure is illustrated in Figure B.5 on the bimodal target PDF of Figure B.4. It can be seen that the stationary distribution of the chain represented by the histogram is closer to the target PDF than in Figure B.4. One increment is also illustrated in order to provide some more insight on Algorithm B.3.

Algorithm B.3 Slice sampler



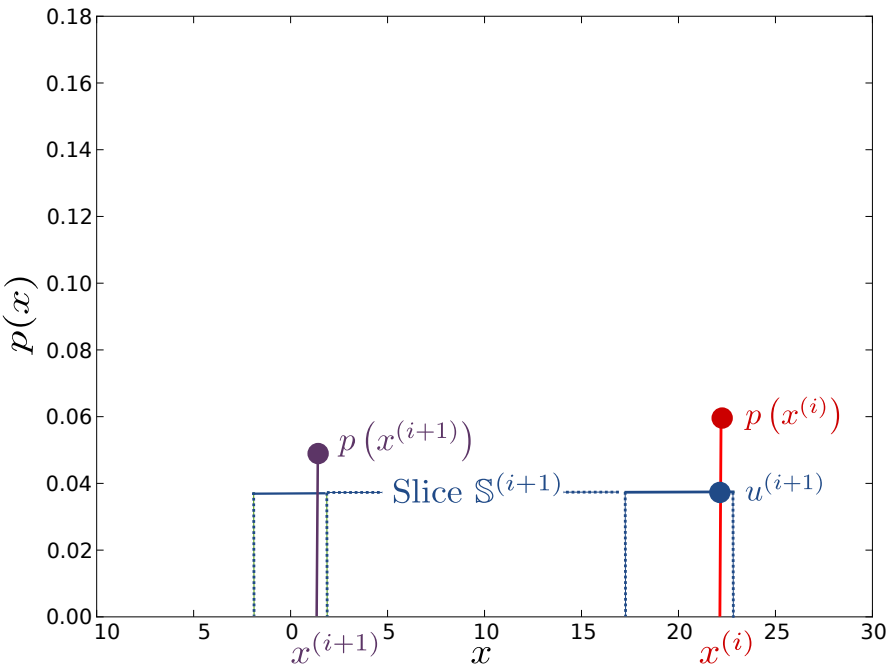


Figure B.5: A bimodal target PDF featuring two distinct modes: $p(x) \propto 0.7\exp(-0.2\,x^2) + 0.3\exp(-0.2\,(x-20)^2)$. This pathological case illustrates the efficiency of the slice sampler described in Algorithm B.3 with respect to the Metropolis-Hastings sampler.

B.3.3.3 Implementation

The actual slice $\mathbb{S}^{(i+1)}$ is not easy to identify in practice, because the problem of finding its boundaries is a non-trivial root finding problem. Although, Neal (2003) came up with a heuristic procedure which makes the approach scalable to a wide variety of target PDFs.

To make it short, his heuristic consists in approximating the slice $\mathbb{S}$ with an hyperrectangle whose widths are first expanded by unit increments w from the current state of the chain $\boldsymbol{x}^{(i)}$. Note that this procedure, called *step-out*, might fail if the actual slice is narrower than the chosen increment w. Then it simply draws a point $\boldsymbol{x}^{\star}$ at random (uniformly) in the previously built hyperrectangle. If the proposed candidate does not lie in the actual slice (*i.e.* if $p(\boldsymbol{x}^{\star}) < u$), then the hyperrectangle bounds is reset with the coordinates of the (rejected) candidate point. This second procedure, called *shrink-in*, might also fail if the target PDF is too nasty (highly multimodal) so that Neal's heuristic is usually aborted as it reaches a maximum number of *shrink-in* iterations (say 100).

The interested reader is referred the original article by Neal (2003) and the book by Robert and Casella (2004, Chapter 8) for a more detailed review of this technique together with a discussion on its convergence properties. In this thesis, slice sampling was only used for sampling from the so-called refinement criteria defined in Chapter 2 on purpose to fill a region of interest with candidate points. Hence, the applicability of Theorem B.2.1 on Markov chains generated with this sampler is beyond the scope of this thesis.

Abstract

This thesis is a contribution to the resolution of the reliability-based design optimization problem. This probabilistic design approach is aimed at considering the uncertainty attached to the system of interest in order to provide optimal and safe solutions. The safety level is quantified in the form of a probability of failure. Then, the optimization problem consists in ensuring that this failure probability remains less than a threshold specified by the stakeholders. The resolution of this problem requires a high number of calls to the limit-state design function underlying the reliability analysis. Hence it becomes cumbersome when the limit-state function involves an expensive-to-evaluate numerical model (*e.g.* a finite element model). In this context, this manuscript proposes a surrogate-based strategy where the limit-state function is progressively replaced by a Kriging meta-model. A special interest has been given to quantifying, reducing and eventually eliminating the error introduced by the use of this meta-model instead of the original model. The proposed methodology is applied to the design of geometrically imperfect shells prone to buckling.

Keywords: adaptive surrogate modelling • Kriging • Gaussian processes for regression and probabilistic classification • reliability analysis • rare event probabilities • importance sampling • reliability-based design optimization • probabilistic buckling • geometrically imperfect shells

Résumé

Cette thèse est une contribution à la résolution du problème d'optimisation sous contrainte de fiabilité. Cette méthode de dimensionnement probabiliste vise à prendre en compte les incertitudes inhérentes au système à concevoir, en vue de proposer des solutions optimales et sûres. Le niveau de sûreté est quantifié par une probabilité de défaillance. Le problème d'optimisation consiste alors à s'assurer que cette probabilité reste inférieure à un seuil fixé par les donneurs d'ordres. La résolution de ce problème nécessite un grand nombre d'appels à la fonction d'état-limite caractérisant le problème de fiabilité sous-jacent. Ainsi, cette méthodologie devient complexe à appliquer dès lors que le dimensionnement s'appuie sur un modèle numérique coûteux à évaluer (*e.g.* un modèle aux éléments finis). Dans ce contexte, ce manuscrit propose une stratégie basée sur la substitution adaptative de la fonction d'état-limite par un méta-modèle par Krigeage. On s'est particulièrement employé à quantifier, réduire et finalement éliminer l'erreur commise par l'utilisation de ce méta-modèle en lieu et place du modèle original. La méthodologie proposée est appliquée au dimensionnement des coques géométriquement imparfaites soumises au flambement.

Mots-clés: méta-modélisation adaptative • Krigeage • régression et classification probabiliste par processus Gaussiens • analyse de fiabilité • probabilités d'évènements rares • échantillonnage préférentiel • optimisation sous contrainte de fiabilité • flambage probabiliste • coques géométriquement imparfaites