Chance constraint optimization of a complex system : Application to the design of a floating offshore wind turbine

Alexis Cousin

▶ **To cite this version:**

Alexis Cousin. Chance constraint optimization of a complex system : Application to the design of a floating offshore wind turbine. Optimization and Control [math.OC]. Institut Polytechnique de Paris, 2021. English. NNT : 2021IPPAX072 . tel-03500604

HAL Id: tel-03500604

https://theses.hal.science/tel-03500604v1

Submitted on 22 Dec 2021

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

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

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2021IPPAX072

Thèse de doctorat

ÉCOLE POLYTECHNIQUE

IP PARIS

Optimisation sous contraintes probabilistes d'un système complexe – Application au dimensionnement d'une éolienne offshore flottante

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°574 Ecole Doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Rueil-Malmaison, le 01/10/2021, par

ALEXIS COUSIN

Composition du Jury :

Bruno Sudret Professor, ETH Zürich (Department of Civil, Environmental and Geomatic Engineering)	Président
Michael Muskulus Professor, Norwegian university of science and technology (Department of Civil and Environmental Engineering)	Rapporteur
Jean-Marc Bourinet Maître de conférences, SIGMA Clermont (Pôle Structures et Mécanique des Matériaux)	Rapporteur
Emmanuel Gobet Professor, Ecole Polytechnique (Centre de Mathématiques Appliquées)	Examinateur
Céline Helbert Maître de conférences, École Centrale de Lyon (Département mathématiques-informatique)	Examinatrice
Josselin Garnier Professor, Ecole Polytechnique (Centre de Mathématiques Appliquées)	Directeur de thèse
Miguel Munoz Zuniga Ingénieur de recherche, IFP Energies Nouvelles (Département mathématiques appliquées)	Co-encadrant de thèse
Martin Guiton Ingénieur de recherche, IFP Energies Nouvelles (Département mécanique des solides)	Co-encadrant de thèse

Acknowledgements

The writing of these acknowledgements marks the end of the PhD and it is with great gratitude that I look back on the last 3 years. Although the format of this PhD has been a bit unusual due to COVID, I am very satisfied with this experience. I therefore begin by thanking the people who contributed the most to the success of this PhD: my supervisor Josselin and my advisors Miguel and Martin. Our way of working was perfect for me: a good balance between autonomy and monitoring of my progress. You have always been very available. Your support and quick feedback made my job easier, especially during the last weeks of writing the manuscript which were a bit busy... You were invested in the success of this work but also in my career by pushing me to participate in conferences and following my job search and I thank you for that too.

I would like to thank all the members of the jury and especially the reviewers Dr. Bourinet and Pr. Muskulus for their careful reading of the manuscript and their suggestions for improvement. Our exchanges have enriched this thesis.

I thank the colleagues at IFPEN for the friendly atmosphere at R11, in particular Delphine for sharing her expertise in optimization and Sylvie Hoguet for her help in all administrative aspects. I also thank Yann Poirette from IFPEN and Jean-Michel Heurtier and Cedric LeCunff from Principia who worked on the DeeplinesWind^{TM} code and the definition of the problem.

I warmly thank Aziz with whom I shared the office during the first half of the thesis, for our interesting and stimulating conversations and who allowed me to get my head out of my thesis.

I would like to thank all the PhD students of R11 that I met during these 3 years: Nicolas, Bastien, Karine, Julien, Zakaria, Sabrina, Thoi, Guissel, Joelle, Karim, and Jingang. I am very happy to have shared these 3 years with you. I also take this opportunity to thank the CMAP PhD students who welcomed me during my few visits to CMAP. Congratulations again to all the new "docteurs" and courage to those who will be soon!

Finally, I deeply thank my friends and family for their support during this PhD and the previous years.

Résumé

Parmi les sources d'énergie marine en cours de développement, les éoliennes offshores flottantes offrent de nombreux avantages. Elles permettent d'étendre la zone d'exploitation, là où le vent est plus fort et plus régulier. Néanmoins, la démonstration de leur faisabilité est complexe et la fiabilité de ces structures doit être garantie.

Cette fiabilité est notamment assurée par le système d'ancrage du support flottant qui limite les mouvements de l'éolienne. Nous cherchons dans cette thèse à proposer une approche permettant d'optimiser la configuration des lignes d'ancrage, en minimisant le coût des matériaux tout en respectant des contraintes d'un état limite de fatigue. Conformément aux normes internationales en état limite de fatigue, le système d'ancrage doit limiter les mouvements du flotteur pour assurer la production de l'éolienne, éviter la compression dans les lignes d'ancrage et résister aux dégâts causés par la fatigue des matériaux avec une probabilité annuelle de défaillance inférieure à 10^{-4}. En effet, ces contraintes héritent du caractère aléatoire des conditions environnementales ainsi que d'incertitudes sur des paramètres du modèle.

Par conséquent, nous sommes confrontés à un problème d'optimisation avec une fonction coût déterministe et des contraintes impliquant des probabilités de dépassement de seuil de maximum et d'intégrale sur une période $[0, T]$ de processus aléatoires dépendant du temps.

La principale difficulté est de devoir évaluer ces probabilités à chaque boucle de l'algorithme d'optimisation. Une approche naïve telle que la méthode de Monte Carlo nécessite de calculer avec un simulateur coûteux en temps de calcul, de nombreuses réalisations des processus aléatoires. Malheureusement, le coût de calcul d'une seule réalisation est trop élevé pour appliquer une telle approche. L'estimation de ces probabilités est d'autant plus difficile que nous sommes confrontés à des événements rares.

Pour résoudre efficacement ce problème, nous proposons une méthodologie en deux étapes. Premièrement, en considérant que T est suffisamment grand, nous utilisons les propriétés des contraintes et les théorèmes limite de la théorie des valeurs extrêmes et de la théorie ergodique pour reformuler les contraintes initiales en contraintes indépendantes du temps. Nous obtenons ainsi un problème équivalent pour lequel les algorithmes classiques sont peu performants. La deuxième étape de notre procédure consiste à résoudre le problème reformulé avec une nouvelle méthode basée sur une stratégie de krigeage adaptative. Cette méthode est appelée AK-ECO pour Adaptive Kriging for Expectation Constraints Optimization et permet de résoudre efficacement des problèmes d'optimisation avec des contraintes faisant intervenir des espérances.

Le cas académique d'un oscillateur harmonique présentant toutes les caractéristiques du

problème industriel est introduit afin d'illustrer notre méthodologie. La procédure est ensuite appliquée avec succès au problème de l'éolienne flottante.

Bien que cette méthodologie soit proposée dans le but de résoudre ce problème, les deux étapes qui la composent sont introduites dans un cadre général afin de pouvoir être appliquées à d'autres problèmes d'optimisation impliquant des contraintes probabilistes dépendant de maximum et d'intégrale de processus aléatoires.

Abstract

Among the marine sources of energy under development, the floating offshore wind turbine (FOWT) solution offers a panel of advantages. It enables to extend the exploitation area where the wind is stronger and steadier. Nevertheless, its feasibility demonstration is complex and the reliability of the FOWT must be guaranteed.

The reliability of the structure is ensured especially by the mooring system of the floating support which restricts the wind turbine motions. The work of this thesis is to propose a feasible methodology to optimize the configuration of the mooring lines by minimizing the material cost while satisfying Fatigue Limit State (FLS) constraints. In accordance with international design standards in FLS, the mooring system must limit the floater movements to ensure the turbine production, avoid compression in the mooring lines and withstand the damage caused by fatigue, with an annual failure probability less than 10^{-4}. Indeed, these constraints inherit the randomness of the marine environment as well as uncertainties on material properties and model parameters.

Therefore, we face an optimization problem with a deterministic cost function and constraints involving probabilities of threshold exceedance of the maximum and the integral over a period $[0, T]$ of time-dependent random processes.

Having to evaluate these failure probabilities at each loop of the optimization algorithm is the main difficulty. A naïve approach such as the Monte Carlo method requires computing with a time-consuming simulator, many realizations of the random processes. Unfortunately, the computation cost of a single realization is too high to apply such basic approaches. The estimation of these probabilities is all the more challenging as we are dealing with rare events.

To solve this problem efficiently, we propose in this thesis a two-step methodology. First, considering that T is sufficiently large, we use the properties of the constraints and limit theorems of the extreme value theory and the ergodic theory to reformulate the original constraints into time-independent ones. We thus obtain an equivalent problem for which classical algorithms perform poorly. The second step of the procedure consists in solving the reformulated problem with a new method based on an adaptive kriging strategy well suited to the reformulated constraints. This method, called AK-ECO for Adaptive Kriging for Expectation Constraints Optimization, enables to efficiently solve optimization problems with constraints involving expectations.

An academic case of a harmonic oscillator presenting all the characteristics of the industrial problem is introduced to illustrate the methodology. The procedure is then applied with success to the FOWT problem.

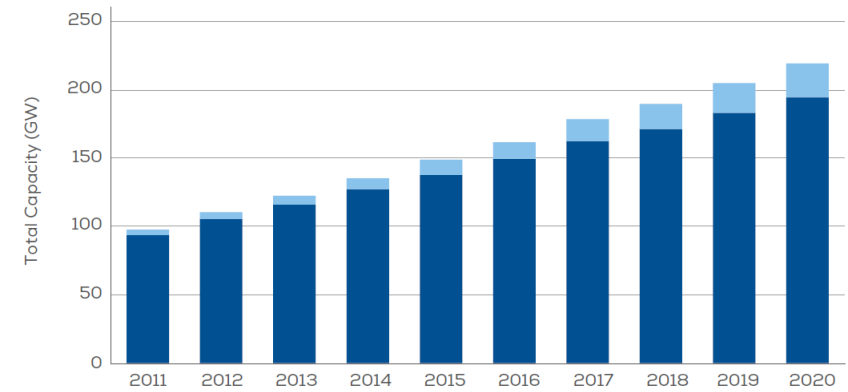
Although this methodology is proposed to solve this problem, the two steps that compose it are introduced in a general framework so that they can be applied to other optimization problems involving probabilistic constraints depending on maximum and integral of random processes.

Contents

Introduction

1 Context

The wind power capacity installed in Europe has doubled over the last decade (see Figure 1) and has reached 220 GW including 25 GW (11,4%) offshore (WindEurope (2020)). Wind energy currently meets 16.4% of electricity demand across EU and UK (WindEurope (2020)).



	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Offshore	4	5	7	8	11	13	16	18	22	25
Onshore	94	105	116	127	137	149	162	171	183	195
Total	**98**	**110**	**123**	**135**	**148**	**162**	**178**	**189**	**205**	**220**

Figure 1: Evolution of the total wind installations in Europe (WindEurope (2020))

Although bottom-fixed offshore wind turbine technology has made it possible to exploit the offshore wind resources, it is limited to land and coastal areas. These structures, directly fixed to the seabed, are not economically attractive for sea depths greater than 60m (WindEurope (2017)).

To extend the exploitation area, the complementary technology of floating offshore wind turbine (FOWT) was then proposed. Underwater foundations are avoided by equipping the wind turbine with a floating structure connected to the seabed by an anchoring system that stabilizes it.

There are currently three dominant structure designs under development (see Figure 2):

- semi-submersible. The floating platform is semi-submerged and is anchored to the seabed by catenary mooring lines composed of steel chains, steel wire ropes or synthetic ropes;

- Spar-buoy. The stability of the structure is ensured by its center of gravity lower in the water than its center of buoyancy. Such a floater is also equipped with catenary mooring lines;
- Tension leg platform. This semi-submerged floating structure is anchored to the seabed by taut mooring lines with high tension.

Figure 2: From left to right: Spar-buoy - Semi-submersible - Tension leg platform (Carbon Trust (2015))

Floating wind turbines can access areas further offshore where wind blows stronger and its flow is more consistent which significantly reduces the cost of electricity production (WindEurope (2017)). In addition, far-from-shore projects reduce the visual and noise footprint for local residents.

This technology has many advantages and its development would be facilitated by reducing its manufacturing costs and ensuring the long-term reliability of the structure.

2 Problem statement

To assess the competitiveness of different energy technologies, the levelized cost of energy (LCOE) measures the cost of a unit of energy produced taking into account all the costs occurring over their lifetimes. For floating wind turbines, the manufacturing cost of the mooring system is a significant contributor to the LCOE (Yu et al. (2018)). Besides, several constraints are usually imposed by international standards (Det Norske Veritas (2013); IEC 61400-3 (2009)) for the design of the mooring system to ensure the reliability of the structure. These constraints involve different type of limit states:

- the Ultimate Limit State (ULS) which considers the capacity of the individual mooring lines to resist the load effect imposed by extreme environmental excitations with 50-year return period;

- the Accidental Limit State (ALS) that ensures that the mooring system is able to withstand the failure of one mooring line or one thruster failure for unknown reasons;
- the Fatigue Limit State (FLS) which deals with the damage accumulation of the individual mooring lines due to cycle loading during the FOWT service life.

In this context, the objective of this thesis is to propose an efficient method to optimize the configuration of the mooring lines by minimizing their material cost while respecting constraints imposed by international standards (Det Norske Veritas (2013)). To identify such a method, we choose to make several simplifications. In this thesis, as a first step towards a complete approach covering all the design limit states, we consider only the FLS constraint and wind loads are represented by constant values, neglecting the turbulence variation. We then consider that an acceptable configuration must limit the floater movements to ensure the turbine production, avoid compression in the mooring lines, and withstand accumulated damage caused by marine conditions, with annual failure probability less than 10^{-4} (Det Norske Veritas (2013)). Indeed, to propose a reliable solution, different sources of uncertainty are considered on model parameters which are represented by random variables. Moreover, the motion of the structure results from waves impacting the platform. These waves are also considered random and are modeled by a time-dependent process which leads to motions of the structure described by random processes. Therefore, we face an optimization problem with a deterministic cost function and probabilistic constraints depending on random variables and time-dependent processes.

The main difficulty to solve this problem is the estimation of these probabilities at each iteration of the optimization algorithm since:

- time-consuming simulations are required to compute the structure motions over time;
- the methods estimating probability need many simulations, especially when the threshold probabilities are low.

3 Content of the thesis

The work of this thesis lies in the research field of reliability optimization. An overview of the current state of the art in this domain is proposed in chapter 1. In this chapter, the mathematical formulation of the problem under study is briefly introduced to highlight its main features.

To solve this problem efficiently, we propose a methodology in two steps. Instead of tackling the problem with constraints involving time-dependent probabilities, we first use the properties of the processes and limit theorems to reformulate the probabilities into time-independent ones that are easier to estimate. This reformulation procedure is detailed in chapter 2. In this chapter, we introduce an academic case based on a harmonic oscillator presenting all the characteristics of the industrial problem.

The second step of the methodology is to solve the reformulated problem. Confronted with limitations of methods of the literature, we introduce in chapter 3 a new optimization algorithm based on an adaptive kriging procedure suited to the reformulated problem and called AK-ECO (for Adaptive Kriging for Expectation Constraints Optimization). This method is then validated on the academic problem and a comparison with state-of-the-art

methods introduced in chapter 1 is carried out.

The two steps of this methodology are introduced in a general framework in order to be applied to other problems presenting the same characteristics as the problem under study in this thesis.

In chapter 4, two analyses are proposed to complete our methodology. A global enrichment procedure is presented. This approach builds accurate initial metamodels for AK-ECO and aims to ensure a fast convergence of the algorithm. Moreover, a new sensitivity analysis adapted to probabilistic constraints is also proposed and provides a better understanding of the influence of each uncertainty on the constraints.

The FOWT problem which considers a case study inspired by the semi-submersible FOWT of (Robertson et al. (2014a)), is presented in more details in chapter 5 and is solved with the two-step methodology introduced in chapters 2 and 3.

Finally, we will highlight the main results and discussed the perspectives of our work in the conclusion of this thesis.

Chapter 1

State of the art

Contents

Due to a lack of knowledge or to inherently random physical phenomena, it is common in engineering to consider uncertainties on the inputs of a numerical model that propagate to its outputs. Several fields of research have emerged with the introduction of uncertainties in models. In **Reliability Analysis** (RA), the objective is to evaluate the probability that a certain event, considered as a failure, occurs. In the context of optimization, we refer to a **Reliability-Based Design Optimization** (RBDO) problem when the constraints involve failure probabilities. The RA and RBDO problems are said to be time-dependent when the model inputs include time-dependent random processes.

Taking uncertainties into account usually leads to many calls to the simulator with different realizations of the random inputs in particular when one deals with probabilities of rare events (Bourinet (2018)). One of the main challenges in RA and RBDO is thus to

solve the considered problem when each simulation is computationally expensive.

The problem under study in this thesis falls within these domains of research and therefore a review of the literature in proposed in this chapter. In the last section of the chapter, we will briefly introduce the FOWT problem to highlight its main characteristics.

1 Reliability Analysis

In the context of reliability analysis, the uncertainties are represented by a random vector $X = (X_1, \cdots, X_{n_X})$ taking values in $\Omega_X \subset \mathbb{R}^{n_X}$. In this chapter, we consider that this vector is composed of n_X continuous random variables with known probability density functions and we assume that its marginal distribution functions are strictly increasing. The quantity of interest is the failure probability defined as:

$$p_f = \mathbb{P}\left(g(X) < 0\right) \tag{1.1}$$

where $g : \mathbb{R}^{n_X} \to \mathbb{R}$ is called the **performance function** and failure occurs when the function is negative. One evaluation of the performance function often requires a call to a time-consuming simulator. Therefore, the methods proposed in RA aim at evaluating the failure probability accurately and with as few calls to g as possible. Since the function g is usually complex, the distribution of $g(X)$ cannot be calculated analytically and a numerical approximation of the failure probability is adopted.

We say that we deal with a **rare event** when p_f is small (in our context $p_f < 10^{-3}$). Estimating probability of occurrence of rare events is especially difficult because it requires many calls to the simulator to be accurate.

Definition 1.1. The performance function g (also called **limit-state function**) defines three domains of Ω_X:

- the **failure domain**: $\{x \in \Omega_X, g(x) < 0\}$;
- the **safe domain** : $\{x \in \Omega_X, g(x) > 0\}$;
- the **limit-state surface**: $\{x \in \Omega_X, g(x) = 0\}$.

The failure probability is therefore the probability that X belongs to the failure domain.

Three main approaches have been proposed to estimate p_f: the **approximation methods**, the **simulation methods** and **metamodel-based** methods. We introduce theses approaches below.

Since the random vector X has a known probability density function, the failure probability can be written as a multidimensional integral. Thus, we could also estimate p_f with a **quadrature** method (chapter 7 of Owen (2013) discusses the most popular ones). However, in practice, this approach is rarely considered since it is too expensive in number of calls to the function g.

1.1 Approximation methods

The approximation methods consider p_f defined in (1.1) as the following probability:

$$p_f = \mathbb{P}(G(U) < 0) \tag{1.2}$$

where U is a vector composed of independent standard Gaussian random variables. To obtain (1.2), we introduce the **isoprobabilistic transformation** T which is a diffeomorphism from Ω_X to $\mathbb{R}^{n_X}$ such that:

$$T^{-1}(U) = X \tag{1.3}$$

and

$$\mathbb{P}\left(g(X) < 0\right) = \mathbb{P}\left(g \circ T^{-1}(U) < 0\right). \tag{1.4}$$

We denote $G(u) = g \circ T^{-1}(u)$. When $X_1, \cdots, X_{n_X}$ are independent random variables with respective cumulative distribution functions $F_1, \cdots, F_{n_X}$, T is defined as follows:

$$T(X) = \left(\Phi^{-1}\left(F_1(x)\right), \cdots, \Phi^{-1}\left(F_{n_X}(x)\right)\right) \tag{1.5}$$

with Φ the standard Gaussian cumulative distribution function. We refer to Bourinet (2018) and Lebrun (2013) for the definition of T when the coordinates of X are not independent.

The approximation methods rely on the identification of the **Most Probable failure Point** (MPP) denoted u^* which is solution (assumed to exist and to be unique) of the following optimization problem:

$$u^* = \underset{\mathbb{R}^{n_X}}{\arg\min} \ \{\|u\|, G(u) \leq 0\}. \tag{1.6}$$

The solution of problem (1.6) cannot be determined analytically. Several algorithms have been proposed to efficiently solve this problem (Bourinet (2018); Yang et al. (2020b)).

Assuming that G is differentiable at u^*, the **First-Order Reliability Method** (FORM) (Hasofer and Lind (1974); Madsen et al. (2006)) considers a linear approximation $\widetilde{G}$ of G at u^* using a first-order Taylor expansion. The failure probability $\mathbb{P}(G(U) < 0)$ is then approximated by $\mathbb{P}(\widetilde{G}(U) < 0)$ since the latter can be deduced analytically (see proposition 1 of Lemaire (2013)):

$$\mathbb{P}\left(\widetilde{G}(U) < 0\right) = \Phi\left(-\|u^*\|\right). \tag{1.7}$$

The norm $\|u^*\|$ is called the **reliability index** and is often denoted β.

Figure 1.1 illustrates the strategy of FORM for a reliable analysis in 2D. The contour lines of the probability density function of X (resp. U) are displayed in blue on the left (resp. on the right). After identifying the MPP, the FORM approximation is performed at this point and the failure probability is estimated from β.

The **Second-Order Reliability Method** (SORM) (Breitung (1984, 1989))is based on the same principal except that G is assumed to be twice differentiable at u^* and a second-order Taylor expansion of G at the MPP is used. We refer to Bourinet (2018) for a detailed description of the FORM and SORM methods.

In the approximation methods, the performance function g is only called during the resolution of problem (1.6) which makes the estimation of the failure probability effective. However, these approaches suffer from two main drawbacks:

1. their accuracy relies on the regularity of the function G. The error committed by replacing G by its approximation $\widetilde{G}$ is not controlled. Therefore, an inaccurate approximation of p_f can be obtained if G is highly nonlinear;

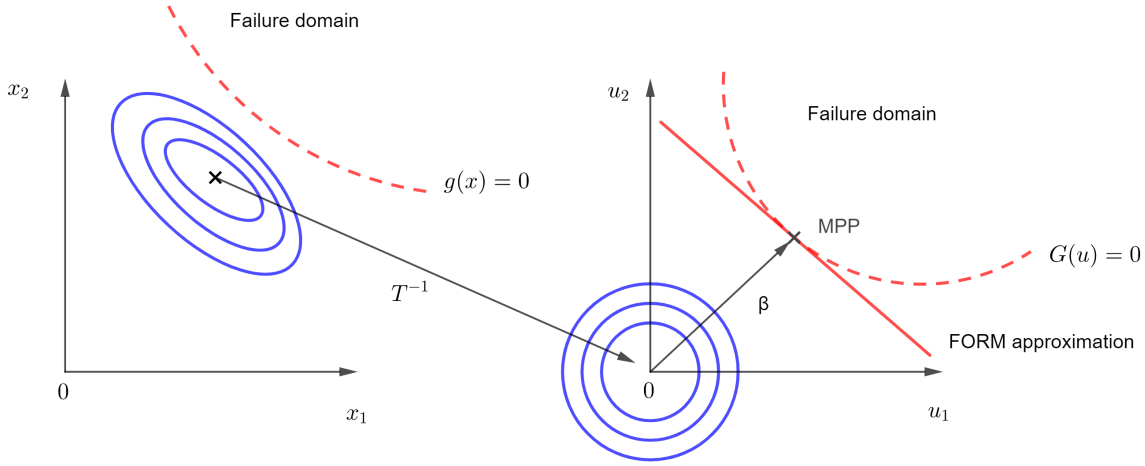


Figure 1.1: FORM approximation (Lopez and Beck (2012))

2. despite the speed of execution of these methods, the resolution of the optimization problem (1.6) sometimes requires a prohibitive number of calls to g.

1.2 Simulation methods

The **Monte Carlo** method enables to estimate an expectation $\mathbb{E}[Y]$ where Y is a random variable from which it is possible to draw realizations. The Monte Carlo estimator I_{MC} is defined as follows:

$$I_{MC} = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} Y_i \tag{1.8}$$

where $(Y_1, \cdots, Y_{n_{MC}})$ is a sequence of i.i.d random variables with the same distribution as Y and n_{MC} is the size of the Monte Carlo sample. The Monte Carlo estimator is unbiased and converges almost surely to $\mathbb{E}[Y]$. Its coefficient of variation $\delta_{n_{MC}}$ is given by the following equation:

$$\delta_{n_{MC}} = \frac{\sqrt{\mathrm{Var}[I_{MC}]}}{\mathbb{E}[I_{MC}]} = \frac{1}{\sqrt{n_{MC}}} \frac{\sqrt{\mathrm{Var}(Y)}}{\mathbb{E}[Y]}. \tag{1.9}$$

In RA, we consider $Y = \mathbb{1}_{g(X)<0}$ and thus, $\delta_{n_{MC}} = \sqrt{\frac{1-p_f}{n_{MC}\, p_f}}$.

The Monte Carlo method is versatile since the only condition required is that Y has a finite variance. It also possesses the good property that, if Y is a function of a random vector, as it is the case in RA, the speed of convergence of the estimator does not depend on the dimension of this vector. However, as Bourinet (2018) points out, the coefficient of variation of the estimator is proportional to $\frac{1}{\sqrt{n_{MC}}}$ which leads to a slow convergence of the Monte Carlo method, especially when one deals with rare events. Hence, for an estimation with a coefficient of variation of 10%, approximately $100/p_f$ realizations are necessary.

Different simulation techniques have been proposed to keep the good properties of the Monte Carlo estimator and increase the rate of convergence. The **Importance Sampling** (IS) (Melchers (1989)) and **Subset Simulation** (SS) (Au and Beck (2001)) are among the most used in RA. We refer to Bourinet (2018) for a detailed description of these techniques. More recent developments are proposed in Munoz Zuniga (2011), Yun et al. (2018), Papaioannou et al. (2019), Papaioannou and Straub (2021), and Rashki

(2021).

Although the improvements of the Monte Carlo method significantly increase the convergence speed of the simulation method, the number of realizations of the variable Y required often remains prohibitive when one evaluation of g is time-consuming.

1.3 Metamodel-based approach

Metamodels

Let consider an expensive model $M : x \in \Omega_x \subset \mathbb{R}^{n_x} \to M(x) \in \mathbb{R}$. When an analysis requires many evaluations of M, the strategy of substituting M by an approximated and fast-to evaluate model $\widetilde{M}$ is often adopted in engineering. This latter model is called **metamodel** (also known as **surrogate model**, **emulator** or **response surface model**). Among the most widespread metamodels, we can refer to the **polynomial approximations** (Myers et al. (2016)), the **Artificial Neural Networks** (ANN) (Lippmann (1987)), the **Radial Basis Functions** (RBF) (Buhmann (2003)), the **Support Vector Machines** (SVM) (Moustapha (2016)), the **Polynomial Chaos Expansion** (PCE) (Wiener (1938)), and the **Gaussian process regression** also called **kriging** (Krige (1951); Rasmussen and Williams (2006)).

These metamodels are defined by a set of hyperparameters that must be calibrated. To do so, the function M is evaluated over a sample of its input space called **Design of Experiment** (DoE) and the responses are used to fit the metamodel.

Design of experiment

Different classes of DoEs exist and differ in the way they span the input space (a review of these classes is provided in Alizadeh et al. (2020)). A DoE optimally covering the entire domain in order to capture the general form of the function M is called **space-filling**. Among the most popular DoE, the **Latin Hypercube Sampling** (LHS) (McKay et al. (1979)) allows to fill the input space such that the one-dimensional projections of the points of experiment on each component are uniformly distributed. To compute a LHS of n_{LHS} points in the space $[0,1]^{n_x}$, for each dimension $i = 1, \cdots, n_x$, the interval $[0,1]$ is divided into n_{LHS} subintervals $I_1^i \cdots, I_{n_{LHS}}^i$ with the same length. The DoE is a LHS if for each dimension i and each interval I_j^i, a unique point of the design of experiments has its i-th coordinate in the interval I_j^i. Figure 1.2 illustrates an example of a 5 points LHS over the space $[0,1] \times [0,1]$.
To ensure the space-filling condition, it is common to select the DoE which optimizes a criterion such as the maximin criterion (Johnson et al. (1990)) aiming to maximize the shortest distance between two points of experiment.

Gaussian process regression or kriging

We present the kriging model in more detail since this technique is at the core of the work of this thesis. The standard method relies on the assumption that the expensive function M is a realization of a stationary Gaussian random process $\widetilde{M}$ defined as:

$$\widetilde{M}(x) = \sum_{j=1}^{p} \beta_j f_j(x) + Z(x) \tag{1.10}$$

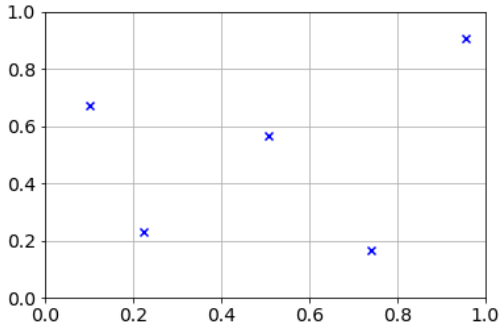


Figure 1.2: LHS of 5 points of $[0,1] \times [0,1]$

where the sum $\sum_{j=1}^{p} \beta_j f_j(x)$ defines the trend of the process characterized by the unknown coefficients $\beta_1, \cdots, \beta_p$ and the known functions $f_1, \cdots, f_p$. Besides, Z is a stationary Gaussian process with zero mean and its covariance function is given by:

$$\mathbb{E}[Z(x)Z(x')] = \sigma^2 C_\theta(x - x'). \tag{1.11}$$

The value of p, the functions f_j $(j = 1, \cdots, p)$ as well as the function C are chosen by the user (the latter can be chosen from a family of parametric correlation functions (Rasmussen and Williams (2006))). The unknown parameters σ^2 and θ are called the **hyper-parameters** and characterize the correlation between two points of the process Z. They represent respectively the variance and the correlation length of Z.

Let $\{x_i, i = 1, \cdots, n_{DoE}\}$ be a DoE of Ω_x. We introduce the following notations:

$$f(x) = [f_j(x)]_{1 \le j \le p} \quad , \quad F = [f_j(x_i)]_{1 \le i \le n_{DoE}, 1 \le j \le p} \tag{1.12}$$

$$r_\theta(x) = [C_\theta(x_i - x)]_{1 \le i \le n_{DoE}} \quad , \quad R_\theta = [C_\theta(x_i - x_j)]_{1 \le i \le n_{DoE}, 1 \le j \le n_{DoE}} \tag{1.13}$$

and m the vector $[M(x_i)]_{1 \le i \le n_{DoE}}$.

The DoE and m are used to fit σ and θ usually with the cross-validation method (Lataniotis et al. (2017)) or the maximum-likelihood (ML) method (Roustant et al. (2012)).

The parameter β can be treated in two ways, both approaches leading to similar results (Helbert et al. (2009)). The first approach is to evaluate β with ML whose solution is given by:

$$\hat{\beta} = \left(F^T R_\theta^{-1} F\right)^{-1} F^T R_\theta^{-1} m. \tag{1.14}$$

On the other hand, the Bayesian approach consists in assuming a prior Gaussian distribution of β with a covariance matrix B and to evaluate the posterior expectation of β conditionally to the DoE and m. In this case and when $\|B^{-1}\| \to 0$, the mean of the posterior distribution tends to $\hat{\beta}$.

Whether the ML or the Bayesian approach is considered for the definition of $\hat{\beta}$, it is shown that the prediction of the metamodel at a new point x_0, conditionally to the DoE and m, follows a normal distribution $\mathcal{N}\left(\mu_{\widehat{M}}(x_0), \sigma^2_{\widehat{M}}(x_0)\right)$ with:

$$\mu_{\widehat{M}}(x_0) = f(x_0)^T \hat{\beta} + r_\theta(x_0)^T R_\theta^{-1}(m - F\hat{\beta}) \tag{1.15}$$

and

$$\sigma^2_{\widehat{M}}(x_0) = \sigma^2 \left(1 - r_\theta(x_0)^T R_\theta^{-1} r_\theta(x_0) + u_\theta(x_0)^T \left(F^T R_\theta^{-1} F\right)^{-1} u_\theta(x_0)\right) \tag{1.16}$$

with $u_\theta(x_0) = F^T R_\theta^{-1} r_\theta(x_0) - f(x_0)$.

The mean $\mu_{\widehat{M}}(x_0)$ of this random variable is used as predictor of M at x_0 while the standard deviation $\sigma_{\widehat{M}}(x_0)$ measures the accuracy of this prediction.

It follows from equations (1.15) and (1.16) that the kriging method is an interpolation technique: for any point x of the DoE used for its calibration, the prediction of the metamodel $\mu_{\widehat{M}}(x)$ equals $M(x)$ and its variance $\sigma^2_{\widehat{M}}(x)$ is zero.

We have introduced in this section the noise-free kriging model since this model will be used throughout this thesis. For the introduction of the kriging model with noisy observations, we refer to Rasmussen and Williams (2006).

Use of metamodels in RA and active learning

In recent years, many articles have adopted the strategy of substituting the costly performance function for a metamodel. Thereby, the failure probability can be estimated with simulation methods since the sample size is no longer an issue. However, this approach can lead to large errors if the metamodel fitting is of poor quality. Thus, an initial calibration of the metamodel is usually followed by an adaptive enrichment procedure (also called **active learning**). To improve the metamodel, new points are selected and added to the DoE. The identification of those points is generally done by maximizing a criterion called **learning function** over the sample space of X or among the sample of the simulation method. The performance function is then evaluated at those points and the responses are used to recalibrate the metamodel. Since the accuracy of the failure probability estimation is mainly due to the ability of the metamodel to correctly predict whether a point belongs to the failure domain, the learning function is usually a tradeoff between the proximity of a point to the limit-state surface and the uncertainty of the metamodel at that point.

In RA, many adaptive methods have been proposed and vary according to the choice of metamodel, the learning function, the enrichment stopping criterion, and the simulation technique used to estimate the failure probability. For instance, SVM (Roy and Chakraborty (2020); Ling and Lu (2021)), PCE (Marelli and Sudret (2018); Cheng and Lu (2020)) and ANN (Xiang et al. (2020)) have been applied to estimate the failure probability with suited active learning methods.

The kriging model is particularly well adapted to active learning since it provides a measure of the prediction uncertainty at each point of the input space. This feature explains the intensive use of this technique in RA. Many learning functions have been proposed, namely the EFF function (Bichon (2010)), the U function in AK-MCS (Echard et al. (2011); Echard (2012)), the H function (Lv et al. (2015)), the REIF2 criterion (Zhang et al. (2019)) or the LIF function (Sun et al. (2017)) with adapted enrichment stopping conditions. In these methods, one point is selected at each enrichment step but multipoint enrichment has also been proposed as in AK-MCSi (Lelièvre et al. (2018)).

An extensive review and a comparison of these adaptive approaches for different metamodels and active learning strategies are detailed in Teixeira et al. (2021).

2 Reliability-Based Design Optimization

A RBDO problem is an optimization problem aiming at minimizing a deterministic cost function and whose constraints involve failure probabilities. It is formulated as follows:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \text{ such that} \\ &\mathbb{P}_{X_d,X_p}\left(g_i\left(X_d,X_p\right)<0\right)<p_{s,i} \quad (i=1,\cdots,n_c) \end{aligned} \tag{1.17}$$

where $\Omega_d \subset \mathbb{R}^{n_d}$ is the design space, d the design variables, and X_d, X_p are two random vectors with respective size n_d and n_p. The distribution of X_d depends on d contrary to the distribution of X_p. The i-th constraint is satisfied if the corresponding failure probability is less than a threshold probability $p_{s,i}$.

Notation. In this thesis, the notations $\mathbb{P}_X$ and $\mathbb{E}_X$ mean that the probability and the expectation are considered with respect to the distribution of X (X can be a random variable, a random vector or a random process).

Many RBDO problems have a cheap-to-evaluate cost function and the difficulty lies in the estimation of the failure probabilities at each iteration of the optimization algorithm. Many methods have been proposed to solve (1.17) while limiting as much as possible the number of evaluations of the performance functions g_i $(i = 1,\cdots,n_c)$. They can be grouped into four different approaches: the **double-loop** methods, the **single-loop** methods, the **decoupled** methods, and methods based on metamodels. In the following sections, we introduce these approaches and we describe in particular the RIA, PMA, SORA and Stieng methods which will be used as comparison methods in chapters 3 and 5.

Benchmarks of the **double-loop**, **single-loop** and **decoupled** methods are available in Aoues and Chateauneuf (2010) and Lopez and Beck (2012).

2.1 Double-loop approach

The intuitive way to solve a RBDO problem is to couple a classical optimization algorithm and a method to estimate the failure probabilities. This approach is called double-loop since two loops are nested: one estimates the failure probabilities, the other updates the design point. The **Reliability Index Approach** (RIA) and the **Performance Measure Approach** (PMA) (Tu et al. (1999)) are among the most popular double-loop approaches. For RIA, the failure probabilities are estimated with FORM. In PMA, the initial constraints are replaced with equivalent ones which are evaluated with an inverse reliability analysis. A double loop strategy can also use a simulation method as in Barrera et al. (2016) and Chaudhuri et al. (2020).

Remark 1.1. In RA, the term "design point" is sometimes used to refer to the Most Probable failure Point (see section 1.1.1). In this thesis, when we consider an optimization problem, the term "design point" refers to a point from the design space of the problem.

Reliability Index Approach (RIA)

The RIA method (Madsen et al. (2006)) follows the double-loop approach and estimates the failure probabilities involved in the constraints of problem (1.17) with FORM. The RBDO problem becomes:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \text{ such that} \\ &\mathbb{P}_U\left(G_{i,d}(U) < 0\right) < p_{s,i} \quad (i = 1, \cdots, n_c) \end{aligned} \tag{1.18}$$

where $G_{i,d} = g_i \circ T_d^{-1}$ and T_d is the isoprobabilistic transformation mapping (X_d, X_p) to U, a vector composed of $n_d + n_p$ standard Gaussian and independent random variables.

Considering the reliability index $\beta_i(d)$ of the failure probability involved in the i-th constraint at d evaluated with FORM, we have $\beta_i(d) \simeq -\Phi^{-1}\left(\mathbb{P}_U\left(G_{i,d}(U) < 0\right)\right)$. Therefore, denoting $\beta_{s,i} = -\Phi^{-1}(p_{s,i})$, problem (1.18) is sometimes also formulated as follows:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \text{ such that} \\ &\beta_i(d) > \beta_{s,i} \quad (i = 1, \cdots, n_c). \end{aligned} \tag{1.19}$$

Performance Measure Approach (PMA)

In RIA, the i-th constraint is satisfied at d if $\beta_i(d)$ is greater than $\beta_{s,i}$ with $\beta_i(d)$ solution of the following problem:

$$\min_{u\in\mathbb{R}^{n_d+n_p}} \{||u||, G_{i,d}(u) \leq 0\}. \tag{1.20}$$

Solving (1.20) can be tedious, therefore it is proposed in the PMA method to reformulate the constraints of problem (1.19). Let consider $G_i^*(d)$ the solution of the following problem:

$$\min_{u\in\mathbb{R}^{n_d+n_p}} \{G_{i,d}(u), ||u|| = \beta_{s,i}\}. \tag{1.21}$$

The solution $G_i^*(d)$ represents the smaller value of $G_{i,d}$ on the sphere of radius $\beta_{s,i}$. Thus, if $G_i^*(d) > 0$ (assuming that the failure domain is a connected space not contained in a sphere of radius smaller than $\beta_{s,i}$), the solution of (1.20) is greater than $\beta_{s,i}$. Therefore, $\beta_i(d) > \beta_{s,i}$ is equivalent to $G_i^*(d) > 0$. Besides, solving problem (1.21) is easier than the resolution of (1.20) since the solution of the former must be searched on the sphere of radius $\beta_{s,i}$. The PMA method then consists in approaching the RBDO problem formulated as follows:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \text{ such that} \\ &G_i^*(d) > 0 \quad (i = 1, \cdots, n_c). \end{aligned} \tag{1.22}$$

The quantity $G_i^*(d)$ is obtained performing an **inverse reliability analysis** which consists in solving (1.21). This resolution can be carried out efficiently with the **Hybrid Mean Value** method (HMV) proposed by Youn et al. (2003) or with more recent techniques (Du et al. (2004); Jung et al. (2020); Wang and Zhang (2020); Keshtegar et al. (2021)).

We denote $u_{i,d}^{RIA}$ and $u_{i,d}^{PMA}$ the points (assumed to be unique) such that $||u_{i,d}^{RIA}|| = \beta_i(d)$ and $G_{i,d}(u_{i,d}^{PMA}) = G_{i,d}^*$. The authors of Lopez and Beck (2012) emphasize that these points are only equal when the i-th reliability constraint is active. Apart from this case, $u_{i,d}^{PMA}$ only represents the point of the sphere of radius $\beta_{s,i}$ which minimizes $G_{i,d}$.

2.2 Single-loop approach

In practice, the nesting of the two loops turns out to be too costly in terms of number of calls to the performance functions. The single-loop approach has been proposed to overcome this drawback. To avoid the failure probability estimation at each iteration, the RBDO problem is transformed into a deterministic problem. In the **Single-Loop Approach** (SLA) (Liang et al. (2008); Li et al. (2019); Yang et al. (2020a)), the constraints of the problem are replaced by the performance functions. These performance functions are evaluated at specific points which are updated at each iteration of the optimization algorithm in order to propose a final design fulfilling the optimality conditions of the reliable version of the problem.

2.3 Decoupled-loop approach

The decoupled approach consists in solving a sequence of deterministic optimization problems. The final design point of an optimization cycle is the starting point of the next one. In the Sequential Approximation Programming (SAP) (Cheng et al. (2006)), problem (1.19) of RIA is solved at each cycle substituting the constraints with local approximations. The **Sequential Optimization and Reliability Assessment** method (SORA) (Du and Chen (2004)) separates the optimization and the reliability loops. A deterministic optimization cycle is carried out for fixed values of the uncertain parameters. The latter are then shifted at the next cycle to make the minimum more reliable. Different methods (Torii et al. (2016); Biswas and Sharma (2021); Jiang et al. (2020); Wang et al. (2020); Zhang et al. (2021b)) use the same approach as SORA but propose a different shifting strategy.

Sequential Optimization and Reliability Assessment (SORA)

In SORA, the distribution of X_d is assumed to be centered at d (i.e. $\mathbb{E}[X_d] = d$). During each cycle of the method, an optimization problem representing a deterministic version of the RBDO problem (1.17) is solved. The optimization problem considered for the first cycle is the following:

$$\begin{aligned} &\min_{d \in \Omega_d} cost(d) \quad \text{such that} \\ &g_i\left(d, \mathbb{E}[X_p]\right) > 0 \quad i = 1, \cdots, n_c. \end{aligned} \tag{1.23}$$

At the end of the first cycle, a design point d^1, solution of (1.23), is obtained. If the i-th deterministic constraint is active at d^1, the solution is not reliable since, considering the uncertainties at this point, the i-th failure probability at d^1 probably exceeds the maximum threshold. The SORA method then proposes to shift the constraint to gain in reliability. A sequence of deterministic optimization problems then begins. The problem solved during the k-th cycle ($k \geq 2$) is:

$$\begin{aligned} &\min_{d \in \Omega_d} cost(d) \quad \text{such that} \\ &g_i(d - s_i^k, x_{p,i}^k) > 0 \quad i = 1, \cdots, n_c \end{aligned} \tag{1.24}$$

starting from d^{k-1}: the design point obtained at the previous cycle. The gain in reliability is achieved by choosing the proper values s_i^k and $x_{p,i}^k$ for each constraint i. In SORA, they are defined as:

$$s_i^k = d^{k-1} - x_{i,MPP}^{k-1} \ , \ x_{p,i}^k = p_{i,MPP}^{k-1}. \tag{1.25}$$

The quantities $x_{i,MPP}^{k-1}$ and $p_{i,MPP}^{k-1}$ are computed with an inverse reliability method at the end of cycle $k-1$:

$$u_{i,MPP}^{k-1} = \underset{u \in \mathbb{R}^{n_d+n_p}}{\arg\min} \{G_{i,d^{k-1}}(u), \|u\| = \beta_{s,i}\}$$

and

$$(x_{i,MPP}^{k-1}, p_{i,MPP}^{k-1}) = T_{d^{k-1}}^{-1}(u_{i,MPP}^{k-1}), \tag{1.26}$$

where $T_{d^{k-1}}^{-1}$ denotes the inverse function of the iso-probabilistic transformation at d^{k-1}.

The cycles of SORA end when the reliability constraints are satisfied and the difference between the cost function evaluations at the design points obtained between two consecutive cycles is small.

Figure 1.3 illustrates the shift carried out at the end of the k-th cycle on the design variables in the case where the design space Ω_d is in 2D and only the uncertainties on the design variables are considered. The design space and the sample space of X_{d_1} and X_{d_2} are superposed. In the figure, the design point obtained at the end of the k-th cycle is denoted (d_1^k, d_2^k), it is represented with a red point and the density of $(X_{d_1^k}, X_{d_2^k})$ at (d_1^k, d_2^k) is indicated. The second red point represents x_{MPP}^k which is obtained with the inverse reliability analysis performed at (d_1^k, d_2^k). It is then used to define the shift (s_1^k, s_2^k) of the design variables during the $(k+1)$-th cycle of SORA.

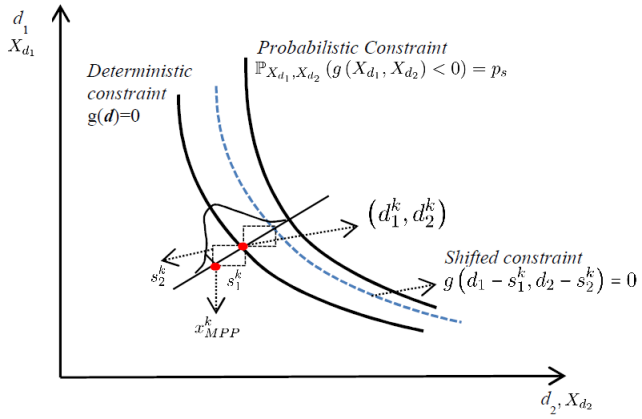


Figure 1.3: SORA shift (Lopez and Beck (2012))

2.4 Metamodel-based approach for RBDO

In RBDO, the failure probabilities depend on the design point d and therefore, multiple metamodel-based approaches can be considered. An overview of the different strategies using metamodels in RBDO is presented in Moustapha and Sudret (2019a).

Usually, as in RA, metamodels replace the performance functions to speed up the estimation of the failure probabilities and active learning strategies are implemented to refine the limit-state surface of each performance function.

A general modular framework is proposed in Moustapha and Sudret (2019a) to solve a RBDO problem with metamodels: the user can choose the adaptive metamodel, the relia-

bility analysis method, and the optimization algorithm. A single metamodel is calibrated for each performance function. In Shang et al. (2021), a combination of the PCE and RBF is chosen. For the methods introduced in Moustapha and Sudret (2019a) and Shang et al. (2021), the enrichment of the metamodels is performed before starting the optimization.

In Dubourg (2011) and Li et al. (2020), kriging models subsituting the performance functions are enriched during the optimization algorithm. In Moustapha et al. (2015), the constraints of problem (1.17) are replaced by constraints involving quantiles and the enrichment of the krigings are performed both before and during optimization.

In the following sections, we introduce the notion of augmented space (Dubourg (2011); Moustapha and Sudret (2019a)) as well as the enrichment stopping condition proposed by Dubourg (2011).

Augmented space

When the approach of building a single metamodel for solving the RBDO problem is adopted, the choice of the domain to be spanned by the initial DoE is crucial. Since the inputs of the i-th performance function g_i are the outcomes of (X_d, X_p), the metamodel replacing g_i must accurately predict the values of g_i for every outcomes of (X_d, X_p) that may be encountered during the optimization resolution and therefore, for different design points d. To restrict this domain, it is interesting to consider the space Ω^{aug} such that, for any $d \in \Omega_d$, it is almost certain that a realization (x_d, x_p) of (X_d, X_p) belongs to Ω^{aug}.

Let $d = (d_1, \cdots, d_{n_d}) \in \Omega_d = \Omega_{d_1} \times \cdots \times \Omega_{d_{n_d}} \subset \mathbb{R}^{n_d}$ be a design point and $X_{d_1}, \cdots, X_{d_{n_d}}$ the random variables with respective quantile functions $F_{d_1}^{-1}, \cdots, F_{d_{n_d}}^{-1}$ describing the uncertainties at this point. The random vector X_p is composed of n_p random variables X_{p_j} $(j = 1, \cdots, n_p)$ with quantile functions denoted $F_{p_j}^{-1}$ $(j = 1, \cdots, n_p)$. The space Ω^{aug} is defined by:

$$\Omega^{aug} = \prod_{i=1}^{n_d} \Omega_{d_i}^{aug} \times \prod_{j=1}^{n_p} \Omega_{p_j}^{aug} \tag{1.27}$$

with

$$\Omega_{d_i}^{aug} = \left[\inf_{d_i \in \Omega_{d_i}} F_{d_i}^{-1}(\alpha), \sup_{d_i \in \Omega_{d_i}} F_{d_i}^{-1}(1-\alpha) \right], \tag{1.28}$$

$$\Omega_{p_j}^{aug} = \left[F_{p_j}^{-1}(\alpha), F_{p_j}^{-1}(1-\alpha) \right] \tag{1.29}$$

where α is a degree of confidence chosen by the user (different values of α could be considered for each set $\Omega_{d_i}^{aug}$ $(i = 1, \cdots, n_d)$ and $\Omega_{p_j}^{aug}$ $(j = 1, \cdots, n_p)$).

Thereby, for all $i \in \{1, \cdots, n_d\}$, $d_i \in \Omega_{d_i}$ and $j \in \{1, \cdots, n_p\}$, the probabilities $\mathbb{P}\left(X_{d_i} \in \Omega_{d_i}^{aug}\right)$ and $\mathbb{P}\left(X_{p_j} \in \Omega_{p_j}^{aug}\right)$ are greater or equal to $1 - 2\alpha$.

The space Ω_{aug} is called the **augmented space**.

Enrichment stopping condition

In the method proposed by Dubourg (2011) to solve a RBDO problem, a kriging model $\widetilde{g}_i$ is built for each performance function g_i in the augmented space such that, for all

$(x_d, x_p) \in \Omega^{aug}$, the distribution of $\widetilde{g_i}(x_d, x_p)$ is $\mathcal{N}\left(\mu_i(x_d, x_p), \sigma_i(x_d, x_p)^2\right)$.

After each iteration of the optimization algorithm, the metamodel is sequentially enriched until a stopping condition is met. This condition checks whether the estimation of the failure probability $\widetilde{p_{f,i}}(d)$ at the current design space d is accurate enough. To do so, the author proposes low and high estimations of $p_{f,i}(d)$ that we denote $\widetilde{p_{f,i}}^-(d)$ and $\widetilde{p_{f,i}}^+(d)$. These quantities are defined as follows:

$$\widetilde{p_{f,i}}^-(d) = \mathbb{P}_{X_d, X_p}\left(\mu_i(X_d, X_p) + k\sigma_i(X_d, X_p) < 0\right), \tag{1.30}$$

$$\widetilde{p_{f,i}}(d) = \mathbb{P}_{X_d, X_p}\left(\mu_i(X_d, X_p) < 0\right), \tag{1.31}$$

$$\widetilde{p_{f,i}}^+(d) = \mathbb{P}_{X_d, X_p}\left(\mu_i(X_d, X_p) - k\sigma_i(X_d, X_p) < 0\right), \tag{1.32}$$

with k a chosen constant and thus:

$$\widetilde{p_{f,i}}^-(d) \leq \widetilde{p_{f,i}}(d) \leq \widetilde{p_{f,i}}^+(d). \tag{1.33}$$

The spread between $\widetilde{p_{f,i}}^-(d)$ and $\widetilde{p_{f,i}}^+(d)$ measures the accuracy of the estimation $\widetilde{p_{f,i}}(d)$. Therefore, the enrichment of $\widetilde{g_i}$ continues until the following condition is met:

$$\log_{10}\left(\frac{\widetilde{p_{f,i}}^+(d)}{\widetilde{p_{f,i}}^-(d)}\right) \leq \epsilon \tag{1.34}$$

where ϵ is an accuracy threshold chosen by the user.

2.5 Hybrid approach

Finally, other approaches combine metamodels with PMA (Du et al. (2020)), SORA (Goswami et al. (2019); Song et al. (2021); Zhang et al. (2020b)) or SLA (Zhang et al. (2020a)). The method proposed in Stieng and Muskulus (2020) that we will call the Stieng method approximates the performance function by the product of two functions: one depending only on the design variables which is the performance function evaluated at the mean value of the uncertainties and the other depending on the uncertain variables. A metamodel is fitted on the second function. The RBDO problem is then solved with sequential cycles of optimization. Each cycle is composed of the update of the metamodels and a resolution with PMA of the problem.

The Stieng method

In the method introduced in Stieng and Muskulus (2020), only parametric uncertainties are considered and the i-th performance function is written as:

$$g_i(d, x_p) = r_i(d, x_p) - q_i(d, x_p) = r_i(d, x_p) - \frac{q_i(d, x_p)}{q_i(d, \mathbb{E}(X_p))} q_i(d, \mathbb{E}(X_p)) \tag{1.35}$$

where only the function q_i is expensive to evaluate and $r_i(d, x_p)$ is generally a threshold. The key idea of the approach is to locally approximate $\dfrac{q_i(d, x_p)}{q_i(d, \mathbb{E}(X_p))}$ by a quantity that only depends on x_p. To do so, at the k-th cycle, the function $y_i^k(x_p)$ is defined as:

$$y_i^k(x_p) = \frac{q_i\left(d^{k-1}, x_p\right)}{q_i\left(d^{k-1}, \mathbb{E}(X_p)\right)} \tag{1.36}$$

where d^0 is chosen by the user and d^k for $k \geq 1$ is the design point obtained at the end of cycle k. Thus, the performance function can be approximated as follows:

$$g_i(d, x_p) \simeq r_i(d, x_p) - y_i^k(x_p)\, q_i(d, \mathbb{E}(X_p)). \tag{1.37}$$

The k-th cycle of the method begins with the sampling of X_p. The function y_i^k is evaluated at these points, a metamodel $\widetilde{y_i^k}$ is built to replace $y_i^k(x_p)$ and $g_i(d, x_p)$ can be approximated by:

$$\widetilde{g_i^k}(d, x_p) = r_i(d, x_p) - \widetilde{y_i^k}(x_p) q_i(d, \mathbb{E}(X_p)). \tag{1.38}$$

Then the following problem is solved:

$$\begin{aligned} &\min_{d \in \Omega_d} cost(d) \quad \text{such that} \\ &\widetilde{G_i^{k,*}}(d) > 0 \quad (i = 1, \cdots, n_c). \end{aligned} \tag{1.39}$$

where

$$\widetilde{G_i^{k,*}}(d) = \min_{u \in \mathbb{R}^{n_d + n_p}} \{\widetilde{G_{i,d}^k}(u),\ \|u\| = \beta_{s,i}\} \tag{1.40}$$

with $\widetilde{G_{i,d}} = \widetilde{g_i^k}(d, x_p) \circ T^{-1}$ and T the isoprobabilistic transformation of X_p.

A new design point d^k is obtained at the end of the cycle. In Stieng and Muskulus (2020), the algorithm stops if the cost function value has converged, otherwise a new cycle starts from d^k.

Remark 1.2. The Stieng approach can be applied with any metamodel. In the paper, the kriging model is chosen and the sampling of X_d is done using a Sobol sequence (Sobol' (1967)) whose size increases with each new cycle.

3 Time-dependent Reliability Analysis

In time-dependent reliability analysis, the performance function involves a time-dependent stochastic process denoted $\mathcal{Y}$. The failure probability is usually written as follows:

$$p_f = \mathbb{P}(\exists t \in [0, T],\ g(X, \mathcal{Y}(t), t) < 0). \tag{1.41}$$

The methods estimating this quantity can be classified into the out-crossing approach and the extreme value approach. Most of them require a failure probability written as:

$$p_f = \mathbb{P}\left(\exists t \in [0, T],\ \hat{g}\left(\hat{X}, t\right) < 0\right) \tag{1.42}$$

where $\hat{X}$ is a random vector. This second formulation can be obtained from the first one with the use of a truncated expansion of the process $\mathcal{Y}$ such as the ones mentioned in section 1.3.1. The methods in t-RA will then be introduced for the second formulation of p_f.

3.1 Approximation of the time-dependent process

Most of the methods in t-RA require to be able to sample realizations of $g(X, \mathcal{Y}(t), t)$ at a fixed time t. To do so, expansion techniques enable to approximate the process $\mathcal{Y}$ with a function f depending on a random vector Y such that:

$$\mathcal{Y}(t) \simeq f(Y, t). \tag{1.43}$$

Among the most widespread expansion methods, we can cite the Expansion Optimal Linear Estimation (EOLE) (Li and Der Kiureghian (1993)), the Karhunen-Loève (KL) expansion (Loeve (1977)), the Orthogonal Series Expansions (OSE) (Zhang and Ellingwood (1994)) and the spectral representation (Shinozuka and Deodatis (1991)) for which realizations of the process $\mathcal{Y}$ are used to identify the function f and the distribution of Y.

Grouping the random vectors in $\hat{X} = (X, Y)$, the failure probability can then be written:

$$p_f = \mathbb{P}\left(\exists t \in [0, T],\, \hat{g}\left(\hat{X}, t\right) < 0\right) \tag{1.44}$$

with $\hat{g}\left(\hat{X}, t\right) = g\left(X, f\left(Y, t\right), t\right)$.

3.2 Out-crossing approach

Let consider $p_{f,i}(t)$ the instantaneous failure probability and the random variable $N^+(T)$ representing the number of crossings from the safe domain to the failure domain in $]0, T]$:

$$p_{f,i}(t) = \mathbb{P}\left(\hat{g}\left(\hat{X}, t\right) < 0\right), \tag{1.45}$$

$$N^+(T) = \mathrm{card}\{\mathrm{t_0} \in]0, \mathrm{T}], \exists \epsilon > 0, \forall \mathrm{t} \in [\mathrm{t_0} - \epsilon, \mathrm{t_0}], \hat{\mathrm{g}}\left(\hat{\mathrm{X}}, \mathrm{t}\right) \geq 0 \text{ and } \forall t \in]t_0, t_0 + \epsilon[, \hat{g}\left(\hat{X}, t\right) < 0\}. \tag{1.46}$$

Then, the failure probability can be bounded (Shinozuka (1964)) as follows:

$$\max_{t \in [0,T]} p_{f,i}(t) \leq p_f \leq p_{f,i}(0) + \mathbb{E}\left[N^+(T)\right]. \tag{1.47}$$

The out-crossing approach seeks to estimate the upper bound of p_f in equation (1.47) by evaluating the **out-crossing rate** $\nu^+(t)$:

$$\nu^+(t) = \lim_{\Delta t \to 0^+} \frac{\mathbb{P}\left(N^+(t + \Delta t) = 1\right)}{\Delta t} \tag{1.48}$$

$$= \lim_{\Delta t \to 0^+} \frac{\mathbb{P}\left(\{\hat{g}\left(\hat{X}, t\right) \geq 0\} \cap \{\hat{g}\left(\hat{X}, t + \Delta t\right) < 0\}\right)}{\Delta t}. \tag{1.49}$$

The mean of out-crossings $\mathbb{E}\left[N^+(T)\right]$ is then computed with:

$$\mathbb{E}\left[N^+(T)\right] = \int_0^T \nu^+(t) dt. \tag{1.50}$$

The out-crossing rate can be obtained in several ways (Hawchar (2017)). In the PHI2 method (Andrieu-Renaud et al. (2004)), $\nu^+(t)$ is estimated by performing two reliability analyses with FORM at t and $t + \Delta t$. When $\mathcal{Y}$ is stationary, the out-crossing rate need to be estimated at only one time t whereas multiple evaluations are required for the non-stationary case.

3.3 Extreme value approach

The extreme value approach directly provides an estimation of the failure probability and generally relies on the use of metamodels. Considering the function $g_{\min}$ such that:

$$g_{\min}(\hat{X}) = \min_{t\in[0,T]} \hat{g}\left(\hat{X}, t\right), \tag{1.51}$$

a first strategy is to estimate the failure probability with a RA method since:

$$p_f = \mathbb{P}\left(g_{min}(\hat{X}) < 0\right). \tag{1.52}$$

In the ePCE and eLRA methods (Hawchar (2017)), the PCE and the Low Rank Approximation (LRA) (Chevreuil et al. (2015)) are respectively used to replace $g_{\min}$. In m-EGO (Hu and Du (2015)), a kriging model is prefered. In these approaches, the enrichments of the metamodel are performed by selecting $\hat{x}$ in the sample space of $\hat{X}$ with different strategies and $g_{\min}(\hat{x})$ is then computed with a dedicated Efficient Global Optimization (EGO) approach (Jones et al. (1998)).

In Ahmadivala et al. (2019), the interval $[0,T]$ is divided into several nodes t_i and the failure probability is approximated by $\mathbb{P}\left(\cup_i \left(\hat{g}\left(\hat{X}, t_i\right) < 0\right)\right)$. For each time t_i, a kriging model replaces the performance function. These metamodels are then enriched with AK-SYS (Fauriat and Gayton (2014)): a method used to carry out a system RA where multiple performance functions are involved in the failure probability.

Other methods based on adaptive kriging of the performance function have been proposed and rely on different metamodel strategies (Hu and Mahadevan (2016); Wang and Chen (2016); Jiang et al. (2019); Hu et al. (2020)).

For all of these methods, a sequential active learning is usually performed to improve the accuracy of the metamodel. The failure probability is then computed with Monte Carlo and the enriched metamodel. In AK-co-IS and AK-co-SS (Ling et al. (2019)), the failure probability is computed with IS and SS and the enrichment of the metamodel is adapted.

In Zhang et al. (2021a), the MPP is computed at each time of a discretization of $[0,T]$. A kriging is then calibrated on the trajectory of these MPPs and enriched to select a new time where to perform a MPP search. A procedure is proposed to estimate the failure probability using the resulting metamodel.

Finally, the NERS method is introduced in Wang and Wang (2012). We denote $t_{min}(\hat{x})$ the function such that $t_{min}(\hat{x}) = \underset{t\in[0,T]}{\arg\min}\ g(\hat{x}, t)$ and it follows $g_{min}(\hat{X}) = g(\hat{X}, t_{min}(\hat{X}))$. The method proposes to substitute the t_{min} function by a kriging model which is sequentially improved. The failure probability can then be estimated with a time-independent RA such as FORM.

4 Time-dependent Reliability-Based Design Optimization

The time-dependent Reliability-Based Design Optimization (t-RBDO) methods seek to solve optimization problems with constraints involving time-dependent performance functions. The most straightforward approach to solve a t-RBDO problem is to couple an optimization algorithm with a t-RA method to estimate the failure probabilities at each

iteration of the optimization problem. Hence, a local eLRA is performed at each iteration in TROL (Hawchar (2017)) whereas in Wang and Wang (2012), the NERS method is used to evaluate the failure probability.

In TROSK (Hawchar et al. (2018)), PSO-t-IRS (Li and Chen (2019)), and in Shi et al. (2020b), a kriging model of each performance function is built and enriched before the optimization and then the resolution of the optimization is done using Monte Carlo with the metamodel.

Finally, t-SORA and t-SLA are introduced in Shi et al. (2020a) and represent the time-dependent versions of SORA and SLA.

5 Introduction of the FOWT problem

We have described in the previous sections the state of the art in RA, RBDO, t-RA and t-RBDO. We now present the main characteristics of the FOWT problem under study in this thesis to see how it falls within these research fields.

As mentioned in the introduction, for FOWTs to be competitive, it is crucial to propose configurations that are economically attractive and reliable. The work of this thesis pursues this objective. For simplicity, we focus especially on the mooring system of a semi-submersible FOWT since it represents an important part of the total cost of the structure and is essential for the station keeping and proper functioning of the turbine (Carbon Trust (2015); Yu et al. (2018)). More precisely, our objective will be to minimize the manufacturing cost of the mooring lines while satisfying several constraints. The proposed configuration should enable the mooring system to limit the movements of the structure to ensure the turbine production, avoid compression in the mooring lines and excessive loading on the electric cable, and to withstand the accumulated damage in the lines. These constraints inherit the randomness of the marine conditions and uncertainties on material properties and model parameters. We are therefore faced with an optimization problem with time-dependent reliability constraints that need to be satisfied with high confidence levels. Indeed, the threshold annual failure probability imposed by internationals standard is 10^{-4} (Det Norske Veritas (2013)).

In this section, we briefly introduce the mathematical formulation of the studied problem to highlight its main characteristics and the difficulties to solve it. The problem will be presented in detail in chapter 5.

We emphasize that the purpose of our work is to identity an efficient strategy to solve this type of problems. We thus consider in this thesis a simplified model. We refer to chapter 5 for the explanation of these simplifications.

The FOWT under study is inspired by the NREL (National Renewable Energy Laboratory) 5MW turbine (Jonkman et al. (2009)) placed on the DeepCwind semi-submersible floating platform (Robertson et al. (2014a)). The floating platform is connected to the seabed by three catenary mooring lines.

5.1 Definition of the design variables

The cost of the mooring lines that we want to minimize depends on three design variables:

- the length d_1 of the mooring line that can be added to, or deducted from, the nominal mooring length. This variable takes values in $[d_1^-, d_1^+]$ (in m);
- the mass per unit length $d_2 \in [d_2^-, d_2^+]$ (in kg/m);
- the position d_3 of the connection of the lines to the columns of the floater. This variable can vary from 0 (which corresponds to a connection at the bottom of the columns) to 1 (top of the columns).

We denote $d = (d_1, d_2, d_3)$ and Ω_d the design space such that $\Omega_d = [d_1^-, d_1^+] \times [d_2^-, d_2^+] \times [0, 1]$. The cost function is denoted *cost*. It is a function from Ω_d to $\mathbb{R}^+$ and is cheap to evaluate. The cost function and the design space will be defined precisely in chapter 5.

5.2 Definition of the time-dependent processes involved in the constraints

The movements of the structure are determined from the environmental loads occurring during the considered period $[0, T]$ (T is equal to one year) and in particular loadings induced by waves and the wind.

Definition of the environmental conditions

The **swell** is modeled as a succession of waves (parallel in our case) meeting the structure and defined by their height at any time at a given point. To account for all the different possible sea state, we discretize the interval of time $[0, T]$ into n_T subintervals I_i of length ΔT. For each interval I_i the sea elevation is represented by a zero-mean stationary Gaussian random process defined by its spectral density. The latter is characterized by parameters called **long term** parameters grouped in s_i which will be called the sea state.

The sea elevation process on I_i is therefore denoted $\eta_i(s_i; .)$. We consider that $\eta_i(s_i; .)$ and $\eta_j(s_j; .)$ are independent processes for $i \neq j$. We denote η the sequence of processes $(\eta_1(s_1; .), \cdots, \eta_{n_T}(s_{n_T}; .))$.

Taking the wind into account in the calculations makes it possible to consider the forces on the rotor-nacelle assembly which affect the movements of the float. It is customary to consider a wind described by particular random processes (see standard IEC 61400-3 (2009)). For simplicity, we will only consider constant wind forces applied on the structure for each sea state (see chapter 5 for the reasons of this choice).

The surge process

The swell and the wind cause motions of the floating platform described, at all times, by the values of six movements called in mechanics **degrees of freedom**. They are represented in the figure 1.4.

We focus especially on the translation along the x-axis called the **surge** of the platform. Since the movement of the platform inherit the randomness of the marine conditions, the

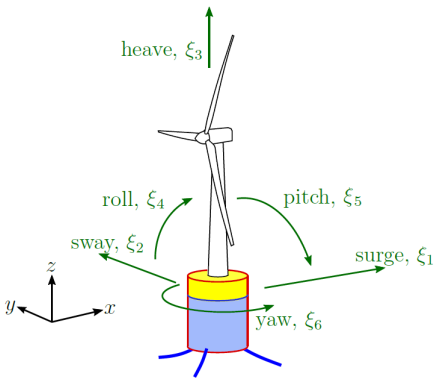


Figure 1.4: Degrees of freedom (Hall (2013))

surge is represented by a time-dependent random process defined as:

$$\mathcal{S}(d, x_p; t) = \sum_{i=1}^{n_T} \mathcal{S}_i(d, x_p, s_i; t)\, \mathbb{1}_{I_i}(t) \tag{1.53}$$

where $\mathcal{S}_i(d, x_p, s_i; t)$ is a process depending on the design d considered, on parametric variables (which will be detailed in chapter 5) grouped in x_p, and on the sea state s_i occurring during I_i.

The tension processes

The movements of the floater cause cycles of **tension** in the mooring lines. For catenary lines, the maximum tension is at the top of the line and this location is considered as critical because of particular end effect on the loading. Depending on the constraint (maximum tension, minimum tension, fatigue) the most critical point could be located elsewhere, for example near the touchdown point. To simplify, we will consider in this thesis only the tension at the top of the lines knowing that this simplification is not conservative for the design. The extension to a more general study with a replacement by the most severe constraint along the line does not change the methodology.

The tension at the top of the l-th mooring line ($l = 1, 2, 3$) is represented by the following random process:

$$\mathcal{T}^l(d, x_p; t) = \sum_{i=1}^{n_T} \mathcal{T}_i^l(d, x_p, s_i; t)\, \mathbb{1}_{I_i}(t) \tag{1.54}$$

where $\mathcal{T}_i^l(d, x_p, s_i; t)$ is a process depending on the design d considered, on the parametric variables grouped in x_p, and on the sea state s_i occurring during I_i.

The instantaneous damage processes and the total damage

When a material is subjected to a high number (millions) of repeated cycles of tension, it accumulates damage also called polycyclic **fatigue** of the material. The marine conditions and the floater movements cause this type of loads on the mooring lines.

As for the tension, we focus on the damage accumulated at the top of each mooring line. We can represent the instantaneous damage occurring at the top of the l-th mooring line ($l = 1, 2, 3$) with the following random process:

$$\mathcal{D}^l\left(d, x_{d_2}, x_p; t\right) = \sum_{i=1}^{n_T} \mathcal{D}_i^l\left(d, x_{d_2}, x_p, s_i; t\right) \mathbb{1}_{I_i}(t) \tag{1.55}$$

where $\mathcal{D}_i^l\left(d, x_{d_2}, x_p, s_i; .\right)$ is a process depending on the design d considered, on a resistance parameter of the model x_{d_2} which depend on d_2 (this parameter will be specified in chapter 5), on the parametric variables x_p, and on the sea state s_i occurring during I_i.

For fixed values of x_{d_2} and x_p, the **total damage** occurring at the top of the line l is the random variable denoted the $D_{[0,T]}^{total,l}(d, x_{d_2}, x_p)$ and defined as the accumulation of the instantaneous damage process over $[0, T]$:

$$D_{[0,T]}^{total,l}(d, x_{d_2}, x_p) = \int_0^T \mathcal{D}^l\left(d, x_{d_2}, x_p; t\right) dt. \tag{1.56}$$

5.3 Definition of the constraints

Introduction of uncertainties

The surge, the tension, and the damage processes depend on parameters denoted x_{d_2} and x_p. We consider that these parameters are subject to uncertainties represented respectively by the random variable X_{d_2} and the random vector X_p.

The surge constraint

To enable the wind turbine to operate correctly and to reduce the loading on the electrical cable, it is necessary to restrict the amplitude of the surge. We therefore consider a threshold $\mathcal{S}_{\max}$ for the maximum of the surge. Taking into account the uncertainties on the parametric variables and the marine conditions therefore leads to the formulation of the first constraint:

$$\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} \left|\mathcal{S}\left(d, X_p; t\right)\right| > \mathcal{S}_{\max}\right) < 10^{-4}. \tag{1.57}$$

The threshold probability of 10^{-4} is recommended by international standards (Det Norske Veritas (2013)) for mooring lines.

Remark 1.3. The surge constraint concerns the absolute displacement of the floater and not the motion amplitude around the mean position of the floater (which varies with the sea state).

The tension constraints

To avoid out-of-plan bending of the line, which would cause the material to break under subsequent tensions, it is important to impose a positive tension at all times on the line. By taking into account the different sources of uncertainties, the tension constraints become:

$$\mathbb{P}_{X_p,\eta}\left(\min_{[0,T]} \mathcal{T}^l\left(d, X_p; t\right) < 0\right) < 10^{-4} \tag{1.58}$$

for $l = 1, 2, 3$.

The fatigue constraints

The last constraints concern the damage accumulated at the top of each line. The fatigue is the accumulation of the instantaneous damages occurring during the considered period. Material failure occurs when the fatigue exceeds a certain threshold. For safety, this threshold is considered uncertain and is represented by the random variable X_R. Taking into account all the uncertainties leads to the following formulation of the fatigue constraints:

$$\mathbb{P}_{X_{d_2},X_p,X_R,\eta}\left(\int_0^T \mathcal{D}^l\left(d,X_{d_2},X_p;t\right)dt > X_R\right) < 10^{-4} \tag{1.59}$$

for $l = 1,2,3$.

5.4 The FOWT time-dependent RBDO problem

In this thesis, we therefore face the following optimization problem:

$$\begin{aligned}
&\min_{d\in\Omega_d} cost(d) \quad \text{such that}\\
&\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} |\mathcal{S}\left(d,X_p;t\right)| > \mathcal{S}_{\max}\right) && < 10^{-4}\\
&\mathbb{P}_{X_p,\eta}\left(\min_{t\in[0,T]} \mathcal{T}^l\left(d,X_p;t\right) < 0\right) && < 10^{-4} \quad , \; l = 1,2,3\\
&\mathbb{P}_{X_{d_2},X_p,X_r,\eta}\left(\int_0^T \mathcal{D}^l\left(d,X_{d_2},X_p;t\right)dt > X_R\right) && < 10^{-4} \quad , \; l = 1,2,3
\end{aligned} \tag{1.60}$$

We consider that all the sources of uncertainties are independent.

Remark 1.4. A formulation of the constraints involving a system failure probability is also proposed by the standards (Det Norske Veritas (2013)) but the failure of a mooring line is critical enough to consider a threshold failure probability for each constraint.

This is a particular t-RBDO problem since the distribution of the processes involved in the constraints depend on the design variables and on the parametric uncertainties. Moreover, the fatigue constraints include the integral of a time-dependent process which is not the usual formulation of constraints in t-RBDO problem.

Since the cost function is cheap to evaluate, the difficulty in solving problem (1.60) lies in the estimation of the constraints. Indeed, evaluating the failure probabilities with a sampling method or with more advanced techniques require to compute many realizations of the time-dependent processes. For design variables d_1, d_2, and d_3 and fixed variables x_p and x_{d_2}, several minutes (for a resolution in the frequency domain (Le Cunff et al. (2008))) to several hours (for a time-domain resolution) of simulation are necessary to compute a realization of the surge and the tension processes.

Although some work has been done in offshore engineering to derive the statistical properties of the output processes from the sea elevation description (as discussed in Sclavounos (2012)), we will see that some results of the Extreme Value Theory have not yet been applied to approximate the threshold exceedance probabilities appearing in the surge and tension constraints.

Concerning the fatigue constraints, different approaches have been proposed to reduce the computational cost of the estimation of the mean total damage. The methods proposed by Müller and Cheng (2018) and Barrera et al. (2020) estimate the total fatigue from a subset of the possible sea states. In Müller and Cheng (2018), a representative set of environmental conditions is selected with a Sobol sequence (Sobol' (1967)) while the maximum dissimilarity algorithm (Camus et al. (2011)) is preferred in Barrera et al. (2020). Adaptive kriging strategies have also been proposed in Huchet et al. (2019) and Teixeira et al. (2019) to efficiently choose the sea states where simulations need to be performed. However, these approaches require to compute many realizations (or one realization over a long period of time) of the tension processes which is computationally expensive. Besides, these approaches consider fixed design variables. Problem (1.60) is more challenging since the fatigue failure probability must be estimated during each iteration of the optimization resolution.

The proposed approach

Under the following simplifications (which will be explained in chapter 5):

- the non-linear hydrodynamic drag forces are linearized (Le Cunff et al. (2008));
- the effects of the second-order hydrodynamic forces on the FOWT are not considered;

it is possible to show that the surge and the tension processes are piece-wise stationary Gaussian processes and their distributions can be deduced from the sea elevation distribution. In this case, the failure probabilities can be estimated without the need to compute realizations of the processes which can be very interesting in terms of computation time. We adopt these simplifications in this thesis and propose to use the resulting properties of the outputs processes to reformulate the constraints of problem (1.60) in a way that is easier to evaluate. The following chapter describes this reformulation of the constraints in a general framework.

Chapter 2

Reformulation of time-dependent failure probabilities

Contents

1 Formulation of a t-RBDO problem with extreme-based and integral-based constraints

Motivated by the problem (1.60) of minimization of the floating offshore wind turbine manufacturing cost introduced in section 1.5, we consider the following t-RBDO problem having the same characteristics as the FOWT problem:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \quad \text{such that} \\ &\mathbb{P}\left(\min_{t\in[0,T]} g_{\mathrm{E}}\left(X_{r_{\mathrm{E}}}, \mathcal{Y}(X_d, X_p, t)\right) < 0\right) < p_s \\ &\mathbb{P}\left(\int_0^T g_{\mathrm{I}}\left(X_{r_{\mathrm{I}}}, \mathcal{Y}(X_d, X_p, t)\right) dt < 0\right) < p_s. \end{aligned} \tag{2.1}$$

In the above equation, the cost function is deterministic and depends on design variables gathered in the vector d. The design space is denoted Ω_d and is a subset of $\mathbb{R}^{n_d}$. The uncertainties on the design variables, on the model and on the resistance thresholds in

the first and second constraints are respectively represented by the random vectors X_d, X_p, X_{r_E} and X_{r_I} which are assumed to have known probability density functions. The performance functions are defined by:

$$g_\mathrm{E}(X_{r_\mathrm{E}}, \mathcal{Y}(X_d, X_p; t)) = X_{r_\mathrm{E}} - \mathcal{Y}(X_d, X_p; t), \tag{2.2}$$

$$g_\mathrm{I}(X_{r_\mathrm{I}}, \mathcal{Y}(X_d, X_p; t)) = X_{r_\mathrm{I}} - f\left(\mathcal{Y}\left(X_d, X_p; t\right)\right), \tag{2.3}$$

where $f : \mathbb{R} \to \mathbb{R}$ is measurable. The time-dependent process is denoted $\mathcal{Y}\left(x_d, x_p; .\right)$ and its distribution depends on the outcomes of the random vectors X_d and X_p. The constraints are satisfied if the failure probabilities do not exceed the failure probability threshold p_s.

We will call **extreme-based** (resp. **integral-based**) constraint, constraints expressed as the first (resp. second) constraint of problem (2.1). The failure probability involved in an extreme-based (resp. integral-based) constraint will be denoted $p_\mathrm{E}(d)$ (resp. $p_\mathrm{I}(d)$) and will be called extreme-based (resp. integral-based) failure probability. Extreme-based failure probability refers to the usual failure probability in t-RBDO except that the process distribution of $\mathcal{Y}$ depends on the random vectors X_d and X_p. The integral-based failure probability is less studied in t-RBDO and represents the probability that the accumulation of a quantity depending on $\mathcal{Y}$ exceeds some threshold over the time interval $[0, T]$.

In the industrial application, the evaluations of the functions g_E and g_I are costly, the time T is large (one year) and the probability threshold p_s is small (10^{-4}).

Methods used to estimate failure probabilities in t-RA and t-RBDO are often time-consuming as they require numerous evaluations of the performance functions and do not exploit any assumption about the process $\mathcal{Y}$. However, in some applications, for fixed values of X_d, X_p, X_{r_E} and X_{r_I} the performance function is a time-dependent process with known distribution. In offshore engineering, the wind speed and sea elevation are usually represented as stationary Gaussian processes (Vorpahl et al. (2013)). When the linearization of the momentum balance equation is a reasonable approximation, quantities of interest such as the displacement of the structure inherit the stationary and Gaussian properties of the input processes (Le Cunff et al. (2008)).

Therefore, instead of following the classical approaches in t-RBDO described in section 1.4, we propose a two-step procedure better suited to the characteristics of the studied problem. The first step of the procedure is to reformulate the extreme-based and integral-based failure probabilities into quantities that are easier to evaluate using spectral properties of the time-dependent processes and limit theorems when T tends to infinity.

This reformulation is described in section 2.2 when $\mathcal{Y}$ is stationary and Gaussian. The more general case when $\mathcal{Y}$ is a piece-wise stationary Gaussian process (which corresponds to the industrial problem with a discretized joint probability of environmental parameters (Vorpahl et al. (2013))) is dealt with in section 2.3. To illustrate the reformulation of extreme-based and integral-based failure probabilities, an academic t-RBDO problem of a harmonic oscillator having the same properties as the industrial problem is introduced and its constraints are reformulated in this chapter.

At the end of this first step, a reformulated optimization problem is obtained. Chapter 3 will focus on the second step of the procedure: solving the reformulated problem with a new efficient method.

2 Constraints defined in terms of a stationary Gaussian process

2.1 Reformulation principle

For fixed values x_d and x_p, we consider in this section that the process $\mathcal{Y}(x_d, x_p; .)$ introduced in section 2.1 is a stationary Gaussian process with zero mean (the consideration of a non-zero mean does not raise any difficulty and is discussed in remarks 2.4 and 2.6). Its distribution is defined by its spectral density $K_{\mathcal{Y}}(x_d, x_p; .)$ which is the Fourier transform of its autocorrelation function $k_{\mathcal{Y}}(x_d, x_p; .)$ that depends on x_d and x_p:

$$K_{\mathcal{Y}}(x_d, x_p; \omega) = \frac{1}{2\pi} \int_{\mathbb{R}} k_{\mathcal{Y}}(x_d, x_p; t)\, e^{-i\omega t} dt. \tag{2.4}$$

The spectral moment of order n of $\mathcal{Y}(x_d, x_p; .)$ is defined as:

$$m_{\mathcal{Y},n}(x_d, x_p) = \int_{\mathbb{R}} \omega^n K_{\mathcal{Y}}(x_d, x_p; \omega)\, d\omega. \tag{2.5}$$

To separate in the calculations, the uncertainties represented by X_d and X_p and the randomness of the time-dependent process, we consider the process $\mathcal{Y}(x_d, x_p; .)$ as output of a linear filter and we introduce the zero-mean stationary Gaussian process η such that:

$$\mathcal{Y}(x_d, x_p; t) = (h_{\mathcal{Y}}(x_d, x_p; .) * \eta)(t). \tag{2.6}$$

Example 2.1. Let us consider a white noise η (i.e. a stationary Gaussian process with zero mean and covariance function $\mathbb{E}[\eta(t)\eta(t')] = \delta(t - t')$) and a function $h_{\mathcal{Y}}(t) = \int_{\mathbb{R}} e^{i\omega t} \sqrt{K_{\mathcal{Y}}(\omega)} d\omega$ such that $h_{\mathcal{Y}} * h_{\mathcal{Y}}(t) = k_{\mathcal{Y}}(t)$ with $k_{\mathcal{Y}}$ a stationary covariance function. It follows that the process $\mathcal{Y}$, defined as $\mathcal{Y}(t) = h_{\mathcal{Y}} * \eta(t)$, is a stationary Gaussian process with zero mean, covariance function $k_{\mathcal{Y}}$, and spectral density $K_{\mathcal{Y}}$ (Lindgren (2010)).

Remark 2.1. In this section, the introduction of the process η is only a convenient trick to simplify the calculations. However, in the FOWT problem under study in this thesis and in the oscillator problem introduced in section 2.2.4, the process η has a physical meaning that we will specify.

Hence, by conditioning, the failure probabilities involved in problem (2.1) can be written:

$$p_{\mathrm{E}}(d) = \mathbb{E}_{X_d, X_p, X_{r_{\mathrm{E}}}} \left[\mathbb{P}_{\eta | X_d, X_p, X_{r_{\mathrm{E}}}} \left(\min_{t \in [0,T]} g_{\mathrm{E}}(X_{r_{\mathrm{E}}}, \mathcal{Y}(X_d, X_p; t)) < 0 \right) \right] \tag{2.7}$$

and

$$p_{\mathrm{I}}(d) = \mathbb{E}_{X_d, X_p, X_{r_{\mathrm{I}}}} \left[\mathbb{P}_{\eta | X_d, X_p, X_{r_{\mathrm{I}}}} \left(\int_0^T g_{\mathrm{I}}(X_{r_{\mathrm{I}}}, \mathcal{Y}(X_d, X_p; t))\, dt < 0 \right) \right]. \tag{2.8}$$

The first contribution of this thesis is to show that, for fixed values x_d, x_p, $x_{r_{\mathrm{E}}}$, $x_{r_{\mathrm{I}}}$ of X_d, X_p, $X_{r_{\mathrm{E}}}$, $X_{r_{\mathrm{I}}}$ and when T is large enough, limit theorems for functionals of a stationary Gaussian process provide good approximations of:

$$\mathbb{P}_\eta \left(\min_{t \in [0,T]} g_{\mathrm{E}}(x_{r_{\mathrm{E}}}, \mathcal{Y}(x_d, x_p; t)) < 0 \right) \text{ and } \mathbb{P}_\eta \left(\int_0^T g_{\mathrm{I}}(x_{r_{\mathrm{I}}}, \mathcal{Y}(x_d, x_p; t))\, dt < 0 \right). \tag{2.9}$$

Therefore, the failure probabilities $p_{\mathrm{E}}(d)$ and $p_{\mathrm{I}}(d)$ can be approximated by expectations depending only on X_d, X_p, $X_{r_{\mathrm{E}}}$ and $X_{r_{\mathrm{I}}}$. The resulting reformulated problem is much easier to solve since the failure probabilities are no longer time-dependent. As we will see, this approximation only uses the spectral properties of $\mathcal{Y}(x_d, x_p; .)$ and no sampling of the time-dependent process is required.

2.2 Approximation of the extreme-based failure probability

For fixed values of X_d, X_p and X_{r_E}, the probability $\mathbb{P}_\eta\left(\min_{t\in[0,T]} g_E\left(x_{r_E}, \mathcal{Y}\left(x_d, x_p; t\right)\right) < 0\right)$ involves the maximum of a stationary Gaussian process since:

$$\mathbb{P}_\eta\left(\min_{t\in[0,T]} g_E\left(x_{r_E}, \mathcal{Y}\left(x_d, x_p; t\right)\right) < 0\right) = \mathbb{P}_\eta\left(\max_{t\in[0,T]} \mathcal{Y}\left(x_d, x_p; t\right) > x_{r_E}\right). \tag{2.10}$$

Thus, the extreme value theory (Leadbetter et al. (1983)) and especially the following theorem are well suited to provide a reformulation of $p_E(d)$.

Extreme value theory for a stationary Gaussian process

We present the theorem 8.2.7 of Leadbetter et al. (1983) for a stationary Gaussian process $\{\xi(t); t \geq 0\}$ with zero mean, autocorrelation function k_ξ, spectral density K_ξ and spectral moment of order n denoted $m_{\xi,n}$.

Theorem 2.1. Suppose that the Gaussian stationary process ξ satisfies the following conditions:

$$k_\xi(\tau) = m_{\xi,0} - \frac{m_{\xi,2}\tau^2}{2} + o(\tau^2) \text{ as } \tau \to 0, \tag{2.11}$$

$$k_\xi(\tau)\log(\tau) \to 0 \text{ as } \tau \to \infty, \tag{2.12}$$

then

$$\mathbb{P}\left(a_T\left(\max_{t\in[0,T]} \frac{\xi(t)}{\sqrt{m_{\xi,0}}} - a_T\right) \leq x\right) \to \exp\left(-\exp(-x)\right) \text{ as } T \to \infty, \tag{2.13}$$

with $a_T = \sqrt{2\log\left(T/T_c\right)}$ and $T_c = 2\pi\sqrt{m_{\xi,0}/m_{\xi,2}}$.

This formulation of theorem 2.1 is obtained by applying theorem 8.2.7 of Leadbetter et al. (1983) to the process $\xi\left(tT_c\right)$.

Remark 2.2. It is possible to give sufficient conditions for theorem 2.1 that are explicit in K_ξ. Denoting K'_ξ the derivative of K_ξ, conditions (2.11) and (2.12) are met if we have:

$$m_{\xi,0} < \infty, m_{\xi,2} < \infty, \tag{2.14}$$

$$K_\xi \in C^1, K_\xi \text{ and } K'_\xi \text{ are integrable.} \tag{2.15}$$

For condition (2.15), we use that if $K_\xi \in C^1$ and K_ξ and K'_ξ are integrable then $\exists c > 0$ such that $\left|k_\xi(\tau)\right| \leq \frac{c}{|\tau|}$ and therefore, condition (2.12) is met. Other sufficient conditions on the spectral density are discussed in Berman (1991).

Reformulation of the extreme-based constraint

For fixed values x_d and x_p, if the process $\mathcal{Y}(x_d, x_p; .)$ meets conditions (2.14) and (2.15), we can apply theorem 2.1 and obtain $\forall x$:

$$\mathbb{P}_\eta\left(a_T\left(x_d, x_p\right)\left(\max_{t\in[0,T]} \frac{\mathcal{Y}\left(x_d, x_p; t\right)}{\sqrt{m_{\mathcal{Y},0}\left(x_d, x_p\right)}} - a_T\left(x_d, x_p\right)\right) \leq x\right) \to \exp\left(-\exp(-x)\right) \text{ as } T \to \infty \tag{2.16}$$

with $a_T(x_d, x_p) = \sqrt{2\log\left(\frac{T}{2\pi}\sqrt{\frac{m_{\mathcal{Y},2}(x_d,x_p)}{m_{\mathcal{Y},0}(x_d,x_p)}}\right)}$. Therefore, it follows from equations (2.7), (2.10) and (2.16) that for T large enough, it is reasonable to make the following approximation:

$$p_{\mathrm{E}}(d) \simeq \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[F_\epsilon\left(\exp\left(a_T\left(X_d,X_p\right)^2 - \frac{a_T\left(X_d,X_p\right)X_{r_{\mathrm{E}}}}{\sqrt{m_{\mathcal{Y},0}\left(X_d,X_p\right)}}\right)\right)\right] \tag{2.17}$$

with $F_\epsilon(x) = 1 - \exp(-x)$. The approximation error made in equation (2.17) can be bounded with classical results (Kratz and Rootzén (1997)) and is discussed in section A.1 of appendix A.

Remark 2.3. The initial failure probability that depends on a random process has been approximated by an expectation which only depends on random vectors. Furthermore, to compute the quantity within the square brackets in (2.17), only two spectral moments of $\mathcal{Y}\left(x_d, x_p; .\right)$ need to be evaluated (for each outcome of X_d and X_p).

Remark 2.4. The extreme-based constraint reformulation can be obtained for a process $\mathcal{Y}\left(x_d, x_p; .\right)$ with a non-zero mean $\mu_{\mathcal{Y}}\left(x_d, x_p\right)$ by applying theorem 2.1 to $\mathcal{Y}\left(x_d, x_p; .\right) - \mu_{\mathcal{Y}}\left(x_d, x_p\right)$. The resulting approximation is the one provided by (2.17) replacing $X_{r_{\mathrm{E}}}$ by $X_{r_{\mathrm{E}}} - \mu_{\mathcal{Y}}\left(X_d, X_p\right)$.

2.3 Approximation of the integral-based failure probability

We focus now on the integral-based failure probability:

$$p_{\mathrm{I}}(d) = \mathbb{P}_{X_d,X_p,X_{r_{\mathrm{I}}},\eta}\left(\int_0^T f\left(\mathcal{Y}\left(X_d,X_p;t\right)\right)dt > X_{r_{\mathrm{I}}}\right). \tag{2.18}$$

For fixed values x_d, x_p, the process $f\left(\mathcal{Y}\left(x_d, x_p; .\right)\right)$ is denoted $\mathcal{F}(x_d, x_p; .)$. Since $\mathcal{Y}\left(x_d, x_p; .\right)$ is stationary, $\mathcal{F}(x_d, x_p; .)$ is also stationary and we denote by $k_{\mathcal{F}}(x_d, x_p; .)$ its autocovariance function.

Ergodicity

Definition 2.1. A time-dependent stationary process ξ is said to be **ergodic** if:

$$\frac{1}{T}\int_0^T \xi(t)dt \xrightarrow[T\to+\infty]{\mathbb{P}} \mathbb{E}\left[\xi(0)\right] \tag{2.19}$$

where $\xrightarrow{\mathbb{P}}$ refers to the convergence in probability. A sufficient condition for the stationary process ξ to be ergodic (cf section 13.1 of Papoulis (1991)) is that its autocovariance function is integrable.

Reformulation of the integral-based constraint

Suppose that $\mathcal{F}\left(x_d, x_p; .\right)$ is ergodic, then, for almost every x (more exactly, for all $x \neq \mathbb{E}_\eta\left[\mathcal{F}\left(x_d, x_p; 0\right)\right]$)

$$\mathbb{P}_\eta\left(\frac{1}{T}\int_0^T \mathcal{F}\left(x_d, x_p; t\right)dt > x\right) \to \mathbb{1}_{\mathbb{E}_\eta[\mathcal{F}(x_d,x_p;0)]>x} \text{ as } T \to \infty. \tag{2.20}$$

Thus, for T large enough, it is reasonable to approximate the integral-based failure probability as follows:

$$p_{\mathrm{I}}(d) \simeq \mathbb{E}_{X_d,X_p}\left[F_{r_{\mathrm{I}}}\left(T\mathbb{E}_\eta\left[\mathcal{F}\left(X_d,X_p;0\right)\right]\right)\right], \tag{2.21}$$

with $F_{r_{\mathrm{I}}}$ the cumulative distribution function of $X_{r_{\mathrm{I}}}$ (which is continuous because $X_{r_{\mathrm{I}}}$ has a density). The approximation error made in equation (2.21) is discussed in section A.2 of appendix A.

Remark 2.5. To compute $\mathbb{E}_\eta\left[\mathcal{F}\left(x_d,x_p;0\right)\right]$, it is necessary to know the distribution of $\mathcal{Y}\left(x_d,x_p;0\right)$. Since the process $\mathcal{Y}\left(x_d,x_p;.\right)$ is Gaussian with zero mean, the variance of $\mathcal{Y}\left(x_d,x_p;0\right)$ determines its distribution. Hence, to compute the quantity within the square brackets in (2.21), we only need to know the variance of $\mathcal{Y}\left(x_d,x_p;0\right)$ for each outcome of X_d and X_p (i.e. $m_{\mathcal{Y},0}\left(x_d,x_p\right)$).

Remark 2.6. The integral-based constraint reformulation is the same for a process $\mathcal{Y}\left(x_d,x_p;.\right)$ with non-zero mean $\mu_{\mathcal{Y}}\left(x_d,x_p\right)$. The value of $\mu_{\mathcal{Y}}\left(x_d,x_p\right)$ is taken into account in the computation of $\mathbb{E}_\eta\left[\mathcal{F}\left(x_d,x_p;0\right)\right]$.

2.4 Optimization problem involving a stationary harmonic oscillator

We present in this section a concrete optimization problem with constraints involving extreme-based and integral-based failure probabilities and we apply the reformulation procedure described in sections 2.2.2 and 2.2.3.

The stationary harmonic oscillator problem

Let us consider a harmonic oscillator on an interval of time $[0,T]$: a spring/mass system. We denote respectively by x_{d_1} the mass of the object, x_{d_2} the spring stiffness and x_p the damping coefficient. An external force is exerted on the system. To account for all the sources of uncertainty of the experiment, the values of x_{d_1}, x_{d_2}, x_p and the external force are considered random. These uncertainties are respectively represented by the random variables X_{d_1}, X_{d_2}, X_p and the stochastic process $\eta(t)$. We denote X_d the random vector (X_{d_1},X_{d_2}) of outcome $x_d = (x_{d_1},x_{d_2})$ whose distribution depends on certain design variables $d = (d_1,d_2)$.

Consequently, for fixed values x_d and x_p, the displacement of the mass with respect to the equilibrium position is expressed by a stochastic process denoted $\mathcal{D}(x_d,x_p;.)$ which is solution of the harmonic oscillator equation:

$$x_{d_1}\mathcal{D}''\left(x_d,x_p;t\right) + x_p\mathcal{D}'\left(x_d,x_p;t\right) + x_{d_2}\mathcal{D}\left(x_d,x_p;t\right) = \eta(t)\,,\ t\in[0,T] \tag{2.22}$$

where $\mathcal{D}'\left(x_d,x_p;.\right)$ and $\mathcal{D}''\left(x_d,x_p;.\right)$ are respectively the velocity and acceleration processes whose sample paths are the first and second time derivatives of the sample path of $\mathcal{D}\left(x_d,x_p;.\right)$.

The optimization problem consists in minimizing a linear function $cost(d_1,d_2)$ while constraints are imposed on the design variables such that:

- the velocity and the acceleration of the oscillator must stay below given thresholds x_{r_1} and x_{r_2} respectively (this is a simplified model for the extreme constraints of the FOWT problem (1.60)).

- the accumulated amount of acceleration of the object exceeding the threshold ρ must remain under a resistance threshold x_{r_3} (this is a simplified model for the fatigue constraint of the FOWT problem (1.60))

The thresholds x_{r_1}, x_{r_2}, and x_{r_3} are also random and therefore outcomes of random variables denoted X_{r_1}, X_{r_2} and X_{r_3}. The optimization problem is formulated as follows:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \quad \text{such that} \\ &\mathbb{P}_{X_d,X_p,X_{r_1},\eta}\left(\max_{t\in[0,T]} \mathcal{D}'\left(X_d,X_p;t\right) > X_{r_1}\right) &< p_s \\ &\mathbb{P}_{X_d,X_p,X_{r_2},\eta}\left(\max_{t\in[0,T]} \mathcal{D}''\left(X_d,X_p;t\right) > X_{r_2}\right) &< p_s \\ &\mathbb{P}_{X_d,X_p,X_{r_3},\eta}\left(\int_0^T \left(\left|\mathcal{D}''\left(X_d,X_p;t\right)\right| - \rho\right)^+ dt > X_{r_3}\right) &< p_s \end{aligned} \tag{2.23}$$

where $\Omega_d \subset \mathbb{R}^2$ is the design space and $x^+ = \max(0,x)$. Thus, the two first constraints of problem (2.23) are extreme-based constraints and the third constraint is integral-based.

The distributions of X_d and X_p are chosen such that the oscillator is underdamped for almost all realizations (i.e. $X_p^2 - 4X_{d_1}X_{d_2} < 0$ almost surely). Moreover, we assume that the process η is stationary, Gaussian with spectral density:

$$K_\eta(\omega) = \frac{\theta}{\sqrt{2\pi}} \exp\left(-\frac{(\theta\omega)^2}{2}\right) \tag{2.24}$$

with $\theta > 0$.

Finally all the sources of uncertainty $(X_{d_1}, X_{d_2}, X_p, X_{r_1}, X_{r_2}, X_{r_3}, \eta)$ are independent.

Useful properties of the velocity and acceleration processes

For fixed values of X_d, X_p, it follows from equation (2.22) and the stationarity of η (see theorem 4.6 of Lindgren (2010)) that the process $\mathcal{D}(x_d,x_p;.)$ is stationary and can be written as the output of a linear filter:

$$\mathcal{D}(x_d,x_p;t) = \left(h_{\mathcal{D}}(x_d,x_p,.) * \eta\right)(t) \tag{2.25}$$

with $h_{\mathcal{D}}$ defined by:

$$H_{\mathcal{D}}(x_d,x_p;\omega) = \mathrm{FT}\left(h_{\mathcal{D}}(x_d,x_p;.)\right)(\omega) = \frac{1}{-\omega^2 x_{d_1} + i\omega x_p + x_{d_2}}$$

where FT refers to the Fourier transformation. It is shown (see theorem 4.1 and comment page 100 of Lindgren (2010)) that the process $\mathcal{D}(x_d,x_p;.)$ is then also Gaussian with zero mean and its spectral density is given by:

$$K_{\mathcal{D}}(x_d,x_p;\omega) = \left|H_{\mathcal{D}}(x_d,x_p;\omega)\right|^2 K_\eta(\omega). \tag{2.26}$$

Furthermore, if the process $\mathcal{D}(x_d,x_p;.)$ has finite spectral moment of order 2 and 4, (see theorem 2.2 and comment page 14 of Lindgren (2010)) the processes $\mathcal{D}'(x_d,x_p;.)$ and $\mathcal{D}''(x_d,x_p;.)$ are also stationary with zero mean and respective spectral densities, denoted $K_{\mathcal{D}'}(x_d,x_p,.)$ and $K_{\mathcal{D}''}(x_d,x_p,.)$, given by:

$$K_{\mathcal{D}'}(x_d,x_p;\omega) = \omega^2 K_{\mathcal{D}}(x_d,x_p;\omega) \text{ and } K_{\mathcal{D}''}(x_d,x_p;\omega) = \omega^4 K_{\mathcal{D}}(x_d,x_p;\omega). \tag{2.27}$$

Remark 2.7. In fact, the result of theorem 2.2 of Lindgren (2010) holds for the first and second order derivatives of $\mathcal{D}(x_d, x_p; .)$ in the quadratic mean sense. But under the supplementary condition that the spectral moment of order 6 of $\mathcal{D}(x_d, x_p; .)$ is also finite, the result holds for sample path derivatives too (see remark 2.3 of Lindgren (2010)).

The spectral moments of order n of $\mathcal{D}'(x_d, x_p; .)$(resp. $\mathcal{D}''(x_d, x_p; .)$), is denoted $m_{\mathcal{D}',n}(x_d, x_p)$ (resp. $m_{\mathcal{D}'',n}(x_d, x_p)$).

We show in appendix B that, for all x_d, x_p, the processes $\mathcal{D}'(x_d, x_p; .)$ and $\mathcal{D}''(x_d, x_p; .)$ meet the sufficient conditions that allow the reformulation of the three constraints of the oscillator problem.

The reformulated oscillator problem

Since all the required conditions are met, for T large enough, we can apply the reformulation steps described in sections 2.2.2 and 2.2.3 to the constraints of problem (2.23) and we obtain the following reformulated problem:

$$\begin{aligned}
&\min_{d\in\Omega_d} cost(d) \quad \text{such that} \\
&\mathbb{E}_{X_d,X_p,X_{r_1}}\left[F_\epsilon\left(\exp\left(a_T^1(X_d,X_p)^2 - \frac{a_T^1(X_d,X_p)\,X_{r_1}}{\sqrt{m_{\mathcal{D}',0}(X_d,X_p)}}\right)\right)\right] &< p_s \\
&\mathbb{E}_{X_d,X_p,X_{r_2}}\left[F_\epsilon\left(\exp\left(a_T^2(X_d,X_p)^2 - \frac{a_T^2(X_d,X_p)\,X_{r_2}}{\sqrt{m_{\mathcal{D}'',0}(X_d,X_p)}}\right)\right)\right] &< p_s \\
&\mathbb{E}_{X_d,X_p}\left[F_{r_3}\left(T\mathbb{E}_\eta\left[\mathcal{F}(X_d,X_p;0)\right]\right)\right] &< p_s
\end{aligned} \tag{2.28}$$

with

$$a_T^1(x_d,x_p) = \sqrt{2\log\left(\frac{T}{2\pi}\sqrt{\frac{m_{\mathcal{D}',2}(x_d,x_p)}{m_{\mathcal{D}',0}(x_d,x_p)}}\right)} \text{ and } a_T^2(x_d,x_p) = \sqrt{2\log\left(\frac{T}{2\pi}\sqrt{\frac{m_{\mathcal{D}'',2}(x_d,x_p)}{m_{\mathcal{D}'',0}(x_d,x_p)}}\right)}, \tag{2.29}$$

$$\mathcal{F}(x_d,x_p;0) = \left(\left|\mathcal{D}''(x_d,x_p;0)\right| - \rho\right)^+ \tag{2.30}$$

and F_{r_3} the cumulative distribution function of X_{r_3}.

3 Constraints defined in terms of a piece-wise stationary Gaussian process

3.1 Definition of a piece-wise stationary process

We consider in this section the optimization problem (2.1) introduced in section 2.1 with extreme-based and integral-based constraints except that the process $\mathcal{Y}$ is now piece-wise stationary. The period $[0,T]$ is decomposed into n_T intervals $I_i = [(i-1)\Delta T, i\Delta T], i = 1,\cdots,n_T$ and for fixed x_d, x_p the process $\mathcal{Y}$ is defined as:

$$\mathcal{Y}(x_d,x_p;t) = \sum_{i=1}^{n_T} \mathcal{Y}_i(x_d,x_p,s_i;t)\,\mathbb{1}_{I_i}(t) \tag{2.31}$$

where $s_1, \cdots, s_{n_T}$ is a sequence of elements of the set $\{s^1, \cdots, s^{n_s}\}$. Thus, on each interval of time I_i the process $\mathcal{Y}(x_d, x_p; .)$ is equal to the process $\mathcal{Y}_i(x_d, x_p, s_i; .)$ whose distribution depends on s_i.

The processes $\mathcal{Y}_1(x_d, x_p, s_1; .), \cdots, \mathcal{Y}_{n_T}(x_d, x_p, s_{n_T}; .)$ are independent stationary Gaussian processes with zero mean. For $i \in \{1, \cdots, n_T\}$, the distribution of $\mathcal{Y}_i(x_d, x_p, s_i; .)$ is defined by its spectral density $K_{\mathcal{Y}}(x_d, x_p, s_i; .)$ and we denote its spectral moment of order n: $m_{\mathcal{Y},n}(x_d, x_p, s_i)$. For $j = 1, \cdots, n_s$, we define n^j and p^j such that:

$$n^j = \text{card}\{i \in \{1, \cdots, n_T\}, s_i = s^j\} \tag{2.32}$$

and

$$p^j = \frac{n^j}{n_T}. \tag{2.33}$$

Finally, we introduce the independent zero-mean stationary Gaussian processes $\eta_1, \cdots, \eta_{n_T}$ such that for $i = 1, \cdots, n_T$:

$$\mathcal{Y}_i(x_d, x_p, s_i; .) = (h_{\mathcal{Y}}(x_d, x_p; .) * \eta_i)(t). \tag{2.34}$$

and the distribution of η_i depends on s_i. We denote η the sequence of processes $(\eta_1, \cdots, \eta_{n_T})$.

Remark 2.8. The definitions of the objects introduced in this section are motivated by the FOWT case introduced in section 1.5. In this application, each process $\mathcal{Y}_i(x_d, x_p, s_i; .)$ can be seen as the displacement over time of the floating platform subjected to environmental conditions characterized by a sea state represented by s_i, for given design x_d and parametric variables x_p. The constraints then relate to the maximum admissible displacement of the structure or to the accumulated damage of the mooring lines. Furthermore, dividing the time interval $[0, T]$ into n_T intervals is standard in offshore models to represent the different sea states encountered during the period $[0, T]$. To stay close to the engineering terminology, s_i will be called the state of the process $\mathcal{Y}_i(x_d, x_p, s_i; .)$ throughout the thesis.

The reasoning for extreme-based and integral-based constraints reformulation is the same as for the stationary case: the purpose is to provide good approximations of

$$\mathbb{P}_\eta\left(\min_{t\in[0,T]} g_{\text{E}}(x_d, x_p, x_{r_{\text{E}}}, \mathcal{Y}(t)) < 0\right) \text{ and } \mathbb{P}_\eta\left(\int_0^T g_{\text{I}}(x_d, x_p, x_{r_{\text{I}}}, \mathcal{Y}(t))\, dt < 0\right) \tag{2.35}$$

that only rely on spectral properties of the processes $\mathcal{Y}_i(x_d, x_p, s_i; .)$, $i = 1, \cdots, n_T$.

3.2 Approximation of the extreme-based failure probability

To approximate the extreme-based failure probability, we claim that, for fixed $x_d, x_p, x_{r_{\text{E}}}$ and for ΔT large enough:

$$\mathbb{P}_\eta\left(\max_{t\in[0,T]} \mathcal{Y}(x_d, x_p; t) \le x_{r_{\text{E}}}\right) \simeq \prod_{j=1}^{n_s} \mathbb{P}_{\eta_1}\left(\max_{t\in[0,Tp^j]} \mathcal{Y}_1\left(x_d, x_p, s^j; t\right) \le x_{r_{\text{E}}}\right). \tag{2.36}$$

This approximation is justified in section C.1 of appendix C and essentially comes from the fact that the processes $\mathcal{Y}_i(x_d, x_p, s_i; .)$ $(i = 1, \cdots, n_T)$ are independent.

Therefore, if for all $x_d, x_p, x_{r_\mathrm{E}}$, and for all states s^j, the process $\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)$ meets the conditions of theorem 2.1, it follows from equation (2.36) and theorem 2.1 that for ΔT sufficiently large, we can make the following approximation:

$$p_\mathrm{E}(d) \simeq \mathbb{E}_{X_d, X_p, X_{r_\mathrm{E}}}\left[F_\epsilon\left(\sum_{j=1}^{n_s} \exp\left(a_{Tp^j}\left(X_d, X_p, s^j\right)^2 - \frac{a_{Tp^j}\left(X_d, X_p, s^j\right) X_{r_\mathrm{E}}}{\sqrt{m_{\mathcal{Y},0}\left(X_d, X_p, s^j\right)}}\right)\right)\right] \tag{2.37}$$

with

$$a_{Tp^j}\left(x_d, x_p, s^j\right) = \sqrt{2\log\left(\frac{Tp^j}{2\pi}\sqrt{\frac{m_{\mathcal{Y},2}\left(x_d, x_p, s^j\right)}{m_{\mathcal{Y},0}\left(x_d, x_p, s^j\right)}}\right)}. \tag{2.38}$$

Bounds on the approximation error of equation (2.37) are proposed in section D.1 of Appendix D.

Remark 2.9. As for the stationary case, the extreme-based constraint reformulation can also be obtained when each process $\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)$ has a non-zero mean $\mu_\mathcal{Y}\left(x_d, x_p, s^j\right)$. The resulting approximation is the one provided by (2.37) replacing X_{r_E} by $X_{r_\mathrm{E}} - \mu_\mathcal{Y}\left(X_d, X_p, s^j\right)$.

3.3 Approximation of the integral-based failure probability

Proposition 2.1. We denote by $\mathcal{F}_1\left(x_d, x_p, s^j; .\right)$ the process $f\left(\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)\right)$. If for all x_d, x_p, s^j, the process $\mathcal{F}_1\left(x_d, x_p, s^j; .\right)$ is ergodic, we have for almost every x:

$$\mathbb{P}_\eta\left(\frac{1}{\Delta T}\int_0^T f\left(\mathcal{Y}\left(x_d, x_p; t\right)\right) dt > x\right) \underset{\Delta T \to +\infty}{\longrightarrow} \mathbb{1}_{\sum_{j=1}^{n_s} n^j \mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(x_d, x_p, s^j; 0\right)\right] > x}. \tag{2.39}$$

The proof of proposition 2.1 is given in section C.2 of appendix C.

Using proposition 2.1, when ΔT is large enough, we can approximate the integral-based failure probability as follows:

$$p_\mathrm{I}(d) \simeq \mathbb{E}_{X_d, X_p}\left[F_{r_\mathrm{I}}\left(T\sum_{j=1}^{n_s} p^j \mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(X_d, X_p, s^j; 0\right)\right]\right)\right], \tag{2.40}$$

with F_{r_I} the cumulative distribution function of X_{r_I}. The approximation error made in equation (2.40) is discussed in section D.2 of appendix D.

3.4 Optimization problem involving a piece-wise stationary harmonic oscillator

The piece-wise stationary oscillator problem

The oscillator problem presented in section 2.2.4 is slightly modified by considering a piece-wise stationary process $\mathcal{D}$ as defined in section 2.3.1. For fixed values x_d, x_p and for $i = 1, \cdots, n_T$, the process $\mathcal{D}_i\left(x_d, x_p, s_i; .\right)$ is solution of the harmonic oscillator equation (2.22) with an external force η_i. The time-dependent process η_i is a zero-mean stationary Gaussian process with spectral density:

$$K_\eta\left(s_i; \omega\right) = \frac{s_i}{\sqrt{2\pi}} \exp\left(-\frac{\left(s_i\omega\right)^2}{2}\right), \quad s_i > 0. \tag{2.41}$$

In the piece-wise stationary problem, the processes $\mathcal{D}'$ and $\mathcal{D}''$ are defined by the following equation:

$$\mathcal{D}'(x_d, x_p; t) = \sum_{i=1}^{n_T} \mathcal{D}'_i(x_d, x_p, s_i; t)\, \mathbb{1}_{I_i}(t) \tag{2.42}$$

and

$$\mathcal{D}''(x_d, x_p; t) = \sum_{i=1}^{n_T} \mathcal{D}''_i(x_d, x_p, s_i; t)\, \mathbb{1}_{I_i}(t) \tag{2.43}$$

with $\mathcal{D}'_i(x_d, x_p, s_i; .)$ and $\mathcal{D}''_i(x_d, x_p, s_i; .)$ the first and second time derivatives of $\mathcal{D}_i(x_d, x_p, s_i; .)$ (in the sample path sense). Hence, it follows from the reasoning of section 2.2.4 that the processes $\mathcal{D}'_i(x_d, x_p, s_i; .)$, and $\mathcal{D}''_i(x_d, x_p, s_i; .)$ are zero-mean stationary Gaussian processes with respective spectral densities:

$$K_{\mathcal{D}'}(x_d, x_p, s_i; \omega) = \omega^2 \left|H_{\mathcal{D}}(x_d, x_p; \omega)\right|^2 K_\eta(s_i; \omega), \tag{2.44}$$

$$K_{\mathcal{D}''}(x_d, x_p, s_i; \omega) = \omega^4 \left|H_{\mathcal{D}}(x_d, x_p; \omega)\right|^2 K_\eta(s_i; \omega) \tag{2.45}$$

where $H_{\mathcal{D}}(x_d, x_p; \omega)$ has been defined in section 2.2.4.

The spectral moments of order n of $\mathcal{D}'_i(x_d, x_p, s_i; .)$ and $\mathcal{D}''_i(x_d, x_p, s_i; .)$ are respectively denoted $m_{\mathcal{D}',n}(x_d, x_p, s_i)$ and $m_{\mathcal{D}'',n}(x_d, x_p, s_i)$. Furthermore, the arguments used in the stationary case allow to show that these processes meet all the conditions to use the reformulation procedure describe in section 2.3.

Reformulated oscillator problem

If ΔT is large enough, it is reasonable to approximate the constraints of problem (2.23) by applying the results of sections 2.3.2 and 2.3.3. We then obtain the following reformulated problem:

$$\begin{aligned}
&\min_{d\in\Omega_d} \; cost(d) \quad \text{such that} \\
&\mathbb{E}_{X_d,X_p,X_{r_1}}\left[F_\epsilon\left(\sum_{j=1}^{n_s}\exp\left(a^1_{Tp^j}(X_d,X_p,s^j)^2 - \frac{a^1_{Tp^j}(X_d,X_p,s^j)X_{r_1}}{\sqrt{m_{\mathcal{D}',0}(X_d,X_p,s^j)}}\right)\right)\right] < p_s \\
&\mathbb{E}_{X_d,X_p,X_{r_2}}\left[F_\epsilon\left(\sum_{j=1}^{n_s}\exp\left(a^2_{Tp^j}(X_d,X_p,s^j)^2 - \frac{a^2_{Tp^j}(X_d,X_p,s^j)X_{r_2}}{\sqrt{m_{\mathcal{D}'',0}(X_d,X_p,s^j)}}\right)\right)\right] < p_s \\
&\mathbb{E}_{X_d,X_p}\left[F_{r_3}\left(T\sum_{j=1}^{n_s}p^j\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(X_d,X_p,s^j;0\right)\right]\right)\right] < p_s
\end{aligned} \tag{2.46}$$

with

$$a^1_{Tp^j}(x_d, x_p, s^j) = \sqrt{2\log\left(\frac{Tp^j}{2\pi}\sqrt{\frac{m_{\mathcal{D}',2}(x_d,x_p,s^j)}{m_{\mathcal{D}',0}(x_d,x_p,s^j)}}\right)}, \tag{2.47}$$

$$a^2_{Tp^j}(x_d, x_p, s^j) = \sqrt{2\log\left(\frac{Tp^j}{2\pi}\sqrt{\frac{m_{\mathcal{D}'',2}(x_d,x_p,s^j)}{m_{\mathcal{D}'',0}(x_d,x_p,s^j)}}\right)}, \tag{2.48}$$

and

$$\mathcal{F}_1(x_d, x_p, s^j; 0) = \left(\left|\mathcal{D}_1''\left(x_d, x_p, s^j; 0\right)\right| - \rho\right)^+. \tag{2.49}$$

The relation between $\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(x_d, x_p, s^j; 0\right)\right]$ and $m_{\mathcal{D}'',0}\left(x_d, x_p, s^j\right)$ is explicitly given in appendix E.

Remark 2.10. To evaluate $a^1_{Tp^j}\left(x_d, x_p, s^j\right)$, $a^2_{Tp^j}\left(x_d, x_p, s^j\right)$ and $\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(x_d, x_p, s^j; 0\right)\right]$, the spectral moments $m_{\mathcal{D}',0}\left(x_d, x_p, s^j\right)$, $m_{\mathcal{D}',2}\left(x_d, x_p, s^j\right)$, $m_{\mathcal{D}'',0}\left(x_d, x_p, s^j\right)$, and $m_{\mathcal{D}'',2}\left(x_d, x_p, s^j\right)$ need to be numerically computed from the integrals $\int_{\mathbb{R}} \omega^n \left|H_{\mathcal{D}}\left(x_d, x_p; \omega\right)\right|^2 K_\eta\left(s^j; \omega\right) d\omega$ for $n = 2, 4, 6$ (see equations (2.44) and (2.45)). Hence, the evaluation of the spectral moments represents the expensive part of the evaluation of the constraints.

Chapter 3

A new active learning Kriging approach for the reformulated optimization problem: AK-ECO

Contents

1 Motivation

We consider in this section the problem (2.1) introduced in section 2.1 involving a piece-wise stationary process defined in section 2.3.1. After the reformulation of the constraints presented in section 2.3, we end up with the following problem:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \quad \text{such that} \\ &\mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[F_\epsilon\left(\sum_{j=1}^{n_s}\exp\left(M_\mathrm{E}\left(X_d,X_p,X_{r_\mathrm{E}},s^j\right)\right)\right)\right] &< p_s \\ &\mathbb{E}_{X_d,X_p}\left[F_{r_\mathrm{I}}\left(\sum_{j=1}^{n_s}Tp^j\left(M_\mathrm{I}\left(X_d,X_p,s^j\right)\right)\right)\right] &< p_s \end{aligned} \tag{3.1}$$

with

$$M_{\mathrm{E}}\left(x_d, x_p, x_{r_{\mathrm{E}}}, s^j\right) = a_{Tp^j}\left(x_d, x_p, s^j\right)^2 - \frac{a_{Tp^j}\left(x_d, x_p, s^j\right) x_{r_{\mathrm{E}}}}{\sqrt{m_{\mathcal{Y},0}\left(x_d, x_p, s^j\right)}}, \tag{3.2}$$

$$M_{\mathrm{I}}\left(x_d, x_p, s^j\right) = \mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(x_d, x_p, s^j; 0\right)\right]. \tag{3.3}$$

All the notations appearing in problem (3.1) have been introduced in section 2.3.

The *cost* function is supposed to be fast to evaluate. We remark that the problem is now a time-independent one. Although the reformulated constraints are easier to evaluate than the initial ones, for each realization x_d, x_p and for each state s^j, the spectral moments of $\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)$ are required to compute the quantities $M_{\mathrm{E}}\left(x_d, x_p, x_{r_{\mathrm{E}}}, s^j\right)$ and $M_{\mathrm{I}}\left(x_d, x_p, s^j\right)$. Since it is the case for the FOWT problem introduced in section 1.5, we consider that the evaluation of those spectral moments requires the call to an expensive simulator. Therefore, an estimation of one of the constraints with the Monte Carlo method and a sample of size n_{MC} would impose $n_{MC} \times n_s$ calls to the simulator. This would be too computationally expensive especially when we deal with rare events (i.e. when $p_s \ll 1$).

We mentioned in section 1.2 existing methods that solve RBDO problems much faster than the brute Monte Carlo. All the effective approaches in the literature of RBDO methods rely on the assumption that the constraints are expressed as probabilities. We can also write the extreme-based and integral-based reformulated constraints of problem (3.1) as constraints involving failure probabilities as follows:

$$\mathbb{P}_{X_d, X_p, X_{r_{\mathrm{E}}}, X_\epsilon}\left(X_\epsilon - \sum_{j=1}^{n_s} \exp\left(M_{\mathrm{E}}\left(X_d, X_p, X_{r_{\mathrm{E}}}, s^j\right)\right) < 0\right) < p_s, \tag{3.4}$$

$$\mathbb{P}_{X_d, X_p, X_{r_{\mathrm{I}}}}\left(X_{r_{\mathrm{I}}} - \sum_{j=1}^{n_s} Tp^j M_{\mathrm{I}}\left(X_d, X_p, s^j\right) < 0\right) < p_s, \tag{3.5}$$

where X_ϵ is a random variable with an exponential distribution of parameter 1.

However, since each evaluation of the functions M_{E} and M_{I} requires a call to an expensive simulator, one evaluation of the performance function of each of the constraints (3.4) and (3.5) would need n_s simulations. When n_s is large, which is the case for offshore applications (Vorpahl et al. (2013)), the double-loop, single-loop and decoupled-loop approaches can be too expensive. This is also the case for the adaptive metamodel approaches since they always replace the whole performance function.

Motivated by the limitation of the existing approaches in RBDO, we propose a new method better suited for the reformulated problem, which we call Adaptive Kriging for Expectation Constraints Optimization (AK-ECO). The procedure of AK-ECO is described in section 3.2. In section 3.3, the oscillator problem introduced in section 2.3 is solved with AK-ECO and state-of-the-art algorithms in RBDO: RIA, PMA, SORA and the Stieng method.

Remark 3.1. Presented as in problem (3.1), the reformulated constraints depend on the same piece-wise stationary process $\mathcal{Y}$. However, the resolution methods that are presented in this chapter can be applied to constraints that depend on different processes. When several constraints depend on the same process, it can be noticed that the outputs of each simulation can be used in the estimation of the different constraints since they depend on the same quantities (i.e. the spectral moments of the process).

Remark 3.2. For simplicity, we present a problem with two constraints (one extreme-based and one integral-based) but the resolution methods that we introduce can be applied to several constraints of each type. Furthermore, a problem with extreme and integral based constraints but with a stationary process $\mathcal{Y}$ can be solved with the same approaches since the reformulated problem would be identical to the piece-wise stationary case considering $n_s = 1$.

2 AK-ECO

In AK-ECO, for each expectation constraint of problem (3.1), a metamodel is built to replace the expensive function involved in the reformulated failure probability. We thus obtain as many metamodels as there are constraints. Then, cycles of optimization are carried out. During each cycle, the metamodels are sequentially enriched and the design vector is updated. The particularity of our approach lies in the metamodel and active learning strategy which are adapted to the reformulated constraints of the studied problem.

2.1 Metamodel strategy: Kriging

In problem (3.1), since the expensive functions of the reformulated extreme-based and integral-based constraints are M_{E} and M_{I}, we propose to build a metamodel for each of these functions. Unlike other metamodel approaches, the metamodels do not replace the performance functions. Thus, for each $(x_d, x_p, x_{r_{\mathrm{E}}})$, the functions M_{E} and M_{I} need to be evaluated only on the relevant states s^j as we will see below. This will reduce the number of calls to the simulator required to obtain accurate estimations of the constraints, especially when n_s is large.

As in Dubourg (2011) and Moustapha and Sudret (2019a), for each constraint, we build the metamodel in a so-called augmented space which makes it possible to use and enrich a single model during the whole procedure of AK-ECO. To do so, the augmented space spans both the design space and the space of uncertainties. The augmented spaces of M_{E} and M_{I} are respectively denoted $\Omega_{\mathrm{E}}^{aug}$ and $\Omega_{\mathrm{I}}^{aug}$. To define precisely those spaces, we need to introduce some notations.

Let $(d_1, \cdots, d_{n_d}) \in \Omega_d = \Omega_{d_1} \times \cdots \times \Omega_{d_{n_d}} \subset \mathbb{R}^{n_d}$ be a design point and $X_{d_1}, \cdots, X_{d_{n_d}}$ the random variables with respective quantile functions $F_{d_1}^{-1}, \cdots, F_{d_{n_d}}^{-1}$ describing the uncertainties at this point. The random vector X_p is composed of n_p random variables $X_{p_i}, (i = 1, \cdots, n_p)$ with quantile function $F_{p_i}^{-1}, (i = 1, \cdots, n_p)$. Finally the quantile function of $X_{r_{\mathrm{E}}}$ is denoted $F_{r_{\mathrm{E}}}^{-1}$. The augmented spaces are then defined as follows:

$$\Omega_{\mathrm{E}}^{aug} = \Omega_d^{aug} \times \Omega_p^{aug} \times \Omega_{r_{\mathrm{E}}}^{aug} \times \Omega_S, \qquad \Omega_{\mathrm{I}}^{aug} = \Omega_d^{aug} \times \Omega_p^{aug} \times \Omega_S, \tag{3.6}$$

with

$$\Omega_d^{aug} = \prod_{i=1}^{n_d} \left[\inf_{d_i \in \Omega_{d_i}} F_{d_i}^{-1}(\alpha), \sup_{d_i \in \Omega_{d_i}} F_{d_i}^{-1}(1-\alpha) \right], \tag{3.7}$$

$$\Omega_p^{aug} = \prod_{i=1}^{n_p} \left[F_{p_i}^{-1}(\alpha), F_{p_i}^{-1}(1-\alpha) \right], \tag{3.8}$$

$$\Omega_{r_{\mathrm{E}}}^{aug} = \left[F_{r_{\mathrm{E}}}^{-1}(\alpha), F_{r_{\mathrm{E}}}^{-1}(1-\alpha) \right], \tag{3.9}$$

$$\Omega_S = \{s^1, \cdots, s^{n_s}\}, \tag{3.10}$$

and α is chosen by the user (different values of α could be considered for each set of the Cartesian products Ω_d^{aug}, Ω_p^{aug}, and for $\Omega_{r_\mathrm{E}}^{aug}$). The value of α should be set so that the augmented space covers a large domain of the uncertainty space. Thus, the metamodel calibrated on this domain can provide an accurate prediction for any likely outcome of the random variables. In this thesis, α is set to 10^{-6}.

The dimensions of Ω_E^{aug} and Ω_I^{aug} are respectively equal to n_d+n_p+1+$\dim_\mathrm{s}$ and n_d+n_p+$\dim_\mathrm{s}$ (with $s^j \in \mathbb{R}^{\dim_\mathrm{s}}$ for $j = 1, \cdots, n_s$). When $n_d + n_p + \dim_\mathrm{s}$ is relatively small (as it is the case for the FOWT problem under study in this thesis for which this sum equals 9), the kriging model is particularly well suited to our approach (we refer the reader to section 1.1 for a description of the kriging method). At each point of the augmented space, the kriging models provide predictions of M_E and M_I in the form of Gaussian random variables. The means of these random variables are used as predictors while the standard deviations measure the accuracy of these predictors. This latter information makes kriging model particularly well suited to active learning and therefore, this metamodel method is chosen in AK-ECO.

We can notice that $\Omega_\mathrm{E}^{aug} = \Omega_\mathrm{I}^{aug} \times \Omega_{r_\mathrm{E}}^{aug}$. Thus, to calibrate the metamodels, we only use one DoE of Ω_E^{aug}. Using one DoE is interesting when the two constraints depend on the same process $\mathcal{Y}$ since each simulation can be used to enrich both metamodels. Indeed, in that case, we recall that to evaluate M_E and M_I at a point $\left(x_d, x_p, x_{r_\mathrm{E}}, s^j\right)$, we only need the spectral moments of the process $\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)$. Although this approach is not optimal to cover the space Ω_I^{aug}, this is a simple way to use a common DoE for all the metamodels.

Remark 3.3. In the stationary case (i.e. when $n_s = 1$), it is not necessary for the augmented spaces to span the space Ω_S. Hence, the respective augmented spaces become $\Omega_\mathrm{E}^{aug} = \Omega_d^{aug} \times \Omega_p^{aug} \times \Omega_{r_\mathrm{E}}^{aug}$ and $\Omega_\mathrm{I} = \Omega_d^{aug} \times \Omega_p^{aug}$.

2.2 Procedure

To solve the reformulated problem (3.1), AK-ECO begins with the initialization of the design point and the kriging models. Then, the reformulated problem is solved through cycles of optimization. The initialization and the optimization cycle structure are described below.

Initialization: the initial design point d^0 is chosen by the user. An initial DoE, denoted DoE^0, is computed and used to calibrate the initial metamodels $\widetilde{M_\mathrm{E}}^0$, $\widetilde{M_\mathrm{I}}^0$ of the functions M_E and M_I with a procedure described below. At the end of the initialization, the first cycle of optimization ($k = 1$) can begin.

Optimization cycle: we respectively denote d^{k-1}, DoE^{k-1}, $\widetilde{M_\mathrm{E}}^{k-1}$ and $\widetilde{M_\mathrm{I}}^{k-1}$, the design point, DoE and kriging models recovered from the initialization if $k = 1$ or from the previous cycle if $k > 1$. Each cycle is numbered k and is decomposed into two steps:

Step 1. Local enrichment at d^{k-1} of the metamodels $\widetilde{M_\mathrm{E}}^{k-1}$ and $\widetilde{M_\mathrm{I}}^{k-1}$:

Step 1.a. An accuracy criterion assesses the precision of each metamodel at d^{k-1} (we detail this step in section 3.2.2).

Step 1.b. For each inaccurate metamodel, one local enrichment is carried out. The local refinement of the metamodel consists in adding to the shared

DoE, the point $x_{enr} = \left(x_{d^{k-1}}, x_p, x_{r_\mathrm{E}}, s^j\right)$ selected by the procedure described below. The simulator is evaluated at this point and the spectral moments obtained are used to recalibrate all the kriging models.

Steps 1.a and 1.b are repeated until each kriging model meets the accuracy condition of Step 1.a. At the end of Step 1, the enriched DoE and kriging models are denoted DoE^k, $\widetilde{M_\mathrm{E}}^k$ and $\widetilde{M_\mathrm{I}}^k$. For each point of their respective augmented spaces, the means of the kriging models are respectively denoted $\mu_\mathrm{E}^k\left(x_d, x_p, x_{r_\mathrm{E}}, s^j\right)$ and $\mu_\mathrm{I}^k\left(x_d, x_p, s^j\right)$ and serve as predictions.

Step 2. The reformulated problem (3.1) is solved using the optimization algorithm chosen by the user starting from d^{k-1}. At each iteration of the optimization, the constraints are estimated through Monte Carlo and the expensive functions are replaced by their current surrogates. For a design d, these estimations, denoted $p_\mathrm{E}^k(d)$ and $p_\mathrm{I}^k(d)$, are given by:

$$p_\mathrm{E}^k(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon\left(\sum_{j=1}^{n_s} \exp\left(\mu_\mathrm{E}^k\left(x_d^i, x_p^i, x_{r_\mathrm{E}}^i, s^j\right)\right)\right), \tag{3.11}$$

$$p_\mathrm{I}^k(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_\mathrm{I}}\left(\sum_{j=1}^{n_s} Tp^j\left(\mu_\mathrm{I}^k\left(x_d^i, x_p^i, s^j\right)\right)\right), \tag{3.12}$$

where $\left\{\left(x_d^i, x_p^i, x_{r_\mathrm{E}}^i\right), i = 1, \cdots, n_{MC}\right\} = \Omega_{MC}(d)$ is the Monte Carlo sample of the random vector $(X_d, X_p, X_{r_\mathrm{E}})$.

Step 2 does not require any call to the expensive simulator. Once the optimization algorithm has converged, a new design point denoted d^k is obtained.

At the end of each cycle k, the following condition is evaluated:

$$\left\|\overline{d^{k-1}} - \overline{d^k}\right\| < \epsilon_d \quad \text{OR} \quad \left|\overline{cost\left(d^{k-1}\right)} - \overline{cost\left(d^k\right)}\right| < \epsilon_{cost} \qquad \text{(stopping condition)}$$

where $\overline{d}$ and $\overline{cost\left(d\right)}$ represent the normalized versions of d and $cost(d)$ in $[0,1]$. If this condition is met, AK-ECO is stopped and the minimum retained, denoted d^{min}, is d^k, otherwise, $k = k+1$ and a new cycle begins from step 1. The stopping criterion of AK-ECO does not include a condition on the satisfaction of the constraints since this point is verified at the end of the optimization during step 2.

The full procedure of AK-ECO is summarized in Figure 3.1.

Remark 3.4. If the random vector X_d depends on d such that $X_d = d + X$ with X a zero-mean random vector, it is possible to use the same Monte Carlo sample Ω_{MC} throughout AK-ECO where $\Omega_{MC} = \left\{\left(x^i, x_p^i, x_{r_\mathrm{E}}^i\right), i = 1, \cdots, n_{MC}\right\}$ is a sample of $(X, X_p, X_{r_\mathrm{E}})$ and thus, $\Omega_{MC}(d) = \left\{\left(d + x^i, x_p^i, x_{r_\mathrm{E}}^i\right), \left(x^i, x_p^i, x_{r_\mathrm{E}}^i\right) \in \Omega_{MC}\right\}$.

Kriging models initialization

The goal of the first kriging models is to provide good predictions of their respective functions over the whole augmented spaces. A space-filling DoE is therefore appropriate. As explained above, we use one DoE for both metamodels: only one sample of Ω_E^{aug} is needed. Therefore a space-filling DoE of n_{DoE}^0 points of $\Omega_d^{aug} \times \Omega_p^{aug} \times \Omega_{r_\mathrm{E}}^{aug}$ is constructed. We then concatenate to this DoE a uniform sample of n_{DoE}^0 points of Ω_S. The resulting DoE is denoted DoE^0. The simulator is evaluated for each point of DoE^0 to calibrate the initial krigings models denoted $\widetilde{M_\mathrm{E}}^0$ and $\widetilde{M_\mathrm{I}}^0$.

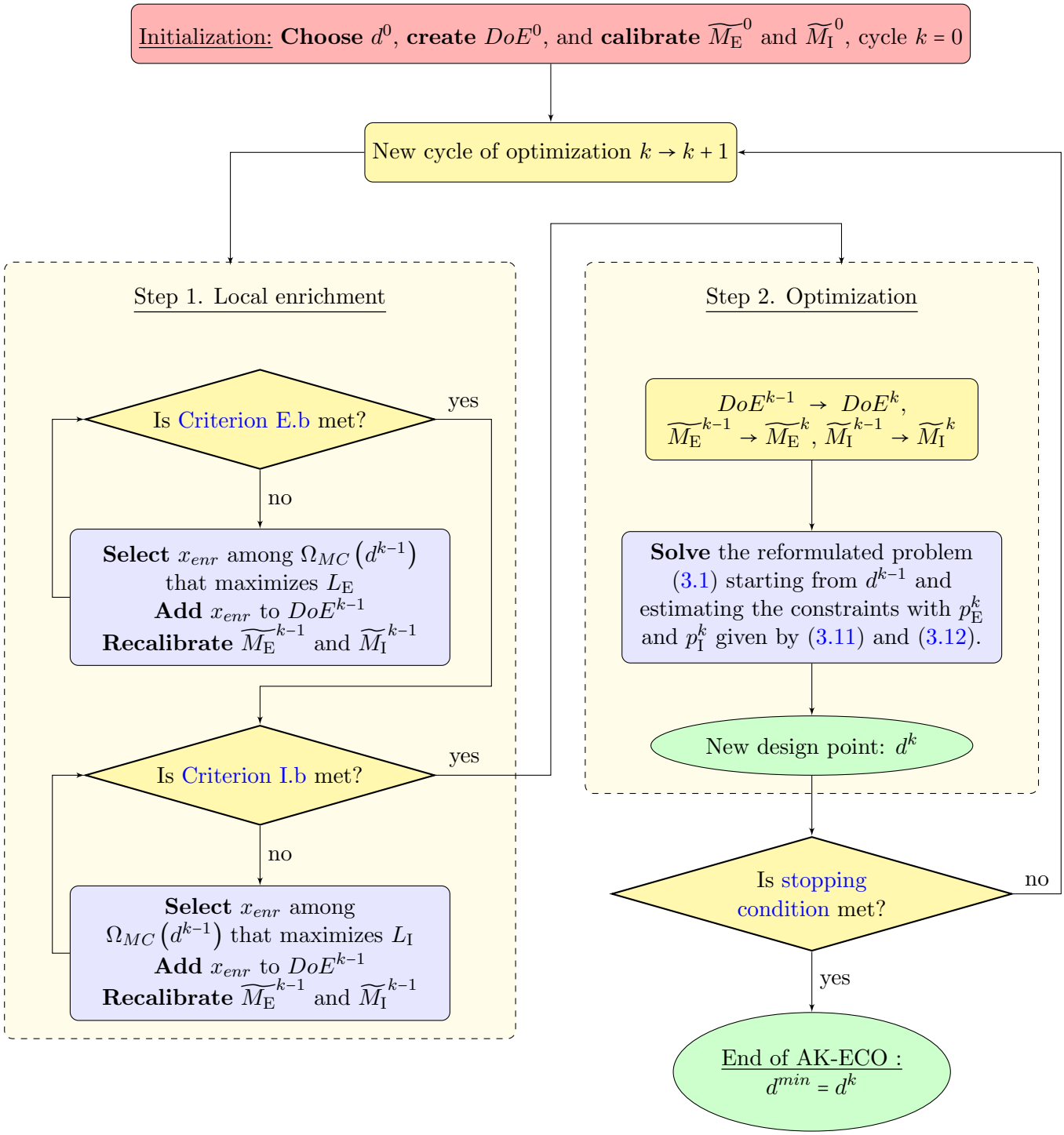


Figure 3.1: Flowchart of AK-ECO

Accuracy criteria

During step 1 of the k-th cycle of optimization, the current kriging model $\widetilde{M_{\mathrm{E}}}^{k-1}$ at $\left(x_d, x_p, x_{r_{\mathrm{E}}}, s^j\right)$ follows a normal distribution with mean and standard deviation denoted $\mu_{\mathrm{E}}^{k-1}\left(x_d, x_p, x_{r_{\mathrm{E}}}, s^j\right)$ and $\sigma_{\mathrm{E}}^{k-1}\left(x_d, x_p, x_{r_{\mathrm{E}}}, s^j\right)$. To evaluate the precision of the approximation $p_{\mathrm{E}}^{k-1}(d)$ of the true reformulated failure probability at d, we adapt the approach proposed by Dubourg (2011) and described in section 1.2. We compute the following quantities:

$$p_{\mathrm{E},+}^{k-1}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon \left(\sum_{j=1}^{n_s} \exp\left(\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right) + 2\sigma_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)\right)\right), \tag{3.13}$$

$$p_{\mathrm{E},-}^{k-1}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon \left(\sum_{j=1}^{n_s} \exp\left(\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right) - 2\sigma_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)\right)\right). \tag{3.14}$$

As the exponential function and F_ϵ are strictly increasing, we have:

$$p_{\mathrm{E},-}^{k-1}(d) < p_{\mathrm{E}}^{k-1}(d) < p_{\mathrm{E},+}^{k-1}(d). \tag{3.15}$$

The distance between $p_{\mathrm{E},-}^{k-1}(d)$ and $p_{\mathrm{E},+}^{k-1}(d)$ is an indicator of the uncertainty of the constraint estimation $p_{\mathrm{E}}^{k-1}(d)$. In Dubourg (2011), a criterion based on the ratio between similar optimistic and pessimistic estimations of the failure probability is proposed. However, since it is useless to know precisely the true failure probability if it is far from p_s, we modify it and measure the accuracy of the metamodel at d with $A_{\mathrm{E}}(d)$ defined as follows:

$$A_{\mathrm{E}}(d) = \frac{\left|p_{\mathrm{E}}^{k-1}(d) - p_s\right|}{p_{\mathrm{E},+}^{k-1}(d) - p_{\mathrm{E},-}^{k-1}(d)}. \tag{3.16}$$

During the k-th cycle of AK-ECO, the metamodel is considered accurate enough if the following condition is met:

$$A_{\mathrm{E}}\left(d^{k-1}\right) > 1. \tag{Criterion E}$$

This criterion is met if the distance between the low and high estimations of the constraint at d^{k-1} is less than the distance between the estimation of the constraint at d^{k-1} and p_s. In this case, we have reasonable grounds to believe that the kriging model accurately predicts whether a point near d^{k-1} belongs or not to the feasible domain.

For the integral-based constraint, a similar criterion is proposed. We consider the kriging model $\widetilde{M_{\mathrm{I}}}^{k-1}$ whose mean and standard deviation at $\left(x_d, x_p, s^j\right)$ are denoted $\mu_{\mathrm{I}}^{k-1}\left(x_d, x_p, s^j\right)$ and $\sigma_{\mathrm{I}}^{k-1}\left(x_d, x_p, s^j\right)$. The function $F_{r_{\mathrm{I}}}$ is also increasing and $p^j > 0$ for $j = 1, \cdots, n_s$. Thus we have, for all d:

$$p_{\mathrm{I},-}^{k-1}(d) < p_{\mathrm{I}}^{k-1}(d) < p_{\mathrm{I},+}^{k-1}(d) \tag{3.17}$$

with

$$p_{\mathrm{I},+}^{k-1}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_{\mathrm{I}}} \left(\sum_{j=1}^{n_s} Tp^j \left(\mu_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right) + 2\sigma_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right)\right)\right), \tag{3.18}$$

$$p_{\mathrm{I},-}^{k-1}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_{\mathrm{I}}} \left(\sum_{j=1}^{n_s} Tp^j \left(\mu_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right) - 2\sigma_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right)\right)\right). \tag{3.19}$$

The accuracy criterion for the integral-based constraint at d^{k-1} is:

$$A_{\mathrm{I}}\left(d^{k-1}\right) > 1 \qquad \text{(Criterion I)}$$

with

$$A_{\mathrm{I}}\left(d\right) = \frac{\left|p_{\mathrm{I}}^{k-1}(d) - p_s\right|}{p_{\mathrm{I},+}^{k-1}(d) - p_{\mathrm{I},-}^{k-1}(d)}. \tag{3.20}$$

To avoid a too large number of enrichment steps during the same cycle, a maximal number $n_{\max}$ of enrichment steps is imposed for each metamodel and cycle. Finally, $\widetilde{M_{\mathrm{E}}}^{k-1}$ is considered accurate enough if

$$p_{\mathrm{E},-}^{k-1}(d^{k-1}) > p_s - \epsilon_p \text{ and } p_{\mathrm{E},+}^{k-1}(d^{k-1}) < p_s + \epsilon_p \tag{3.21}$$

where ϵ_p is chosen by the user (we proceed similarly for $\widetilde{M_{\mathrm{I}}}^{k-1}$). Hence, $\widetilde{M_{\mathrm{E}}}^{k-1}$, respectively $\widetilde{M_{\mathrm{I}}}^{k-1}$, is enriched if Criterion E.b, respectively Criterion I.b, is not met and these criteria are defined as:

Criterion E is met OR $\left(n_{\mathrm{E}}^k \geq n_{\max}\right)$ OR $\left(p_s - \epsilon_p < p_{\mathrm{E},-}^{k-1}(d^{k-1}) < p_{\mathrm{E},+}^{k-1}(d^{k-1}) < p_s + \epsilon_p\right)$ (Criterion E.b)

Criterion I is met OR $\left(n_{\mathrm{I}}^k \geq n_{\max}\right)$ OR $\left(p_s - \epsilon_p < p_{\mathrm{I},-}^{k-1}(d^{k-1}) < p_{\mathrm{I},+}^{k-1}(d^{k-1}) < p_s + \epsilon_p\right)$ (Criterion I.b)

where n_{E}^k and n_{I}^k are the numbers of enrichment steps of each metamodel during cycle k. Therefore, during each cycle of AK-ECO, the number of enrichments is at most equal to $n_{\max}$ multiplied by the number of metamodels.

When both criteria Criterion E.b and Criterion I.b are met, step 1 ends and step 2 begins.

Selection of the enrichment point

In the metamodel-based approaches aiming to estimate a failure probability, the purpose of the metamodel is to predict whether or not a point belongs to the safe space defined by the performance function. The learning function then seeks to select enrichment points close to the limit state surface. However, in problem (3.1), the reformulated failure probabilities are expressed as expectations. For the first constraint, improving the estimation of this expectation is not equivalent to enriching the metamodel of M_{E} such that it correctly predicts a specific contour line of M_{E}. The same goes for the second constraint. Therefore, the usual learning functions are not suitable for AK-ECO and we propose in this section new ones adapted to our metamodel strategy.

During the k-th cycle, if Criterion E.b is not met, the model $\widetilde{M_{\mathrm{E}}}^{k-1}$ is not considered sufficiently accurate at d^{k-1}. To improve its precision, a point x_{enr} maximizing a learning function L_{E} is selected among the Monte Carlo sample $\Omega_{MC}\left(d^{k-1}\right)$ used in equation (3.11) to estimate the extreme-based constraint at d^{k-1}:

$$x_{enr} = \underset{\substack{\left(x_{d^{k-1}}^i, x_p^i, x_{r_{\mathrm{E}}}^i\right) \in \Omega_{MC}\left(d^{k-1}\right) \\ s^j \in \{s^1, \cdots, s^{n_s}\}}}{\operatorname{argmax}} L_{\mathrm{E}}(x_{d^{k-1}}^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j). \tag{3.22}$$

The goal of criterion L_{E} is to favor points where the uncertainty of prediction of $\widetilde{M_{\mathrm{E}}}^{k-1}$ implies important uncertainties on the constraint estimation at d^{k-1}. Hence, we define L_{E} as follows:

$$\begin{aligned} L_{\mathrm{E}}(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j) &= f_{(X_d, X_p, X_{r_{\mathrm{E}}})}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i\right) \\ &\times \left[F_\epsilon\left(e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right) + 2\sigma_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)} + \sum_{j' \neq j} e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^{j'}\right)} \right) \right. \\ &\left. - F_\epsilon\left(e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right) - 2\sigma_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)} + \sum_{j' \neq j} e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^{j'}\right)} \right) \right], \end{aligned} \tag{3.23}$$

with $f_{(X_d, X_p, X_{r_{\mathrm{E}}})}$ the probability density function of the random vector $(X_d, X_p, X_{r_{\mathrm{E}}})$. Once x_{enr} is selected, it is added to the current DoE, DoE^{k-1}, and one call to the simulator is made at this point.

Let us explain this learning function. We recall that, during the k-th cycle, the Monte Carlo estimation of the considered constraint at d is:

$$p_{\mathrm{E}}^{k-1}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon\left(\sum_{j=1}^{n_s} \exp\left(\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)\right) \right). \tag{3.24}$$

The learning function L_{E} at $(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j)$ considers the uncertainties on the term

$$F_\epsilon\left(\sum_{j=1}^{n_s} \exp\left(\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)\right) \right) \tag{3.25}$$

of the sum in (3.24) accounting only for the uncertainty of the kriging at $(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i)$ and for the state s^j for a given j. Therefore, high and low estimations of (3.25), considering only the uncertainty at $(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j)$, are respectively given by:

$$F_\epsilon\left(e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right) + 2\sigma_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)} + \sum_{j' \neq j} e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^{j'}\right)} \right) \tag{3.26}$$

and

$$F_\epsilon\left(e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right) - 2\sigma_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)} + \sum_{j' \neq j} e^{\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^{j'}\right)} \right). \tag{3.27}$$

The difference between (3.26) and (3.27) indicates the uncertainties on the term (3.25) due to the uncertainty of the kriging at $(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j)$. The learning function L_{E} thus selects the point which maximizes this distance. Finally, we multiply this distance by $f_{(X_d, X_p, X_{r_{\mathrm{E}}})}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i\right)$ to favor points where the probability density function of $(X_d, X_p, X_{r_{\mathrm{E}}})$ is high.

For integral-based constraint, the idea is the same, if Criterion I.b is not met, a new point x_{enr} is selected as follows:

$$x_{enr} = \underset{\substack{\left(x_{d^{k-1}}^i, x_p^i, x_{r_{\mathrm{E}}}^i\right) \in \Omega_{MC}\left(d^{k-1}\right) \\ s^j \in \{s^1, \cdots, s^{n_s}\}}}{\operatorname{argmax}} L_{\mathrm{I}}(x_{d^{k-1}}^i, x_p^i, s^j) \tag{3.28}$$

with:

$$L_{\mathrm{I}}(x_d^i, x_p^i, s^j) = f_{(X_d,X_p)}\left(x_d^i, x_p^i\right) \times \left[F_{r_{\mathrm{I}}}\left(Tp^j\left(\mu_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right) + 2\sigma_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right)\right) + \sum_{j'\neq j} Tp^{j'}\mu_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^{j'}\right)\right) - F_{r_{\mathrm{I}}}\left(Tp^j\left(\mu_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right) - 2\sigma_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^j\right)\right) + \sum_{j'\neq j} Tp^{j'}\mu_{\mathrm{I}}^{k-1}\left(x_d^i, x_p^i, s^{j'}\right)\right)\right], \quad (3.29)$$

with $f_{(X_d,X_p)}$ the probability density function of the random vector (X_d, X_p).

Remark 3.5. In equation (3.28), we note that L_{I} does not depend on $x_{r_{\mathrm{E}}}^i$. The coordinates of $x_{enr} = \left(x_{d^{k-1}}^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)$ are such that $\left(x_{d^{k-1}}^i, x_p^i, s^j\right)$ maximizes L_{I} and $x_{r_{\mathrm{E}}}^i$ corresponds to $\left(x_{d^{k-1}}^i, x_p^i\right)$ in the Monte Carlo sample $\Omega_{MC}\left(d^{k-1}\right)$.

Remark 3.6. In the procedure of AK-ECO, we choose to rely on the standard deviation of the kriging model to select a new enrichment point. Other adaptive kriging strategies consider the observed error of the metamodel as well in the selection of the enrichment point such as the method proposed in Le Gratiet and Cannamela (2015) which adjusts the variance with the observed leave-one-out cross-validation (LOO-CV) error. Following the same idea, the learning functions L_E and L_I can be weighted by the LOO-CV error in AK-ECO to give more importance to areas of the augmented space where the metamodel previously gave poor predictions.

3 Validation of AK-ECO on the harmonic oscillator problem

To validate the method proposed in section 3.2, we study the resolution of the reformulated problem of the piece-wise stationary harmonic oscillator described in section 2.3. First, a resolution of the problem with a double loop approach using the Monte Carlo method to estimate the constraints is used as a reference. Then, the problem is solved with AK-ECO and the resolution of its probabilistic formulation (see equations (3.4) and (3.5) for the probabilistic formulations of extreme and integral based reformulated constraints) is carried out with our implementation of RIA, PMA, SORA and the Stieng approach presented in section 1.2.

3.1 Cost function, sources of uncertainties and parameters of the problem

Here the cost function is:

$$cost\left(d_1, d_2\right) = d_2 - 10d_1. \quad (3.30)$$

The optimization problem is solved on the design space $\Omega_d = [1,5] \times [20,50]$ and for the parameters ρ, n_T, ΔT, T, n_s and p_s given in Table 3.1.

Parameter	ρ	n_T	ΔT	T	n_s	p_s
Value	1	100	216	21600	7	10^{-4}

Table 3.1: Parameters of the problem

The distributions considered for the random variables X_{d_1}, X_{d_2}, X_p, X_{r_1}, X_{r_2}, and X_{r_3} are given in Table 3.2 and the couples $\left(s^j, p^j\right)$, $j = 1, \cdots, 7$ in Table 3.3.

Uncertainty	Distribution	Uncertainty	Distribution
X_{d_1}	$\mathcal{U}\left[d_1 - 0.3, d_1 + 0.3\right]$	X_{r_1}	$\mathcal{N}\left(1, 0.1^2\right)$
X_{d_2}	$\mathcal{U}\left[d_2 - 1, d_2 + 1\right]$	X_{r_2}	$\mathcal{N}\left(2.5, 0.25^2\right)$
X_p	$\mathcal{U}[0.5, 1.5]$	X_{r_3}	$\mathcal{N}\left(15, 3^2\right)$

Table 3.2: Distributions of $X_{d_1}, X_{d_2}, X_p, X_{r_1}, X_{r_2}, X_{r_3}$

The notations $\mathcal{U}[a, b]$ and $\mathcal{N}\left(\mu, \sigma^2\right)$ refer respectively to the uniform distribution on $[a, b]$ and the normal distribution of mean μ and standard deviation σ.

j	1	2	3	4	5	6	7
s^j	1.20	1.16	1.10	1.05	0.99	0.95	0.90
p^j	0.21	0.17	0.18	0.16	0.13	0.09	0.06

Table 3.3: Couples $\left(s^j, p^j\right)$

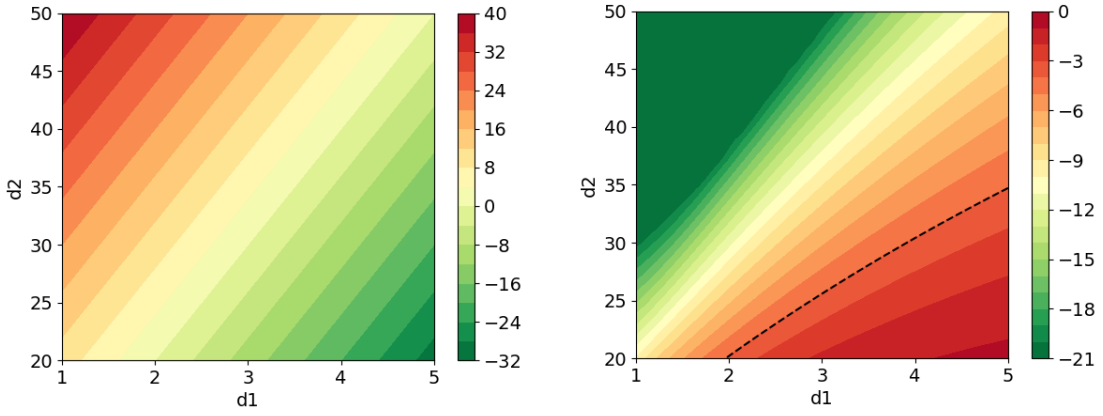
3.2 Graphs of the cost function and the failure probabilities

Here, the functions involved in the reformulated constraints are actually not very expensive. Thus, for the i-th constraint ($i = 1, 2, 3$), the reformulated failure probabilities at any design point d can be estimated by Monte Carlo with a sample of size 30000 (the result is denoted $p_i^{MC}(d)$). So we are able to check the performances of the different optimization methods. The contour lines of the cost function and of $\log\left(p_1^{MC}\right)$, $\log\left(p_2^{MC}\right)$, $\log\left(p_3^{MC}\right)$ are displayed in Figure 3.2. The black dotted line indicates the design points d such that $p_i^{MC}(d) = 10^{-4}$ ($i = 1, 2, 3$).

3.3 Implementations

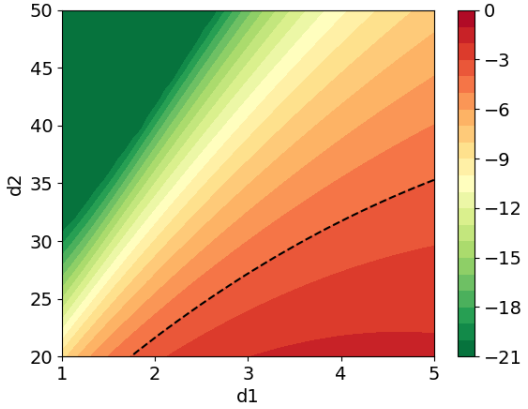
The reference results are obtained using the COBYLA (Powell (1994)) optimization algorithm and a massive Monte Carlo method to estimate the reformulated failure probabilities (this approach is denoted MC). The COBYLA algorithm, which is a local derivative-free optimization solver, is also used for the other methods. This optimizer has been chosen since in the FOWT problem, the derivative of the cost function is not known. Moreover, all the comparison methods converge correctly with COBYLA.

The FORM method in RIA is performed with the Abdo-Rackwitz algorithm (Abdo and Rackwitz (1991)) available in the OpenTURNS Python package (Baudin et al. (2016)). The HMV algorithm (Youn et al. (2003)) is implemented to solve the inverse reliability analysis in PMA. In SORA and Stieng, the SQP (Kraft (1988)) algorithm is chosen instead of HMV since it performs better on the studied case. The size n_{MC} of the sample used in the MC method is 30000.

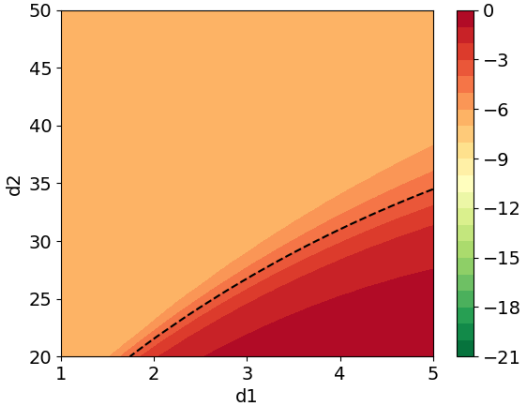


(a) cost function

(b) $\log\left(p_1^{MC}\right)$



(c) $\log\left(p_2^{MC}\right)$



(d) $\log\left(p_3^{MC}\right)$

Figure 3.2: Contour lines of the cost function and of $\log\left(p_1^{MC}\right)$, $\log\left(p_2^{MC}\right)$, and $\log\left(p_3^{MC}\right)$

In Stieng and AK-ECO, the kriging implementation of OpenTURNS is used with a constant trend and an anisotropic 3/2-Matérn covariance kernel. Moreover, the hyperparameters are updated at each calibration of the metamodels. The points of the successive DoEs are normalized in $[0, 1]$ and each hyperparameter is selected in $[10^{-5}, 10]$ using the multistart Truncated Newton Constrained solver implemented in OpenTURNS from 20 initial points. In AK-ECO, the states s_i are treated as continuous variables by the kriging kernel since, in the oscillator application, the $s^j, j = 1, \cdots, n_s$ are real numbers. In Stieng, as proposed by the authors in Stieng and Muskulus (2020), the initial DoE is a Sobol sequence (Sobol' (1967)). The authors do not provide guidelines for the size of this DoE but since the metamodel is calibrated in a 4-dimensional space, an initial DoE of size 12 is used for the first cycle, 40 for the second cycle, 160 for the third one and 400 for the next cycles. For AK-ECO, the initial space-filling DoE is a Latin Hypercube Sampling (LHS) maximin (Santner et al. (2018)) of size 50. The maximum number of enrichment steps per cycle and per constraint n_{max} is set to 15.

For Stieng, SORA and AK-ECO, the cycles of optimization stop if the stopping condition introduced in section 3.2.2 is met for ϵ_d and ϵ_{cost} equal to 10^{-3}.

3.4 Numerical results

The problem under study is solved with each approach starting from the center of the design space $(3, 35)$. The results are displayed in Table 3.4. The first and second rows indicate the design point d^{min} obtained by each approach and the value of the cost function at this point. For $i = 1, 2, 3$, the i-th reformulated failure probability at d^{min} is then estimated with a massive Monte Carlo of 30000 points. This estimation is denoted $p_i^{MC}(d^{min})$.

The results displayed in Table 3.4 with the MC approach, with AK-ECO and the computations of $p_i^{MC}(d^{min})$ have all been done using the same Monte Carlo sample. This is possible since, in the oscillator problem, the random vector X_d is centered in d (see remark 3.4).

As explained in remark 2.10, the expensive part of the constraints is the evaluation of the spectral moments of $\mathcal{D}_1\left(x_d, x_p, s^j; .\right)$. Therefore, during the resolution of the problem, one estimation of the spectral moments for one point $\left(x_d, x_p, s^j\right)$ is considered as one call to the expensive simulator. The number of calls to the simulator by each method is denoted n_{call}.

It is important to notice that, unlike usual papers in reliability analysis, the number n_{call} does not refer to the number of calls to the performance functions but to the number of simulations. Hence, for RIA, PMA, SORA and Stieng, the number of calls to the performance functions is equal to $n_{call}/7$ since $n_s = 7$.

	MC	RIA	PMA	SORA	Stieng	AK-ECO
d^{min}	(5.0, 35.74)	(5.0, 35.22)	(5.0, 35.04)	(5.0, 36.77)	(5.0, 37.29)	(5.0, 35.73)
$cost(d^{min})$	-14.26	-14.78	-14.96	-13.23	-12.70	-14.27
$p_1^{MC}\left(d^{min}\right)$	0.8×10^{-4}	1.0×10^{-4}	1.0×10^{-4}	0.4×10^{-4}	0.3×10^{-4}	0.8×10^{-4}
$p_2^{MC}\left(d^{min}\right)$	1.0×10^{-4}	1.3×10^{-4}	1.3×10^{-4}	0.6×10^{-4}	0.5×10^{-4}	1.0×10^{-4}
$p_3^{MC}\left(d^{min}\right)$	0.1×10^{-4}	0.8×10^{-4}	0.4×10^{-4}	0	0	0.1×10^{-4}
n_{call}	3.57×10^6	791175	29393	15722	53200	252

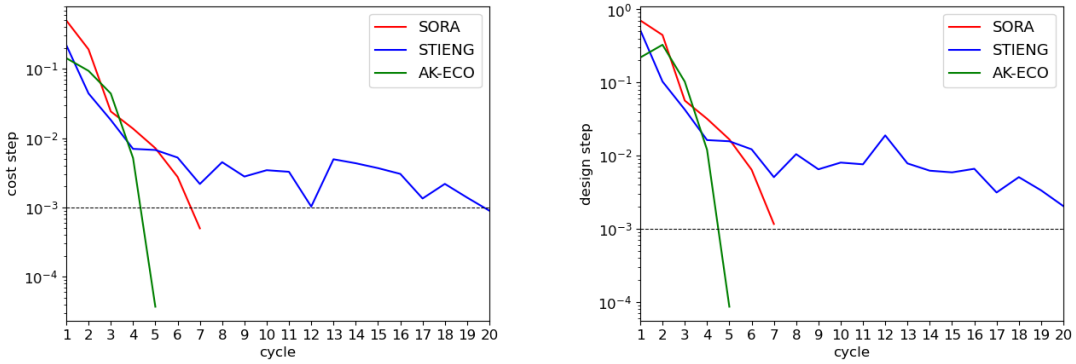
Table 3.4: Results of AK-ECO and the comparison methods

The execution time of every method is discussed in section F.1 of appendix F.

We observe that all the methods converge towards the same design point. However, AK-ECO provides the closest design point to the reference point obtained with MC and requires far fewer calls to the expensive simulator than the comparison methods: only 252 calls are required (50 for the initial DoE and 202 for the local enrichments of the metamodels during the optimization cycles). This is due to the fact that AK-ECO is well adapted to the reformulated problem: each call to the simulator allows to enrich every kriging models and the simulation is performed only at the relevant states s^j. Furthermore, among the comparison methods and AK-ECO only RIA and AK-ECO provide an estimation of the failure probabilities. At the design point obtained with RIA, the first, second and third failure probabilities are estimated with FORM as $1.1 \times 10^{-4}, 1.3 \times 10^{-4}$, 0.8×10^{-4}. The

reformulated failure probabilities estimated with AK-ECO at the design point obtained with this algorithm are 0.8×10^{-4}, 1.0×10^{-4}, 0.1×10^{-4}. Hence, with AK-ECO, we observe a good approximation of the failure probabilities since they are close to p_1^{MC}, p_2^{MC}, and p_3^{MC} obtained with Monte Carlo on the real functions.

For SORA, the Stieng approach and AK-ECO, at the end of cycle k a design point d^k is obtained. The evolution of $\log\left(\left|\overline{cost}\left(d^{k-1}\right)-\overline{cost}\left(d^{k}\right)\right|\right)$ for each method and each cycle k is displayed in Figure 3.3a and the evolution of $\log\left(\left\|\overline{d^k}-\overline{d^{k-1}}\right\|\right)$ in Figure 3.3b where $\overline{cost}\left(d^{k-1}\right)$ and $\overline{d^k}$ refer respectively to the normalized versions of $cost\left(d^k\right)$ and d^k in $[0,1]$.



(a) Evolution of $\log\left(\left|\overline{cost}\left(d^{k-1}\right)-\overline{cost}\left(d^{k}\right)\right|\right)$ (b) Evolution of $\log\left(\left\|\overline{d^k}-\overline{d^{k-1}}\right\|\right)$

Figure 3.3: Evolution of the stopping condition for SORA, Stieng and AK-ECO

We observe that the resolution of the studied problem takes 5 cycles for AK-ECO to converge while 7 and 20 cycles are necessary for SORA and Stieng to meet the stopping condition. In AK-ECO, the closer the design point is to the true infeasible domain boundary, the more enrichment steps are performed. During the first cycle, 22 points are added to the DoE while 45 simulations are performed during the second, third, fourth and last ones.

Moreover, p_1^{MC}, p_2^{MC}, p_3^{MC} have been evaluated with a Monte Carlo of 30000 points at the design point obtained at the end of each cycle of SORA, Stieng and AK-ECO and their evolution is displayed in Figure 3.4. We can see that with Stieng and AK-ECO, the true constraints are satisfied at the end of each cycle while it takes 4 cycles for SORA (actually, with AK-ECO, p_2^{MC} is slightly above the threshold 10^{-4} at the end of the third cycle).

The resolution of the problem has been repeated with AK-ECO from 20 different starting design points selected with a LHS of the design space. Each time, the initial kriging models are calibrated with a new DoE. The results, displayed in section G.1 of appendix G, show that the performance of AK-ECO is not affected by the initial DoE or the initial design point since the algorithm converges towards the good design point each time and with a number of simulations varying from 174 to 416 with a mean number of calls at 229.1.

The oscillator problem has also been solved using kriging models with Gaussian and 5/2-Matérn covariance kernels: each time, AK-ECO converges towards the same design point within a similar number of calls to the simulator. The mutlistart resolutions with these kernels result in an average number of simulations of 190.5 and 191.1, respectively.

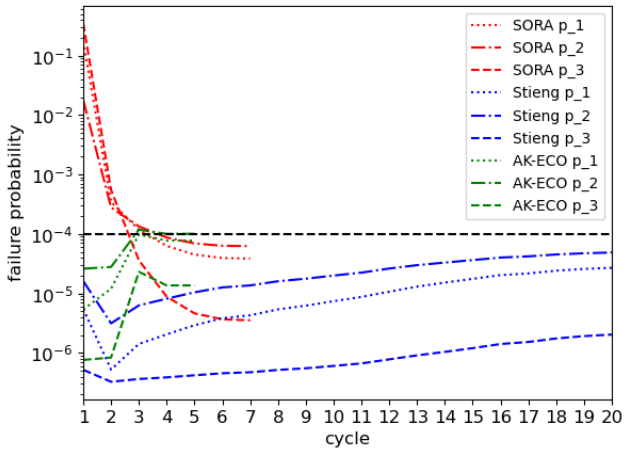


Figure 3.4: Evolution of the Monte Carlo estimation of the failure probabilities at d^k for SORA, Stieng and AK-ECO

3.5 Visualization of the oscillator problem resolution with AK-ECO

We illustrate in Figure 3.5, the first cycle steps of AK-ECO applied to the oscillator problem. The contour lines of the cost function are displayed and the infeasible domain defined by the three constraints and computed with Monte Carlo is the black hatched area. After initialization of the metamodels, the first cycle of AK-ECO begins with step 1.a by evaluating their accuracy at the initial design point d^0 ★. Since they are not accurate enough, enrichment candidates of the augmented space are considered (projections onto the design space of several candidates are represented with transparent blue crosses ×). Enrichment points are then selected among these candidates until the precision criteria are met (step 1.b). The projection of 5 of these points are indicated by blue crosses × in Figure 3.5. Once the accuracy criteria are met, step 2 begins and the optimization problem is solved with COBYLA using the enriched metamodels: therefore, no calls to the simulator is needed during this step. The iterations of this optimization are represented by grey triangles ▼. This resolution provides a new design point d_1 ★ which will be the starting point of the next cycle of AK-ECO.

We now focus on the second constraint since, in the resolution of the oscillator problem, even though every reformulated failure probabilities are close to 10^{-4} at the final design point d^{min}, only the second constraint is active. Using the notation introduced in section 2.3, the second reformulated constraint of the oscillator problem is:

$$\mathbb{E}_{X_d,X_p,X_{r_2}}\left[F_\epsilon\left(\sum_{j=1}^{n_s}\exp\left(a_{Tp^j}^2(X_d,X_p,s^j)^2-\frac{a_{Tp^j}^2(X_d,X_p,s^j)X_{r_2}}{\sqrt{m_{\mathcal{D}'',0}(X_d,X_p,s^j)}}\right)\right)\right]<10^{-4}. \tag{3.31}$$

We denote by M_2, the expensive function of this constraint and $\widetilde{M_2}^k$ its kriging model obtained at the end of the k-th cycle of AK-ECO:

$$M_2\left(x_d,x_p,x_{r_2},s^j\right)=a_{Tp^j}^2(x_d,x_p,s^j)^2-\frac{a_{Tp^j}^2(x_d,x_p,s^j)x_{r_2}}{\sqrt{m_{\mathcal{D}'',0}(x_d,x_p,s^j)}}. \tag{3.32}$$

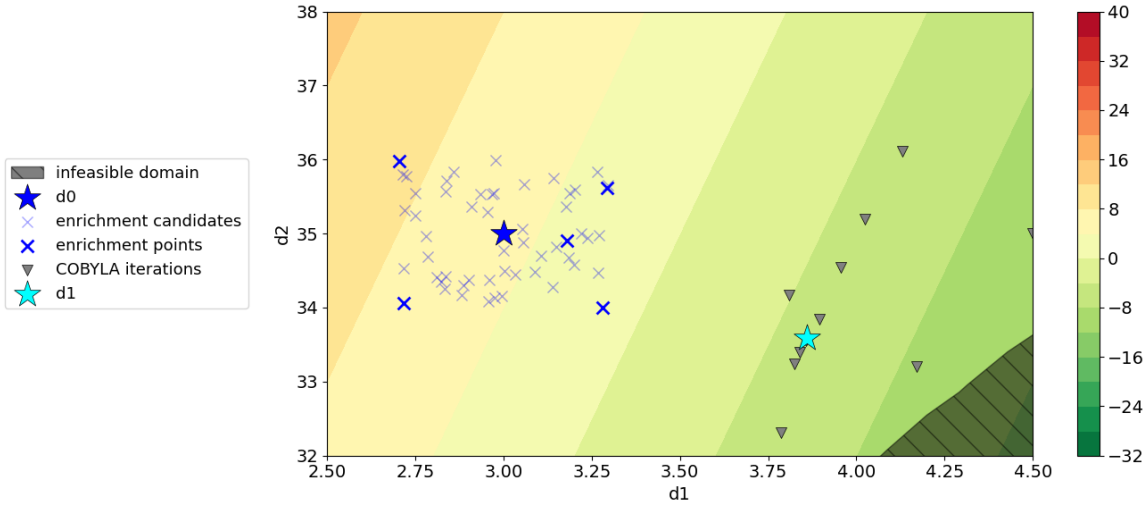


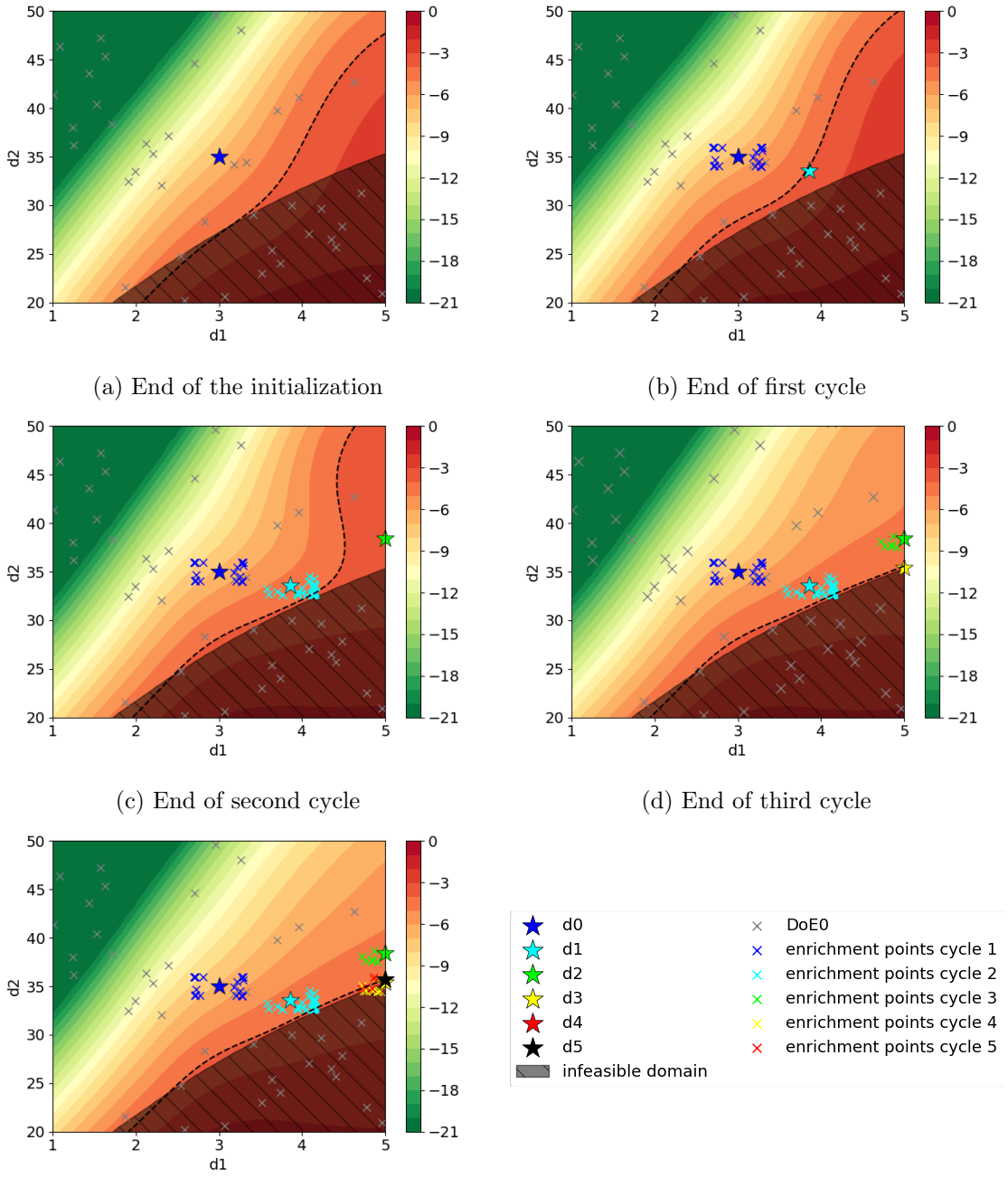
Figure 3.5: Visualization of the first cycle of AK-ECO

In Figure 3.6, we can see the results obtained at the end of the initialization (Figure 3.6a), of the first cycle (Figure 3.6b), the second cycle (Figure 3.6c), the third cycle (Figure 3.6d) and when AK-ECO stops (Figure 3.6e). On each of these figures, we display:

- the real infeasible domain defined by the second constraint (3.31). This domain is obtained by computing the second reformulated failure probability with Monte Carlo and the real function M_2. It is represented by the black hatched area in the figures;
- the contour lines of the reformulated failure probability estimated with the current metamodel $\widetilde{M_2}^k$ (the contour line equal to 10^{-4} is indicated with black dotted lines);
- the projections onto the design space of the initial DoE used to calibrate the initial metamodel $\widetilde{M_2}^0$ (they are represented by grey crosses);
- for each cycle k, the design point from which the local enrichment is performed and the optimization algorithm starts is denoted d^{k-1} and the design point obtained at the end of the optimization algorithm is d^k (stars indicate their positions in the figures);
- the projections onto the design space of the enrichment points (denoted x_{enr} in section 3.2.2) selected during each cycle of AK-ECO are represented by crosses. The color of the enrichment points selected during cycle k corresponds to the color used for d^{k-1}.

Note that at each cycle k, the design point d^k provided at the end of COBYLA is an approximation of the reference design point obtained with MC, and that this approximation becomes better when k increases. Moreover, with each enrichment of the metamodel, the approximation with the kriging model $\widetilde{M_2}^k$ of the real infeasible domain improves locally around d^{k-1}.

In Figure 3.6e, the design point d^3, d^4 and d^5 are so close that they are superposed. The figure corresponding to the end of the fourth cycle of AK-ECO is not displayed since there is no visible difference with Figure 3.6e which represents the end of the fifth cycle.



(a) End of the initialization

(b) End of first cycle

(c) End of second cycle

(d) End of third cycle

(e) End of AK-ECO

(f) Legend

Figure 3.6: Resolution of the oscillator problem with AK-ECO: focus on the second constraint

4 Discussion and perspectives for AK-ECO

We have seen that since AK-ECO enables the simulator to be called only at the relevant states s^j and uses the outputs of the simulations to enrich all the metamodels, AK-ECO performs better than state-of-the-art methods in RBDO for a problem with reformulated extreme and integral based constraints. In this section, we discuss guidelines for the user and possible improvements of the method.

4.1 Discussion on the choice of AK-ECO parameters

During the definition of the augmented space in section 3.2, the parameter α is set to 10^{-6}. Different values of this parameter are considered in Dubourg (2011) and Moustapha (2016). In this thesis, little work has been devoted to the choice of α. Although, with the current implementation of AK-ECO, this parameter only impacts the initial DoE, it would be interesting to analyse the sensitivity of AK-ECO to α.

It is possible for the user to increase the desired accuracy of the estimation of the reformulated failure probabilities with AK-ECO by changing the following quantities:

- the value 2 used in the estimations of $p_{\mathrm{E},+}^{k-1}$, $p_{\mathrm{E},-}^{k-1}$, $p_{\mathrm{I},+}^{k-1}$, and $p_{\mathrm{I},-}^{k-1}$ (see equations (3.13), (3.14), (3.18), and (3.19)). If this value gets larger, the low estimations will get smaller and the high estimations larger. Therefore, more enrichments of the metamodels will be necessary to meet Criterion E and Criterion I;
- the parameters ϵ_p and n_{max} (see equation (3.21) and Criterion E.b and Criterion I.b).

4.2 Application of AK-ECO to other problems

The procedure of AK-ECO described in section 3.2.2 is introduced in the context of reformulated extreme and integral-based constraints but can easily be generalized to any optimization problem with constraints expressed as:

$$\mathbb{E}_{X_d,X_p,X_r}\left[F_0\left(\sum_{i=1}^{n} F_i\left(M\left(X_d,X_p,X_r\right)\right)\right)\right] < p_s, \tag{3.33}$$

if F_0 and F_i $(i = 1,\cdots,n)$ are monotonic functions and M a function whose evaluation requires a call to an expensive simulator.

4.3 Limitations of AK-ECO and perspectives

Although AK-ECO provides much better results than the comparison methods for the studied problem, several points can still be improved and are discussed in this section. We recall that n_{MC} and n_s correspond respectively to the size of the Monte Carlo sample used in AK-ECO and to the number of different states considered.

Limitation due to the Monte Carlo method

When p_s gets smaller than 10^{-4}, using the Monte Carlo method can become cumbersome even used with the metamodels since a greater sample size is required for the estimations of the reformulated failure probabilities to be accurate. Indeed, one evaluation of p_{E}^{k}, p_{I}^{k}, $p_{\mathrm{E},+}^{k-1}$, $p_{\mathrm{E},-}^{k-1}$, $p_{\mathrm{I},+}^{k-1}$, or $p_{\mathrm{I},-}^{k-1}$ requires $n_{MC} \times n_s$ calls to the corresponding kriging model. For a large sample, their estimations are time-consuming as well as the selection of the enrichment points (see equations (3.22) and (3.28)).

To avoid this issue, AK-ECO would benefit to be coupled with a variance reduction technique such as Importance Sampling (Melchers (1989)).

Another issue to consider is the error made with the Monte Carlo estimation of the constraints. We discuss this point below in section 3.4.3.

Limitation when the number of different states is large

When n_s becomes large, the same problem as presented in the previous section will appear and could be solved with a variance reduction technique.

The selection of the enrichment points is particularly sensitive to n_s. Solving problems (3.22) and (3.28) require $n_{MC} \times n_s^2$ evaluations of the corresponding metamodel. It would then be advisable to solve (3.22) and (3.28) with an optimization algorithm instead of evaluating the learning functions for each state and each point of the sampling method.

Limitation due to the kriging method

The effectiveness of AK-ECO may also suffer from the limitations of the kriging technique:

- if the stationary assumption of the kriging method is not verified, the metamodel can perform poorly leading to inaccurate predictions of the expensive function or to an excessive number of simulations;
- another problem is the time of calibration of the metamodel when the size of the DoE gets large. This problem is particularly relevant dealing with high-dimension (when the dimension of the augmented space is large).

Different approaches have been proposed for non-stationary kriging models (Xiong et al. (2007); Toal and Keane (2012); Moustapha and Sudret (2019b); Risser and Turek (2020)) and in high dimension (Bouhlel et al. (2016); Zhao et al. (2020); Zhou and Lu (2020); Zhou and Peng (2020)) and could help AK-ECO to cope with these difficulties.

A multipoint enrichment as proposed by Lelièvre et al. (2018) could also decrease the number of kriging calibrations and thus reduce the post-processing time.

Improving the low and high estimations of the reformulated failure probabilities and of the active learning procedure

During the $(k+1)$-th cycle of AK-ECO, the high and low estimations of the reformulated integral-based failure probability at the design point d are given by:

$$p_{\mathrm{I},+}^k(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_\mathrm{I}} \left(\sum_{j=1}^{n_s} Tp^j \left(\mu_\mathrm{I}^k \left(x_d^i, x_p^i, s^j \right) + 2\sigma_\mathrm{I}^k \left(x_d^i, x_p^i, s^j \right) \right) \right), \tag{3.34}$$

$$p_{\mathrm{I},-}^k(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_\mathrm{I}} \left(\sum_{j=1}^{n_s} Tp^j \left(\mu_\mathrm{I}^k \left(x_d^i, x_p^i, s^j \right) - 2\sigma_\mathrm{I}^k \left(x_d^i, x_p^i, s^j \right) \right) \right), \tag{3.35}$$

where $\widetilde{M_\mathrm{I}}^k \left(x_d^i, x_p^i, s^j \right) \sim \mathcal{N} \left(\mu_\mathrm{I}^k \left(x_d^i, x_p^i, s^j \right), \sigma_\mathrm{I}^k \left(x_d^i, x_p^i, s^j \right)^2 \right)$ is the kriging model of the expensive function M_I at the point $\left(x_d^i, x_p^i, s^j \right)$.

In practice, since the uncertainties add up for each state and each sample in $p_{\mathrm{I},+}^k(d)$ and $p_{\mathrm{I},-}^k(d)$, these estimations can become quiet pessimistic and remain distant even for an accurate kriging model $\widetilde{M_\mathrm{I}}^k$.

In order to avoid an enrichment stopping criterion Criterion I depending only on these quantities, the user can indicate a minimum and a maximum number of enrichments of the metamodel at each cycle of AK-ECO as it is proposed in the section 3.2.2. The constant 2 can also be reduced to this end.

A perspective for AK-ECO is to provide better estimations for high and low estimations of the reformulated failure probability. Indeed, considering the uncertainty of the kriging model, we denote:

$$\widetilde{p}_{\mathrm{I}}^{k}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_{\mathrm{I}}} \left(\sum_{j=1}^{n_s} Tp^j \left(\widetilde{M}_{\mathrm{I}}^{k} \left(x_d^i, x_p^i, s^j \right) \right) \right). \tag{3.36}$$

Using the properties of the kriging model, the random variable

$$\Sigma_{\mathrm{I}}^{k} \left(x_d^i, x_p^i \right) = \sum_{j=1}^{n_s} Tp^j \left(\widetilde{M}_{\mathrm{I}}^{k} \left(x_d^i, x_p^i, s^j \right) \right) \tag{3.37}$$

follows a normal distribution $\mathcal{N}\left(\mu_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)}, \sigma^2_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)}\right)$ with a mean and a standard deviation that can be computed. Therefore, high and low estimations of the reformulated failure probability could be:

$$p_{\mathrm{I},+,\mathrm{bis}}^{k}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_{\mathrm{I}}} \left(\mu_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} + 2\sigma_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} \right), \tag{3.38}$$

$$p_{\mathrm{I},-,\mathrm{bis}}^{k}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_{r_{\mathrm{I}}} \left(\mu_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} - 2\sigma_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} \right). \tag{3.39}$$

The idea could also be applied to the learning function and the new procedure of enrichment could be:

1. select the point (x_d, x_p) from the Monte Carlo sample maximizing:

$$\left[F_{r_{\mathrm{I}}} \left(\mu_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} + 2\sigma_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} \right) - F_{r_{\mathrm{I}}} \left(\mu_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} - 2\sigma_{\Sigma_{\mathrm{I}}^{k}(x_d^i,x_p^i)} \right) \right] \times f_{(X_d,X_p)} \left(x_d^i, x_p^i \right)$$

   with $f_{(X_d,X_p)}$ the probability density function of the random vector (X_d, X_p).

2. then, select the state $s^j \in \{s^1, \cdots, s^{n_s}\}$ reducing the variance of $\Sigma_{\mathrm{I}}^{k}(x_d, x_p)$.

The resulting enrichment point would be $\left(x_d, x_p, s^j\right)$.

This approach could also be applied to the reformulated extreme-based failure probability. However, in this case, we have to be able to deduce from the properties of the kriging model the distribution of the random variable:

$$\Sigma_{\mathrm{E}}^{k} \left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i \right) = \sum_{j=1}^{n_s} \exp \left(\widetilde{M}_{\mathrm{E}}^{k} \left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j \right) \right) \tag{3.40}$$

Even if the distribution of $\Sigma_{\mathrm{E}}^{k} \left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i \right)$ cannot be computed analytically, work has been done to estimate the distribution of the sum of dependent random variables with log-normal distributions (Mehta et al. (2007); Asmussen et al. (2011)).

Concerning the choice of the optimization algorithm

Since AK-ECO only performs local enrichment of the metamodel during each cycle, it has been implemented to be coupled with a local optimization algorithm and thus provides a local optimum. To obtain a global optimum, it is possible to perform multistart optimizations with AK-ECO from several initial design points.

If the user wishes to use global optimization algorithms, it will be necessary to adapt the AK-ECO enrichment strategy.

On the error made during the estimation of the reformulated failure probabilities

The estimations of the reformulated failure probabilities are subject to two sources of error: the error made with Monte Carlo to approximate the expectations and the errors of the kriging models.

Usually in RA, the Monte Carlo method is used to estimate a failure probability p_f and thus, the estimation consists in sampling a Bernoulli random variable whose parameter is p_f. When p_f is small, a large sample is required to provide an accurate estimation of the failure probability. However, with our methodology, the reformulated constraints involve expectations of continuous random variables with small coefficients of variation. Therefore, these expectations can be accurately estimated using Monte Carlo with a reasonable sample size. Indeed, in the oscillator problem, the Monte Carlo estimation of the i-th reformulated failure probability at $d^{\min} = [5.0, 35.73]$ (i.e. the design point obtained with AK-ECO) is denoted $p_i^{MC}\left(d^{\min}\right)$ and we have:

- $p_1^{MC}\left(d^{\min}\right) = 0.8 \times 10^{-4}$, $\sigma_1^{MC} = 19.2 \times 10^{-4}$;
- $p_2^{MC}\left(d^{\min}\right) = 1.0 \times 10^{-4}$, $\sigma_2^{MC} = 15.7 \times 10^{-4}$;
- $p_3^{MC}\left(d^{\min}\right) = 0.1 \times 10^{-4}$, $\sigma_3^{MC} = 1.1 \times 10^{-4}$;

where σ_i^{MC} is the standard deviation of the Monte Carlo sample of 30000 points obtained for the i-th constraint. Denoting p_i the i-th reformulated failure probability, we thus obtain the following confidence intervals:

$$\begin{aligned}
&\mathbb{P}\left(p_1\left(d^{\min}\right) \in \left[0.6 \times 10^{-4}, 1.0 \times 10^{-4}\right]\right) \simeq 95\%,\\
&\mathbb{P}\left(p_2\left(d^{\min}\right) \in \left[0.8 \times 10^{-4}, 1.2 \times 10^{-4}\right]\right) \simeq 95\%,\\
&\mathbb{P}\left(p_3\left(d^{\min}\right) \in \left[0.9 \times 10^{-5}, 1.1 \times 10^{-5}\right]\right) \simeq 95\%.
\end{aligned}$$

Therefore, we consider in this thesis that the sample size is large enough that the Monte Carlo estimation error is negligible. However, we could take into account the uncertainty of the method by simply replacing the Monte Carlo approximation in the constraint with the upper bound of the confidence interval.

Concerning the kriging error, let us recall that, during the k-th cycle of AK-ECO, the estimation of the reformulated extreme-based failure probability is written as follows:

$$p_{\mathrm{E}}^k(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon\left(\sum_{j=1}^{n_s} \exp\left(\mu_{\mathrm{E}}^k\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j\right)\right)\right) \tag{3.41}$$

where μ_{E}^k is the mean of the kriging model $\widetilde{M_{\mathrm{E}}}^k$. Therefore, we could incorporate the kriging uncertainty by considering the following constraint:

$$\mathbb{P}\left(\widetilde{p_{\mathrm{E}}^k}(d) < p_s\right) > p_M \tag{3.42}$$

where p_M is a probability threshold chosen by the user and

$$\widetilde{p_{\mathrm{E}}^k}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon \left(\sum_{j=1}^{n_s} \exp\left(\widetilde{M_{\mathrm{E}}}^k \left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j \right) \right) \right). \tag{3.43}$$

However the distribution of $\widetilde{p_{\mathrm{E}}^k}$ is hard to estimate so it is difficult to compute the probability in equation (3.42). In AK-ECO, a simpler approach is adopted: we consider that the errors due to the kriging models are sufficiently small if the kriging models satisfy the accuracy criteria introduced in section 3.2.2.

Problem encountered with the scikit-learn Python package

A problem occurs using the kriging model implemented in the scikit-learn Python package (Pedregosa et al. (2011)). For several DoEs encountered during the enrichment procedure, the standard deviations of the kriging model were poorly computed. In particular, the standard deviations of the metamodel evaluated at points belonging to the DoE were large (although they should have been zero). This led to problems during the learning procedure. Upon further investigation, the problem arises from the approximation error in estimating the inversion of a matrix when calculating the posterior variance of the kriging model.

Another problem occurred during the calibration of the kriging models with the scikit-learn Python package: during some enrichments, the hyperparameters changed abruptly from their values obtained at the end of the previous enrichment, which led to poor quality kriging models. To solve this problem, we have added a prior information on the hyperparameters as described in Santner et al. (2018). More precisely, considering a hyperparameter θ, during the k-th cycle of enrichment, a uniform distribution $\mathcal{U}[\theta_{k-1} - 1, \theta_{k+1} + 1]$ is assumed for the prior where θ_{k-1} is the value of θ obtained at the end of the $k-1$-th cycle.

Both of these problems were not encountered with the OpenTURNS Python package.

Limit due to the quality of the initial DoE

If the initial DoE used during the initialization of AK-ECO to calibrate the initial metamodels is not good enough, the minimum obtained at the end of this first cycle of AK-ECO may be far from the true optimum due to a bad estimation of the reformulated failure probabilities. This can increase the number of optimization cycles of AK-ECO and of calls to the simulator.

To adress this issue, we propose in the next chapter a procedure of global enrichment of the metamodels ensuring the accuracy of the kriging models used during the first cycle of AK-ECO.

Chapter 4

A global enrichment procedure and a RBDO-oriented sensitivity analysis

Contents

Let us recall that this thesis is motivated by the resolution of an optimization problem with extreme and integral based constraints as introduced in section 2.1. After the reformulation of the constraints described in section 2.3, we obtain the following problem:

$$\begin{aligned} &\min_{d\in\Omega_d} cost(d) \quad \text{such that} \\ &\mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[F_\epsilon\left(\sum_{j=1}^{n_s}\exp\left(M_\mathrm{E}\left(X_d,X_p,X_{r_\mathrm{E}},s^j\right)\right)\right)\right] &< p_s \\ &\mathbb{E}_{X_d,X_p}\left[F_{r_\mathrm{I}}\left(\sum_{j=1}^{n_s}Tp^j\left(M_\mathrm{I}\left(X_d,X_p,s^j\right)\right)\right)\right] &< p_s. \end{aligned} \tag{4.1}$$

We have introduced in chapter 3 a new method to solve problem (4.1) requiring few calls to the expensive functions M_E and M_I. A DoE in the augmented space Ω_E^{aug} defined in section 3.2 is created and kriging models of M_E and M_I are calibrated during the initialization step. Optimization cycles combining local enrichments of the metamodels and a resolution of problem (4.1) using the Monte Carlo method with the kriging models are then carried out.

In this chapter, we introduce a so-called global enrichment procedure of the metamodels before starting AK-ECO as well as a new sensitivity analysis evaluating the influence of each input random variable on the constraints.

1 Global Enrichment

1.1 Motivation for the global enrichment of the metamodels

During the first optimization cycle of AK-ECO, the initial krigings are built from a DoE of the augmented space Ω_E^{aug}. As mentioned in section 3.4, if this DoE is of poor quality, the minimum obtained at the end of this first cycle of AK-ECO may be far from the true optimum due to a bad estimation of the reformulated failure probabilities using those metamodels. This can lead to a large number of optimization cycles before convergence of AK-ECO and thus slow down the resolution of the problem. Therefore, we propose to improve the metamodels before starting AK-ECO in order to ensure their quality during the first optimization cycle. In opposition to the local enrichments in AK-ECO aiming at improving the constraints estimations around the current design point, this enrichment procedure will be called **global** since its purpose is to improve the estimation of the feasible domain boundary over the whole design space. Once this global enrichment procedure is completed, the obtained kriging models will be considered as the initial metamodels of AK-ECO.

Similar global enrichment strategies, performed before solving an optimization problem, have been proposed in Moustapha et al. (2015), Moustapha and Sudret (2019a), and Shang et al. (2021). These methods aim to select points near the limit-state surface of the performance function to improve the metamodels accuracy. However, due to the particular constraints of problem (4.1) involving expectations instead of failure probabilities and our metamodel strategy, these enrichment procedures are not adapted to problem (4.1). We therefore present in section 4.1 a new procedure of global enrichment (GE) of the metamodels adapted to the kriging strategy introduced with AK-ECO.

1.2 Global enrichment procedure

A new enrichment procedure of the metamodels of M_E and M_I called global enrichment (GE) is presented in this section. The purpose is to improve the quality of the metamod-

els in order to estimate with accuracy the boundary of the feasible domain defined by each reformulated failure probability over all the design space. The GE is composed of cycles of enrichments. During each cycle, one enrichment of each inaccurate metamodel is performed. The steps of the procedure are described below:

Initialization: This step is similar to the initialization of AK-ECO described in section 3.2. An initial DoE, denoted DoE^0, is computed and used to calibrate the initial metamodels $\widetilde{M_{\mathrm{E}}}^0$ and $\widetilde{M_{\mathrm{I}}}^0$ of the functions M_{E} and M_{I}. At the end of the initialization, the first cycle of enrichment ($k = 1$) begins.

Enrichment cycle: We denote DoE^{k-1}, $\widetilde{M_{\mathrm{E}}}^{k-1}$, and $\widetilde{M_{\mathrm{I}}}^{k-1}$, the DoE and kriging models recovered from the initialization if $k = 1$ or from the previous cycle if $k > 1$. For each inaccurate kriging model, an enrichment is performed following steps 1 to 3:

Step 1: Calibration of the metamodel from DoE^{k-1};

Step 2: A design point $d_{enr} \in \Omega_d$ is selected with the procedure described in section 4.1.3. This point is chosen if the metamodel cannot predict with enough certainty whether or not it belongs to the feasible domain;

Step 3: A local enrichment of the metamodel is performed at d_{enr} (we detail this step in section 4.1.4). A point x_{enr} is selected from the augmented space improving the accuracy of the reformulated failure probability estimation at d_{enr}. The simulator is evaluated at x_{enr} and DoE^{k-1} is updated.

At the end of the enrichment cycle, every inaccurate metamodel has been enriched once and all of the metamodels are recalibrated from the last DoE. The resulting DoE and kriging models are denoted respectively DoE^k, $\widetilde{M_{\mathrm{E}}}^k$, and $\widetilde{M_{\mathrm{I}}}^k$. For each enriched metamodel, an accuracy criterion introduced in section 4.1.5 is then evaluated. The metamodels considered accurate enough will not be enriched during the following cycles. The global enrichment ends when each metamodel satisfies the accuracy criterion or if $n_{\mathrm{tot}} > n_{\mathrm{tot,max}}$ where n_{tot} is the total number of every metamodel enrichments performed since the beginning of the GE and $n_{\mathrm{tot,max}}$ is the maximum of enrichments allowed by the user. If the condition is not met, a new enrichment cycle begins.

The full procedure of global enrichment is summarized in Figure 4.1.

1.3 Selection of the design point to perform the local enrichment

At the beginning of the k-th cycle of enrichment, the purpose of the metamodel $\widetilde{M_E}^{k-1}$ is to be accurate enough in order to obtain a good estimation of the feasible domain defined by the reformulated extreme-based constraint and computed with Monte Carlo and the metamodel. Therefore, a local enrichment of the metamodel needs to be performed at a design point d, if p_{E}^{k-1} wrongly predicts that d belongs to the feasible domain with:

$$p_{\mathrm{E}}^{k-1}(d) = \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} F_\epsilon \left(\sum_{j=1}^{n_s} \exp\left(\mu_{\mathrm{E}}^{k-1}\left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j \right) \right) \right) \tag{4.2}$$

and where $\left\{ \left(x_d^i, x_p^i, x_{r_{\mathrm{E}}}^i \right), i = 1, \cdots, n_{MC} \right\} = \Omega_{MC}(d)$ is the Monte Carlo sample of the random vector $(X_d, X_p, X_{r_{\mathrm{E}}})$.

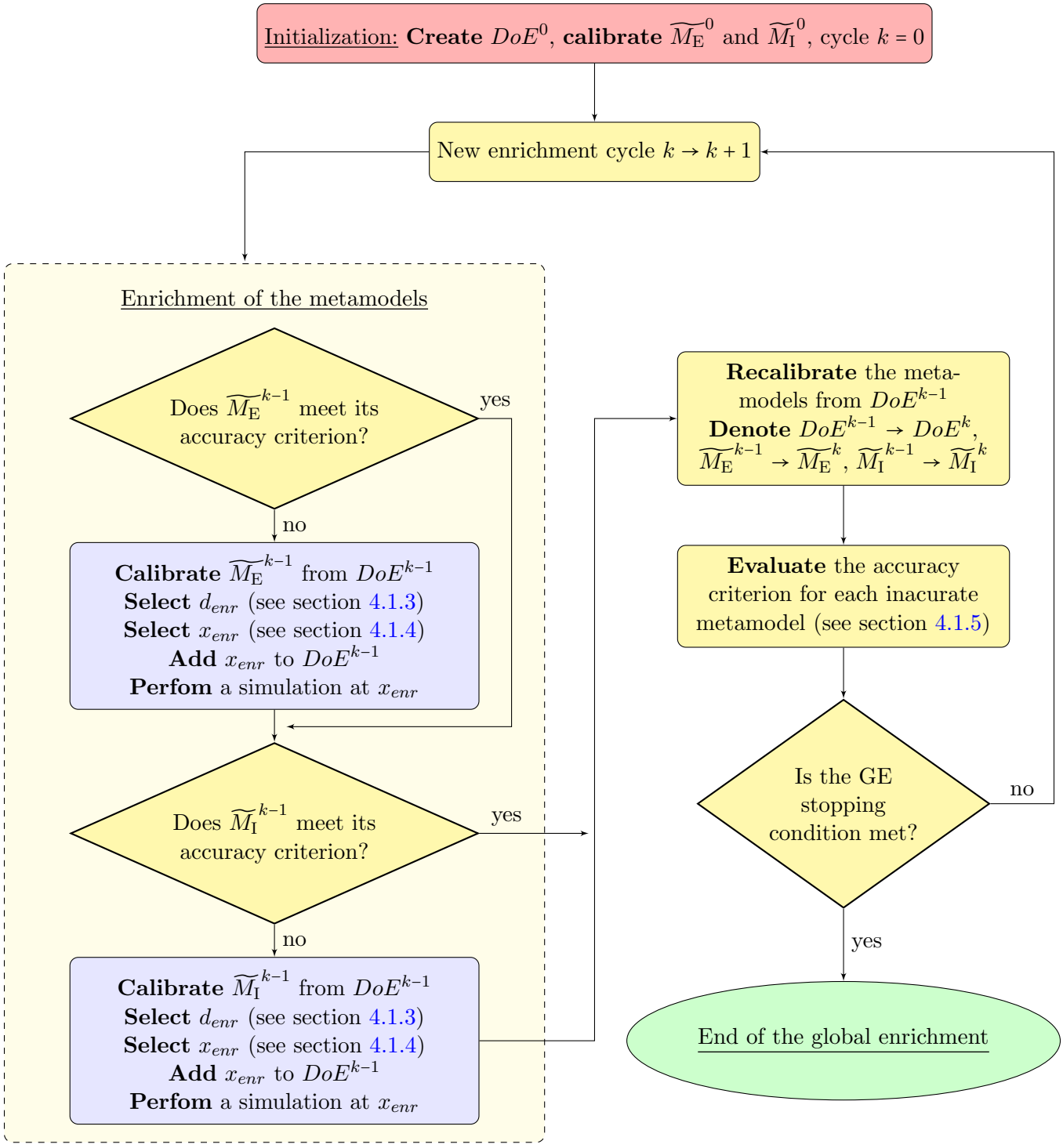


Figure 4.1: Flowchart of the global enrichment procedure

To select a design point, we use the high and low estimations of the reformulated failure probability introduced in section 3.2.2 and denoted respectively $p_{\mathrm{E},+}^{k-1}$ and $p_{\mathrm{E},-}^{k-1}$. The design point d_{enr} is selected by minimizing a criterion on the design space as follows:

$$d_{enr} = \underset{d \in \Omega_d}{\arg\min}\, A_{\mathrm{E}}(d) \tag{4.3}$$

with

$$A_{\mathrm{E}}(d) = \frac{\left|p_{\mathrm{E}}^{k-1}(d) - p_s\right|}{p_{\mathrm{E},+}^{k-1}(d) - p_{\mathrm{E},-}^{k-1}(d)}. \tag{4.4}$$

We notice that the criterion A_{E} is used to evaluate the accuracy of the metamodel in AK-ECO (see section 3.2.2). It is used in GE to select the design point d_{enr} since A_{E} decreases when $p_{\mathrm{E}}^{k-1}(d)$ is close to p_s and when the uncertainty measured by the distance between $p_{\mathrm{E},+}^{k-1}(d)$ and $p_{\mathrm{E},-}^{k-1}(d)$ is high. Hence, we consider that a local enrichment of the metamodel is needed at d if the distance between $p_{\mathrm{E}}^{k-1}(d)$ and p_s is small and if the uncertainty measured by $p_{\mathrm{E},+}^{k-1}(d) - p_{\mathrm{E},-}^{k-1}(d)$ is large.

In a same way, to select the design point where the local enrichment of $\widetilde{M_{\mathrm{I}}}^{k-1}$ is needed, the following problem is solved:

$$d_{enr} = \underset{d \in \Omega_d}{\arg\min}\, A_{\mathrm{I}}(d) \tag{4.5}$$

with

$$A_{\mathrm{I}}(d) = \frac{\left|p_{\mathrm{I}}^{k-1}(d) - p_s\right|}{p_{\mathrm{I},+}^{k-1}(d) - p_{\mathrm{I},-}^{k-1}(d)} \tag{4.6}$$

and $p_{\mathrm{I}}^{k-1}(d)$, $p_{\mathrm{I},-}^{k-1}(d)$, and $p_{\mathrm{I},+}^{k-1}(d)$ are defined in section 3.2.2.

To solve problems (4.3) and (4.5), new DoEs of Ω_d are created during each cycle and are denoted $DoE_{\mathrm{E,enr}}^{k}$ and $DoE_{\mathrm{I,enr}}^{k}$. The solution of (4.3) (resp. (4.5)) is the point of $DoE_{\mathrm{E,enr}}^{k}$ (resp. $DoE_{\mathrm{I,enr}}^{k}$) minimizing A_{E} (resp. A_{I}).

1.4 Selection of the enrichment point in the augmented space

Once the point in the design space d_{enr} is selected, we need to determine the point x_{enr} in the augmented space for which the metamodel needs to be specified in order to improve the estimation of the reformulated failure probability at d_{enr}. To do so, the learning functions of AK-ECO, L_{E} and L_{I} defined respectively in equations (3.23) and (3.29) of section 3.2.2 are used.

Thus, during the k-th cycle, $\widetilde{M_{\mathrm{E}}}^{k-1}$ is enriched by selecting x_{enr} such that:

$$x_{enr} = \underset{\substack{\left(x_{d_{enr}}^i, x_p^i, x_{r_{\mathrm{E}}}^i\right) \in\, \Omega_{MC}\,(d_{enr}) \\ s^j \in \{s^1, \cdots, s^{n_s}\}}}{\operatorname{argmax}} L_{\mathrm{E}}(x_{d_{enr}}^i, x_p^i, x_{r_{\mathrm{E}}}^i, s^j) \tag{4.7}$$

where $\Omega_{MC}\,(d_{enr})$ is the Monte Carlo sample used in the estimation of $p_E^{k-1}(d_{enr})$.

For the enrichment of $\widetilde{M_{\mathrm{I}}}^{k-1}$, once d_{enr} is identified, x_{enr} is selected by solving the following problem:

$$x_{enr} = \underset{\substack{\left(x^i_{d_{enr}}, x^i_p, x^i_{r_{\mathrm{E}}}\right) \in \Omega_{MC}\left(d_{enr}\right) \\ s^j \in \{s^1, \cdots, s^{n_s}\}}}{\operatorname{argmax}} L_{\mathrm{I}}(x^i_{d_{enr}}, x^i_p, s^j). \tag{4.8}$$

Then, the simulator is evaluated at point x_{enr} and the results are added to the shared DoE DoE^{k-1}.

1.5 Accuracy criteria

When the GE starts, every metamodel is considered inaccurate. At the end of each enrichment cycle, every inaccurate metamodel are recalibrated with the shared DoE. Then, for each of them, a criterion assesses the stability of the feasible domain of the corresponding constraint estimated from the current metamodel.

At the end of the k-th cycle, to evaluate the stability of the feasible domain estimated with $\widetilde{M_{\mathrm{E}}}^k$, the following quantity is computed:

$$n^k_{\mathrm{E,stop}} = \mathrm{Card}\left\{d \in DoE_{\mathrm{stop}} \text{ s.t. } \left(p^k_{\mathrm{E}}(d) - p_s\right) \times \left(p^{k-1}_E(d) - p_s\right) < 0\right\} \tag{4.9}$$

where:

- DoE_{stop} is a space-filling DoE of Ω_d created at the beginning of the GE;
- $p^k_{\mathrm{E}}(d)$ is computed with the metamodel obtained at the end of the k-th cycle;
- $p^{k-1}_{\mathrm{E}}(d)$ is computed with the metamodel obtained at the end of the $(k-1)$-th cycle.

Hence, $n^k_{\mathrm{E,stop}}$ corresponds to the number of points in DoE_{stop} that belong to the feasible domain estimated during the $(k-1)$-th cycle and to the infeasible domain at the next one (or vice versa).

Similarly, we measure the stability of the feasible domain estimated with $\widetilde{M_{\mathrm{I}}}^k$ thanks to the following quantity:

$$n^k_{\mathrm{I,stop}} = \mathrm{Card}\left\{d \in DoE_{\mathrm{stop}} \text{ s.t } \left(p^k_{\mathrm{I}}(d) - p_s\right) \times \left(p^{k-1}_I(d) - p_s\right) < 0\right\}. \tag{4.10}$$

The feasible domain estimated with the metamodel of M_{E} (resp. M_{I}) is considered stable when $n^k_{\mathrm{E,stop}} \leq n_{\mathrm{max,stop}}$ (resp. $n^k_{\mathrm{I}} \leq n_{\mathrm{max,stop}}$) over nb_{stop} consecutive enrichment cycles. The choice of these parameters is discussed in section 4.4.1.

These criteria are inspired by the criterion proposed in Moustapha and Sudret (2019a) but are adapted to our kriging strategy.

Once the feasible domain estimated from a metamodel is considered stable, this metamodel will not be enriched during the following enrichment cycles and its accuracy criterion will not be evaluated again.

2 RBDO-oriented GSA

2.1 Motivation for a new sensitivity analysis approach

A brief state of the art in sensitivity analysis

When the outputs of a model depend on multiple random input parameters, the concern arises of determining the respective contribution of each parameter to the output uncertainties. **Sensitivity Analysis** (SA) (Saltelli et al. (2008)) methods have been proposed to address this issue and provide engineers a better understanding of the model. The purpose of SA is often to identify the most influential parameters: a more thorough study of these parameters leads to a better accuracy of the model. One may also wish to identify the non-influential inputs. A model dimension reduction would then be possible by fixing these parameters at nominal values without affecting the model outputs. We refer to Chabridon (2018) for a more extensive list of objectives pursued in SA.

Most of the methods in SA provide an index for each of the input variables revealing a hierarchy among them. They are usually divided into two categories:

- local SA: based on the gradient of the model with respect to its parameters, this approach studies the influence of small input perturbations on the outputs at a nominal value;
- global SA (GSA): these methods consider the effects of the input parameters on the model when they vary in their entire domain. The screening techniques (Morris (1991)) aggregate the results of local analyses performed throughout the definition domain of the input parameters. The variance-based methods rely on the decomposition of the model output variance as a sum of terms reflecting the contribution of each input parameter. We present in more detail variance-based indices in section 4.2. Finally, for each input, dissimilarity-based methods (Borgonovo (2007); Da Veiga (2015)) compare the distribution of the output considering every input uncertainties and the conditional distribution of the output given that one of the parameters is fixed at a given value. A review of GSA methods is provided in Iooss and Lemaître (2015) and Spagnol (2020).

To reduce the computational burden of classical SA method, several approaches propose to take advantages of metamodels. A screening GSA is performed in Shang et al. (2020) coupled with a RBF model. In Sudret (2008), Bhattacharyya (2020), and Ehre et al. (2020) the variance-based indices are efficiently computed using PCE whereas an adaptive kriging strategy is preferred in Guo and Dias (2020).

The SA is said **goal-oriented** when the quantity of interest is not directly the model output (Fort et al. (2016)).

Hence, **reliability-oriented** SA methods have been introduced to account for the influence of the input variables on a failure probability. Importance factors interpreted as the contribution of each input to the variance of the performance function are obtained as a by-product of FORM (Madsen et al. (2006)). Reliability-oriented indices have also been proposed by performing a variance-based method on the indicator of the performance function (Cui et al. (2010); Luyi et al. (2012)). These indices are computed efficiently with sampling techniques in Wei et al. (2012), Zhang et al. (2020c), and Liu and Li (2020) or with adaptive metamodels strategies (He et al. (2020); Lei et al. (2021)). Several

reliability-oriented indices are reviewed in Kala (2020b).

An **optimization-oriented** SA is introduced in Spagnol (2020). The sublevel set of interest is defined as a subset of the design space including the design points where the function to minimize is low and where the constraints are satisfied. Then, the influence of the input on specific outputs are measured with kernel-based indices (Gretton et al. (2005)). These outputs take into account the value of the cost function and the sublevel set of interest. This analysis makes the optimization problem easier to resolve by reducing the dimension of the design space which also leads to a degradation of the optimum.

Finally, to enhance the efficiency of RBDO methods, several approaches have been proposed to evaluate the local sensitivity of the probabilistic constraints with respect to the design variables during the resolution of the problem. In RIA (Hou (2004)), the importance factors mentioned above are used for this purpose and similar sensitivity measures are analytically provided by Lee et al. (2010a) for PMA. In Lee et al. (2010b) and Lee et al. (2011), score functions are used to express the sensitivity of the failure probabilities to the design variables as expectations that can be estimated with Monte Carlo. The performance functions are replaced by metamodels in the latter papers.

Limitation with current approaches

Many methods have been proposed in SA and applied to various problems. However, their application to RBDO problems is limited to local SA performed during the optimization to evaluate the sensitivity of the constraints to the design variables.

In this thesis, we are interested in knowing what are the influence of each source of uncertainty involved in the constraints on their satisfaction. This SA encounters a new difficulty since the sensitivity of the constraints in relation to each input depends on the design point considered.

In this chapter, we tackle this problem and propose a new RBDO-oriented GSA approach in section 4.2.

2.2 RBDO-oriented GSA procedure

We propose in this section a RBDO-oriented GSA approach aiming at identifying, for each constraint, the contributions of the input uncertainties on the constraint satisfaction.

We face a difficulty since the contribution of the model inputs may differ depending on the design point considered. A random variable may have a strong impact on the constraint at a specific design point while it is of small influence at another design point. Since the indices proposed in the SA literature are performed at fixed design, we propose to carry out multiple GSAs for different points and to gather this information into new cumulative indices.

We present our GSA approach in a general framework for a RBDO problem with constraints expressed as follows:

$$\mathbb{E}_U\left[H_d\left(U\right)\right] < p_s \tag{4.11}$$

where $U = (U_1, \cdots, U_{n_u})$ is a random vector grouping all the random inputs whose influence on the constraint is to be evaluated. The components of U are assumed to be independent

with known probability density functions. Moreover, for all design point $d \in \Omega_d$, $H_d : \mathbb{R}^{n_u} \to \mathbb{R}$ is a square-integrable function. We denote:

$$p_f(d) = \mathbb{E}_U\left[H_d(U)\right] \tag{4.12}$$

and

$$Y_d = H_d(U). \tag{4.13}$$

The analysis presented in this section must be performed for each constraint of the considered RBDO problem independently.

2.3 GSA at a fixed design point with the Sobol indices

We detail here one of the most used GSA methods: the functional decomposition of the variance and the Sobol indices. We will see later how they can be adapted to fit our need.

Variance decomposition and Sobol indices in SA

It is shown that if a function H defined on $[0,1]^{n_u}$ is square-integrable, it can be decomposed into a sum of increasing dimension functions (Hoeffding (1992)). Let us consider a random vector $U = (U_1, \cdots, U_{n_u})$ with independent components and known probability density function, it follows from the decomposition of H that the variance of $Y = H(U)$ can be expressed as (Efron and Stein (1981); Sudret (2007)):

$$\mathrm{Var}(Y) = \sum_{s=1}^{n_u} \sum_{i_1<\cdots<i_s}^{n_u} D_{i_1\cdots i_s} = \sum_{i=1}^{n_u} D_i + \sum_{i<j}^{n_u} D_{ij} + \cdots + D_{1\cdots n_u} \tag{4.14}$$

with

$$D_i = \mathrm{Var}\left(\mathbb{E}\left[Y|U_i\right]\right), \tag{4.15}$$

$$D_{ij} = \mathrm{Var}\left(\mathbb{E}\left[Y|U_i, U_j\right]\right) - D_i - D_j, \tag{4.16}$$

and with the recursive definition:

$$D_{i_1\cdots i_s} = \mathrm{Var}\left(\mathbb{E}\left[Y|U_{i_1}, \cdots, U_{i_s}\right]\right) - \sum_{j=1}^{s-1} \sum_{\substack{k_1 < \cdots < k_j \\ \{k_1,\cdots,k_j\} \subset \{i_1,\cdots,i_s\}}} D_{k_1\cdots k_j}. \tag{4.17}$$

The decomposition of equation (4.14) is called the functional decomposition of variance or ANOVA (for ANalysis Of VAriance).

The **Sobol indices** introduced in Sobol' (2001) and denoted $S_{i_1\cdots i_s}$ are defined by normalizing the terms $D_{i_1\cdots i_s}$ with $\mathrm{Var}(Y)$:

$$S_{i_1\cdots i_s} = \frac{D_{i_1\cdots i_s}}{\mathrm{Var}(Y)} \tag{4.18}$$

and it follows that $\sum_{s=1}^{n_u} \sum_{i_1<\cdots<i_s}^{n_u} S_{i_1\cdots i_s} = 1$.

In particular the **first-order** Sobol index S_i of U_i is:

$$S_i = \frac{D_i}{\mathrm{Var}(Y)} = \frac{\mathrm{Var}\left(\mathbb{E}\left[Y|U_i\right]\right)}{\mathrm{Var}(Y)}. \tag{4.19}$$

Thus, the index $S_{i_1\ldots i_s}$ represents the proportion of the variance of Y due to the interaction between the variables $U_{i_1}, \cdots, U_{i_s}$. The first-order index considers the influence of U_i on

$\mathrm{Var}(Y)$ without accounting for its interactions with the other input parameters.

The **total-order** Sobol index S_i^T has been proposed in Homma and Saltelli (1996) to accumulate the effects of all the combinations of input parameters involving U_i:

$$S_i^T = \sum_{s=1}^{n_u} \sum_{\substack{i_1 < \cdots < i_s \\ i \in \{i_1, \cdots, i_s\}}} S_{i_1 \cdots i_s}. \tag{4.20}$$

Hence, using the total variance law, we obtain:

$$S_i^T = 1 - \frac{\mathrm{Var}\left(\mathbb{E}\left[Y|U_{\sim i}\right]\right)}{\mathrm{Var}(Y)} = \frac{\mathbb{E}\left[\mathrm{Var}\left(Y|U_{\sim i}\right)\right]}{\mathrm{Var}(Y)} \tag{4.21}$$

with $U_{\sim i} = (U_1, \cdots, U_{i-1}, U_{i+1}, \cdots, U_{n_u})$.

The Sobol indices enable to evaluate between 0 and 1, the contribution of each input parameter to the variance of the output. The closer the index of a parameter is to 1, the more influential it is.

These indices can be computed with Monte Carlo (Saltelli (2002)) or with more efficient methods (Saltelli et al. (2008); Iooss and Lemaître (2015)) such as the Sobol pick-freeze approach (Sobol' (2001); Janon et al. (2014)) or the Fourier Amplitude Sensitivity Testing (FAST) method (Cukier et al. (1978); Saltelli et al. (1999)) which relies on a Fourier expansion of the function H_d.

Application of Sobol indices for expectation constraints

We consider from now on the notations introduced in the beginning of section 4.2.

At a fixed design point $d \in \Omega_d$, we are interested in understanding the impact of the uncertainty of every random variable U_i $(i = 1, \cdots, n_u)$ on $p_f(d)$. Sobol indices naturally adapt to this need by considering the Sobol indices of Y_d defined in equation (4.13).

For each random variable U_i the first-order Sobol index, denoted $S_i(d)$ and the total-order index $S_i^T(d)$ are defined as:

$$S_i(d) = \frac{\mathrm{Var}_{U_i}\left(\mathbb{E}_{U_{\sim i}}[Y_d|U_i]\right)}{\mathrm{Var}(Y_d)} \tag{4.22}$$

and

$$S_i^T(d) = 1 - \frac{\mathrm{Var}_{U_{\sim i}}\left(\mathbb{E}_{U_i}\left[Y_d|U_{\sim i}\right]\right)}{\mathrm{Var}(Y_d)}. \tag{4.23}$$

As explained in section 4.2.3, these indices measure the proportion of the variance of Y_d due to each input parameters at a fixed design point.

These indices are adapted to the expectation constraint defined in equation (4.11) since they also reflect the influence of each input uncertainties on $p_f(d)$. The larger S_i and S_i^T are, the more influential the i-th parameter is on $p_f(d)$.

Remark 4.1. This use of Sobol indices has been inspired from the reliability-oriented SA proposed in Cui et al. (2010) and Luyi et al. (2012). Indeed, we find the same indices when $H_d(U) = \mathbb{1}_{g(d,U)\leq 0}$ with $g(d, U)$ the performance function of the considered failure probability.

2.4 Definition of the critical domain and selection of the design points where to perform GSA

The Sobol indices being calculated for fixed design points, we have to select the designs for which we are interested to know the results of the GSA. Since it would be too expensive to perform a GSA for every points of the design space, we restrain the analysis to a subdomain of Ω_d. The influence of the input parameters on $p_f(d)$ are particularly important if $p_f(d)$ is close to p_s. Indeed, if d belongs to this region, variations of $p_f(d)$ are directly linked to the satisfaction of the constraint at d. Hence, a parameter having a strong impact on $p_f(d)$ can determine whether $p_f(d) > p_s$ or $p_f(d) < p_s$.

We thus define the **critical domain** $\Omega_{\text{critical}} \subset \Omega_d$ such that:

$$\Omega_{\text{critical}} = \{d \in \Omega_d,\, p_f(d) \in [p_s^-, p_s^+]\} \tag{4.24}$$

where p_s^-, p_s^+ are parameters chosen by the user such that $p_s^- < p_s < p_s^+$ (we discuss this choice in section 4.4.2).

Our RBDO-GSA approach relies on the two following assumptions:

Assumption 1. If $d \in \Omega_{\text{critical}}$ and U_i is an influential parameter on $p_f(d)$ then U_i is an influencial parameter on the satisfaction of the constraint at d.

Assumption 2. If $d \notin \Omega_{\text{critical}}$, the influence of the different sources of uncertainty on the satisfaction of the constraint at d are negligible.

Hence, whether the objective of the RBDO-GSA is to identify the influential sources of uncertainties on the definition of the feasible domain or to search for the non-influential ones in order to reduce the dimension of the input space, it follows from assumptions 1 and 2 that only the GSA performed in the critical domain are relevant.

We therefore propose to select several points in the critical domain with the following procedure:

1. we consider an empty set Ω_{SA} and $k = 1$;
2. a space-filling DoE of size n_k of the design space is created;
3. for each point d of the DoE, we compute $p_f(d)$. If $d \in \Omega_{\text{critical}}$, d is added to Ω_{SA}. A metamodel strategy can be used to estimate $p_f(d)$ as we will see in section 4.2.6;
4. if $card\,(\Omega_{\text{SA}}) \geq n_{\text{SA,min}}$ or if $k > k_{\max}$, the procedure stops, otherwise, $k = k + 1$ and go back to step 2.

The parameters $n_1, \cdots, n_{k_{\max}}$, and $n_{\text{SA,min}}$ are chosen by the user. The parameter $n_{\text{SA,min}}$ must be determined such that, at the end of the procedure, enough points belong to Ω_{SA} to correctly span Ω_{critical}. If $card\,(\Omega_{\text{SA}})$ is too large, it can result in a high concentration of points of Ω_{SA} in the same area which would increase the computation time without providing much information. This is why the size $n_1, \cdots, n_{k_{\max}}$ is sequentially increased and these parameters must be chosen to avoid such a phenomenon.

A GSA is then performed for each point of Ω_{SA}: for each $d \in \Omega_{\text{SA}}$ and each random variable U_i $(i = 1, \cdots, n_u)$, the first-order and total-order indices defined in section 4.2.3 are computed.

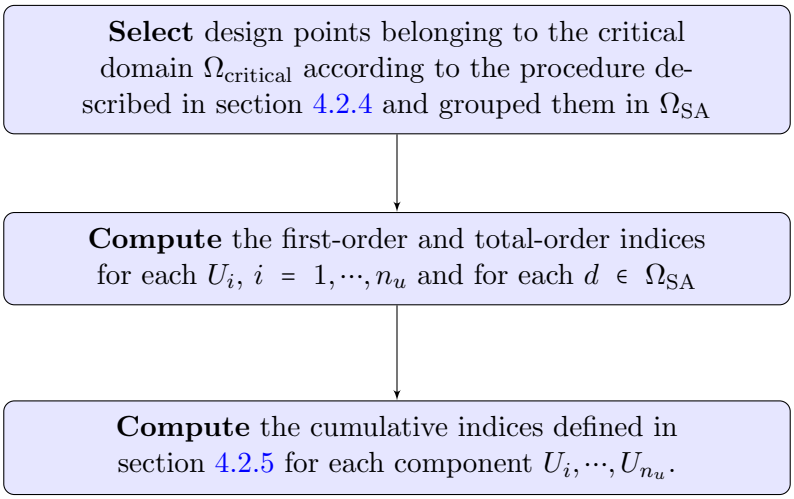


Figure 4.2: Flowchart of the RBDO-oriented GSA

2.5 Introduction of the cumulative indices

The design points have now been selected, grouped in Ω_{SA} and the Sobol indices of the considered uncertainties have been evaluated at each of these design points. To aggregate these results, we introduced for each random variable U_i ($i = 1, \cdots, n_u$), the **first-order cumulative Sobol index** $\overline{S_i}$ and the **total-order cumulative Sobol index** $\overline{S_i^T}$ defined as follows:

$$\overline{S_i} = \frac{1}{n_{\text{SA}}} \sum_{d \in \Omega_{\text{SA}}} S_i(d), \tag{4.25}$$

$$\overline{S_i^T} = \frac{1}{n_{\text{SA}}} \sum_{d \in \Omega_{\text{SA}}} S_i^T(d), \tag{4.26}$$

with $n_{\text{SA}} = Card\,(\Omega_{\text{SA}})$.

Hence, $\overline{S_i}$ (resp. $\overline{S_i^T}$) represents the average of the first-order (resp. total-order) indices of the i-th input parameter over Ω_{SA}.

We propose also to consider the maximum of these indices over Ω_{SA} and defined two other cumulative indices:

$$S_{i,\max} = \max_{d \in \Omega_{\text{SA}}} S_i(d) \tag{4.27}$$

and

$$S_{i,\max}^T = \max_{d \in \Omega_{\text{SA}}} S_i^T(d). \tag{4.28}$$

These four cumulative indices allow to have a better understanding of the influence of each uncertainty on $p_f(d)$ for different points belonging to Ω_{critical} and consequently its influence on the feasible domain. The closer the cumulative indices of a parameter are to one, the greater the impact on the definition of the feasible domain of the RBDO problem.

The procedure of the RBDO-oriented GSA is summarized in Figure 4.2.

2.6 RBDO-oriented GSA for the reformulated problem

The RBDO-oriented GSA approach proposed in this chapter has been introduced in a general framework for RBDO problems with constraint expressed as in equation (4.11). To apply our method to the reformulated problem (4.1), the notations of section 4.2 with the reformulated extreme-based constraint give:

$$U = (X, X_p, X_{r_\mathrm{E}}) \text{ with, } \forall d, X_d = d + X \tag{4.29}$$

and

$$Y_d = H_d(U) = F_\epsilon\left(\sum_{j=1}^{n_s} \exp\left(M_\mathrm{E}\left(d + X, X_p, X_{r_\mathrm{E}}, s^j\right)\right)\right). \tag{4.30}$$

For the reformulated integral-based constraint, we also want to evaluate the sensitivity of the constraint to the random variable X_{r_I}. Therefore, the reformulated failure probability is written as follows:

$$\mathbb{E}_{X_d,X_p}\left[F_{r_\mathrm{I}}\left(\sum_{j=1}^{n_s} Tp^j\left(M_\mathrm{I}\left(X_d, X_p, s^j\right)\right)\right)\right] = \mathbb{E}_{X_d,X_p,X_{r_\mathrm{I}}}\left[\mathbb{1}_{X_{r_\mathrm{I}} \leq \sum_{j=1}^{n_s} Tp^j(M_\mathrm{I}(X_d,X_p,s^j))}\right]. \tag{4.31}$$

In this case,

$$U = (X, X_p, X_{r_\mathrm{I}}) \tag{4.32}$$

and

$$Y_d = H_d\left(U\right) = \mathbb{1}_{X_{r_\mathrm{I}} \leq \sum_{j=1}^{n_s} Tp^j(M_\mathrm{I}(d+X,X_p,s^j))}. \tag{4.33}$$

For simplicity, we have considered the special case where the uncertainties on d represented by X_d can be expressed as $d+X$ but the GSA could also be applied to a more general case.

Moreover, the cumulative indices measure the influence of uncertainties on the reformulated failure probabilities in the critical domain: where p_f is close to p_s. When a GE procedure is carried out before the RBDO-oriented GSA, it is also in this region that the metamodels have been enriched. Therefore, the evaluation of p_f during the creation of Ω_{SA} described in section 4.2.4 and the evaluation of the cumulative indices can be performed from these metamodels and the complete RBDO-oriented GSA procedure can be carried out without any call to the expensive simulator.

3 Validation of the global enrichment and the sensitivity analysis on the harmonic oscillator problem

The different analyses introduced in this chapter are applied to the reformulated problem of the harmonic oscillator presented in section 2.3.4 with the parameters detailed in section 3.3 as follows:

1. metamodels are calibrated and enriched according to the global enrichment procedure described in section 4.1;

2. using the enriched metamodels obtained at the end of the GE, the sensitivity analysis defined in section 4.2 is performed for each constraint. Cumulative sensitivity indices are obtained for each source of uncertainty and each constraint. These analyses do not require any call to the expensive simulators since they are completely performed with the metamodels;

3. finally, the oscillator problem is solved with AK-ECO introduced in section 3.2. The metamodels used during the first cycle of optimization are the ones obtained at the end of the GE.

3.1 Global enrichment applied to the oscillator problem

For the i-th constraint of the oscillator problem, we denote M_i the expensive function, $\widetilde{M}_i^k$ its kriging model after k cycles of global enrichments and p_i^k the estimation of the reformulated failure probability using Monte Carlo with $\widetilde{M}_i^k$. The corresponding low and high estimations are denoted $p_{i,-}^k$ and $p_{i,+}^k$.

Implementation

For the initialization of GE, the DoE used to calibrate the initial metamodels is a LHS maximin of the augmented space of size 50. The kriging implementation of OpenTURNS is used with a constant trend and a 3/2-Matèrn covariance kernel.

During each cycle and for the i-th constraint, p_i^k, $p_{i,-}^k$, and $p_{i,+}^k$ are estimated with a Monte Carlo sample of size 30000.

The local enrichments are performed at design points selected among a DoE of the 2D design space. This DoE (denoted $DoE_{\mathrm{E,enr}}^k$ or $DoE_{\mathrm{I,enr}}^k$ in section 4.1.3) is a LHS of size 40. At each enrichment of the metamodels, the hyperparameters are updated.

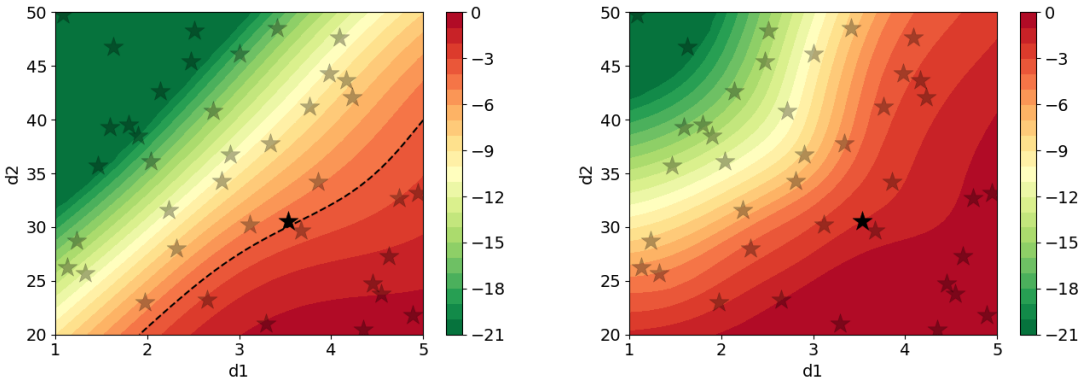
The parameters $n_{\max,\mathrm{stop}}$ and nb_{stop} involved in the accuracy criteria are set respectively to 0 and 2. To avoid increasing the computation time, these criteria are evaluated every 5 enrichment cycles.

Visualization of the first global enrichment cycle

To visualize the GE procedure, we display in Figures 4.3 and 4.4, the different steps of enrichment of the kriging model $\widetilde{M}_1^0$ corresponding to the first constraint of the studied problem, during the first cycle of global enrichment.

Figure 4.3 illustrates the first step of this enrichment: the selection of d_{enr}. The contour lines of $\log\left(p_1^0\right)$ (which is the estimation of the first reformulated failure probability with $\widetilde{M}_1^0$) are displayed in Figure 4.3a and the level set equal to $\log\left(10^{-4}\right)$ is indicated with a black dotted line. In Figure 4.3b, the contour lines of $\log\left(p_{1,+}^0 - p_{1,-}^0\right)$ are shown. In both figures, the candidates for d_{enr} are indicated with transparent black stars ★. The point d_{enr}, represented by a black star ★, is selected among them by minimizing the criterion defined in section 4.1.3. We notice that d_{enr} is such that the numerator of this criterion is small ($p_1^0\left(d_{enr}\right)$ is close to 10^{-4}) and its denominator ($p_{1,+}^0\left(d_{enr}\right) - p_{1,-}^0\left(d_{enr}\right)$) is large.

Once d_{enr} is selected, a local enrichment of $\widetilde{M}_1^0$ is performed by choosing a point x_{enr} from the Monte Carlo sample $\Omega_{MC}\left(d_{enr}\right)$. This step is illustrated in Figure 4.4 which shows the contour lines of $\log\left(p_1^0\right)$. The projections onto the design space of several enrichment candidates are indicated by transparent black crosses ×. The enrichment point x_{enr} is represented with a black cross ✕. The simulator is then evaluated at x_{enr} and the DoE is



(a) Contour lines of $\log\left(p_1^0\right)$ (b) Contour lines of $\log\left(p_{1,+}^0 - p_{1,-}^0\right)$

Figure 4.3: Selection of d_{enr}

updated. The first global enrichment cycle continues with the enrichment of the second and third metamodels denoted respectively $\widetilde{M_2}^0$ and $\widetilde{M_3}^0$.

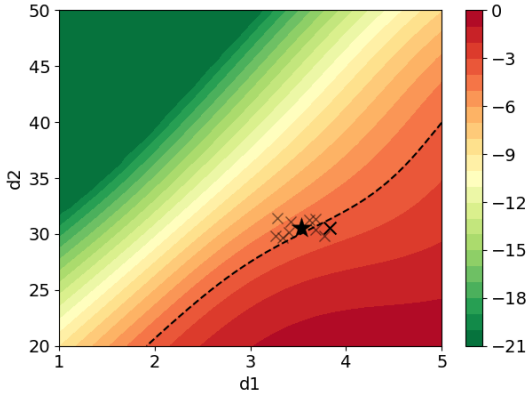


Figure 4.4: Selection of x_{enr}

Results of the global enrichment

The global enrichment procedure applied to the oscillator problem ends after 30 cycles. The accuracy criterion is satisfied after 10 cycles by the metamodel of M_1, after 25 cycles for the metamodel of M_2, and after 30 cycles for the one of M_3. Therefore, 65 calls to the expensive simulator were performed. At the end of the 30-th cycle, all the metamodels are recalibrated and are denoted $\widetilde{M_1}^{30}$, $\widetilde{M_2}^{30}$, and $\widetilde{M_3}^{30}$.

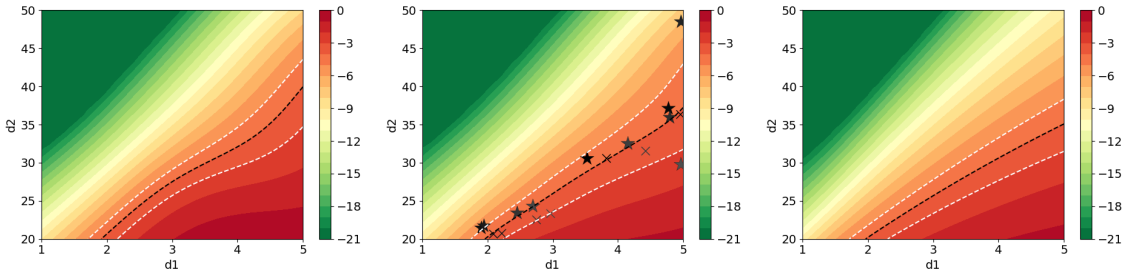
In Figures 4.5, 4.6, and 4.7, we display:

- the contour lines of $\log\left(p_1^0\right)$, $\log\left(p_2^0\right)$, and $\log\left(p_3^0\right)$ (Figures 4.5a, 4.6a, and 4.7a) computed from the initial metamodels with a massive Monte Carlo method;
- the contour lines of $\log\left(p_1^{30}\right)$, $\log\left(p_2^{30}\right)$, and $\log\left(p_3^{30}\right)$ (Figures 4.5b, 4.6b, 4.7b) computed from the enriched metamodels obtained at the end of the global enrichment. In these figures, the stars (★ → ☆) and crosses (× →) represent the design points

d_{enr} and the enrichment points x_{enr} selected during all the enrichment cycles for the considered metamodel. The darker the color of the symbol, the earlier the point is selected during the GE procedure;

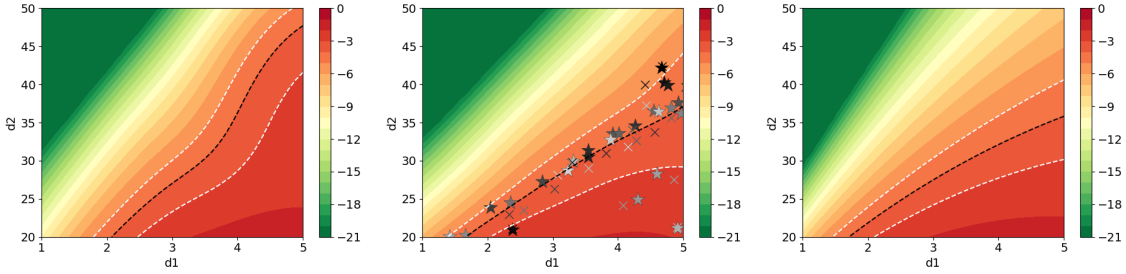
- the contour lines of $\log\left(p_1^{MC}\right)$, $\log\left(p_2^{MC}\right)$, and $\log\left(p_3^{MC}\right)$ computed from the real functions M_1, M_2, and M_3 (Figures 4.5c, 4.6c, and 4.7c).

In each figure, the level set equal to $\log\left(10^{-4}\right)$ of the estimated reformulated failure probabilities is indicated by a black dotted line while white lines are used for the levels sets $\log\left(10^{-5}\right)$ and $\log\left(10^{-3}\right)$.



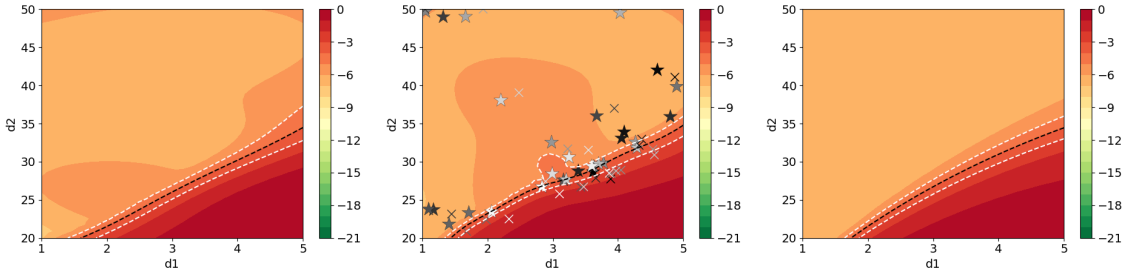
(a) $\log\left(p_1^0\right)$ (b) $\log\left(p_1^{30}\right)$ (c) $\log\left(p_1^{MC}\right)$

Figure 4.5: First constraint



(a) $\log\left(p_2^0\right)$ (b) $\log\left(p_2^{30}\right)$ (c) $\log\left(p_2^{MC}\right)$

Figure 4.6: Second constraint

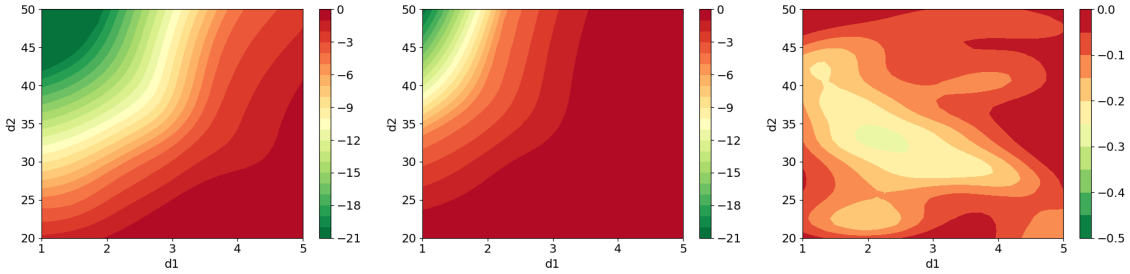


(a) $\log\left(p_3^0\right)$ (b) $\log\left(p_3^{30}\right)$ (c) $\log\left(p_3^{MC}\right)$

Figure 4.7: Third constraint

We observe that when the GE ends, the infeasible domain defined by each constraint is well estimated by the enriched kriging models. The local enrichments are often carried out

in the vicinity of the feasible domain boundary but several points far from the boundary are selected due to the uncertainty of the metamodel (especially for the third constraint). The countour lines of $\log\left(p^0_{1,+} - p^0_{1,-}\right)$, $\log\left(p^0_{2,+} - p^0_{2,-}\right)$, and $\log\left(p^0_{3,+} - p^0_{3,-}\right)$ representing the uncertainties of the first, second and third constraint estimations using the initial metamodels are displayed in Figure 4.8



(a) First constraint: $i = 1$ (b) Second constraint: $i = 2$ (c) Third constraint: $i = 3$

Figure 4.8: countour lines of $\log\left(p^0_{i,+} - p^0_{i,-}\right)$ $(i = 1, 2, 3)$

We could be surprised that the metamodel of M_3 required more enrichments than the other ones since the estimation of the infeasible domain with the initial metamodel seems better than for the other constraints. Two reasons can explain this behavior:

- the first and second constraints are sensitive to the same states (i.e. the states such that s^j and p^j are small). Therefore, each enrichment of the metamodel of M_1 also improves the accuracy of the metamodel of M_2 and vice versa. On the other hand, the third constraint is sensitive to the states such that s^j and p^j are large. Thus, the points selected during the enrichments of the metamodels of M_1 and M_2 do not improve a lot the accuracy of the metamodel of M_3. This could explain why more enrichments were needed for this metamodel;
- Moreover, even though for the third constraint, the approximation of the infeasible domain boundary is good from the start, the uncertainties on the estimations of the reformulated failure probability are important and the level set 10^{-4} estimated with the kriging model varies during the different enrichment cycles.

The global enrichment procedure has been repeated from 2 other initial DoEs. The results are similar. The GE procedure stopped after 70 calls to the simulator with the second DoE and with the third DoE as well. The metamodels $\widetilde{M_1}^0$, $\widetilde{M_2}^0$, $\widetilde{M_3}^0$ met the accuracy criterion respectively after 15, 25, and 30 cycles with the second DoE and after 15, 15, and 40 cycles for the third DoE. With both DoEs, the infeasible domain of each constraint is well estimated.

3.2 RBDO-oriented GSA applied to the oscillator problem

For the i-th $(i = 1, 2, 3)$ constraint of the oscillator problem, as explained in section 4.2.6, the GSA approach described in section 4.2 is carried out using the metamodels $\widetilde{M_i}^{30}$ obtained at the end of the GE procedure detailed in section 4.3.1. Therefore, the analysis does not require any call to the expensive simulator. The design points selected during the first step of the analysis are gathered in Ω^i_{SA}. We then evaluate the cumulative indices of the sources of uncertainties grouped in U^i with:

- $U^1 = (X_1, X_2, X_p, X_{r_1})$ for the first constraint;
- $U^2 = (X_1, X_2, X_p, X_{r_2})$ for the second constraint;
- $U^3 = (X_1, X_2, X_p, X_{r_3})$ for the third constraint;

and $X_d = d + X = (d_1 + X_1, d_2 + X_2)$.

Implementation

In the application of our SA approach to the oscillator problem, we define the critical domain for the i-th constraint as follows:

$$\Omega^i_{\text{critical}} = \left\{ d \in \Omega_d,\, p_i^{30}(d) \in \left[10^{-5}, 10^{-3}\right] \right\}, \tag{4.34}$$

thus $p_s^- = 10^{-5}$ and $p_s^+ = 10^{-3}$.

During the creation of the set Ω^i_{SA} described in section 4.2.4, we consider that the number of points in Ω^i_{SA} must be larger than $n_{\text{SA,min}} = 10$. To select these points, the size of the consecutive DoEs of the design space are 20, 40, 80, 120, and 200 (for n_1 to n_5 and thus $k_{max} = 5$).

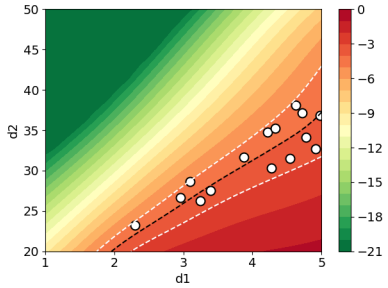
The Fourier Amplitude Sensitivity Testing (FAST) (Cukier et al. (1978); Saltelli et al. (1999)) method implemented in the OpenTURNS (Baudin et al. (2016)) Python package is used to evaluate the Sobol indices. The FAST parameters (denoted N, M and N_r) have been chosen following the recommendations of the OpenTURNS documentation and such that the result of two consecutive FAST analyses performed at the same design point provided the same results. We used $N = 10000$, $M = 10$ and $N_r = 8$.

Results of the RBDO-oriented GSA

The results of the RBDO-oriented GSA applied to the oscillator problem are shown in Figures 4.9, 4.10 and 4.11:

- In Figures 4.9a, 4.10a, and 4.11a, the countour lines of the logarithm of each reformulated failure probability estimated with the enriched metamodels are displayed. The set Ω^i_{SA} of design points selected during the first step of the RBDO-oriented GSA and where the Sobol indices are computed is represented with white dots ○;
- The cumulative Sobol indices evaluated with the enriched metamodels are indicated in Tables 4.9b, 4.10b, and 4.11b;
- To verify the accuracy of the Sobol indices computed from the metamodels, the indices have also been calculated with the real functions M_i at each point of Ω^i_{SA} ($i = 1, 2, 3$) and the cumulative indices are diplayed in Tables 4.9c, 4.10c, and 4.11c.

Let us recall that, for the i-th constraint, the first-order cumulative indices $\overline{S}$ and $S_{\max}$ respectively represent the average and the maximum of the first-order indices of a particular parameter calculated for the different design points of Ω^i_{SA}. The total-order cumulative indices $\overline{S^T}$ and $S^T_{\max}$ are the average and the maximum of the total-order indices calculated for the different design points of Ω^i_{SA}. We have omitted the subscript in the notation of $\overline{S}$, $S_{\max}$, $\overline{S^T}$, and $S^T_{\max}$ to avoid confusion with the number of the considered constraint.



(a) Contour lines of $\log\left(p_1^{30}\right)$ and Ω_{SA}^1

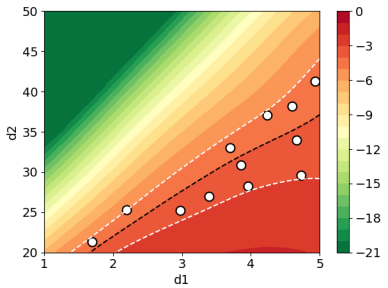
	$\overline{S}$	$S_{\max}$	$\overline{S^T}$	$S_{\max}^T$
X_1	0.003	0.006	0.869	0.948
X_2	0.002	0.002	0.843	0.942
X_p	0.009	0.191	0.906	0.945
X_{r_1}	0.019	0.041	0.944	0.950

(b) Cumulative indices estimated with $\widetilde{M_1}^{30}$

	$\overline{S}$	$S_{\max}$	$\overline{S^T}$	$S_{\max}^T$
X_1	0.003	0.005	0.808	0.922
X_2	0.002	0.003	0.788	0.938
X_p	0.013	0.031	0.913	0.946
X_{r_1}	0.020	0.048	0.941	0.950

(c) Cumulative indices computed with M_1

Figure 4.9: Results of the RBDO-oriented GSA applied to the first constraint of the oscillator problem



(a) Contour lines of $\log\left(p_2^{30}\right)$ and Ω_{SA}^2

	$\overline{S}$	$S_{\max}$	$\overline{S^T}$	$S_{\max}^T$
X_1	0.002	0.004	0.866	0.945
X_2	0.001	0.002	0.861	0.939
X_p	0.010	0.023	0.928	0.948
X_{r_2}	0.017	0.041	0.945	0.950

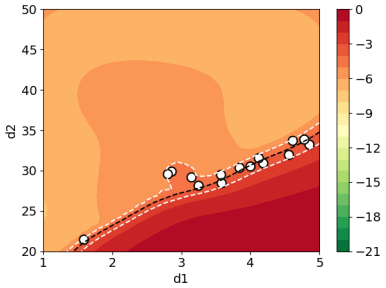
(b) Cumulative indices estimated with $\widetilde{M_2}^{30}$

	$\overline{S}$	$S_{\max}$	$\overline{S^T}$	$S_{\max}^T$
X_1	0.003	0.008	0.786	0.908
X_2	0.002	0.003	0.756	0.921
X_p	0.020	0.053	0.911	0.949
X_{r_2}	0.025	0.073	0.938	0.949

(c) Cumulative indices computed with M_2

Figure 4.10: Results of the RBDO-oriented GSA applied to the second constraint of the oscillator problem

We notice that the cumulative indices estimated from the metamodels are close to those calculated with the real functions which validate our approach. With the enriched metamodels obtained from the GE, the sensitivity of each of the constraints with respect to each source of uncertainty can be accurately assessed without additional call to the ex-



(a) Contour lines of $\log\left(p_3^{30}\right)$ and Ω_{SA}^3

	$\overline{S}$	$S_{\max}$	$\overline{S^T}$	$S_{\max}^T$
X_1	0.003	0.005	0.858	0.950
X_2	0.003	0.005	0.813	0.950
X_p	0.004	0.012	0.950	0.951
X_{r_3}	0.003	0.008	0.940	0.950

(b) Cumulative indices estimated with $\widetilde{M_3}^{30}$

	$\overline{S}$	$S_{\max}$	$\overline{S^T}$	$S_{\max}^T$
X_1	0.002	0.004	0.902	0.950
X_2	0.002	0.004	0.832	0.951
X_p	0.005	0.013	0.948	0.950
X_{r_3}	0.006	0.009	0.939	0.950

(c) Cumulative indices computed with M_3

Figure 4.11: Results of the RBDO-oriented GSA applied to the third constraint of the oscillator problem

pensive simulator.

The results of the RBDO-oriented GSA show that, for each constraint and each input parameter, the first-order cumulative indices are small while the total-order cumulative indices are large. This means that the variability of each constraint is mainly explained by the interactions between the different random variables. It is therefore necessary to consider all these uncertainties in the RBDO problem resolution.

3.3 Resolution of the oscillator problem with global enrichment and AK-ECO

The reformulated oscillator problem described in sections 2.3.4 and 3.3 is solved in this section. After 30 cycles of global enrichment of the metamodels whose results are discussed in section 4.3.1, we obtained three enriched metamodels $\widetilde{M_1}^{30}, \widetilde{M_2}^{30}, \widetilde{M_3}^{30}$. At this stage, 115 calls to the expensive simulator have been carried out (50 for the initial DoE and 65 during the enrichment cycles). The oscillator problem is solved with AK-ECO except that the metamodels used during the first optimization cycle are the ones provided by the GE procedure. We denote this method GE+AK-ECO.

The results are displayed in Table 4.1. They are compared with the reference results obtained with MC discussed in the section 3.3, as well as with the AK-ECO resolution without preliminary global enrichment of the metamodel. The initial DoE of AK-ECO is the same as the initial DoE used in GE for the GE+AK-ECO method. The execution time of GE+AK-ECO is discussed in section F.1.2 of appendix F.

The implementation used for MC and AK-ECO are given is section 3.3 and the resolution

of the oscillator problem starts from the center of the design space with coordinate $(3, 35)$.

	MC	AK-ECO	GE + AK-ECO
d^{min}	(5.0, 35.74)	(5.0, 35.73)	(5.0, 35.74)
$cost(d^{min})$	-14.26	-14.27	-14.26
$p_1^{MC}\left(d^{min}\right)$	0.8×10^{-4}	0.8×10^{-4}	0.8×10^{-4}
$p_2^{MC}\left(d^{min}\right)$	1.0×10^{-4}	1.0×10^{-4}	1.0×10^{-4}
$p_3^{MC}\left(d^{min}\right)$	0.1×10^{-4}	0.1×10^{-4}	0.1×10^{-4}
n_{call}	3.57×10^6	252	223

Table 4.1: Result of GE+AK-ECO

The resolution with GE+AK-ECO required 3 cycles of optimization and 108 additional simulations. Therefore a total of 223 calls were necessary for the full procedure (115 for GE and 108 to solve the optimization problem). When no preliminary global enrichment is performed, AK-ECO requires 252 calls to the simulator and converges within 5 cycles of optimization. Moreover, the design point obtained with GE+AK-ECO is closer to the reference result obtained with MC.

Finally, a multistart resolution of the oscillator problem has been carried out with GE+AK-ECO. The oscillator problem is solved with AK-ECO starting from 20 different design points. The metamodels used during the first cycle of AK-ECO for each resolution are $\widetilde{M_1}^{30}, \widetilde{M_2}^{30}, \widetilde{M_3}^{30}$. For each starting point, GE+AK-ECO converges toward the minimum obtained with MC. The number of additional calls to the simulator varies from 90 to 179 with an average number of calls at 127.75. We have to add the 115 calls performed for GE to obtain the total number of simulations which equals 2670 for the whole multistart procedure GE+AK-ECO.

The multistart resolution has also been performed with AK-ECO and the same initial DoE but without the preliminary GE. In this case, 2968 calls to the simulator were necessary. Consequently, for this problem, better results are obtained regarding the number of calls to the expensive code and the accuracy of the results with a preliminary global enrichment of the metamodels.

The results of the multistart resolutions are displayed in section G.2 of appendix G.

4 Discussion and perspectives

We present in this section some guidance for the choice of different parameters appearing in the GE and the RBDO-oriented GSA approaches. We also discuss limitations of these methods as well as possible improvements.

4.1 Concerning the global enrichment procedure

Remarks specific to GE

In the GE, several parameters involved in the stopping criteria of the procedure and introduced in section 4.1.5 have to be chosen by the user. We recall that a metamodel is considered accurate in GE when the boundary of the feasible domain estimated with the metamodel remains stable over nb_{stop} cycles of enrichment. To evaluate this stability, the reformulated failure probabilities are evaluated at each point of a DoE in the design space denoted DoE_{stop}. In order for the metamodel to be accurate enough at the end, it is important that:

- the size a space-filling DoE_{stop} is large enough size to span the entire design space;
- the number $n_{max,\text{stop}}$ of points is small enough to ensure that a large proportion of the points in DoE_{stop} remains in the same feasibility domain during the last cycle of GE;
- the number of cycles nb_{stop} is large enough to guarantee that the stability is achieved.

The accuracy of the metamodels at the end of the GE is mainly driven by these parameters.

Furthermore, as discussed in the implementation of the GE applied to the oscillator problem, since the evaluation of the accuracy criteria at each cycle can be cumbersome, they can be evaluated only every n cycles.

The enrichment cycle of GE begins with the selection of d_{enr} described in section 4.1.3 by solving an optimization problem on the design space (equations (4.3) and (4.5)). The resolution is done by evaluating candidates among a DoE in the design space. This approach is reasonable when the dimension of the design space is small since a small DoE can cover Ω_d. For a higher dimensional space, an optimization algorithm might be more efficient.

Remarks common to AK-ECO and GE

Since GE and AK-ECO are based on similar kriging strategies, several remarks and limitations of AK-ECO are also relevant for GE:

- the GE procedure can be applied for the more general formulation of the constraints presented in equation (3.33);
- when p_s become small or n_s large, the estimation of the reformulated failure probabilities with Monte Carlo may become cumbersome as well as the selection of x_{enr}. This problem is especially true in GE since many evaluations of p_{E}^k, $p_{\text{E},-}^k$, $p_{\text{E},+}^k$, p_{I}^k, $p_{\text{I},-}^k$, and $p_{\text{I},+}^k$ are required for the selection of d_{enr} and the evaluation of the accuracy criteria;
- the GE also faces the limitations of the kriging technique.

We refer to section 3.4 for a detailed presentation of these issues and the solutions that we proposed for each. In section 3.4, new estimations of low and high estimations of the reformulated failure probabilities are proposed and could benefit the GE especially for the selection of d_{enr}.

4.2 Concerning the RBDO-oriented GSA

We have shown that the RBDO-oriented GSA proposed in this chapter can provide sensitivity indices for the different sources of uncertainty revealing the influence of each on the satisfaction of the constraints. However, the method was applied to a simple problem with few input variables and it would take further investigation on high dimensional problems to validate the approach. Nevertheless, we can already discuss the choice of different parameters and some improvements that could be made to the method.

On the critical domain definition

Our approach relies on assumptions 1 and 2 which state that the relevant GSAs are those performed at design points in Ω_{critical} with:

$$\Omega_{\text{critical}} = \left\{ d \in \Omega_d,\, p_f(d) \in [p_s^-, p_s^+] \right\}. \tag{4.35}$$

The choice of p_s^- and p_s^+ should be driven by the requirement that the uncertainties at a design point d should not affect the order of magnitude of $p_f(d)$. Hence, when $p_s = 10^{-4}$, we suggest to choose $p_s^- = 10^{-5}$ and $p_s^+ = 10^{-3}$.

Defining the critical domain remains a tricky task and further investigation would be necessary to obtain precise guidance. In particular, the following questions should be considered:

- Do the indices calculated for different points in Ω_{critical} vary greatly or are the influences of the parameters of the same order?
- How are the cumulative indices impacted by the choice of p_s^- and p_s^+?

Additional work is needed to answer these questions.

It is also necessary to take into account in the definition of the critical domain that, if the user wishes to use a metamodel for this analysis, the quality of this metamodel must be assessed in the appropriate domain.

Taking into account the cost function in the definition of the critical space

In the work of Spagnol (2020), the cost function is taken into account in the definition of the sublevel set of interest. In the same way, the cost function could be considered in the definition of the critical domain. A simple way to do so is to consider:

$$\Omega_{\text{critical}} = \{ d \in \Omega_d,\, p_f(d) \in [p_s^-, p_s^+] \text{ and } cost(d) \leq c_{\max} \} \tag{4.36}$$

where $c_{\max}$ is a parameter chosen by the user.

The goal of the RBDO-oriented GSA would then be the analysis of the influence of each source of uncertainty on the feasible domain but only where the solution of the optimization problem could be found.

Taking into account the definition of the critical domain in the accuracy criteria of GE

Since the goal of GE is to obtain metamodels allowing to correctly estimate the feasible space boundary, the accuracy criteria currently only assess the stability of the feasible domain. These criteria could easily be adapted to RBDO-oriented GSA by evaluating the stability of the critical domain.

SA dealing with rare event probabilities

During the application of the RBDO-oriented GSA to the oscillator problem, the estimation of the Sobol indices with FAST provided several times a "nan" result. When these results were obtained, they were simply not included in the computation of the cumulative indices. This problem might be due to a variance of Y_d evaluated at 0. To avoid this problem, it would be interesting to investigate the use of methods in GSA dealing with rare events (Kala (2020a)).

Chapter 5

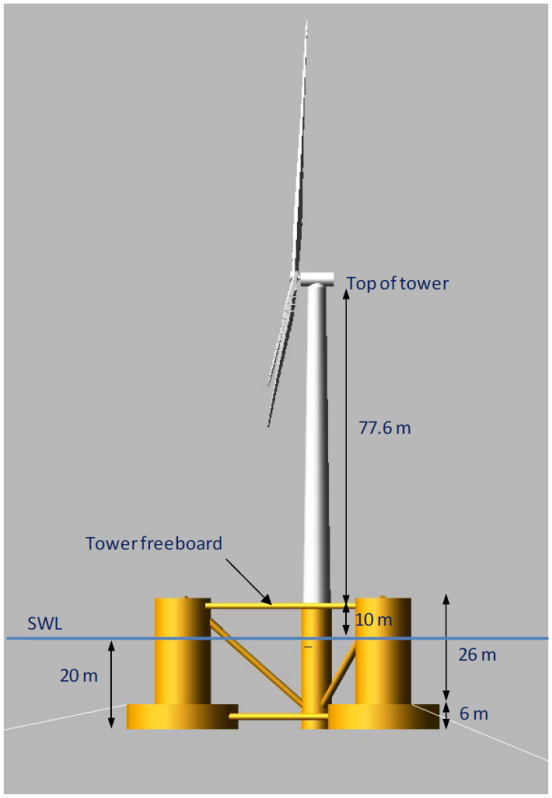
Resolution of the industrial problem

Contents

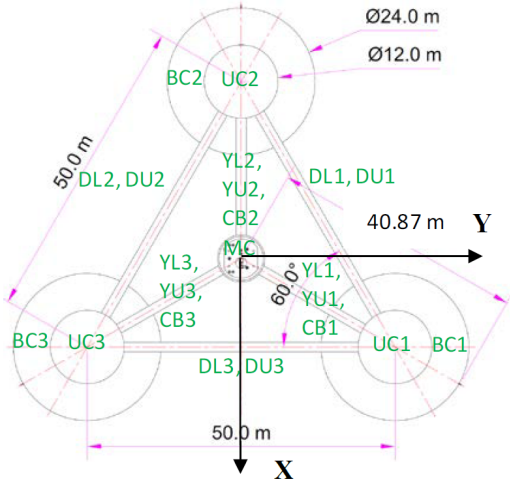
The work of this thesis was driven by an optimization problem seeking to minimize the cost of the mooring lines of a floating offshore wind turbine (FOWT) under Fatigue Limit State constraints. This problem has been briefly introduced in chapter 1 to highlight its main characteristics. We will now present in more details the studied FOWT and the framework used to model the processes involved in the constraints. Then, the optimization problem will be solved by applying the methodology introduced in chapters 2 and 3.

1 The studied floating offshore wind turbine

The floating offshore wind turbine under study is inspired by the NREL (National Renewable Energy Laboratory) 5MW turbine placed on the DeepCwind semi-submersible floating platform (Robertson et al. (2014a)). Figure 5.1 indicates the dimensions of the structure.



(a) DeepCwind floating wind system design



(b) Plan view of the DeepCwind semisubmersible platform

Figure 5.1: DeepCwind FOWT (Robertson et al. (2014a))

The floating platform is composed of four columns: a central one and three offset columns. The role of these columns is to support the weight of the turbine and to stabilize the structure. They are connected through a system of pontoons and cross braces, which allows the structure to have low exposure to waves.

The platform is connected to the seabed by three catenary lines which are composed of chains and are intended to hold the floater in place. The distance between the projection

of the center of the floater to the seabed and the anchors is 837.6m. The angle between two adjacent mooring lines is 120°. In this thesis, a difference with the structure described in Robertson et al. (2014a) is that each mooring line is connected either to the outer surface of the upper part of an external column or to the downward prolongation of this column. In the second case, it is assumed that the heave plate (i.e. the larger surface at the bottom of each offset column which aims at damping heave motion) has a hole through which the mooring line passes. This modification facilitates the change of position of the fairleads, which are the connections where the mooring lines are attached to the columns, during the optimization. This leads to a nominal length of the mooring lines of 841.56m. The axial stiffness, the hydrodynamic diameter, and the breaking load are calculated from linear relations (determined by expert opinion) depending on the mass per unit length of the mooring lines.

Finally we considered a water depth (i.e. the distance between the mean sea level and the seabed) of 200m.

2 Floating offshore wind turbine modeling

2.1 Physics of FOWT modeling

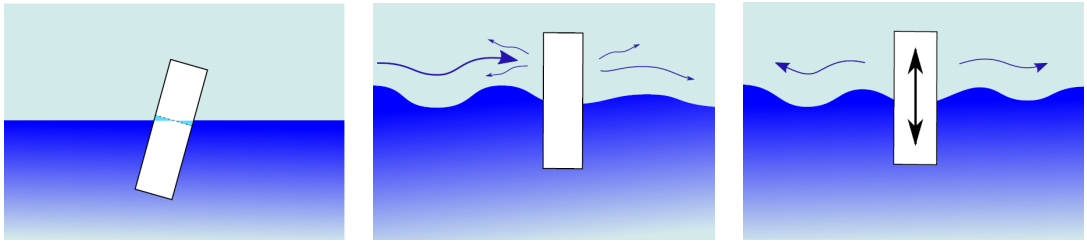
The floating platform is a partially immersed body and its movements result from hydrodynamic and aerodynamic forces applied to the structure. The dynamic response of a floating wind turbine, under wave and wind loading, can be numerically simulated by different multiphysics (i.e. aero-servo-hydro-elastic) codes which can differ according to their level of coupling (a comparison of these codes on the structure OC4 is provided in Robertson et al. (2014b)).

Regarding the hydrodynamic forces, the floater is considered as a large structure (Molin (2002)) compared to the amplitude of the waves and its size is comparable to their wavelength: the flow remains attached to the body and the viscous effects are considered negligible. The resulting hydrodynamic forces are composed of:

- the hydrostatic forces which refer to the static restoring forces due to buoyancy and weight when the platform is moved from the equilibrium position;
- the diffraction forces which correspond to the excitation forces exerted by incident and diffracted waves on the structure at rest;
- the radiation forces which represent the pressure forces applied to the structure by the radiated waves created by the platform motion.

In addition to these effects, we consider an additional quadratic drag as recommended in Robertson et al. (2014a). Finally, the second-order wave loads will not be considered in our study which may lead to understimate the floater motion.

We briefly present the modeling of these efforts in a simplified framework and considering a regular wave represented by a sinusoidal wave. Understanding how the structure reacts to regular waves will allow to model the motion of the platform when subjected to a random loading with a frequency domain solver (see section 5.2.5).



(a) Hydrostatic (b) Diffraction (c) Radiation

Figure 5.2: Hydrodynamic forces (Hall (2013))

2.2 Degrees of freedom of the platform

The movement of the floating platform can be described by six degrees of freedom (DoF) represented in Figure 5.3: the surge, the sway, the heave, the roll, the pitch, and the yaw. They are indicated in the coordinate system $Oxyz$ where:

- O is the intersection of the vertical tower axis and the mean sea level;
- x, y and z are orthonormal vectors;
- the xy-plane represents the mean sea level;
- the x-axis is collinear and in opposite direction with the nominal wind and wave direction;
- the z-axis is directed upward (opposite to gravity).

In this coordinate system, the surge, which will be the DoF of interest in our study, is defined as the horizontal shift along the x-axis. We draw the reader's attention to the fact that the orientation of the vector x in our study is opposite to the one in Figure 5.3. Therefore, a negative surge corresponds to a displacement of the platform in the wind direction.

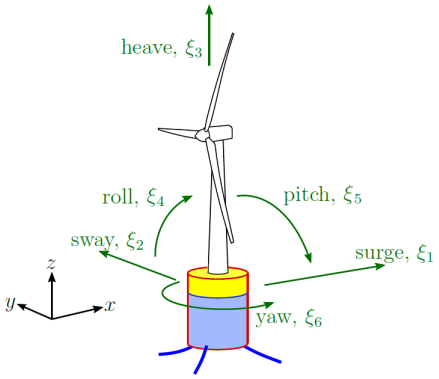


Figure 5.3: Degrees of freedom (Hall (2013))

2.3 Linearized equation of motion of the platform considering a regular wave

We present in this section the equation of motion of the structure subjected to a regular deterministic wave which means that the wave is represented by a sinusoidal function with

angular frequency ω, amplitude A_{wave} and an angle with respect to the x-axis called the **wave heading** and denoted by x_{p_1} (we refer to Molin (2002) for more details).

The movements of the platform at t are described by the DoFs grouped in the vector $Y(t)$:

$$Y(t) = \begin{pmatrix} surge(t) \\ sway(t) \\ heave(t) \\ roll(t) \\ pitch(t) \\ yaw(t) \end{pmatrix} \tag{5.1}$$

and it follows from the Newton's second law, that Y is solution of the motion equation:

$$M\ddot{Y}(t) = F^{hydro}(t) + F^{diff}(t) + F^{rad}(t) + F^{quad}(t) + F^{moor}(t) + F^{wind}(t) \tag{5.2}$$

with

- M the mass matrix of the platform,
- $F^{hydro}(t)$ the hydrostatic force,
- $F^{diff}(t)$ the diffraction force,
- $F^{rad}(t)$ the radiation force,
- $F^{quad}(t)$ the quadratic damping,
- $F^{moor}(t)$ the mooring force which is a tension at the top of each mooring line,
- $F^{wind}(t)$ the main component forces on the Rotor Nacelle Assembly in operational mode: the thrust on the blades and the reaction torque of the turbine.

Hydrostatic

The i-th ($i = 1, \cdots, 6$) coordinate of the hydrostatic force applied to the structure is given by (Robertson et al. (2014a)):

$$F_i^{hydro}(t) = \rho g V_0 \delta_{i3} - \sum_{j=1}^{6} C_{ij}^{hydro} Y_j(t) \tag{5.3}$$

with ρ the water density, g the gravitational acceleration constant, V_0 the displaced volume of fluid when the platform is in its undisplaced position and δ the Kronecker delta function.

This formulation decomposes the hydrostatic forces into two components. The first term represents the buoyancy force from the Archimedes' principle (i.e. the force directed vertically upward and equal to the weight of the displaced fluid when the platform is in its undisplaced position). The term C_{ij}^{hydro} represents the change in the hydrostatic force and moment as the platform is displaced. The matrix C^{hydro} is called the stiffness matrix.

Diffraction

Using the potential-flow theory (Molin (2002)), the diffraction forces can be written (Duarte et al. (2014)):

$$F^{diff}(t) = Re\left(A_{wave}\hat{X}(\omega, x_{p_1})e^{i\omega t}\right) \tag{5.4}$$

where A_{wave} is the wave amplitude and $\hat{X}(\omega, x_{p_1})$ is the first-order wave load per unit amplitude which depends on the wave angular frequency ω and the wave heading x_{p_1}.

Radiation

For each DoF, the radiation forces are decomposed into two components: one in phase with the body acceleration (the added mass term) and one in phase with the body velocity (the damping term). We denote $M_a(\omega)$ the added mass matrix and $B^{rad}(\omega)$ the damping matrix. The potential-flow theory is used to compute these matrices (Robertson et al. (2014a)).

Quadratic drag

As recommended in Robertson et al. (2014a), a quadratic drag is considered to accurately represent the damping in real system. This extra damping represents the dissipation by vortex shedding at the corner of column heave plates and Morison forces contribution on thinner elements like braces. The i-th coordinate of the resulting force is given by the following equation:

$$F_i^{quad}(t) = -\sum_{j=1}^{6} B_{ij}^{quad}|\dot{Y}_j(t)|\dot{Y}_j(t). \tag{5.5}$$

The matrix B^{quad} is referred as the additional quadratic drag matrix and we denote the surge and pitch components $x_{p_2} = B_{11}^{quad}$ and $x_{p_3} = B_{55}^{quad}$. The estimation of these terms generally requires a comparison to complex CFD simulations or to a scaled basin experiment (Burmester et al. (2020)).

Equation (5.5) is linearized following the principle of quasi-linearization described in Borgman (1967) and Le Cunff et al. (2008) to obtain:

$$F_i^{quad}(t) \simeq -\sum_{j=1}^{6} B_{ij}^{quad,lin}(\omega)\dot{Y}_j(t). \tag{5.6}$$

Linearized Mooring Model

All the mooring lines contribute to the load on the support. These forces are described by the following equation (Robertson et al. (2014a)):

$$F_i^{moor}(t) = F_i^{moor,0} - \sum_{j=1}^{6} C_{ij}^{moor}Y_j(t) \tag{5.7}$$

where $F_i^{lines,0}$ is the i-th component of the total mooring system load acting on the support platform in its undisplaced position (pre-tension at the fairlead from the weight of the mooring lines not resting on the seafloor). The matrix C^{moor} is the linearized restoring matrix which combines the elastic stiffness of the mooring lines and the effective geometric stiffness due to the weight of the lines in the water.

Wind loads

For simplicity, the wind is considered in the simulations only by means of constant thrust and torque applied to the floating platform. Their values correspond to the mean of dynamic simulations with DeeplinesWind^{TM} including the turbine in operation and the controller with a turbulence corresponding to a class B wind turbine (IEC 61400-3 (2009)).

Linear damping coefficients due to the turbine are also considered. The values were provided by time domain simulations with an operating turbine. This simplification also comes from the fact that in the current version of the DeeplinesWind^{TM} software we are using, the aerodynamic simulation of wind turbine is only possible with a time domain solver.

Linearized equation of motion and RAO

The linearized equation of motion can thus be formulated as:

$$M^{total}(\omega)\ddot{Y}(t) + B^{total}(\omega)\dot{Y}(t) + C^{total}Y(t) = Re(A_{wave}\hat{X}(\omega, x_{p_1})e^{i\omega t}) + F^{static} \tag{5.8}$$

with $M^{total}(\omega) = M + M_a(\omega)$, $B^{total}(\omega) = B^{rad}(\omega) + B^{quad,lin}(\omega)$, and $C^{total} = C^{hydro} + C^{moor}$. The constant forces due to the wind, the hydrostatic and the mooring loads are considered in F^{static}.

We can decompose the vector $Y(t)$ into two components as $Y(t) = \overline{Y}(t) + Y^{static}$ with:

$$M^{total}(\omega)\ddot{\overline{Y}}(t) + B^{total}(\omega)\dot{\overline{Y}}(t) + C^{total}\overline{Y}(t) = Re(A_{wave}\hat{X}(\omega, x_{p_1})e^{i\omega t}) \tag{5.9}$$

and

$$C^{total}Y^{static} = F^{static}. \tag{5.10}$$

Denoting $\overline{Y}(t) = Re\left(\hat{\overline{Y}}(\omega)e^{i\omega t}\right)$, it follows from equation (5.9):

$$\frac{\hat{\overline{Y}}(\omega)}{A_{wave}} = (-\omega^2 M^{total}(\omega) + i\omega B^{total}(\omega) + C^{total})^{-1}\hat{X}(\omega, x_{p_1}). \tag{5.11}$$

The term on right-hand side is called the Response Amplitude Operator (RAO) (Molin (2002)) with:

$$\begin{pmatrix} RAO_1(\omega) \\ . \\ . \\ . \\ RAO_6(\omega) \end{pmatrix} = (-\omega^2 M^{total}(\omega) + i\omega B^{total}(\omega) + C^{total})^{-1}\hat{X}(\omega, x_{p_1}) \tag{5.12}$$

and RAO_1 corresponds to the surge RAO. It depends on the design variables d that will be specified later, on the wave heading x_{p_1}, the surge and pitch drag coefficients x_{p_2} and x_{p_3}, and the mean wind speed denoted u of the time series used to estimate the constant thrust and torque as explained in section 5.2.3.

2.4 Motion of the mooring lines and tension

Some of the constraints of the optimization problem considered in this thesis depend on the tension in the mooring lines. As explained in section 1.5, we are particularly interested in the tension at the top of each line l ($l = 1, 2, 3$). The lines are composed of chains modeled by 1D 2-node elements with linear elasticity. The tension in each element is derived from the usual Finite Element approximation with linear functions of element node displacements.

As for the platform DoFs, an equation of motion is considered to determine the motions of the mooring lines nodes subjected to a regular wave. The external forces applied to these lines are thus the gravity force, the floater reaction forces, the forces resulting from Archimedes' principle and the seabed, and the wave loads. For the latter, the mooring lines are considered as small structures compared to the wave amplitude and the hydrodynamic forces on the lines are described by the empirical formula of Morison (Robertson et al. (2014a)).

The Morison forces are then linearized (Le Cunff et al. (2008)) enabling to obtain a linear relation between the regular wave and the motion of the nodes. From hereabove mentioned post-processing of the tension and from the linearization of the Morison forces, it follows a linear relation between the tension and the regular wave. Thus, similarly to what was done for the DoF of the floating platform, the tension at the top of the line l is decomposed into a static component and a dynamic component. The latter is solution of a linear equation which enables us to obtain the RAO of the tension at the top of the l-th line denoted $RAO_{\mathcal{T}^l}$. This RAO depends on d, x_{p_1}, x_{p_2}, x_{p_3}, and u as well. We denote x_p the vector $(x_{p_1}, x_{p_2}, x_{p_3})$.

2.5 From a regular wave to a random sea elevation process

The sea elevation process

We have presented how the DoFs of the platform and the tension are obtained when the structure is subjected to a deterministic regular wave. However, to account for the randomness of marine conditions, it is usual to represent the waves as a random stationary process called the **sea elevation process**. When the water depth is important, the process is considered Gaussian since its realizations can be seen as the superposition of a large number of regular waves with different angular frequencies ω (Molin (2002)).

To represent the sea elevation at each instant, the considered period of time $[0, T]$ is divided into n_T subintervals I_i ($i = 1, \cdots, n_T$) of length ΔT (by convention $\Delta T = 3$ hours). During each interval on time I_i, the sea elevation is represented by a **stationary Gaussian** process with **zero mean**. Its distribution is defined by its spectral density (see section 2.2.1 for the definition of a spectral density) which depends on the **long-term** parameters $h_{s,i}$ and $t_{p,i}$ where:

- $h_{s,i}$ is the significant waveheight during I_i;
- $t_{p,i}$ is the peak period during I_i;
- usually the spectral density is also characterized by a long-term parameter γ quantifying the spread of the spectrum but we consider this parameter equal to 3 (which is a value close to the value 2.87 taken in Robertson et al. (2014b)).

We denote $K_\eta\left(h_{s,i}, t_{p,i}; .\right)$ the spectral density of the i-th stationary Gaussian process. The one-sided spectral density $K_\eta^+\left(h_{s,i}, t_{p,i}; .\right)$ is the function such that:

$$\begin{aligned} K_\eta^+\left(h_{s,i}, t_{p,i}; \omega\right) &= 2K_\eta(h_{s,i}, t_{p,i}; \omega) && \text{if } \omega \geq 0 \\ K_\eta^+\left(h_{s,i}, t_{p,i}; \omega\right) &= 0 && \text{if } \omega < 0. \end{aligned}$$

In this study, we consider the **JONSWAP** one-sided spectral density (Hasselmann et al. (1980)) defined as follows:

$$K_\eta^+\left(s_i; \omega\right) = \alpha\, h_{s,i}^2\, \omega_{p,i}^4\, \omega^{-5} \exp\left(-\frac{5}{4}\left(\frac{\omega}{\omega_{p,i}}\right)^{-4}\right) \gamma^{\beta(\omega)} \tag{5.13}$$

with $\omega_{p,i} = 2\pi / t_{p,i}$,

$$\beta(\omega) = \exp\left(-\frac{(\omega - \omega_{p,i})^2}{2\,\omega_p^2\, s(\omega)^2}\right) \tag{5.14}$$

and $s(\omega) = 0.07$ if $\omega \leq \omega_{p,i}$ and $s(\omega) = 0.09$ otherwise. Finally, the constant α is chosen such that:

$$h_{s,i}^2 = 16 \int_0^\infty K_\eta^+\left(h_{s,i}, t_{p,i}; ; \omega\right) d\omega. \tag{5.15}$$

The JONSWAP one-sided spectral density considering $h_{s,i} = 2.37$m and $t_{p,i} = 8.03$s is displayed in Figure 5.4.

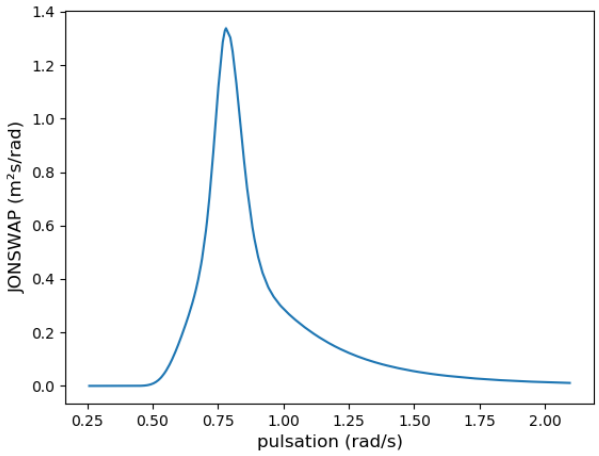


Figure 5.4: JONSWAP one-sided spectral density

We denote $\eta_i\left(h_{s,i}, t_{p,i}; .\right)$ the sea elevation process such that, for each instant $t \in I_i$, the sea elevation is $\eta_i\left(h_{s,i}, t_{p,i}; t\right)$. Moreover, η will represent the sequence of independent processes $\left(\eta_1\left(h_{s,1}, t_{p,1}; .\right), \cdots, \eta_{n_T}\left(h_{s,n_T}, t_{p,n_T}; .\right)\right)$.

Definition of the sea states

For each long-term parameter combination $(h_{s,i}, t_{p,i})$ $(i = 1, \cdots, n_T)$, we associate a mean wind speed u_i. We will call the **sea state** during I_i, the vector $s_i = (h_{s,i}, t_{p,i}, u_i)$. Each sea state $s_1, \cdots, s_{n_T}$ is equal to an element of the set $\{s^1, \cdots, s^{n_s}\}$.

Usually, the set of possible sea states $\{s^1, \cdots, s^{n_s}\}$ encountered during $[0, T]$ as well as their probability of occurrence $\{p^1, \cdots, p^{n_s}\}$ are given by a histogram called a scatter diagram which groups several thousand or even tens of thousands of combinations of long-term parameters (Vorpahl et al. (2013)).

In order to solve a simplified problem, only seven combinations of long-term parameters are considered in this study. They are obtained by massively grouping adjacent blocks in the scatter diagram created from buoy data collected during several decades on the US east coast (Stewart et al. (2016)). This scatter diagram contains many combinations of long-term parameters. The blocking is performed by considering only 7 wind speeds $u^1, \cdots, u^7$ among the wind speeds appearing in the scatter diagram. Then, for each u^j $(j = 1, \cdots, 7)$, we associate the maximum significant height h_s^j and the mean peak period t_p^j of the combinations in the scatter diagram for which the wind speed is equal to u^j.

For each sea state s^j $(j = 1, \cdots, 7)$, we denote n^j the number of intervals I_i in $[0, T]$ such that $s_i = s^j$ and we define $p^j = \frac{n^j}{n_T}$. To estimate the parameters p^j, we consider a random variable U with a Weibull distribution (with a shape $= 2.128$ and a scale $= 9.495$) and the p^j $(j = 1, \cdots, 7)$ are computed as follows:

- $p^1 = \mathbb{P}\left(U \leq u^1 + 1\right)$;
- $p^j = \mathbb{P}\left(U \in \left[u^j - 1, u^j + 1\right]\right)$ for $j = 2, \cdots, 6$;
- $p^7 = \mathbb{P}\left(U \geq u^7 - 1\right)$.

The resulting sea states s^j and the respective p^j $(j = 1, \cdots, 7)$ are presented in Table 5.1.

Sea state $s^j = (h_s^j, t_p^j, u^j)$	p^j
$s^1 = (2.37, 8.03, 4)$	$p^1 = 0.23$
$s^2 = (2.51, 8.09, 6)$	$p^2 = 0.18$
$s^3 = (2.86, 8.12, 8)$	$p^3 = 0.18$
$s^4 = (3.47, 7.61, 10)$	$p^4 = 0.16$
$s^5 = (4.04, 7.12, 12)$	$p^5 = 0.11$
$s^6 = (4.3, 6.50, 14)$	$p^6 = 0.07$
$s^7 = (5.08, 6.64, 16)$	$p^7 = 0.07$

Table 5.1: Couples $\left(s^j, p^j\right)$ with h_s^j (m), t_p^j (s), and u^j (m/s)

It is to be noted that in buoy data the mean wind speed is given for 8 minutes. As the wind stationary length of about 10 minutes is much shorter than that of wave, the international standard of FOWT design usually recommend using a conservative transformation into a mean wind speed over longer duration. For simplicity, this point is neglected in the following of this thesis.

Remark 5.1. We recall that although p^j is usually considered as an occurrence probability in offshore engineering, we have defined the p^j in the previous chapters as the proportion of intervals I_i such that $s_i = s^j$ (see section 2.3.1 for more details).

To simplify the notation, we will denote from now on, $\eta_i(s_i;.)$ the sea elevation process during I_i with spectral density $K_\eta(s_i;.)$ even though these objects do not depend on u_i.

The surge process

For a given sea state s_i, considering a sea elevation random process during I_i implies that the surge becomes a random process as well. Since the sea elevation process is a superposition of regular waves, the surge process during I_i denoted $\mathcal{S}_i(d, x_p, s_i;.)$ can be decomposed as:

$$\mathcal{S}_i(d, x_p, s_i; t) = \mu_{\mathcal{S}}(d, x_p, s_i) + \overline{\mathcal{S}_i}(d, x_p, s_i; t) \ , \ \forall t \in I_i \tag{5.16}$$

where $\mu_{\mathcal{S}}(d, x_p, s_i)$ is given solving (5.10) and $\overline{\mathcal{S}_i}(d, x_p, s_i;.)$ is the superposition of elementary motions (Molin (2002)) solutions of equation (5.9) considering different angular frequencies. It follows that :

$$\overline{\mathcal{S}_i}(d, x_p, s_i;.) = h_{\mathcal{S}}(d, x_p, s_i;.) * \eta_i(s_i;.) \tag{5.17}$$

with, denoting TF the Fourier transformation,

$$\mathrm{TF}\left(h_{\mathcal{S}}(d, x_p, s_i;.)\right)(\omega) = RAO_1(\omega). \tag{5.18}$$

To remain consistent with the notation used from the beginning of the thesis, RAO_1 will be denoted $H_{\mathcal{S}}(d, x_p, s_i;.)$. It follows from equation (5.17) and the properties of the sea elevation process that $\mathcal{S}_i(d, x_p, s_i;.)$ is a stationary Gaussian process with mean $\mu_{\mathcal{S}}(d, x_p, s_i)$ and spectral density:

$$K_{\mathcal{S}}(d, x_p, s_i;.) = \left|H_{\mathcal{S}}(d, x_p, s_i;.)\right|^2 K_\eta(s_i;.). \tag{5.19}$$

Remark 5.2. We point out that $H_{\mathcal{S}}$ represents the surge RAO while h_s is a long-term parameter on which the distribution of the sea elevation process depends.

The spectral moment of order n of $\mathcal{S}_i(d, x_p, s_i;.)$ is therefore defined as follows:

$$m_{\mathcal{S},n}(d, x_p, s_i) = \int_{\mathbb{R}} \omega^n K_{\mathcal{S}}(d, x_p, s_i; \omega)\, d\omega. \tag{5.20}$$

To take into account the different sea states, the surge process over $[0, T]$ is defined as:

$$\mathcal{S}(d, x_p; t) = \sum_{i=1}^{n_T} \mathcal{S}_i(d, x_p, s_i; t)\, \mathbb{1}_{I_i}(t) \ , \ \forall t \in [0, T] \tag{5.21}$$

It is therefore a piece-wise stationary Gaussian process.

The tension processes

The reasoning presented for the definition of the surge process holds for the tension: for a given sea state s_i, when we consider a random sea elevation process, the tension at the top of the l-th line ($l = 1, 2, 3$) is represented by a random process such that:

$$\mathcal{T}_i^l(d, x_p, s_i; t) = \mu_{\mathcal{T}^l}(d, x_p, s_i) + \overline{\mathcal{T}_i^l}(d, x_p, s_i; t) \ , \ \forall t \in I_i \tag{5.22}$$

with $\overline{\mathcal{T}_i^l}(d, x_p, s_i;.) = h_{\mathcal{T}^l}(d, x_p, s_i;.) * \eta_i(s_i;.)$ and $\mathrm{TF}\left(h_{\mathcal{T}^l}(d, x_p, s_i;.)\right)(\omega) = RAO_{\mathcal{T}^l}(\omega)$.

The RAO $RAO_{\mathcal{T}^l}$ will be denoted $H_{\mathcal{T}^l}(d, x_p, s_i; .)$. It follows from the properties of the sea elevation process that $\mathcal{T}_i^l(d, x_p, s_i; .)$ is a stationary Gaussian process with mean $\mu_{\mathcal{T}^l}(d, x_p, s_i)$ and spectral density:

$$K_{\mathcal{T}^l}(d, x_p, s_i; .) = |H_{\mathcal{T}^l}(d, x_p, s_i; .)|^2 K_\eta(s_i; .). \tag{5.23}$$

The spectral moment of order n of $\mathcal{T}_i^l(d, x_p, s_i; .)$ is therefore defined as follows:

$$m_{\mathcal{T}^l,n}(d, x_p, s_i) = \int \omega^n K_{\mathcal{T}^l}(d, x_p, s_i; \omega)\, d\omega \tag{5.24}$$

and the tension at the top of the line l ($l = 1, 2, 3$) over $[0, T]$ is defined as the following piece-wise stationary Gaussian process:

$$\mathcal{T}^l(d, x_p; t) = \sum_{i=1}^{n_T} \mathcal{T}_i^l(d, x_p, s_i; t)\, \mathbb{1}_{I_i}(t)\ ,\ \forall t \in [0, T]. \tag{5.25}$$

Deeplines$^{\text{TM}}$ simulation outputs

We emphasize that to apply the methodology presented in this thesis to solve the FOWT optimization problem, only the distributions of the surge and tension processes are required. As these processes are piece-wise stationary Gaussian, we are interested in their means and spectral moments. The former are obtained by solving the static equation while for the latter, the RAOs and the JONSWAP spectral density are used.

All these quantities are computed with the Deeplines$^{\text{TM}}$ software. Our script takes as input a sea state s_i, a design point d and parameters x_{p_1}, x_{p_2}, and x_{p_3}, calls a Deeplines$^{\text{TM}}$ executable and returns the means and the spectral moments of the processes $\mathcal{S}_i(d, x_p, s_i; .)$ and $\mathcal{T}_i^l(d, x_p, s_i; .)$.

Hence, for our approach, it is not necessary to sample realizations of the surge and tension processes.

2.6 Modeling simplifications

Let us recap the modeling simplifications introduced in the section:

- first, during each interval of time I_i, the wind loads are replaced by a constant thrust and a constant torque which leads to underestimate the fatigue on the mooring lines by eliminating the variations due to aerodynamics;
- we do not consider the effects of the second-order hydrodynamic forces on the FOWT. Despite the fact that Duarte et al. (2014) has shown their influence on the surge, considering only the first-order wave forces is common in many FOWT studies as can been shown in Robertson et al. (2014b).
- the wave heading x_{p_1} is considered constant over $[0, T]$;
- all the non-linear forces (the quadratic drag and the Morison forces) are linearized. Although the international design standards (Det Norske Veritas (2013)) recommend to simulate in the time domain the behavior of the floating wind turbine to take into account the non-linearities, we assume that a frequency calculation can provide a good proxy for estimating the fatigue.

An analysis has been conducted to determine the influence of these simplifications on the surge and tension processes. More precisely, time series of these processes have been computed considering the model framework introduced in this section and compared with time series generated with external forces that are not linearized and wind loads that are not constant. The results are displayed in section H.1 of appendix H.

The linear equations of motion obtained from these simplifications enable us to obtain Gaussian stationary output processes. Finally, only 7 sea states were considered in the modeling of the marine conditions.

3 Fatigue

When a component of a structure is subjected to repeated stress variations, damage accumulates and a gradual degradation of the material is observed: this is called the material fatigue. In metallic structures, this process results from the development of dislocations within the polycristalline aggregate which creates surface irregularities. Subjected to millions of cycles, a crack can appear and propagate in the material due to the polycyclic fatigue. Once this crack is large enough, the component is no longer able to withstand the load imposed on it and failure occurs. Fatigue is a common cause of failure and is particularly difficult to estimate. Therefore, it is important that the design of the structure enables to resist these cycles over a long period of time.

In offshore engineering and particularly for mooring lines, the common practice (Det Norske Veritas (2005)) is to estimate the total damage accumulated by independent loading events which is assumed to provide a reasonable estimate for stationary cases.

Thus, we can represent the **total fatigue damage** at the top of the line l ($l = 1, 2, 3$) over the period $[0, T]$, denoted $D^{total,l}_{[0,T]}(d, x_{d_2}, x_p)$, as the accumulation of the **instantaneous damage** $\mathcal{D}^l(d, x_{d_2}, x_p; t)$ over $[0, T]$ and thus write it as the following integral:

$$D^{total,l}_{[0,T]}(d, x_{d_2}, x_p) = \int_0^T \mathcal{D}^l(d, x_{d_2}, x_p; t)\, dt. \tag{5.26}$$

Since the damage at the top of the line depends on the tension at this point, when this tension is represented by a random process $\mathcal{T}^l(d, x_p; .)$, the instantaneous damage $\mathcal{D}^l(d, x_{d_2}, x_p; .)$ is also a random process and the total damage $D^{total,l}_{[0,T]}(d, x_{d_2}, x_p)$ is a random variable. Their distributions depend on the distribution of the tension process as well as on the parameter x_{d_2} that is introduced below. Furthermore, the instantaneous process over $[0, T]$ is defined as:

$$\mathcal{D}^l(d, x_{d_2}, x_p; t) = \sum_{i=1}^{n_T} \mathcal{D}^l_i(d, x_{d_2}, x_p; t)\, \mathbb{1}_{I_i}(t)\ ,\ \forall t \in [0, T] \tag{5.27}$$

where $\mathcal{D}^l_i(d, x_{d_2}, x_p; t)$ is the instantaneous damage caused by $\mathcal{T}^l_i(d, x_p, s_i; t)$, $(i = 1, \cdots, n_T)$. Finally, we denote $D^{total,l}_{I_i}(d, x_{d_2}, x_p, s_i)$ the total damage occurring during I_i.

In practice, engineers do not have access to realizations of the instantaneous damage process but use methods to estimate the total damage. Two approaches can be distinguished to evaluate the mean total damage:

- the time-domain approach which deduces the mean total damage from a unique realization considered over a large enough period of time of the tension process;
- the spectral approach which uses the spectral moments of the tension process.

We will now present the time-domain method to extract the damage caused by a tension history which will enable us to introduce different elements used in frequency-domain methods.

3.1 Fatigue evaluation from a tension history

Let us consider a realization of the tension process $\mathcal{T}_i^l(d, x_p, s_i; .)$ over I_i. The total damage over I_i is then a scalar that we denote $d_i^{total,l}$.

Cycle counting methods

Polycyclic fatigue occurs as a result of millions of cycles of loading (stress or force like tension) in the time history. The main influence of these cycles comes from the amplitudes of the loading variations and to a lesser extent from the average tension which will be neglected in this thesis. The first step of the total damage estimation is to extract from the tension history the characteristics of these cycles. This is performed by a **cycle counting** method. The rainflow counting method (Matsuishi and Endo (1968); Downing and Socie (1982)) is considered as the most accurate counting procedure (an overview of cycle counting methods is provided in Dirlik (1985)). From a tension time series, a counting method provides a histogram. For each cycle, the tension range and the mean tension are evaluated and the histogram indicates the number of cycles contained in the tension history for each tension range and mean tension.

We will denote n_c the number of cycles counted from the tension realization considered in this section and r_k the tension range of the k-th cycle.

The T-N curve

Once the tension cycles of the tension history have been characterized, it is necessary to know the damage caused by each cycle on the material. The relation between a tension range r and the number of cycles n_r with this tension range that the material can withstand is provided by the T-N curve. This relation is obtained by experimentally repeating cycles of tension on the considered material with different tension range until failure. Fitting parameters are then adjusted (see Figure 5.5) to obtain the following relation:

$$\log(n_r) = \log(K_D) - m \log\left(\frac{r}{BL(d_2)}\right) \tag{5.28}$$

where K_D and m are given by the experiments and the breaking load BL depends on the material considered.

Denoting $x_{d_2} = K_D BL(d_2)^m$, we have:

$$n_r = \frac{x_{d_2}}{r^m}. \tag{5.29}$$

In this study, we consider the Stiff's fatigue curve for chain with $m = 3$ (Rossi (2005)). The breaking load is given by $BL(d_2) = c_1 \times d_2 - c_2$, where c_1 and c_2 are values determined by expert opinion, and d_2 is the mass per unit length of the mooring lines.

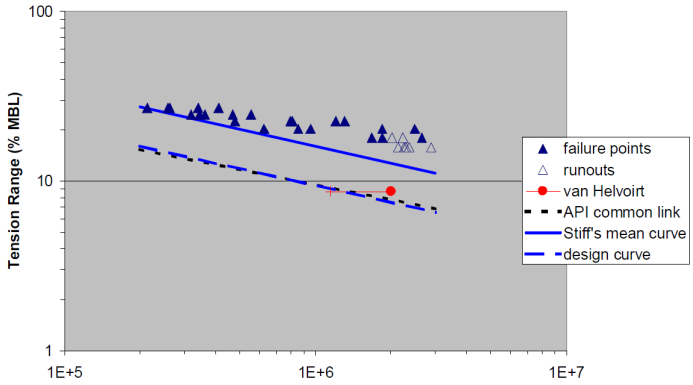


Figure 5.5: Fitting of the T-N curve (Rossi (2005))

The Palmgren-Miner rule

The total damage caused by the time history is finally obtained by summing the damage caused by each cycle shown in the tension range histogram. This approach is known as the Palmgren-Miner accumulated damage rule (Miner (2021)) and the total damage caused by the considered tension history is given by:

$$d_i^{total,l} = \sum_{k=1}^{n_c} \frac{1}{n_{r_k}} = \sum_{k=1}^{n_c} \frac{r_k^m}{x_{d_2}}. \tag{5.30}$$

The Palmgren-Miner rule assumes that the order of occurrence of cycles has no influence. This is considered as a limitation since cycles of low stress followed by high stress cause more damage than would be predicted by the rule (Eskandari and Kim (2017)).

The time-domain approach enables to obtain a realization of the total damage random variable from a realization of the tension process. For a FOWT application, a time series of tension of one hour (discarding the first 200s of transition) is considered sufficiently long to assume the ergodicity of the instantaneous damage process. Therefore, the mean total damage caused during I_i can be deduced from one time series of tension.

3.2 Fatigue evaluation with a spectral approach: the Dirlik method

Contrary to the time-domain approach, the spectral methods directly consider the properties of the tension process to deduce the mean of the total fatigue damage random variable. Considering not a realization but the complete tension process $\mathcal{T}_i^l(d, x_p, s_i; .)$, the number of cycle n_c and the range of tension of the k-th cycle denoted r_k introduced in section 5.3.1 become random variables that we denote respectively N_c and R_k. It follows from the Palmgren-Miner rule:

$$\mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}(d, x_{d_2}, x_p, s_i)\right] = \mathbb{E}_{\eta_i}\left[\sum_{k=1}^{N_c} \frac{R_k^m}{x_{d_2}}\right] \tag{5.31}$$

where N_c and R_k depend on d, x_p, and s_i. Under the assumption that the sequence of R_k are i.i.d (Benasciutti and Tovo (2004)), we have:

$$\mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}\left(d,x_{d_2},x_p,s_i\right)\right] = \sum_{n_c\geq 0}\mathbb{P}\left(N_c=n_c\right)\sum_{k=1}^{n_c}\mathbb{E}\left[\frac{R_k^m}{x_{d_2}}\right] \tag{5.32}$$

$$= \sum_{n_c\geq 0}\mathbb{P}\left(N_c=n_c\right)n_c\mathbb{E}\left[\frac{R_1^m}{x_{d_2}}\right] \tag{5.33}$$

$$= \frac{1}{x_{d_2}}\mathbb{E}\left[N_c\right]\mathbb{E}\left[R_1^m\right]. \tag{5.34}$$

To simplify the notation, the spectral moment of order n of $\mathcal{T}_i^l\left(d,x_p,s_i;.\right)$ will be denoted m_n in this section (i.e. $m_n = m_{\mathcal{T}^l,n}\left(d,x_p,s_i\right)$).

As explained in Benasciutti and Tovo (2004), for a stationary Gaussian tension process $\mathcal{T}_i^l\left(d,x_p,s_i;.\right)$:

$$\mathbb{E}[N_c] = \Delta T\frac{1}{2\pi}\sqrt{\frac{m_4}{m_2}} \tag{5.35}$$

and

$$\mathbb{E}\left[R_1^m\right] = \int_{\mathbb{R}^+} r^m p_{R_1}(r)dr \tag{5.36}$$

with p_{R_1} the probability density function of the tension range of the first cycle.

The density p_{R_1} cannot be deduced analytically but several methods have been proposed to approach this function (Benasciutti and Tovo (2004)). In particular, the Dirlik method (Dirlik (1985)) proposes an approximation denoted $p_{R_1}^{Dir}$ when considering a Gaussian tension process based on the combination of an exponential and two Rayleigh densities:

$$p_R^{Dir}(r) = \frac{1}{2\sqrt{m_0}}\left(\frac{D_1}{Q}e^{-\frac{Z}{Q}} + \frac{D_2 Z}{R^2}e^{-\frac{Z^2}{2R^2}} + D_3 Z e^{-\frac{Z^2}{2}}\right) \tag{5.37}$$

with

$$Z = \frac{r}{2\sqrt{m_0}} \quad , \quad \alpha_2 = \frac{m_2}{\sqrt{m_0 m_4}} \quad , \quad m_1^+ = 2\int_0^\infty \omega K_{\mathcal{T}^l}\left(d,x_p,s_i;\omega\right)d\omega, \tag{5.38}$$

$$x_m = \frac{m_1^+}{m_0}\sqrt{\frac{m_2}{m_4}} \quad , \quad D_1 = \frac{2(x_m-\alpha_2^2)}{1+\alpha_2^2} \quad , \quad R = \frac{\alpha_2 - x_m - D_1^2}{1-\alpha_2-D_1+D_1^2}, \tag{5.39}$$

$$D_2 = \frac{1-\alpha_2-D_1+D_1^2}{1-R} \quad , \quad D_3 = 1-D_1-D_2 \quad , \quad Q = \frac{1.25(\alpha_2-D_3-(D_2R)}{D_1}. \tag{5.40}$$

Using this approximation, if follows that:

$$\mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}\left(d,x_{d_2},x_p,s_i\right)\right] = \frac{\Delta T}{x_{d_2}}\frac{1}{2\pi}\sqrt{\frac{m_4}{m_2}}\left(2\sqrt{m_0}\right)^m\left(D_1 Q^m\Gamma(1+m)\right.$$
$$\left. + \sqrt{2}^m\Gamma\left(1+\frac{m}{2}\right)\left(D_2|R|^m+D_3\right)\right) \tag{5.41}$$

where Γ is the gamma function.

We will denote $E_{\mathcal{T}^l}(d, x_p, s_i)$ the quantity such that:

$$\mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}(d, x_{d_2}, x_p, s_i)\right] = \frac{\Delta T}{x_{d_2}} E_{\mathcal{T}^l}(d, x_p, s_i)\,. \tag{5.42}$$

Besides, under the assumption that the instantaneous damage process $\mathcal{D}_i^l(d, x_{d_2}, x_p, s_i; .)$ is stationary, we obtain:

$$\mathbb{E}\left[\int_{I_i} \mathcal{D}_i^l(d, x_{d_2}, x_p, s_i; t)\, dt\right] = \Delta T \mathbb{E}\left[\mathcal{D}_i^l(d, x_{d_2}, x_p, s_i; 0)\right] = \mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}(d, x_{d_2}, x_p, s_i)\right] \tag{5.43}$$

and it follows:

$$\mathbb{E}\left[\mathcal{D}_i^l(d, x_{d_2}, x_p, s_i; 0)\right] = \frac{1}{x_{d_2}} E_{\mathcal{T}^l}(d, x_p, s_i)\,. \tag{5.44}$$

A comparison of the time-domain and the spectral approaches have been carried out on the nominal configuration of the studied case and the results are available in section H.2 of appendix H. It shows that the Dirlik method provides a good approximation of the fatigue which is why this method is chosen in this thesis. We refer the reader to Rychlik (1996) for a more theoretical background on the damage due to continuous stochastic processes irregular or smooth (e.g. stationary Gaussian processes with finite intensity of local maxima).

4 Problem formulation

We now present in more details the cost function and the uncertainties of the FOWT optimization problem introduced in chapter 1. We seek to minimize the manufacturing cost of the mooring lines under constraints involving the surge of the platform, the tension, and the fatigue at the top of each line over the considered period. These constraints take into account several sources of uncertainties and therefore are expressed as failure probabilities.

4.1 Cost function and design variables

In the considered optimization problem, the cost of materials used to manufacture the mooring lines depends on three design variables:

- the length d_1 of the mooring line that can be added to, or deducted from, the nominal mooring length of 841.56 meters. This variable takes values in $[-0.5, 2]$ (in m);
- the mass per unit length $d_2 \in [70, 180]$ (in kg/m);
- the position of the fairlead d_3 (i.e. the location where each mooring line is attached to the matching column). This variable can vary from 0 (which corresponds to a connection at the bottom of the columns) to 1 (top of the columns).

The nominal configuration proposed in Robertson et al. (2014a) is $d_1 = 0$, $d_2 = 113.35$ and $d_3 = 3/16$.

The ranges of variation of the design variables have been determined by expert opinion. We denote $d = (d_1, d_2, d_3)$ and the design space $\Omega_d = [-0.5, 2] \times [70, 180] \times [0, 1]$.

Remark 5.3. The range of variation for d_1 may seem small (considering that the nominal length is 841.56 meters). But we cannot significantly reduce the length of the mooring lines since the tension in the lines at the nominal configuration is already high. Moreover, increasing d_1 quickly activates the surge constraint. Therefore, we only allow a maximum of 2 meters to be added.

The cost function is defined as follows:

$$\begin{array}{rccc} cost & : & \Omega_d = [-0.5, 2] \times [70, 180] \times [0, 1] & \rightarrow & \mathbb{R}^+ \\ & & (d_1, d_2, d_3) & \mapsto & L(d_1, d_3) \times c(d_2) \end{array} \tag{5.45}$$

where:

- $L(d_1, d_3)$ is the length in meters of the mooring line:

$$L(d_1, d_3) = L_{init}(d_3) + d_1, \tag{5.46}$$

  where the function $L_{init}(d_3)$ represents the initial length of the mooring line. When the platform is at rest, each line is composed of a part lying on the seabed and a part linking the fairlead to the ground contact point. Initially, the first part measures 601.76m while the length of the latter depends on d_3 and is computed with the catenary equation (Chakrabarti (1987)). The initial mooring length $L_{init}(d_3)$ is the sum of these lengths to which we add d_1 meters to obtain $L(d_1, d_3)$;

- $c(d_2)$ is the cost in euros of one meter of mooring line:

$$c(d_2) = \begin{cases} a_1(a_2 \times d_2 - a_3) - a_4 & \text{if } a_2 \times d_2 - a_3 < b \\ a_5(a_2 \times d_2 - a_3) - a_6 & \text{if } a_2 \times d_2 - a_3 > b \end{cases} \tag{5.47}$$

  where $a_1, a_2, a_3, a_4, a_5, a_6$, and b are confidential constants.

4.2 Definition of the model uncertainties and other parameters of the problem

We consider that the values of the parameters x_{p_1}, x_{p_2}, x_{p_3}, and x_{d_2} introduced in section 5.2 are uncertain. These uncertainties are represented by random variables denoted respectively X_{p_1}, X_{p_2}, X_{p_3}, X_{d_2} and are defined as follows:

- the uncertainties on the wave heading is represented by the random variable X_{p_1} uniformly distributed between plus and minus 10° around the wind turbine axis;

- to estimate the quadratic viscous drag coefficients for the surge and pitch, decay tests are performed numerically with different codes in Robertson et al. (2020) considering the OC5-DeepCwind semi-submersible. The uncertainties resulting from these simulations are transposed to the floater under study in this thesis to obtain the distribution of X_{p_2} and X_{p_3} which follow uniform laws respectively on $[10^5, 10^6]$ (in N.s^2.m^{-2}) and $[3 \times 10^{10}, 7 \times 10^{10}]$ (in N.m.s^2.rad^{-2});

- we consider uncertainties on the fatigue law parameter K_D introduced in section 5.3.1 which are represented by a random variable with distribution $\mathcal{N}\left(\log(830), 0.8^2\right)$ whose parameters are chosen after the Stiff's fatigue curve for chain (Rossi (2005)) and 0.8 is equal to twice the difference between the logarithm of the mean curve intercept and the logarithm of the intercept of the fatigue design curve. Therefore, the uncertainties on x_{d_2} follow a distribution $\mathcal{LN}(\log(830 \times BL(d_2), 0.8^2)$ since $x_{d_2} = K_D BL(d_2)^m$.

- for safety, the resistance threshold involved in the fatigue constraints is also considered uncertain and is represented by a random variable X_R with $X_R \sim \mathcal{LN}(\mu_R, \sigma_R^2)$, $\mu_R = 1$ and $\sigma_R = 0.3$ (Leira et al. (2005)).

Moreover, the random variables X_{p_1}, X_{p_2}, X_{p_3}, X_{d_2}, and X_R are independent and we denote X_p the random vector $(X_{p_1}, X_{p_2}, X_{p_3})$. Their distributions are gathered in Table 5.2.

Uncertainty	Distribution	Unit
X_{p_1} (wave heading)	$\mathcal{U}[-10, 10]$	degree
X_{p_2} (surge quadratic viscous drag coefficient)	$\mathcal{U}\left[10^5, 10^6\right]$	$\mathrm{N.s^2.m^{-2}}$
X_{p_3} (pitch quadratic viscous drag coefficient)	$\mathcal{U}\left[3 \times 10^{10}; 7 \times 10^{10}\right]$	$\mathrm{N.m.s^2.rad^{-2}}$
X_{d_2} (fatigue law parameter)	$\mathcal{LN}\left(\mu_{d_2}, 0.8^2\right)$	
X_R (resistance threshold)	$\mathcal{LN}\left(1, 0.3^2\right)$	

Table 5.2: Distributions of X_{p_1}, X_{p_2}, X_{p_3}, X_{d_2}, and X_R

Finally, we display in Table 5.3, the other parameters involved in the optimization problem including the threshold $\mathcal{S}_{\max}$ appearing in the surge constraint, which is set to a conservative value of 5% of the water depth.

Parameter	Value
T (duration of the studied period)	$365 \times 24 \times 3600$(s)
ΔT (duration of each subinterval of time I_i)	3×3600 (s)
n_T (number of subintervals I_i)	2920
n_s (number of sea states)	7
$\mathcal{S}_{\max}$ (surge threshold)	10 (m)

Table 5.3: Parameters of the problem

4.3 Formulation of the FOWT time-dependent RBDO problem

As explained in chapter 1, constraints are imposed on the surge, the tension and the fatigue in the lines to restrict the movements of the platform and to avoid risky configurations or rupture of the mooring lines. The resulting time-dependent RBDO problem is formulated as follows:

$$\begin{aligned}
&\min_{d \in \Omega_d} cost(d) \quad \text{such that} \\
&\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} |\mathcal{S}(d, X_p; t)| > \mathcal{S}_{\max}\right) && < 10^{-4} \\
&\mathbb{P}_{X_p,\eta}\left(\min_{t\in[0,T]} \mathcal{T}^l(d, X_p; t) < 0\right) && < 10^{-4} \quad , \ l = 1, 2, 3 \\
&\mathbb{P}_{X_{d_2},X_p,X_R,\eta}\left(\int_0^T \mathcal{D}^l(d, X_{d_2}, X_p; t)\, dt > X_R\right) && < 10^{-4} \quad , \ l = 1, 2, 3.
\end{aligned} \tag{5.48}$$

We recall that the failure probability thresholds of 10^{-4} are imposed by international standards (Det Norske Veritas (2013)).

Using the definition introduced in section 2.1, the surge and tension constraints are thus extreme-based constraints while the fatigue constraints are integral-based.

We will show in the next section that under the assumptions introduced in this chapter, the constraint reformulation described in chapter 2 can be applied to problem (5.48). This will allow us to solve in section 5.6 the reformulated problem with AK-ECO. Thus, no realization of the surge, the tension or the instantaneous damage processes will be necessary to solve the problem under study.

5 Reformulation of the time-dependent constraints

The failure probabilities appearing in the surge, the tension and the damage constraints on line l ($l = 1, 2, 3$) and at the design point d will be denoted respectively $p_{\mathcal{S}}(d)$, $p_{\mathcal{T}^l}(d)$, and $p_{\mathcal{D}^l}(d)$.

5.1 Properties of the surge and tension processes

As mentioned in section 5.2.5, under the assumptions outlined in section 5.2.6, for fixed values of x_{p_1}, x_{p_2}, and x_{p_3}, the surge process is a piece-wise stationary Gaussian process. Over each interval I_i, it is equal to the process $\mathcal{S}_i(d, x_p, s_i; .)$ with mean $\mu_{\mathcal{S}}(d, x_p, s_i)$ and spectral moment of order n denoted $m_{\mathcal{S},n}(d, x_p, s_i)$.

In the same way, the tension process at the top of the line l over I_i is represented by a Gaussian process $\mathcal{T}_i^l(d, x_p, s_i; .)$ with mean $\mu_{\mathcal{T}^l}(d, x_p, s_i)$ and spectral moment of order n denoted $m_{\mathcal{T}^l,n}(d, x_p, s_i)$ (see section 5.2.5).

5.2 Reformulation of the extreme-based constraints

Reformulation of the surge constraint

The failure probability involved in the surge constraint is written as

$$p_{\mathcal{S}}(d) = \mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} |\mathcal{S}(d, X_p; t)| > \mathcal{S}_{\max}\right). \tag{5.49}$$

The probability $\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} \mathcal{S}(d, X_p; t) > \mathcal{S}_{\max}\right)$ corresponds to the probability of a horizontal shift of the platform greater than $S_{\max}$ meters in the opposite direction of the wind propagation. Thus, we consider this probability negligible. Indeed, due to the wind loads, the mean of the surge process is negative and besides, its variance is small. Therefore, $p_{\mathcal{S}}(d)$ can be approximated as:

$$\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} |\mathcal{S}(d, X_p; t)| > \mathcal{S}_{\max}\right) \simeq \mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} -\mathcal{S}(d, X_p; t) > \mathcal{S}_{\max}\right). \tag{5.50}$$

This approximation is justified in more details in appendix I.

Moreover, we prove in appendix J that the surge process meets the sufficient conditions to apply the reformulation of extreme-based constraints described in section 2.3.2. Therefore,

considering that ΔT is large enough, we can approximate the right-hand side of equation (5.50) as follows:

$$p_{\mathcal{S}}(d) \simeq \mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(a_{Tp^j}(d,X_p,s^j)^2 - \frac{a_{Tp^j}(d,X_p,s^j)\left(\mathcal{S}_{\max}+\mu_{\mathcal{S}}\left(d,X_p,s^j\right)\right)}{\sqrt{m_{\mathcal{S},0}\left(d,X_p,s^j\right)}}\right)\right)\right] \tag{5.51}$$

with $a_{Tp^j}(d,x_p,s^j) = \sqrt{2\log\left(\frac{Tp^j}{2\pi}\sqrt{\frac{m_{\mathcal{S},2}(d,x_p,s^j)}{m_{\mathcal{S},0}(d,x_p,s^j)}}\right)}$.

Reformulation of the tension constraints

The failure probability involved in the tension constraint of the l-th mooring line is defined as:

$$p_{\mathcal{T}^l}(d) = \mathbb{P}_{X_p,\eta}\left(\min_{t\in[0,T]}\mathcal{T}^l(d,X_p;t) < 0\right) = \mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]} -\mathcal{T}^l(d,X_p;t) > 0\right). \tag{5.52}$$

We prove in appendix J that the tension process meets all the conditions to reformulate the constraint. For ΔT large enough, the failure probability can thus be approximated as follows:

$$p_{\mathcal{T}^l}(d) \simeq \mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(b^l_{Tp^j}(d,X_p,s^j)^2 - \frac{b^l_{Tp^j}(d,X_p,s^j)\mu_{\mathcal{T}^l}\left(d,X_p,s^j\right)}{\sqrt{m_{\mathcal{T}^l,0}\left(d,X_p,s^j\right)}}\right)\right)\right] \tag{5.53}$$

with $b^l_{Tp^j}(d,x_p,s^j) = \sqrt{2\log\left(\frac{Tp^j}{2\pi}\sqrt{\frac{m_{\mathcal{T}^l,2}(d,x_p,s^j)}{m_{\mathcal{T}^l,0}(d,x_p,s^j)}}\right)}$.

5.3 Reformulation of the integral-based constraints

For a fixed sea state, it is commonly accepted that the instantaneous damage process is ergodic. Thus, we will assume that for every design point d, parameters x_p and x_{d_2}, state s^j and line l, the instantaneous process $\mathcal{D}^l_i\left(d,x_p,x_{d_2},s^j\right)$ is ergodic. We can thus apply proposition 2.1 allowing to reformulate integral-based constraints and we obtain the following approximation:

$$p_{\mathcal{D}^l}(d) \simeq \mathbb{P}_{X_{d_2},X_p,X_R}\left(X_R \leq T\sum_{j=1}^{7} p^j \mathbb{E}_{\eta_1}\left[\mathcal{D}^l_1\left(d,X_{d_2},X_p,s^j;0\right)\right]\right). \tag{5.54}$$

The expectation $\mathbb{E}_{\eta_1}\left[\mathcal{D}^l_1\left(d,x_{d_2},x_p,s^j;0\right)\right]$ can be approached using the Dirlik method described in section 5.3.2 as:

$$\mathbb{E}_{\eta_1}\left[\mathcal{D}^l_1\left(d,x_{d_2},x_p,s^j;0\right)\right] \simeq \frac{1}{x_{d_2}} E_{\mathcal{T}^l}\left(d,x_p,s^j\right) \tag{5.55}$$

where $E_{\mathcal{T}^l}\left(d,x_p,s^j\right)$ is a quantity which depends on the spectral moments of the tension process $\mathcal{T}^l_1\left(d,x_p,s^j;.\right)$.

Calculation of the conditional probabilities considering the uncertainties on X_{d_2} and X_R

We can simplify the estimation of the fatigue failure probabilities by calculating analytically the probabilities on X_{d_2} and X_R as follows:

$$\mathbb{P}_{X_{d_2},X_p,X_R}\left(X_R \leq T\sum_{j=1}^{7} p^j \frac{1}{X_{d_2}} E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right) \tag{5.56}$$

$$=\mathbb{P}_{X_{d_2},X_p,X_R}\left(X_R X_{d_2} \leq T\sum_{j=1}^{7} p^j E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right) \tag{5.57}$$

$$=\mathbb{P}_{X_{d_2},X_p,X_R}\left(\log(X_R)+\log(X_{d_2}) \leq \log\left(T\sum_{j=1}^{7} p^j E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right)\right) \tag{5.58}$$

$$=\mathbb{E}_{X_p}\left[\mathbb{P}_{X_{d_2},X_R|X_p}\left(\log(X_R)+\log(X_{d_2}) \leq \log\left(T\sum_{j=1}^{7} p^j E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right)\right)\right] \tag{5.59}$$

$$=\mathbb{E}_{X_p}\left[\Phi\left(\frac{\log\left(T\sum_{j=1}^{7} p^j E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right)-(\mu_R+\mu_{d_2})}{\sqrt{\sigma_R^2+\sigma_{d_2}^2}}\right)\right]. \tag{5.60}$$

We used in the last equation that X_R and X_{d_2} are both log-normal random variables with $X_R \sim \mathcal{LN}(\mu_R,\sigma_R^2)$ and $X_{d_2} \sim \mathcal{LN}(\mu_{d_2},\sigma_{d_2}^2)$.

5.4 Reformulated problem

The properties of the studied problem and the reformulation procedure enable to reformulate the FOWT time-dependent RBDO problem (5.48) as follows:

$$\begin{aligned}
&\min_{d\in\Omega_d} cost(d) \quad \text{such that}\\
&\mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(a_{Tp^j}(d,X_p,s^j)^2-\frac{a_{Tp^j}(d,X_p,s^j)\left(\mathcal{S}_{\max}+\mu_{\mathcal{S}}\left(d,X_p,s^j\right)\right)}{\sqrt{m_{\mathcal{S},0}\left(d,X_p,s^j\right)}}\right)\right)\right] < 10^{-4}\\
&\mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(b^l_{Tp^j}(d,X_p,s^j)^2-\frac{b^l_{Tp^j}(d,X_p,s^j)\mu_{\mathcal{T}^l}\left(d,X_p,s^j\right)}{\sqrt{m_{\mathcal{T}^l,0}\left(d,X_p,s^j\right)}}\right)\right)\right] < 10^{-4}\ ,\ l=1,2,3\\
&\mathbb{E}_{X_p}\left[\Phi\left(\frac{\log\left(T\sum_{j=1}^{7} p^j E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right)-(\mu_R+\mu_{d_2})}{\sqrt{\sigma_R^2+\sigma_{d_2}^2}}\right)\right] < 10^{-4}\ ,\ l=1,2,3.
\end{aligned} \tag{5.61}$$

We recall that to evaluate the quantities $a_{Tp^j}\left(d,x_p,s^j\right)$, $\mu_{\mathcal{S}}\left(d,x_p,s^j\right)$, $b^l_{Tp^j}\left(d,x_p,s^j\right)$, $\mu_{\mathcal{T}^l}\left(d,x_p,s^j\right)$, and $E_{\mathcal{T}^l}\left(d,x_p,s^j\right)$, the spectral moments of the processes $\mathcal{S}_1\left(d,x_p,s^j,.\right)$ and $\mathcal{T}_1^l\left(d,x_p,s^j,.\right)$ are required. All these spectral moments are output of a single simulation. More precisely, this simulation provides the RAOs and the means of the processes from which the spectral moments and the Dirlik fatigue approximation are computed as explained in sections 5.2.5 and 5.3.2.

It takes about 30 seconds to run the code providing these spectral moments. It is still considered too computationally expensive since one estimation of the constraints with Monte

Carlo and a sample size of 30000 points would require several days ($30000\times7\times0.5/60/24 > 72$ days). Even a variance reduction technique would be too expensive. Indeed, assuming that a sample of only 100 points would be enough, one evaluation of one constraint would still take about 6 hours ($100 \times 7 \times 0.5/60 \simeq 5.8$). Thus, considering 7 constraints and assuming that the optimizer needs about 50 evaluations of each constraint to converge, it would take more than 85 days ($100 \times 7 \times 0.5/60 \times 7 \times 50 > 85$) to solve the reformulated problem.

We therefore solve the reformulated problem with AK-ECO in the next section.

6 Resolution of the reformulated problem

We solve in this section the reformulated FOWT problem (5.61) with AK-ECO and three comparison methods.

6.1 Implementation of AK-ECO

Definition of the expensive functions

The expensive functions involved in the reformulated constraints are $M_{\mathcal{S}}$, $M_{\mathcal{T}^l}$, and $M_{\mathcal{D}^l}$ ($l = 1, 2, 3$) with:

$$M_{\mathcal{S}}\left(d, x_p, s^j\right) = a_{Tp^j}(d, x_p, s^j)^2 - \frac{a_{Tp^j}(d, x_p, s^j)\left(\mathcal{S}_{\max} + \mu_{\mathcal{S}}\left(d, x_p, s^j\right)\right)}{\sqrt{m_{\mathcal{S},0}\left(d, x_p, s^j\right)}} \tag{5.62}$$

$$M_{\mathcal{T}^l}\left(d, x_p, s^j\right) = b^l_{Tp^j}(d, x_p, s^j)^2 - \frac{b^l_{Tp^j}(d, x_p, s^j)\mu_{\mathcal{T}^l}\left(d, x_p, s^j\right)}{\sqrt{m_{\mathcal{T}^l,0}\left(d, x_p, s^j\right)}}, \qquad l = 1, 2, 3 \tag{5.63}$$

$$M_{\mathcal{D}^l}\left(d, x_p, s^j\right) = E_{\mathcal{T}^l}\left(d, x_p, s^j\right), \qquad l = 1, 2, 3. \tag{5.64}$$

Indeed, as mentioned in the previous section, evaluating all of these functions for a set of inputs d, x_p, and s^j requires a call to a simulator which provides the RAOs and the means of the surge and tension processes.

Definition of the augmented space and dimension reduction

We observe that the functions $M_{\mathcal{S}}$, $M_{\mathcal{T}^l}$ and $M_{\mathcal{D}^l}$ ($l = 1, 2, 3$) are defined on the same input space. We thus can consider a common augmented space for all the constraints. As explained in the introduction of AK-ECO, this enables us to create a single DoE of the augmented space and to use the outputs of every simulation to calibrate and enrich all the metamodels during AK-ECO.

A first option for the definition of the augmented space is the following:

$$\Omega^{aug} = \Omega_d \times \Omega_p^{aug} \times \Omega_{h_s} \times \Omega_{t_p} \times \Omega_u \tag{5.65}$$

where Ω_d is the design space and, denoting $F^{-1}_{p_i}$ the quantile function of X_{p_i} ($i = 1, 2, 3$),

$$\Omega_p^{aug} = \prod_{i=1}^{3}\left[F^{-1}_{p_i}(\alpha), F^{-1}_{p_i}(1-\alpha)\right] \tag{5.66}$$

with α a degree of confidence chosen by the user and

$$\Omega_{h_s} = \{h_s^j,\, j = 1, \cdots, 7\}\ ,\ \Omega_{t_p} = \{t_p^j,\, j = 1, \cdots, 7\}\ ,\ \Omega_u = \{u^j,\, j = 1, \cdots, 7\}. \tag{5.67}$$

This augmented space Ω_{aug} is therefore of dimension 9: $\Omega_d \subset \mathbb{R}^3$, $\Omega_p^{aug} \subset \mathbb{R}^3$, and Ω_{h_s}, Ω_{t_p}, and Ω_u are subsets of $\mathbb{R}$.

However, we recall that only 7 sea states $s^1, \cdots, s^7$ are considered in our problem and we can observe in Table 5.1 that these sea states are lumped such that h_s determines t_p and u. Thus, it seems natural to consider the following functions:

$$M_{\mathcal{S}}^{bis}\left(d, x_p, h_s^j\right) = M_{\mathcal{S}}\left(d, x_p, s^j\right) \tag{5.68}$$

$$M_{\mathcal{T}^l}^{bis}\left(d, x_p, h_s^j\right) = M_{\mathcal{T}^l}\left(d, x_p, s^j\right), \quad l = 1, 2, 3 \tag{5.69}$$

$$M_{\mathcal{D}^l}^{bis}\left(d, x_p, h_s^j\right) = M_{\mathcal{D}^l}\left(d, x_p, s^j\right), \quad l = 1, 2, 3 \tag{5.70}$$

and the augmented space $\Omega_{bis}^{aug} = \Omega_d \times \Omega_p^{aug} \times \Omega_{h_s}$. This permits to reduce the dimension of the augmented space from 9 to 7.

Remark 5.4. This dimension reduction of the augmented space is reasonable because the 7 sea states considered are well ordered. In the case where we would consider a full scatter diagram, this approach would not make sense and it would be more appropriate to conserve the augmented space in dimension 9.

To calibrate the first kriging models of $M_{\mathcal{S}}^{bis}$, $M_{\mathcal{T}^l}^{bis}$, $M_{\mathcal{D}^l}^{bis}$ $(l = 1, 2, 3)$, a LHS maximin of 60 points of $\Omega_d \times \Omega_p^{aug}$ is carried out. A uniform sample of 60 points of Ω_{h_s} is concatenated to the LHS which provides the initial DoE of the augmented space.

Enrichment procedure for the fatigue constraints

Except that no uncertainties are considered on the design variables, the reformulated extreme-based constraints of the FOWT problem have exactly the same formulation as the one given in section 2.3.2. The enrichment procedure can thus be implemented as described in section 3.2.2.

For the reformulated fatigue constraints however, the analytic calculation of the probabilities on X_{d_2} and X_R presented in section 5.5.3 leads to a reformulation of these constraints different from the reformulation of the integral-based constraint described in section 2.3.3. Indeed, the reformulated fatigue constraints of the FOWT problem are given by:

$$\mathbb{E}_{X_p}\left[\Phi\left(\frac{\log\left(T\sum_{j=1}^{7} p^j M_{\mathcal{D}^l}\left(d, x_p, s^j\right)\right) - \left(\mu_R + \mu_{d_2}\right)}{\sqrt{\sigma_R^2 + \sigma_{d_2}^2}}\right)\right] < 10^{-4} \tag{5.71}$$

while the reformulated integral-based constraint in the introduction of AK-ECO was presented as follows:

$$\mathbb{E}_{X_d, X_p}\left[F_{r_{\mathrm{I}}}\left(\sum_{j=1}^{n_s} T p^j \left(M_{\mathrm{I}}\left(X_d, X_p, s^j\right)\right)\right)\right] < p_s. \tag{5.72}$$

This difference requires a slight modification of the enrichment procedure described in section 3.2.2. Indeed, as indicated in the perspectives of chapter 3, the AK-ECO procedure adapts easily to any constraint being written:

$$\mathbb{E}_{X_d,X_p,X_r}\left[F_0\left(\sum_{i=1}^{n_s} F_i\left(M\left(X_d,X_p,X_r\right)\right)\right)\right] < p_s, \tag{5.73}$$

if F_0 and F_i $(i = 1,\cdots,n_s)$ are monotonic functions and M a function whose evaluation requires a call to an expensive simulator.

6.2 The reference method: MC+K1600

Since it would be too expensive to solve the reformulated problem using Monte Carlo directly with the costly functions $M_{\mathcal{S}}$, $M_{\mathcal{T}^l}$, and $M_{\mathcal{D}^l}$ $(l = 1,2,3)$, the method providing the reference results consist in calibrating kriging models of the costly functions from a large DoE.

To build these metamodels, a LHS maximin of 1600 points of Ω_{bis}^{aug} is created following the procedure for the initial DoE of AK-ECO. Kriging models of $M_{\mathcal{S}}^{bis}$, $M_{\mathcal{T}^l}^{bis}$, and $M_{\mathcal{D}^l}^{bis}$ $(l = 1,2,3)$ are then calibrated. The reformulated problem is then solved with an optimization algorithm and the constraints are evaluated with Monte Carlo on the kriging models during each iteration. No enrichment of the metamodels is performed in this method. Thus, 1600 calls to the expensive simulator are needed for the resolution of the reformulated problem. We will denote this method MC+K1600.

6.3 Implementation of SORA and the Stieng method

The resolution of the reformulated problem with RIA and PMA being too computationally expensive, we will only present the results obtained with SORA and Stieng.

We recall that to apply these approaches, one must consider a RBDO problem with constraints expressed in terms of failure probabilities. We thus reformulate the problem (5.61) as follows:

$$\begin{aligned}
&\min_{d\in\Omega_d} cost(d) \quad \text{such that}\\
&\mathbb{P}_{X_p,X_\epsilon}\left(X_\epsilon \le \sum_{j=1}^{7}\exp\left(a_{Tp^j}\left(d,X_p,s^j\right)^2 - \frac{a_{Tp^j}\left(d,X_p,s^j\right)\left(\mathcal{S}_{\max}+\mu_{\mathcal{S}}\left(d,X_p,s^j\right)\right)}{\sqrt{m_{\mathcal{S},0}\left(d,X_p,s^j\right)}}\right)\right) < 10^{-4}\\
&\mathbb{P}_{X_p,X_\epsilon}\left(X_\epsilon \le \sum_{j=1}^{7}\exp\left(b^l_{Tp^j}\left(d,X_p,s^j\right)^2 - \frac{b^l_{Tp^j}\left(d,X_p,s^j\right)\mu_{\mathcal{T}^l}\left(d,X_p,s^j\right)}{\sqrt{m_{\mathcal{T}^l,0}\left(d,X_p,s^j\right)}}\right)\right) < 10^{-4}\ ,\ l=1,2,3\\
&\mathbb{P}_{X_p,X_r}\left(X_r \le \frac{\log\left(T\sum_{j=1}^{7} p^j E_{\mathcal{T}^l}\left(d,X_p,s^j\right)\right) - \left(\mu_R+\mu_{d_2}\right)}{\sqrt{\sigma_R^2+\sigma_{d_2}^2}}\right) < 10^{-4}\ ,\ l=1,2,3
\end{aligned} \tag{5.74}$$

where X_ϵ is a random variable with an exponential distribution of parameter 1 and $X_r \sim \mathcal{N}(0,1)$.

We draw the reader's attention to the fact X_r represents a random variable introduced artificially to obtain a formulation of the constraints with failure probabilities whereas X_R

is the random variable which models the uncertainties on the fatigue threshold.

We use SQP (Kraft (1988)) for solving the inverse reliability analyses in SORA and Stieng because we observed better performance of this approach compared to HMV (introduced in section 1.2.1) on the oscillator problem.

A Sobol sequence (Sobol' (1967)) of 30 points of the 3D output space of X_p is used to calibrate the kriging models of the first cycle of Stieng. For the second cycle, the size of the sequence is 90 and 300 for the next ones.

6.4 Numerical results

For AK-ECO, MC+K1600, and Stieng, a kriging with constant trend and an anisotropic 3/2-Matérn covariance kernel is used for every metamodel. At each calibration of a metamodel during AK-ECO and Stieng, its hyperparameters are updated. The points of the successive DoEs are normalized in $[0,1]$ and each hyperparameter is selected in $[10^{-5},10]$ using the multistart Truncated Newton Constrained solver implemented in OpenTURNS from 20 initial points. In AK-ECO and MC+K1600, the variable h_s is considered as continuous by the kriging model.

The optimization algorithm applied for each approach is COBYLA (Powell (1994)) since the derivative of the cost function is not known and this optimizer provides good results when solving the oscillator problem.

For AK-ECO, SORA and Stieng, the cycles of optimization stop if the stopping condition introduced in section 3.2.2 is satisfied for ϵ_d and ϵ_{cost} equal to 10^{-3}.

The same Monte Carlo sample of X_p of size 30000 is used in MC+K1600 and AK-ECO to estimate the reformulated failure probabilities.

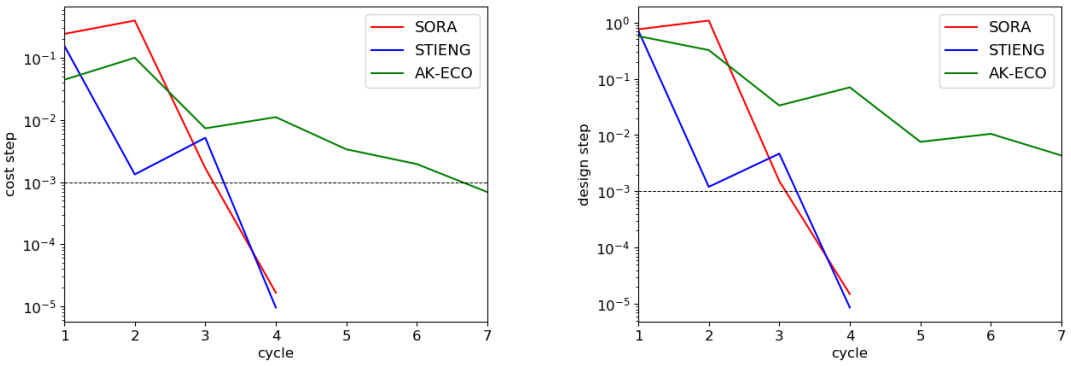
The results obtained by each approach, considering an initial design point at the center of the design space $(0.75, 125, 0.5)$ are given in Table 5.4. The design point obtained by each method is denoted d^{min} and $\overline{cost}(d^{min})$ is the cost function evaluated at this point normalized between $[0,1]$. The surge, tension, and fatigue failure probabilities at d^{min} obtained by each method have been evaluated with Monte Carlo and the kriging models of the MC+K1600 method. The results are denoted $p_{\mathcal{S}}^{K1600}\left(d^{min}\right)$, $p_{\mathcal{T}^l}^{K1600}\left(d^{min}\right)$ and $p_{\mathcal{D}^l}^{K1600}\left(d^{min}\right)$ $(l = 1,2,3)$. Finally, the number of simulations performed during the resolution of the studied problem with each method is denoted n_{call}. The execution time of every method is discussed in section F.2 of appendix F.

	MC+K1600	SORA	Stieng	AK-ECO
d^{min}	$(0.862, 108.86, 0)$	$(2, 141.22, 0)$	$(2, 140.58, 0)$	$(0.947, 109.52, 0)$
$\overline{cost(d^{min})}$	0.2818	0.5883	0.5819	0.2867
$p_{\mathcal{S}}^{K1600}(d^{min})$	1.0×10^{-4}	0	0	0.9×10^{-4}
$p_{\mathcal{T}^l}^{K1600}(d^{min})$ $(l = 1, 2, 3)$	0	0	0	0
$p_{\mathcal{D}^l}^{K1600}(d^{min})$ $(l = 1, 2)$	0	0	0	0
$p_{\mathcal{D}^3}^{K1600}(d^{min})$	1.0×10^{-4}	0.2×10^{-4}	0.2×10^{-4}	0.9×10^{-4}
n_{call}	1600	16394	5754	305

Table 5.4: Results of AK-ECO and the comparison methods

We observe that each method provides a reliable optimum. However, the design proposed by AK-ECO is much closer to the reference result obtained with MC+K1600 than the ones proposed by SORA and Stieng. This difference is due to the inverse reliability analyses performed during these methods which underestimate the reliability associated with the last constraint. This leads to a sub-optimal design point provided by these approaches. Furthermore, AK-ECO requires much less evaluations of the simulator since only 305 calls were needed (60 for the initial DoE and 245 for the enrichment procedure during the optimization cycles).

For SORA, Stieng and AK-ECO, Figure 5.6a shows the evolution over the consecutive cycles of optimization of $\left|\overline{cost(d^{k-1})} - \overline{cost(d^{k})}\right|$ where d^k is the design point obtained at the end of the k-th cycle. Figure 5.6b displays the evolution of $\log\left(\left\|\overline{d^k} - \overline{d^{k-1}}\right\|\right)$ where $\overline{d^k}$ refers to the normalization of d^k. For SORA and Stieng, it took 4 optimization cycles to converge whereas 7 cycles were necessary for AK-ECO. During every cycle of AK-ECO, 35 enrichments were performed.



(a) Evolution of $\log\left(\left|\overline{cost\,(d^{k-1})} - \overline{cost\,(d^{k})}\right|\right)$ (b) Evolution of $\log\left(\left\|\overline{d^k} - \overline{d^{k-1}}\right\|\right)$

Figure 5.6: Evolution of the stopping condition for SORA, Stieng and AK-ECO

At the end of each cycle of SORA, Stieng and AK-ECO, the failure probabilities at d^k

have been evaluated with MC and the kriging models used in MC+K1600. Only the estimations of the first failure probability and the last one are displayed in Figure 5.7. They are denoted respectively p_S and p_D3 in the legend of the figure. All the other failure probabilities were always null.

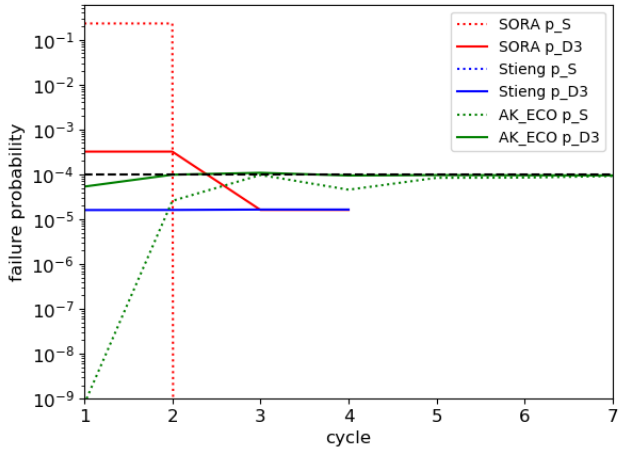


Figure 5.7: Evolution of the Monte Carlo estimation of the failure probabilities for SORA, Stieng and AK-ECO

We observe that the design points obtained at the end of every cycle of AK-ECO and Stieng are reliable while it took 3 cycles for SORA to provide a reliable design. The surge failure probability estimated at the design point proposed at the end of each cycle of Stieng is zero. Therefore, it does not appear on the figure.

To ensure that the minimum provided by MC+K1600 is a global minimum, the resolution of the problem has been repeated from a LHS of 9 points of the design space. The resulting designs are always the same.

The multistart resolution has also been conducted with AK-ECO and the results are displayed in section G.3 of appendix G. The same initial DoE of 60 points is considered for every resolution. The design points obtained by all the resolutions are close and the number of additional calls to the simulator varies from 193 to 476 with a mean number per resolution at 294.6.

To validate the estimation of the reformulated failure probabilities computed with the kriging models of MC+K1600, the reformulated failure probabilities have also been evaluated from the real expensive functions $M_{\mathcal{S}}$, $M_{\mathcal{T}^l}$, $M_{\mathcal{D}^l}$ ($l = 1, 2, 3$) and the same Monte Carlo sample of X_p of size 30000. Theses estimations are denoted $p_{\mathcal{S}}^{MC}$, $p_{\mathcal{T}^l}^{MC}$, and $p_{\mathcal{D}^l}^{MC}$ ($l = 1, 2, 3$). Due to the limited computing time budget, only the failure probabilities at the design points obtained with MC+K1600 and AK-ECO have been estimated and the results are shown in Table 5.5.

	MC+K1600	AK-ECO
$p_{\mathcal{S}}^{MC}\left(d^{min}\right)$	1.06×10^{-4}	0.97×10^{-4}
$p_{\mathcal{T}^l}^{MC}\left(d^{min}\right)$ $(l = 1, 2, 3)$	0	0
$p_{\mathcal{D}^l}^{MC}\left(d^{min}\right)$ $(l = 1, 2)$	0	0
$p_{\mathcal{D}^3}^{MC}\left(d^{min}\right)$	1.06×10^{-4}	0.98×10^{-4}

Table 5.5: Estimation of the reformulated failure probabilities at the design points obtained with MC+K1600 and AK-ECO

We observe from the estimations displayed in Table 5.5, that the threshold probability of 10^{-4} is slightly crossed at the design point obtained with MC+K1600 whereas AK-ECO stopped at a reliable point. The estimations of probabilities from the final kriging models of AK-ECO (calibrated from a DoE of 305 points) are therefore more accurate than those from the krigings used in MC+K1600 (calibrated from a DoE of 1600 points). These results also assure that even if the design points provided by SORA and Stieng are reliable, they are not optimal.

Remark 5.5. The reformulated failure probabilities involved in problem (5.61) are written as expectations of strictly positive random variables. However, in Table 5.4 and Table 5.5, for different design points, some of these expectations are estimated to be zero. We remark that these values correspond to numerical estimations, they are the results of numerical approximations which depend on the machine precision.

6.5 Interpretation of the results

We recall that the coordinates of the design point $d = (d_1, d_2, d_3)$ correspond to:

- the length d_1 of the mooring line that can be added to, or deducted from, the nominal mooring length of 841.56 meters;
- the mass per unit length d_2;
- the position of the fairlead d_3. It is equal to 0 if the connection is at the bottom of the columns and to 1 if it is at the top.

The nominal configuration proposed in Robertson et al. (2014a) is $(0, 113.35, 3/16)$ and the normalized cost function is equal to 0.316 at this point. The configuration proposed by AK-ECO is $(0.947, 109.52, 0)$ and the normalized cost function is equal to 0.287 at this point. Therefore, the optimal configuration is obtained when 0.947m are added to the mooring lines, with a mass per unit length of 109.52 kg/m and when the mooring lines are connected to the bottom of the columns. Moreover, considering the different sources of uncertainty of the studied problem, the nominal configuration is conservative since a reliable and less expensive configuration is obtained with AK-ECO.

This conclusion is however relative since only a very limited number of sea states has been considered. Moreover, the much larger Design Load Cases required by the standards, included that of Ultimate Limit States have not been taken into account. However, this study case has illustrated how the new approach proposed in this thesis could be used for a reliable estimation of the fatigue risk with a limited CPU cost.

7 Perspectives

In this chapter, we apply the methodology introduced in chapter 2 and 3 to the FOWT problem (5.48). First, the constraints are reformulated and then we solve the reformulated problem with AK-ECO. The implementation of this approach relies on modeling choices detailed in section 5.2.6 leading to Gaussian output processes. We discuss in this section the perspectives of work when a full scatter diagram is considered to represent the environmental conditions and when the modeling framework implies non-Gaussian output processes.

7.1 Concerning the number of sea states considered

The decision to model the environmental conditions with only 7 sea states was made to simplify the problem studied.

Besides the fact that the approximations described in section 5.5 become less precise when the number of sea states increases (since this precision increases w.r.t Tp^j), the reformulation of the constraints does not depend on the number of sea states considered. Regarding AK-ECO, our approach is already implemented to handle large number of sea states even if some improvements could be considered to decrease the execution time (we refer to section 3.4 for more details).

Moreover the simple dimension reduction proposed in section 5.6.1 is not reasonable when dealing with a full scatter diagram. If a dimension reduction was considered on the sea state variables, the latent variable method described in Zhang et al. (2020d) could for instance be investigated.

7.2 Considering non-constant wind forces

In this chapter, the loads on the structure due to the wind are replaced by constant forces. This simplification is made because the current version of the Deeplines$^{\text{TM}}$ software does not provide the transfer functions considering a turbulent wind. A perspective for the FOWT problem is therefore to implement a spectral approach enabling to consider non-constant wind loads as proposed in Pegalajar-Jurado et al. (2018).

7.3 When the output processes are not Gaussian

The linearization of the equation of motion described in section 5.2.3 and leading to Gaussian output processes is not always feasible. In particular, the outputs are no longer Gaussian when:

- considering extreme conditions where the linearization is too rough to be acceptable;
- second-order wave forces are considered.

In this case, the consequences are different for the extreme and the integral based constraints (i.e. for the surge or tension constraints and the fatigue constraints).

Concerning the extreme-based constraints

The reformulation of extreme-based constraints is presented in this thesis when a piecewise stationary Gaussian process is involved in the failure probability. If the process is no

longer Gaussian, theorem 2.1 cannot be applied. However, the reformulation is based, for the surge constraint, on the approximation of the conditional probability:

$$\mathbb{P}_\eta\left(\max_{t\in[0,Tp^j]} -\mathcal{S}_1\left(d, x_p, s^j; t\right) > \mathcal{S}_{\max}\right) \tag{5.75}$$

given x_p. Thus, if one is able to approximate this probability for any d, x_p, s^j, the reformulation of the constraint could be carried out and the AK-ECO enrichment procedure would need to be adapted to the resulting reformulated constraint. The same reasoning could be applied to the tension constraints as well. This means that extreme value theory for stationary non-Gaussian processes should be considered (Azaïs and Wschebor (2009)).

Concerning the integral-based constraints

Our approach concerning the fatigue constraints is mainly based on the ergodicity of the instantaneous damage process. The Gaussian property of the tension processes is only used to apply the Dirlik method and calculate the total damage over I_i:

$$\mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}\left(d, x_{d_2}, x_p, s_i\right)\right] \tag{5.76}$$

at (d, x_{d_2}, x_p, s_i) as explained in section 5.3.2.

This method enables us to quickly evaluate this expectation but our methodology can already be applied with any method being able to evaluate $\mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}\left(d, x_{d_2}, x_p, s_i\right)\right]$.

Concerning the error of the Dirlik method

Although the error of the Dirlik method is not large compared to the time-domain approach (see section H.2 in appendix H), it is possible to consider uncertainties on the Dirlik approximation $E_{\mathcal{T}^l}\left(d, x_p, s_i\right)$ introduced in section 5.3.2. A simple way to do so, is to add a new random variable denoted X_E, for example $X_E \sim \mathcal{U}[0.9, 1.1]$, and to consider $X_E \times E_{\mathcal{T}^l}\left(d, x_p, s_i\right)$ which would be in this case a random variable uniformly distributed between $0.9 \times E_{\mathcal{T}^l}\left(d, x_p, s_i\right)$ and $1.1 \times E_{\mathcal{T}^l}\left(d, x_p, s_i\right)$.

Conclusion and perspectives

8 Conclusion

The work of this thesis was motivated by the need for an efficient method to identify a configuration of a semi-submersible FOWT mooring system that minimizes its cost. This configuration must also restrict the movements of the structure, avoid compression in the mooring lines and withstand fatigue damage. In this thesis, the hydrodynamic forces acting on the structure are linearized. Moreover, constant wind forces are considered and the second-order wave effects are not taken into account. We consider that for FLS constraints, a frequency calculation provides a good proxy for the fatigue estimation. As it is usual in offshore engineering, we consider a random stationary Gaussian sea elevation process. To ensure the reliability of the solution, we also represent uncertainties on parameters of the model with random variables.

As a result, the constraints of the optimization problem depend on random variables and on piece-wise stationary Gaussian processes. Thus, we face an optimization problem with a fast-to-evaluate cost function and probabilistic constraints involving time-dependent processes. We classify these constraints into two categories: the extreme-based constraints and the integral-based constraints involving respectively the maximum and the integral over a period $[0, T]$ of a random process. Finally, the threshold failure probability imposed by international standards is 10^{-4}. Therefore, the probability estimations need to be very accurate to assure the reliability of the proposed configuration.

We have noticed from the state of the art presented in chapter 1, that the methods enabling to solve this problem require to sample realizations of the processes involved in the constraints. We decide that this approach is too computationally expensive in our case and, instead of solving the problem in its initial formulation, we introduce a two-step methodology.

First, we use the characteristics of the constraints to reformulate them in a way that is easier to estimate. The period $[0, T]$ is divided into subintervals of length ΔT during which the output processes involved in the constraints are Gaussian stationary processes. The reformulation consists in applying limit theorems which provide the asymptotic distributions of the maximum and the integral of these processes when ΔT is large. In chapter 2, we use the extreme value theory to reformulate the extreme-based constraints. The reformulation of integral-based constraints relies on the ergodicity of the process involved.

The reformulated problem is simpler to solve since it involves only constraints with time-independent expectations depending on a small number of random variables. To estimate these expectations, it is necessary to evaluate the spectral moments of the processes in-

volved in the initial constraints. Since the evaluation of these spectral moments requires the use of a time-consuming simulation, estimating these expectations is still expensive.

Therefore, we propose in chapter 3 a new efficient approach to solve the optimization problem with reformulated extreme and integral based constraints that relies on an adaptive kriging strategy. This method is called AK-ECO for Adaptive Kriging for Expectation Constraints Optimization. It starts with the calibration of initial metamodels of the expensive functions involved in the reformulated constraints. Then, optimization cycles are carried out. During each cycle, an enrichment of the metamodels is performed and the reformulated problem is then solved starting from the design point obtained at the previous cycle. The expectations in the constraints are estimated by using Monte Carlo simulations and by replacing the expensive functions by their enriched metamodels.

This two-step methodology combining the constraints reformulation and AK-ECO is first applied to an academic problem involving a harmonic oscillator and presenting all the characteristics of the FOWT problem in chapters 2 and 3 and then to the FOWT problem in chapter 5. For the latter, our approach provides a configuration that is less expensive than the nominal configuration and that satisfies the reliability constraints. In both cases, AK-ECO is compared with state-of-the-art algorithms and proves to be very efficient.

Several features of this method contribute to this performance:

- first, the metamodel strategy is adapted to the reformulated constraints. Indeed, the reformulated constraints are written as:

$$\mathbb{E}_{X_d,X_p,X_r}\left[F_0\left(\sum_{i=1}^{n_s} F_i\left(M\left(X_d,X_p,X_r\right)\right)\right)\right] < p_s, \tag{5.77}$$

  where F_0 and F_i $(i = 1,\cdots,n_s)$ are monotonic functions, M is a function whose evaluation requires a call to an expensive simulator, and p_s is the failure probability threshold. Equation (5.77) corresponds to the general reformulation introduced in chapter 2 where X_d, X_p, and X_r are random vectors representing respectively the uncertainties on the design variables, on model parameters, and on thresholds. In the FOWT problem, n_s corresponds to the number of sea states considered and we only consider uncertainties on model parameters.

  In AK-ECO, the metamodel is built to replace M. Hence, one enrichment of the metamodel corresponds to one call to the simulator. The approaches of the literature are based on a formulation of the constraints involving a failure probability written as:

$$\mathbb{P}_{X_d,X_p,X_r}\left(g\left(X_d,X_p,X_r\right) < 0\right) < p_s. \tag{5.78}$$

  The expectation of equation (5.77) can be written as a probability. However, with these approaches n_s simulations are needed for one evaluation of g;

- in the FOWT problem, each simulation provides information about all the costly functions involved in each constraint. Hence, a common DoE is shared by the metamodels in AK-ECO and each simulation is used to enrich all of them.

To improve the accuracy of the metamodels used during the first cycle of AK-ECO, a procedure of global enrichment (GE) is introduced in chapter 4. Contrary to the enrichments performed during the cycles of AK-ECO which aim at improving the estimation of the reformulated constraints only locally, the global enrichment improves the metamodels

such that the boundary of the feasible domain over the entire design space is accurately predicted. The benefits of this global enrichment have been demonstrated on the oscillator problem.

Moreover, a new sensitivity analysis is proposed in 4 to provide a better understanding of the influence of the input parameter uncertainties on the constraints and is called RBDO-oriented GSA. Since the impact of these uncertainties can vary depending on the design point considered, a set of design points is first selected in a sub-domain (called critical domain) of the design space containing the feasible domain boundary. For each selected design point, the sensitivity of the constraints to each input random variable is measured with Sobol indices. We then aggregate the information of these sensitivity analyses with the introduction of cumulative indices. To reduce the number of calls to the simulator, it is recommended to perform the RBDO-oriented GSA after the global enrichment phase with the enriched kriging models (once the metamodels enriched, the GSA method does not require any additional expensive simulation). This approach is applied to the oscillator problem to reveal that all the input random variables involved have a strong influence on the constraints.

9 Perspectives

9.1 Concerning the reformulation of the constraints

The approximation error made with the reformulation of the extreme and integral based failure probabilities is discussed in appendix. However, we have not conducted numerical experiments to compare the values of the initial failure probabilities and of the reformulated ones. The results of this analysis for the oscillator and the FOWT problems would provide a better understanding on how the error varies with respect to the different parameters of the problem and in particular w.r.t ΔT.

Although the reformulation of the constraints is introduced primarily to solve the FOWT problem, its presentation in chapter 2 is formulated in such a way that it can be applied to many time-dependent failure probabilities involving maximums and integrals of piece-wise stationary Gaussian processes.

Regarding the reformulation of extreme-based constraints, it would be interesting to extend the reformulation to non-Gaussian processes. The resulting reformulation could then be applied to the FOWT problem when non-Gaussian output processes are involved. Concerning the integral-based constraints, the reformulation relies on the ergodicity of the time-dependent process and can therefore be generalized to non-Gaussian processes. However, in this case, the Dirlik method will no longer be suitable for the damage assessment and another approach (such as the time-domain approach described in section 5.3.1) will have to be considered for the fatigue estimation.

Our methodology could also be applied to other fields of research on optimization problems with constraints involving the maximum or the integral of random processes over a long period of time. We can refer for instance, to the evaluation of the risk of failure of structures due to earthquake (Zentner (2010)).

9.2 Concerning AK-ECO and the global enrichment

As explained in section 3.4, the enrichment strategy of AK-ECO and GE is introduced for reformulated extreme and integral based constraints but can be applied to any constraint formulated as in equation (5.77).

Although these methods are ready to handle problems with large n_s, they currently rely on a coupling with Monte-Carlo. A variance reduction technique such as Importance Sampling would accelerate the enrichment procedure and the evaluation of the constraints. Besides, a multipoint enrichment would reduce the number of calibrations to be performed. The latter represents an important part of the post-processing time especially when the DoE size becomes large. In addition, a multipoint procedure would take advantage of parallel computing. We also propose in section 3.4, perspectives to improve the high and low estimations of the reformulated constraints and we discuss different kriging strategy improvements that could be used to deal with large dimension problems.

Finally, the enrichment procedure of AK-ECO is implemented to be coupled with a local optimization algorithm since the enrichments aim to improve locally the constraint estimation at the design point obtained with the previous optimization cycle. To identify a global optimum, a multistart resolution is performed in this thesis. If one wishes to couple AK-ECO with a global optimization algorithm (Picheny (2014)), the enrichment procedure will need to be adapted.

We refer to sections 3.4 and 4.4.1 for more details.

9.3 Concerning the RBDO-oriented GSA

Different improvements of the sensitivity analysis proposed in this thesis are discussed in section 4.4.2. Among the possible perspectives, the cost function could be taken into account in the definition of the critical domain. Besides, the critical domain definition could also be involved in the accuracy criteria of the GE to provide more suitable metamodels for the RBDO-oriented GSA.

Due to time constraints, the RBDO-oriented GSA proposed in this thesis could not be tested on other cases than the oscillator problem. To better understand the validity domain of this method, it would be interesting to apply it to other cases and in particular to problems depending on a larger number of random variables.

9.4 Concerning the FOWT problem

The objective of this thesis was to tackle a simplified FOWT problem to identify an efficient methodology that could then be improved to deal with complex industrial problems. Some modeling choices have been made in this sense. The continuity of this thesis work would thus be to study the performance of our approach to more realistic problems.

In particular, the first step would be to consider a full scatter diagram and evaluate the effectiveness and limitations of our methodology in this case.

The next step would be to consider non-constant wind forces, non-linearized hydrodynamic forces and second-order wave loads which would leads to non-Gaussian output processes.

We would then be confronted with the real complexity of the industrial case. As we have mentioned, dealing with Fatigue Limit State constraints with our methodology is already feasible in this case. On the other hand, extreme-based constraints, such as ULS constraints, would require more effort in order to evaluate the failure probability involving maximums of non-Gaussian processes.

Appendix A

Approximation error made with the reformulation of failure probabilities involving stationary processes

Contents

1 Extreme-based failure probability reformulation error

Bounds can be obtained on the approximation (2.17) using the theorem 2.2 of Kratz and Rootzén (1997). We present here the theorem for a process ξ satisfying the conditions of theorem 2.1 and we show how it can be applied to control the reformulation error.

1.1 Rate of convergence of theorem 2.1

Theorem A.1 (Theorem 2.2 of Kratz and Rootzén (1997))**.** Let ξ be the process introduced in theorem 2.1 satisfying condition (2.11) and the following conditions:

$$\mathbb{E}\left[\left(\xi'(t)-\xi'(0)\right)\right]^2 = 2\left(k_\xi''(\tau)-k_\xi''(0)\right) \le c\tau^2, \quad \tau \ge 0 \tag{A.1}$$

$$\left|k_\xi(\tau)\right| \le Ct^{-\alpha}, \quad \left|k_\xi(\tau)\right| + k_\xi'(\tau)^2 \le C_0 t^{-\alpha}, \quad \tau \ge 0 \tag{A.2}$$

for some $\alpha > 2$ and constants c, C, C_0. Then there is a constant K which depends on k_ξ but not on u or T such that, for $T \ge T_0 > 1$,

$$\left|\mathbb{P}\left(\max_{t\in[0,T]} \frac{\xi(t)}{\sqrt{m_{\xi,0}}} \le u\right) - \exp\left(-\sqrt{\frac{m_{\xi,2}}{m_{\xi,0}}}T\mu(u)\right)\right| \le K\frac{\log\left(\sqrt{\frac{m_{\xi,2}}{m_{\xi,0}}}T\right)^{1+1/\alpha}}{\left(\sqrt{\frac{m_{\xi,2}}{m_{\xi,0}}}T\right)^{\delta}} \tag{A.3}$$

with $\mu(u) = \frac{1}{2\pi}e^{-\frac{u^2}{2}}$, $\delta = \min\{1/2, \inf_{\tau\geq 0}\rho(\tau)\}$ and $\rho(\tau) = \frac{\left(1-\frac{k_\xi(\tau)}{m_{\xi,0}}\right)^2}{1-\frac{1}{m_{\xi,0}^2}k_\xi(\tau)^2+\frac{1}{m_{\xi,0}m_{\xi,2}}k'_\xi(\tau)\left|k'_\xi(\tau)\right|}$.

When the process ξ is known through its spectral density K_ξ, conditions (A.1) and (A.2) are met if we have:

$$m_{\xi,4} < \infty, \tag{A.4}$$

$$K_\xi \in C^3, K_\xi^{(i)} \text{ and } \omega K_\xi^{(j)} \text{ are integrable for } i = 0,1,2,3 \text{ and } j = 0,1,2 \tag{A.5}$$

with $K_\xi^{(i)}$ the i-th derivative of K_ξ.

Proposition A.1. If a process ξ meets all the conditions of theorem A.1, it follows that for all $x_r \in \mathbb{R}$:

$$\left|\mathbb{P}\left(\max_{t\in[0,T]}\xi(t)\leq x_r\right) - \exp\left(-e^{a_T^2-\frac{a_T x_r}{\sqrt{m_{\xi,0}}}}\right)\right| \leq K\frac{\log\left(2\pi\frac{T}{T_c}\right)^{1+1/\alpha}}{\left(2\pi\frac{T}{T_c}\right)^{\delta}} + \exp\left(-\frac{T}{T_c}\exp^{-\frac{x_r^2}{2m_{\xi,0}}}\right). \tag{A.6}$$

with $a_T = \sqrt{2\log\left(\frac{T}{T_c}\right)}$ and $T_c = 2\pi\sqrt{\frac{m_{\xi,0}}{m_{\xi,2}}}$.

Proof A.1 (Proof of proposition A.1)**.** Let us prove the preliminary result:

$$\exp\left(-e^{a_T^2-\frac{a_T x_r}{\sqrt{m_{\xi,0}}}}\right) \leq \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right),\ \forall x_r \in \mathbb{R}. \tag{A.7}$$

We have, $\forall x_r \in \mathbb{R}$:

$$a_T^2 - a_T\frac{x_r}{\sqrt{m_{\xi,0}}} + \frac{x_r^2}{2m_{\xi,0}} \geq \min_{x'_r\in\mathbb{R}} a_T^2 - a_T\frac{x'_r}{\sqrt{m_{\xi,0}}} + \frac{x_r'^2}{2m_{\xi,0}} = \frac{a_T^2}{2} \tag{A.8}$$

$$\Rightarrow \exp\left(a_T^2 - a_T\frac{x_r}{\sqrt{m_{\xi,0}}} + \frac{x_r^2}{2m_{\xi,0}}\right) \geq \frac{T}{T_c} \tag{A.9}$$

$$\Rightarrow \exp\left(a_T^2 - a_T\frac{x_r}{\sqrt{m_{\xi,0}}}\right) \geq \frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right) \tag{A.10}$$

$$\Rightarrow \exp\left(-\exp\left(a_T^2 - a_T\frac{x_r}{\sqrt{m_{\xi,0}}}\right)\right) \leq \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right) \tag{A.11}$$

It follows from this result and theorem A.1 with $u = \frac{x_r}{\sqrt{m_{\xi,0}}}$ that:

$$\left|\mathbb{P}\left(\max_{t\in[0,T]}\xi(t)\leq x_r\right) - \exp\left(-e^{a_T^2-\frac{a_T x_r}{\sqrt{m_{\xi,0}}}}\right)\right| \tag{A.12}$$

$$\leq \left|\mathbb{P}\left(\max_{t\in[0,T]}\xi(t)\leq x_r\right) - \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right)\right| + \left|\exp\left(-e^{a_T^2-\frac{a_T x_r}{\sqrt{m_{\xi,0}}}}\right) - \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right)\right| \tag{A.13}$$

$$\leq \left|\mathbb{P}\left(\max_{t\in[0,T]}\xi(t)\leq x_r\right) - \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right)\right| + \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right) \tag{A.14}$$

$$\leq K\frac{\log\left(2\pi\frac{T}{T_c}\right)^{1+1/\alpha}}{\left(2\pi\frac{T}{T_c}\right)^{\delta}} + \exp\left(-\frac{T}{T_c}\exp\left(-\frac{x_r^2}{2m_{\xi,0}}\right)\right) \tag{A.15}$$

□

1.2 Application of proposition A.1 to extreme-based failure probability approximation

Proposition A.2. We denote by f_{r_E} the probability density function of X_{r_E} and $T_c(x_d, x_p) = 2\pi\sqrt{\frac{m_{\mathcal{Y},0}(x_d,x_p)}{m_{\mathcal{Y},2}(x_d,x_p)}}$. We assume the following conditions:

(1) $\exists K, \alpha, \delta$ such that $\forall x_d, x_p$, the process $\mathcal{Y}(x_d, x_p; .)$ satisfies the conditions of theorem A.1,

(2) $\exists T_1 > 0, T_2 > 0, m_1 > 0$ such that $\forall x_d, x_p$, $T_1 \le T_c(x_d, x_p) \le T_2$ and $m_1 \le m_{\mathcal{Y},0}(x_d, x_p)$,

(3) $\exists c_1 > 0, c_r > 0$ such that $\forall x, |x| \ge c_1, f_{r_\mathrm{E}}(x) \le c_r \exp\left(-\frac{x^2}{m_1}\right)$.

Then, if $\sqrt{m_1 \log(T)} > c_1$, the error made in the approximation (2.17) can be bounded as follows:

$$\left| p_\mathrm{E}(d) - \mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[F_\epsilon\left(e^{a_T(X_d,X_p)^2 - \frac{a_T(X_d,X_p)X_{r_\mathrm{E}}}{\sqrt{m_{\mathcal{Y},0}(X_d,X_p)}}}\right)\right]\right| \le K\frac{\log\left(2\pi\frac{T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{T}{T_2}\right)^{\delta}} + \frac{T_2 + c_r\sqrt{\pi m_1}}{\sqrt{T}}. \tag{A.16}$$

Proof A.2 (Proof of proposition A.2)**.**

$$\left| p_\mathrm{E}(d) - \mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[F_\epsilon\left(e^{a_T(X_d,X_p)^2 - \frac{a_T(X_d,X_p)X_{r_\mathrm{E}}}{\sqrt{m_{\mathcal{Y},0}(X_d,X_p)}}}\right)\right]\right| \tag{A.17}$$

$$= \left| \mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[\mathbb{P}_\eta\left(\max_{t\in[0,T]} \mathcal{Y}(X_d, X_p; t) > X_{r_\mathrm{E}}\right) - F_\epsilon\left(e^{a_T(X_d,X_p)^2 - \frac{a_T(X_d,X_p)X_{r_\mathrm{E}}}{\sqrt{m_{\mathcal{Y},0}(X_d,X_p)}}}\right)\right]\right| \tag{A.18}$$

$$\le \mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[\left|\mathbb{P}_\eta\left(\max_{t\in[0,T]} \mathcal{Y}(X_d, X_p; t) \le X_{r_\mathrm{E}}\right) - \exp\left(-e^{a_T(X_d,X_p)^2 - \frac{a_T(X_d,X_p)X_{r_\mathrm{E}}}{\sqrt{m_{\mathcal{Y},0}(X_d,X_p)}}}\right)\right|\right] \tag{A.19}$$

$$\le \mathbb{E}_{X_d,X_p,X_{r_\mathrm{E}}}\left[\left|K\frac{\log\left(2\pi\frac{T}{T_c(X_d,X_p)}\right)^{1+1/\alpha}}{\left(2\pi\frac{T}{T_c(X_d,X_p)}\right)^{\delta}} + \exp\left(-\frac{T}{T_c(X_d,X_p)}\exp\left(-\frac{X_{r_\mathrm{E}}^2}{2m_{\mathcal{Y},0}(X_d,X_p)}\right)\right)\right|\right] \tag{A.20}$$

$$\le K\frac{\log\left(2\pi\frac{T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{T}{T_2}\right)^{\delta}} + \mathbb{E}_{X_{r_\mathrm{E}}}\left[\exp\left(-\frac{T}{T_2}\exp\left(-\frac{X_{r_\mathrm{E}}^2}{2m_1}\right)\right)\right] \tag{A.21}$$

The two last equations are obtained using successively proposition A.1 and assumption (2). Denoting $\alpha_T = \sqrt{m_1 \log(T)}$, it follows from assumption (3):

$$\mathbb{E}_{X_{r_\mathrm{E}}}\left[\exp\left(-\frac{T}{T_2}\exp\left(-\frac{X_{r_\mathrm{E}}^2}{2m_1}\right)\right)\right] \tag{A.22}$$

$$= \int_{-\alpha_T}^{\alpha_T} \exp\left(-\frac{T}{T_2}\exp\left(-\frac{x^2}{2m_1}\right)\right) f_{r_\mathrm{E}}(x)dx + \int_{\mathbb{R}/[-\alpha_T,\alpha_T]} \exp\left(-\frac{T}{T_2}\exp\left(-\frac{x^2}{2m_1}\right)\right) f_{r_\mathrm{E}}(x)dx \tag{A.23}$$

$$\leq \exp\left(-\frac{T}{T_2}\exp\left(-\frac{\alpha_T^2}{2m_1}\right)\right) + \int_{\mathbb{R}/[-\alpha_T,\alpha_T]} f_{r_\mathrm{E}}(x)dx \tag{A.24}$$

$$\leq \exp\left(-\frac{\sqrt{T}}{T_2}\right) + c_r \int_{\mathbb{R}/[-\alpha_T,\alpha_T]} \exp\left(-\frac{x^2}{m_1}\right)dx \tag{A.25}$$

$$\leq \frac{T_2}{\sqrt{T}} + c_r\left(\int_{\mathbb{R}} \exp\left(-\frac{x^2}{m_1}\right)dx - 2\int_0^{\alpha_T} \exp\left(-\frac{x^2}{m_1}\right)dx\right) \tag{A.26}$$

$$\leq \frac{T_2}{\sqrt{T}} + c_r\left(\sqrt{\pi m_1} - \sqrt{\pi m_1}\sqrt{1-\exp\left(-\frac{\alpha_T^2}{m_1}\right)}\right) \tag{A.27}$$

$$\leq \frac{T_2}{\sqrt{T}} + c_r\sqrt{\pi m_1}\left(1-\sqrt{1-\frac{1}{T}}\right) \tag{A.28}$$

$$\leq \frac{T_2}{\sqrt{T}} + c_r\sqrt{\frac{\pi m_1}{T}}. \tag{A.29}$$

Equation (A.27) is obtained considering the integral $I = \int_0^{\alpha_T} \exp\left(-\frac{x^2}{m_1}\right)dx$, then I^2 can be bounded working in polar coordinates. □

2 Integral-based failure probability reformulation error

Proposition A.3. We denote by f_{r_I} the probability density function of X_{r_I} and for fixed values x_d, x_p, $Z_T(x_d, x_p)$ the random variable $\frac{1}{T}\int_0^T \mathcal{F}(x_d, x_p; t)\,dt$. We assume the following conditions:

(1) $\exists c_1 > 0, c_r > 0$ such that $\forall x \geq c_1$, $f_{r_\mathrm{I}}(x) \leq \frac{c_r}{x}$,

(2) $\exists T_0, c_2 > 0$ such that $\forall x_d, x_p$ and $\forall T > T_0$, $Z_T(x_d, x_p) \geq c_2$ almost surely,

(3) $\exists c_\mathcal{F} > 0$ such that for all x_d, x_p:

$$\int_{\mathbb{R}} |k_\mathcal{F}(x_d, x_p; \tau)|\,d\tau < c_\mathcal{F}. \tag{A.30}$$

Then, if $T > T_0$ and $T > \frac{c_1}{c_2}$, we have:

$$\left|p_\mathrm{I}(d) - \mathbb{E}_{X_d,X_p}\left[F_{r_\mathrm{I}}\left(T\mathbb{E}_\eta\left[\mathcal{F}(X_d, X_p; 0)\right]\right)\right]\right| \leq \frac{c_r}{c_2}\sqrt{\frac{2c_\mathcal{F}}{T}}. \tag{A.31}$$

Proof A.3 (Proof of proposition A.3)**.**

$$\left|p_{\mathrm{I}}(d) - \mathbb{E}_{X_d,X_p}\left[F_{r_{\mathrm{I}}}\left(T\mathbb{E}_\eta\left[\mathcal{F}\left(X_d,X_p;0\right)\right]\right)\right]\right| \tag{A.32}$$

$$= \left|\mathbb{P}_{X_d,X_p,X_{r_{\mathrm{I}}},\eta}\left(\int_0^T f\left(\mathcal{Y}\left(X_d,X_p;t\right)\right)dt > X_{r_{\mathrm{I}}}\right) - \mathbb{E}_{X_d,X_p}\left[F_{r_{\mathrm{I}}}\left(T\mathbb{E}_\eta\left[\mathcal{F}\left(X_d,X_p;0\right)\right]\right)\right]\right| \tag{A.33}$$

$$= \left|\mathbb{E}_{X_d,X_p,\eta}\left[F_{r_{\mathrm{I}}}\left(TZ_T(X_d,X_p)\right)\right] - \mathbb{E}_{X_d,X_p}\left[F_{r_{\mathrm{I}}}\left(T\mathbb{E}_\eta\left[Z_T\left(X_d,X_p\right)\right]\right)\right]\right| \tag{A.34}$$

$$\leq \mathbb{E}_{X_d,X_p,\eta}\left[\left|F_{r_{\mathrm{I}}}\left(TZ_T(X_d,X_p)\right) - F_{r_{\mathrm{I}}}\left(T\mathbb{E}_\eta\left[Z_T\left(X_d,X_p\right)\right]\right)\right|\right] \tag{A.35}$$

$$\leq \mathbb{E}_{X_d,X_p,\eta}\left[\left|TZ_T(X_d,X_p) - T\mathbb{E}_\eta\left[Z_T\left(X_d,X_p\right)\right]\right| \max\left(\frac{c_r}{\left(TZ_T\left(X_d,X_p\right)\right)}, \frac{c_r}{\left(T\mathbb{E}_\eta\left[Z_T\left(X_d,X_p\right)\right]\right)}\right)\right] \tag{A.36}$$

$$\leq \mathbb{E}_{X_d,X_p,\eta}\left[T\left|Z_T(X_d,X_p) - \mathbb{E}_\eta\left[Z_T\left(X_d,X_p\right)\right]\right| \frac{c_r}{c_2 T}\right] \tag{A.37}$$

$$\leq \frac{c_r}{c_2}\,\mathbb{E}_{X_d,X_p}\left[\sqrt{\mathrm{Var}\left(Z_T\left(X_d,X_p\right)\right)}\right] \tag{A.38}$$

$$\leq \frac{c_r}{c_2}\sqrt{\frac{2c_{\mathcal{F}}}{T}} \tag{A.39}$$

Conditions (1), (2) and $T \geq \frac{c_1}{c_2}$ imply equation (A.36). In (A.38), we apply the Cauchy–Schwarz inequality and equation (A.39) follows from:

$$\mathrm{Var}(Z_T\left(x_d,x_p\right)) = \frac{1}{T^2}\int_0^T\int_0^T k_{\mathcal{F}}\left(x_d,x_p;|t-t'|\right)dtdt' \tag{A.40}$$

$$= \frac{2}{T}\int_0^T\left(1-\frac{\tau}{T}\right)k_{\mathcal{F}}\left(x_d,x_p;\tau\right)d\tau \tag{A.41}$$

$$\leq \frac{2}{T}\int_0^T \left|k_{\mathcal{F}}\left(x_d,x_p;\tau\right)\right|d\tau \tag{A.42}$$

□

Appendix B

Proof of the sufficient conditions to reformulate the stationary harmonic oscillator problem

Contents

1 Sufficient conditions for the extreme-based constraints

To apply the approximation (2.17) to the first and second constraints of the oscillator problem, it is sufficient to show that for all x_d, x_p, the following conditions are satisfied:

$$m_{\mathcal{D}',0}(x_d,x_p)<\infty, m_{\mathcal{D}',2}(x_d,x_p)<\infty, m_{\mathcal{D}'',0}(x_d,x_p)<\infty, m_{\mathcal{D}'',2}(x_d,x_p)<\infty \tag{B.1}$$

$$K_{\mathcal{D}'}(x_d,x_p,.)\in C^1, K_{\mathcal{D}'}(x_d,x_p,.) \text{ and } K'_{\mathcal{D}'}(x_d,x_p,.) \text{ are integrable,} \tag{B.2}$$

$$K_{\mathcal{D}''}(x_d,x_p,.)\in C^1, K_{\mathcal{D}''}(x_d,x_p,.) \text{ and } K'_{\mathcal{D}''}(x_d,x_p,.) \text{ are integrable.} \tag{B.3}$$

As we have the relations:

$$K_{\mathcal{D}'}\left(x_d,x_p;\omega\right)=\omega^2 K_{\mathcal{D}}\left(x_d,x_p;\omega\right) \tag{B.4}$$

and

$$K_{\mathcal{D}''}\left(x_d,x_p;\omega\right)=\omega^4 K_{\mathcal{D}}\left(x_d,x_p;\omega\right), \tag{B.5}$$

to show (B.1), (B.2) and (B.3), it is sufficient to prove that $\omega^i K_{\mathcal{D}}^{(j)}$ is integrable for the appropriate values of i and j where $K_{\mathcal{D}}^{(j)}$ is the j-th derivative of $K_{\mathcal{D}}$. In fact, in our case it is true for all i and j. Indeed, it follows from relation (2.26) that $K_{\mathcal{D}}\left(x_d,x_p;.\right)$ is the product of a rational function (with no real pole) and a Gaussian function. Hence, we can demonstrate that $\forall i\in\mathbb{N}, \forall j\in\mathbb{N}$, $\omega^i K_{\mathcal{D}}^{(j)}\left(x_d,x_p;.\right)$ is integrable.

2 Sufficient condition for the integral-based constraint

To apply the approximation (2.21) to the third constraint of the oscillator problem, we need to prove that for all x_d, x_p, the process $(|\mathcal{D}''(x_d, x_p; .)| - \rho)^+$ is ergodic. To do so, we show in this section that its covariance function is integrable. Introducing the following notation:

$$\begin{aligned} f(x) &= (|x| - \rho)^+ \quad , \quad U_t = f\left(\mathcal{D}''(x_d, x_p; t)\right), \\ \sigma^2 &= \mathrm{Var}\left(\mathcal{D}''(x_d, x_p; 0)\right) \quad , \quad k_t = \frac{\mathrm{Cov}(\mathcal{D}''(x_d,x_p;0),\mathcal{D}''(x_d,x_p;t))}{\sigma^2}, \end{aligned} \tag{B.6}$$

we want to prove that $\int_{\mathbb{R}} |\mathrm{Cov}(U_0, U_t)|\, dt < \infty$ with $\mathrm{Cov}(U_0, U_t) = \mathbb{E}[U_0 U_t] - \mathbb{E}[U_0]^2$ and

$$\mathbb{E}[U_0 U_t] = \frac{1}{2\pi\sigma^2\sqrt{1-k_t^2}} \iint f(u) f(v) \exp\left(-\frac{u^2 + v^2 - 2k_t uv}{2\sigma^2(1-k_t^2)}\right) du dv. \tag{B.7}$$

Let us consider the function g defined as

$$g(k, u, v) = \frac{1}{2\pi\sigma^2\sqrt{1-k^2}} f(u) f(v) \exp\left(-\frac{u^2 + v^2 - 2kuv}{2\sigma^2(1-k^2)}\right). \tag{B.8}$$

Since g and its two first derivatives w.r.t k are dominated $\forall k \in [-\frac{1}{2}, \frac{1}{2}]$ by a integrable function $h(u, v)$, it follows that

$$\iint g(k, u, v) du dv = \iint g(0, u, v) du dv + k \iint \frac{\partial g(0, u, v)}{\partial k} du dv + O(k^2). \tag{B.9}$$

Using that $\iint g(0, u, v) du dv = \mathbb{E}[U_0]^2$, we deduce that $Cov(U_0, U_t) = O(k_t)$. Therefore, $\mathrm{Cov}(U_0, U_t)$ is integrable because the covariance function of the process $\mathcal{D}''(x_d, x_p; .)$ is integrable.

Appendix C

Reformulation of failure probabilities involving piece-wise stationary processes

Contents

1 Extreme-based failure probability reformulation

We recall that the notation n^j $(j = 1, \cdots, n_s)$ refers to the number of intervals I_i such that the sea state s_i considered during I_i is equal to s^j. Besides, we denote $p^j = \frac{n^j}{n_T}$ $(j = 1, \cdots, n_s)$ with n_T is the number of subintervals I_i in $[0, T]$.

For $j = 1, \cdots, n_s$, let us denote $I^j = \bigcup_{i, s_i = s^j} I_i$. Hence, I^j is the union of n^j intervals of length ΔT. For fixed $x_d, x_p, x_{r_\mathrm{E}}$, we have:

$$\mathbb{P}_\eta\left(\max_{t\in[0,T]} \mathcal{Y}(x_d, x_p; t) \leq x_{r_\mathrm{E}}\right) \tag{C.1}$$

$$= \mathbb{P}_\eta\left(\max_{t\in[0,T]} \sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d, x_p, s_i; t\right) \mathbb{1}_{I_i}(t) \leq x_{r_\mathrm{E}}\right) \tag{C.2}$$

$$= \mathbb{P}_\eta\left(\max_{t\in I^1} \sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d, x_p, s^1; t\right) \mathbb{1}_{I_i}(t) \leq x_{r_\mathrm{E}}, \cdots, \max_{t\in I^{n_s}} \sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d, x_p, s^{n_s}; t\right) \mathbb{1}_{I_i}(t) \leq x_{r_\mathrm{E}}\right) \tag{C.3}$$

$$= \prod_{j=1}^{n_s} \mathbb{P}_{\eta_1, \cdots, \eta_{n_T}}\left(\max_{t\in I^j} \sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d, x_p, s^j; t\right) \mathbb{1}_{I_i}(t) \leq x_{r_\mathrm{E}}\right) \tag{C.4}$$

$$= \prod_{j=1}^{n_s} \mathbb{P}_{\eta_1, \cdots, \eta_{n^j}}\left(\max_{t\in[0, n^j \Delta T]} \sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d, x_p, s^j; t\right) \mathbb{1}_{I_i}(t) \leq x_{r_\mathrm{E}}\right). \tag{C.5}$$

The independence of $\mathcal{Y}_i\left(x_d, x, s_i; .\right)$ and $\mathcal{Y}_{i'}\left(x_d, x_p, s_{i'}; .\right)$ for all $i \neq i'$ is used to obtain equation (C.4). The last equation results from the fact that $\mathcal{Y}_i\left(x_d, x_p, s^j; .\right)$ and $\mathcal{Y}_{i'}\left(x_d, x_p, s^j; .\right)$

for $i \neq i'$ are i.i.d processes.

Finally, when ΔT is large, we consider that each term of the product appearing in equation (C.5) can be approached by $\mathbb{P}_{\eta_1}\left(\max_{t\in[0,n^j\Delta T]} \mathcal{Y}_1\left(x_d, x_p, s^j; t\right) \leq x_{r_E}\right)$ which leads to approximation (2.36) since $n^j \Delta T = T p^j$.

2 Integral-based failure probability reformulation

Proof C.1 (Proof of proposition 2.1)**.** For fixed values x_d, x_p and x, we have:

$$\mathbb{P}_\eta\left(\frac{1}{\Delta T}\int_0^T f\left(\sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d, x_p, s_i; t\right) \mathbb{1}_{I_i}(t)\right) dt > x\right) \tag{C.6}$$

$$= \mathbb{P}_\eta\left(\frac{1}{\Delta T}\sum_{i=1}^{n_T}\int_{I_i} f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; t\right)\right) dt > x\right) \tag{C.7}$$

$$= \mathbb{P}_\eta\left(\sum_{i=1}^{n_T}\frac{1}{\Delta T}\int_0^{\Delta T} f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; t\right)\right) dt > x\right). \tag{C.8}$$

The last equality is obtained using the stationarity of the processes $f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; .\right)\right)$ and the independence between $f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; .\right)\right)$ and $f\left(\mathcal{Y}_{i'}\left(x_d, x_p, s_{i'}; .\right)\right)$ for $i \neq i'$.
Let $U_{i,\Delta T}$ be the random variable $\frac{1}{\Delta T}\int_0^{\Delta T} f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; t\right)\right) dt$. It follows from the assumption that $f\left(\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)\right)$ is ergodic $\forall s^j$:

$$U_{i,\Delta T} \xrightarrow[\Delta T\to+\infty]{\mathbb{P}} u_i \tag{C.9}$$

with $u_i = \mathbb{E}_{\eta_i}\left[f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; 0\right)\right)\right]$. Using the independence of $U_{i,\Delta T}$ and $U_{i',\Delta T}$ for $i \neq i'$, we deduce that:

$$\sum_{i=1}^{n_T} U_{i,\Delta T} \xrightarrow[\Delta T\to+\infty]{\mathbb{P}} \sum_{i=1}^{n_T} u_i. \tag{C.10}$$

Therefore, we have the convergence in distribution:

$$\mathbb{P}_\eta\left(\sum_{i=1}^{n_T}\frac{1}{\Delta T}\int_0^{\Delta T} f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; t\right)\right) dt > x\right) \xrightarrow[\Delta T\to+\infty]{} \mathbb{1}_{\sum_{i=1}^{n_T}\mathbb{E}_{\eta_i}\left[f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; 0\right)\right)\right] > x} \tag{C.11}$$

for all $x \neq \sum_{i=1}^{n_T}\mathbb{E}_{\eta_i}\left[f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; 0\right)\right)\right]$, with

$$\sum_{i=1}^{n_T}\mathbb{E}_{\eta_i}\left[f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; 0\right)\right)\right] = \sum_{i=1}^{n_T}\mathbb{E}_{\eta_1}\left[f\left(\mathcal{Y}_1\left(x_d, x_p, s_i; 0\right)\right)\right] \tag{C.12}$$

$$= \sum_{j=1}^{n_s} n^j \mathbb{E}_{\eta_1}\left[f\left(\mathcal{Y}_1\left(x_d, x_p, s^j; 0\right)\right)\right]. \tag{C.13}$$

□

Appendix D

Approximation error made with the reformulation of failure probabilities involving piece-wise stationary processes

Contents

1 Extreme-based failure probability reformulation error

Proposition D.1. We denote $f_{r_{\mathrm{E}}}$ the probability density function of $X_{r_{\mathrm{E}}}$ and $T_c\left(x_d, x_p, s^j\right) = 2\pi\sqrt{\frac{m_{\mathcal{Y},0}(x_d,x_p,s^j)}{m_{\mathcal{Y},2}(x_d,x_p,s^j)}}$. We assume the following conditions:

(1) $\exists K, \alpha, \delta$ such that $\forall x_d, x_p, s^j$, the process $\mathcal{Y}_1\left(x_d, x_p, s^j; .\right)$ satisfies the conditions of theorem A.1,

(2) $\exists T_1 > 0, T_2 > 0, m_1 > 0$ such that $\forall x_d, x_p, s^j$, $T_1 \leq T_c\left(x_d, x_p, s^j\right) \leq T_2$ and $m_1 \leq m_{\mathcal{Y},0}\left(x_d, x_p, s^j\right)$,

(3) $\exists c_1 > 0, c_r > 0$ such that $\forall x, |x| \geq c_1, f_{r_{\mathrm{E}}}(x) \leq c_r \exp\left(-\frac{x^2}{m_1}\right)$.

Then, if $\sqrt{m_1 \log(\Delta T)} > c_1$ the error made in equation (2.37) can be bounded as follows:

$$\left| p_{\mathrm{E}}(d) - \mathbb{E}_{X_d, X_p, X_{r_{\mathrm{E}}}}\left[F_\epsilon\left(\sum_{j=1}^{n_s} e^{a_{Tp^j}\left(X_d,X_p,s^j\right)^2 - \frac{a_{Tp^j}\left(X_d,X_p,s^j\right)X_{r_{\mathrm{E}}}}{\sqrt{m_{\mathcal{Y},0}\left(X_d,X_p,s^j\right)}}} \right) \right] \right|$$
$$\leq n_s \left(n_T K \frac{\log\left(2\pi\frac{\Delta T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{\Delta T}{T_2}\right)^{\delta}} + (n_T+1)\frac{T_2 + c_r\sqrt{\pi m_1}}{\sqrt{\Delta T}} \right). \quad \text{(D.1)}$$

Proof D.1 (Proof of proposition D.1)**.** We introduce the following notations:

$$P_j(T) = \mathbb{P}_{\eta_1}\left(\max_{t\in[0,T]} \mathcal{Y}_1\left(x_d, x_p, s^j; t\right) \leq x_{r_{\mathrm{E}}}\right), \tag{D.2}$$

$$E_j(T) = \exp\left(-e^{a_T\left(x_d,x_p,s^j\right)^2 - \frac{a_T\left(x_d,x_p,s^j\right)x_{r_{\mathrm{E}}}}{\sqrt{m_{\mathcal{Y},0}\left(x_d,x_p,s^j\right)}}}\right), \tag{D.3}$$

$$E_-(T) = \min_{j=1,\cdots,n_s} E_j(T)\ ,\ E_+(T) = \max_{j=1,\cdots,n_s} E_j(T), \tag{D.4}$$

$$p^- = \min_{j=1,\cdots,n_s} p^j\ ,\ p^+ = \max_{j=1,\cdots,n_s} p^j. \tag{D.5}$$

For fixed values of $x_d, x_p, x_{r_{\mathrm{E}}}$, we have $\forall j \in \{1,\cdots,n_s\}$:

$$\left|P_j\left(\Delta T\right)^{n^j} - E_j\left(Tp^j\right)\right| \leq \left|P_j\left(\Delta T\right)^{n^j} - E_j\left(\Delta T\right)^{n^j}\right| + \left|E_j\left(\Delta T\right)^{n^j} - E_j\left(Tp^j\right)\right|. \tag{D.6}$$

Besides, using $P_j(\Delta T) \in [0,1]$, $E_j(\Delta T) \in [0,1]$, proposition A.1 and assumption (2), we obtain:

$$\left|P_j\left(\Delta T\right)^{n^j} - E_j\left(\Delta T\right)^{n^j}\right| \leq n^j \left|P_j\left(\Delta T\right) - E_j\left(\Delta T\right)\right| \tag{D.7}$$

$$\leq n_T\left(K\frac{\log\left(2\pi\frac{\Delta T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{\Delta T}{T_2}\right)^{\delta}} + \exp\left(-\frac{\Delta T}{T_2}\exp\left(-\frac{x_{r_{\mathrm{E}}}^2}{2m_1}\right)\right)\right). \tag{D.8}$$

Since $E_j(T)$ is a decreasing function w.r.t to T and that its images are in $[0,1]$, we have for $n^j > 0$:

$$-E_+\left(Tp^-\right) \leq E_j\left(\Delta T\right)^{n^j} - E_j\left(Tp^j\right) \leq E_+\left(\Delta T\right). \tag{D.9}$$

It follows, with $Tp^- > 1$ (i.e. if $n^j > 0, \forall j$) and applying the preliminary result of proof A.1:

$$\left|E_j\left(\Delta T\right)^{n^j} - E_j\left(Tp^j\right)\right| \leq E_+(\Delta T) \leq \exp\left(-\frac{\Delta T}{T_2}\exp\left(-\frac{x_{r_{\mathrm{E}}}^2}{2m_1}\right)\right). \tag{D.10}$$

We therefore can use equations (D.8) and (D.10) to bound the approximation error:

$$\left| p_{\mathrm{E}}(d) - \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[1 - \exp\left(-\sum_{j=1}^{n_s} e^{a_{Tp^j}\left(X_d,X_p,s^j\right)^2 - \frac{a_{Tp^j}\left(X_d,X_p,s^j\right)X_{r_{\mathrm{E}}}}{\sqrt{m_{\mathcal{Y},0}\left(X_d,X_p,s^j\right)}}}\right)\right]\right| \tag{D.11}$$

$$= \left| \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[1 - \prod_{j=1}^{n_s} P_j\left(\Delta T\right)^{n^j}\right] - \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[1 - \prod_{j=1}^{n_s} E_j\left(Tp^j\right)\right]\right| \tag{D.12}$$

$$\leq \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[\left|\prod_{j=1}^{n_s} P_j\left(\Delta T\right)^{n^j} - \prod_{j=1}^{n_s} E_j\left(Tp^j\right)\right|\right] \tag{D.13}$$

$$\tag{D.14}$$

$$\leq n_s \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[n_T K \frac{\log\left(2\pi\frac{\Delta T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{\Delta T}{T_2}\right)^{\delta}} + \left(n_T+1\right)\exp\left(-\frac{\Delta T}{T_2}\exp\left(-\frac{X_{r_{\mathrm{E}}}^2}{2m_1}\right)\right)\right] \tag{D.15}$$

$$\leq n_s \left(n_T K \frac{\log\left(2\pi\frac{\Delta T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{\Delta T}{T_2}\right)^{\delta}} + \left(n_T+1\right)\mathbb{E}_{X_{r_{\mathrm{E}}}}\left[\exp\left(-\frac{\Delta T}{T_2}\exp\left(-\frac{X_{r_{\mathrm{E}}}^2}{2m_1}\right)\right)\right]\right) \tag{D.16}$$

$$\leq n_s \left(n_T K \frac{\log\left(2\pi\frac{\Delta T}{T_1}\right)^{1+1/\alpha}}{\left(2\pi\frac{\Delta T}{T_2}\right)^{\delta}} + \left(n_T+1\right)\frac{T_2 + c_r\sqrt{\pi m_1}}{\sqrt{\Delta T}}\right) \tag{D.17}$$

To get equation (D.12), we use the following equalities:

$$p_{\mathrm{E}}(d) = \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[1 - \prod_{j=1}^{n_s} \mathbb{P}_{\eta_1,\cdots,\eta_{n^j}}\left(\max_{t\in[0,n^j\Delta T]} \sum_{i=1}^{n_T} \mathcal{Y}_i\left(x_d,x_p,s^j;t\right)\mathbb{1}_{I_i}(t) \leq x_{r_{\mathrm{E}}}\right)\right] \tag{D.18}$$

$$= \mathbb{E}_{X_d,X_p,X_{r_{\mathrm{E}}}}\left[1 - \prod_{j=1}^{n_s} \mathbb{P}_{\eta_1}\left(\max_{t\in[0,\Delta T]} \mathcal{Y}_1\left(x_d,x_p,s^j;t\right) \leq x_{r_{\mathrm{E}}}\right)^{n^j}\right] \tag{D.19}$$

Equation (D.13) holds since, for $a_i, b_i \in [0,1], n \in \mathbb{N}^*$, if $|a_i - b_i| < c, i = 1,\cdots,n$ then $\left|\prod_i^n a_i - \prod_i^n b_i\right| < nc$. Finally, in equation (D.17), the quantity $\mathbb{E}_{X_{r_{\mathrm{E}}}}\left[\exp\left(-\frac{\Delta T}{T_2}\exp\left(-\frac{X_{r_{\mathrm{E}}}^2}{2m_1}\right)\right)\right]$ is bounded applying the reasoning of proof A.2. □

2 Integral-based failure probability reformulation error

Proposition D.2. We denote $f_{r_{\mathrm{I}}}$ the probability density function of $X_{r_{\mathrm{I}}}$ and for fixed values x_d, x_p:

$$Z_{\Delta T}(x_d,x_p) = \frac{1}{\Delta T}\int_0^T f\left(\mathcal{Y}\left(x_d,x_p;t\right)\right)dt = \frac{1}{\Delta T}\sum_{i=1}^{n_T}\int_{I_i} f\left(\mathcal{Y}_i\left(x_d,x_p,s_i;t\right)\right)dt, \tag{D.20}$$

with $k_{\mathcal{F}_1}\left(x_d,x_p,s_i;.\right)$ the autocovariance function of the process $\mathcal{F}_1\left(x_d,x_p,s_i;.\right)$. We assume that the following assumptions are valid:

(1) $\exists c_1 > 0, c_r > 0$ such that $\forall x \geq c_1$, $f_{r_{\mathrm{I}}}(x) \leq \frac{c_r}{x}$,

(2) $\exists \Delta T_0, \exists c_2 > 0$ such that $\forall x_d, x_p$ and $\Delta T > \Delta T_0$, $Z_{\Delta T}\left(x_d,x_p\right) \geq c_2$ almost surely,

(3) for each state s^j, $\exists c_{\mathcal{F},s^j} > 0$ such that for all x_d, x_p:

$$\int_{\mathbb{R}} \left| k_{\mathcal{F}_1}\left(x_d, x_p, s^j; \tau\right)\right| d\tau < c_{\mathcal{F},s^j}. \tag{D.21}$$

Then, if $\Delta T > T_0$ and $\Delta T \geq \frac{c_1}{c_2}$, the approximation error made in equation (2.40) can be bounded as follows:

$$\left| p_{\mathrm{I}}(d) - \mathbb{E}_{X_d,X_p}\left[F_{r_{\mathrm{I}}}\left(T \sum_{j=1}^{n_s} p^j \mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(X_d, X_p, s^j; 0\right)\right]\right)\right]\right| \leq \frac{c_r}{c_2}\sqrt{\frac{2n_T}{\Delta T}\sum_{j=1}^{n_s} p^j c_{\mathcal{F},s^j}}. \tag{D.22}$$

Proof D.2. [Proof of proposition D.2]
We denote in this proof, for fixed values x_d, x_p:

$$A_i = \int_{I_i} f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; t\right)\right) dt, \tag{D.23}$$

$$\mathcal{F}_i\left(s_i; t\right) = f\left(\mathcal{Y}_i\left(x_d, x_p, s_i; t\right)\right) \tag{D.24}$$

and

$$k_{\mathcal{F}_i,\mathcal{F}_j}(s_i, s_j; t, t') = \mathbb{E}_{\eta_i,\eta_j}\left[\mathcal{F}_i\left(s_i; t\right)\mathcal{F}_j\left(s_j; t'\right)\right] - \mathbb{E}_{\eta_i}\left[\mathcal{F}_i\left(s_i; t\right)\right]\mathbb{E}_{\eta_j}\left[\mathcal{F}_j\left(s_j; t'\right)\right]. \tag{D.25}$$

We start by calculating $\mathbb{E}_\eta\left[A_i\right], \mathbb{E}_\eta\left[A_i A_j\right]$ and $\mathrm{Var}_\eta\left(Z_{\Delta T}\left(x_d, x_p\right)\right)$ which are used further down in the proof. For the first quantity, we use Fubini's theorem to obtain:

$$\mathbb{E}_\eta\left[A_i\right] = \Delta T \mathbb{E}_{\eta_i}\left[\mathcal{F}_i\left(s_i; 0\right)\right]. \tag{D.26}$$

Besides,

$$\mathbb{E}_\eta\left[A_i A_j\right] = \mathbb{E}_\eta\left[\int_{I_i}\int_{I_j} \mathcal{F}_i\left(s_i; t\right)\mathcal{F}_j\left(s_j; t'\right) dtdt'\right] \tag{D.27}$$

$$= \int_{I_i}\int_{I_j}\left(k_{\mathcal{F}_i,\mathcal{F}_j}(s_i, s_j, t, t') + \mathbb{E}_{\eta_i}\left[\mathcal{F}_i\left(s_i; t\right)\right]\mathbb{E}_{\eta_j}\left[\mathcal{F}_j\left(s_j; t'\right)\right]\right) dtdt' \tag{D.28}$$

- if $i \neq j$: by independence of $\mathcal{F}_i\left(s_i; .\right)$ and $\mathcal{F}_j\left(s_j; .\right)$ we have $k_{\mathcal{F}_i,\mathcal{F}_j}\left(s_i, s_j; t, t'\right) = 0$ and it follows that $\mathbb{E}_\eta\left[A_i A_j\right] = \mathbb{E}_\eta\left[A_i\right]\mathbb{E}_\eta\left[A_j\right]$,
- if $i = j$: we use the stationarity of $\mathcal{F}_i\left(s_i, .\right)$ to obtain the following equalities:

$$\mathbb{E}_\eta\left[A_i^2\right] = \int_{I_i}\int_{I_i}\left(k_{\mathcal{F}_i,\mathcal{F}_i}(s_i, s_i, t, t') + \mathbb{E}_{\eta_i}\left[\mathcal{F}_i\left(s_i; t\right)\right]^2\right) dtdt' \tag{D.29}$$

$$= \int_{I_i}\int_{I_i} k_{\mathcal{F}_i}(s_i, |t - t'|) dtdt' + \Delta T^2 \mathbb{E}_{\eta_i}\left[\mathcal{F}_i\left(s_i; 0\right)\right]^2 \tag{D.30}$$

Let us calculate $\mathrm{Var}_\eta\left(Z_{\Delta T}\left(x_d, x_p\right)\right)$:

$$\mathrm{Var}_\eta\left(Z_{\Delta T}\left(x_d, x_p\right)\right) \tag{D.31}$$

$$= \mathbb{E}_\eta\left[Z_{\Delta T}\left(x_d, x_p\right)^2\right] - \mathbb{E}_\eta\left[Z_{\Delta T}\left(x_d, x_p\right)\right]^2 \tag{D.32}$$

$$= \frac{1}{\Delta T^2}\left(\sum_{i=1}^{n_T}\mathbb{E}_\eta\left[A_i^2\right] + 2\sum_{i<j}\mathbb{E}_\eta\left[A_i A_j\right] - \sum_{i=1}^{n_T}\mathbb{E}_\eta\left[A_i\right]^2 - 2\sum_{i<j}\mathbb{E}_\eta\left[A_i\right]\mathbb{E}_\eta\left[A_j\right]\right) \tag{D.33}$$

$$= \frac{1}{\Delta T^2}\left(\sum_{i=1}^{n_T}\mathbb{E}_\eta\left[A_i^2\right] - \mathbb{E}_\eta\left[A_i\right]^2\right) \tag{D.34}$$

$$= \frac{1}{\Delta T^2}\left(\sum_{i=1}^{n_T}\int_{I_i}\int_{I_i} k_{\mathcal{F}_i}(s_i, |t-t'|)dtdt'\right) \tag{D.35}$$

$$= \frac{1}{\Delta T^2}\left(\sum_{i=1}^{n_T} 2\Delta T\int_0^{\Delta T}\left(1-\frac{\tau}{\Delta T}\right)k_{\mathcal{F}_i}(s_i,\tau)d\tau\right) \tag{D.36}$$

$$\leq \frac{2}{\Delta T}\sum_{i=1}^{n_T}\int_0^{\Delta T} k_{\mathcal{F}_i}\left(s_i;\tau\right)d\tau \tag{D.37}$$

$$\leq \frac{2}{\Delta T}\sum_{j=1}^{n_s} n^j c_{\mathcal{F},s^j} \tag{D.38}$$

$$\leq \frac{2n_T}{\Delta T}\sum_{j=1}^{n_s} p^j c_{\mathcal{F},s^j} \tag{D.39}$$

Equation (D.38) is obtained using assumption (3) of proposition D.2.
To bound the approximation error made in equation (2.40), the reasoning used in proof A.3 in the stationary case is applied here: assumption (1), (2) and $\Delta T > \frac{c_1}{c_2}$ imply equation (D.43) while the Cauchy-Schwarz inequality is used to obtain equation (D.44):

$$\left|p_{\mathrm{I}}(d) - \mathbb{E}_{X_d,X_p}\left[F_{r_{\mathrm{I}}}\left(T\sum_{j=1}^{n_s} p^j\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(X_d, X_p, s^j; 0\right)\right]\right)\right]\right| \tag{D.40}$$

$$= \left|\mathbb{E}_{X_d,X_p,\eta}\left[F_{X_{r_{\mathrm{I}}}}\left(\Delta T Z_{\Delta T}(X_d,X_p)\right)\right] - \mathbb{E}_{X_d,X_p}\left[F_{X_{r_{\mathrm{I}}}}\left(T\sum_{j=1}^{n_s} p^j\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(X_d, X_p, s^j; 0\right)\right]\right)\right]\right| \tag{D.41}$$

$$\leq \mathbb{E}_{X_d,X_p,\eta}\left[\left|F_{X_{r_{\mathrm{I}}}}\left(\Delta T Z_{\Delta T}(X_d,X_p)\right) - F_{X_{r_{\mathrm{I}}}}\left(\Delta T\mathbb{E}_\eta\left[Z_{\Delta T}\left(X_d,X_p\right)\right]\right)\right|\right] \tag{D.42}$$

$$\leq \frac{c_r}{c_2}\mathbb{E}_{X_d,X_p,\eta}\left[\left|Z_{\Delta T}(X_d,X_p) - \mathbb{E}_\eta\left[Z_{\Delta T}\left(x_d,x_p\right)\right]\right|\right] \tag{D.43}$$

$$\leq \frac{c_r}{c_2}\mathbb{E}_{X_d,X_p}\left[\sqrt{\mathrm{Var}_\eta\left(Z_{\Delta T}\left(X_d,X_p\right)\right)}\right] \tag{D.44}$$

$$\leq \frac{c_r}{c_2}\sqrt{\frac{2n_T}{\Delta T}\sum_{j=1}^{n_s} p^j c_{\mathcal{F},s^j}}. \tag{D.45}$$

□

Appendix E

Computation of the expectation involved in the integral-based constraint of the oscillator problem

We recall that the random variable $\mathcal{D}_1''\left(x_d, x_p, s^j; 0\right)$ follows a normal distribution with zero mean and standard deviation $\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)$. Thus, we have for the oscillator problem:

$$\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(x_d, x_p, s^j; 0\right)\right] \tag{E.1}$$

$$= \mathbb{E}_{\eta_1}\left[\left(\left|\mathcal{D}_1''\left(x_d, x_p, s^j; 0\right)\right| - \rho\right)^+\right] \tag{E.2}$$

$$= \int_{\mathbb{R}} \left(|u| - \rho\right)^+ \frac{1}{\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)\sqrt{2\pi}} \exp\left(-\frac{u^2}{2\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)^2}\right) du. \tag{E.3}$$

We replace $\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)$ by σ to simplify the following calculations:

$$\int_{\mathbb{R}} \left(|u| - \rho\right)^+ \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u \tag{E.4}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left(\int_{-\infty}^{-\rho} (-u - \rho) \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u + \int_{\rho}^{\infty} (u - \rho) \exp\left(-\frac{u^2}{2\sigma^2}\mathrm{d}u\right)\right) \tag{E.5}$$

$$\begin{aligned} = \frac{1}{\sigma\sqrt{2\pi}}\Bigg(&- \int_{-\infty}^{-\rho} u \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u - \rho \int_{-\infty}^{-\rho} \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u \\ &+ \int_{\rho}^{\infty} u \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u - \rho \int_{\rho}^{\infty} \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u\Bigg) \end{aligned} \tag{E.6}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left(2\int_{\rho}^{\infty} u \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u - 2\rho \int_{\rho}^{\infty} \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u\right) \tag{E.7}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left(2\sigma^2 \exp\left(\frac{-\rho^2}{2\sigma^2}\right) - 2\rho\sqrt{2\pi}\sigma\left(1 - \Phi\left(\frac{\rho}{\sigma}\right)\right)\right) \tag{E.8}$$

with Φ the cumulative distribution function of the standard normal distribution and it follows:

$$\mathbb{E}_{\eta_1}\left[\mathcal{F}_1\left(x_d, x_p, s^j; 0\right)\right] = \sqrt{\frac{2}{\pi}}\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)\exp\left(-\frac{\rho^2}{2\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)^2}\right) + 2\rho\left(\Phi\left(\frac{\rho}{\sigma_{\mathcal{D}''}\left(x_d, x_p, s^j\right)}\right) - 1\right) \quad \text{(E.9)}$$

Appendix F

Execution time of AK-ECO and the comparison methods

Contents

1 Execution time to solve the oscillator problem

We display in Table F.1 the execution times of the methods introduced in section 3.3 of chapter 3 to solve the reformulated oscillator problem. We recall that n_{call} corresponds to the number of simulations required to solve the problem with each approach. For the oscillator problem, the CPU time of one simulation is about 5 milliseconds.

	MC	RIA	PMA	SORA	Stieng	AK-ECO
n_{calls}	3.57×10^6	791175	29393	15722	53200	252
Execution time	4h 50m 51s	1h 03m 37s	3m 12s	1m 34s	6h 13m 34s	3h 0m 35s

Table F.1: Execution time of AK-ECO and the comparison methods

For MC, RIA, PMA, and SORA, the simulations represent the main contribution to the total execution time.

During the solution of the problem with the Stieng method, 33 minutes are spent on the metamodel calibrations. The resolution of the optimization problem during each cycle represents the major part of the total execution time (these resolutions take 5h 36m 16s).

1.1 Details of the execution time of AK-ECO

In Table F.2, we display the time dedicated to each step of AK-ECO to solve the oscillator problem.

Step of AK-ECO	Execution time	Proportion
Calibrations	56m 41s	31.4%
Selections of x_{enr}	44m 56s	24.9%
COBYLA resolutions	28m 45s	15.9%
Stopping condition evaluations	48m 02s	26.6%
Simulations	1.5s	0.0%
Total	3h 0m 35s	100%

Table F.2: Execution time of AK-ECO

We note that the execution time is mainly allocated to the calibrations of the metamodels (the calibration of the first metamodel takes 5 seconds but 30 seconds are required for the last calibration), the selections of the enrichment points (the selection of the first and last enrichment points take respectively 4 and 20 seconds), the resolution of the optimization problem during each cycle of optimization (the resolutions take between 3m 53s and 8m 14s), and the evaluation of the stopping conditions (5s are required for the first evaluation and 20s for the last one). The time dedicated to the simulations is negligible.

1.2 Global enrichment with AK-ECO

We indicate in this section the execution times to solve the oscillator problem by combining the global enrichment and AK-ECO as described in section 4.3.2 of chapter 4.

The global enrichment takes 6h 36m 12s. It starts with an initial DoE of 50 points and then performs 65 enrichments. Most of the execution time is devoted to the selections of the design points where the local enrichments of the metamodels need to be carried out (225 seconds are needed for the selection of the first d_{enr} and 405 seconds for the selection of the last one). During the first cycle of enrichment, the calibration, the selection of x_{enr}, and the evaluation of the stopping condition take respectively 5s, 5s, and 121s while they take respectively 8s, 10s, and 171s during the last cycle.

Step of GE	Execution time	Proportion
Calibrations	6m 58s	1.8 %
Selections of d_{enr}	5h 41m 53s	86.3%
Selections of x_{enr}	8m 37s	2.2%
Stopping condition evaluations	38m 43s	9.8%
Simulation	0.6s	0.0%
Total	6h 36m 12s	100%

Table F.3: Execution time of the global enrichment

The reformulated oscillator problem is then solved with AK-ECO in 1h 44m 20s using the metamodels obtained after GE and performing 108 supplementary simulations. The details of the execution time is displayed in Table F.4. The procedure GE+AK-ECO thus takes 8h 20m 32s.

Step of AK-ECO	Execution time	Proportion
Calibrations	35m 39s	34.2%
Selections of x_{enr}	24m 59s	23.9%
COBYLA resolutions	15m 07s	14.5%
Stopping condition evaluations	26m 58s	25.8%
Simulations	0.5s	0.0%
Total	1h 44m 20s	100%

Table F.4: Execution time of AK-ECO after GE

2 Execution time to solve the FOWT problem

In Table F.5 we display the execution times of the methods introduced in section 5.6 of chapter 5 to solve the reformulated FOWT problem. We recall that n_{call} corresponds to the number of simulations required to solve the problem with each approach. For the FOWT problem, the CPU time of one simulation is about 30s.

	K1600	SORA	Stieng	AK-ECO
n_{calls}	0	16394	5754	305
Time	3h 28m 56s*	25h 40m 13s**	7h 53m 50s**	9h 35m 44s

Table F.5: Execution time of AK-ECO and the comparison methods

* The execution time indicated for K1600 takes into account only the resolution of the reformulated problem when the DoE of 1600 points is already computed and the metamodels are already calibrated. To obtain the DoE, 2h 10m 7s are necessary using 36 processors in parallel. The calibrations of the kriging models then take 35m 42s.

** For SORA and Stieng, to stay under the limitation of 5 days of computation allowed with the supercomputer used in IFPEN, we use 7 processors in parallel to evaluate the performance function.

For SORA, the simulations represent the main part of the total execution time. For Stieng, 23s are used for the calibrations of the metamodels, 1h 17m for the resolutions of the optimization problems, and the rest is devoted to the simulations used for the DoEs.

For AK-ECO, we do not use parallel computing. In Table F.6, we display the time dedicated to each step of AK-ECO to solve the FOWT problem.

Step of AK-ECO	Execution time	Proportion
Calibrations	1h 43m 22s	18.0%
Selections of x_{enr}	55m 03s	9.6%
COBYLA resolutions	3h 24m 33s	35.5 %
Stopping condition evaluations	55m 09s	9.6 %
Simulations	2h 32m 30s	26.5%
Total	9h 35m 44s	100%

Table F.6: Execution time of AK-ECO

3 Conclusion on the execution time

We observe that, to solve the oscillator problem, some of the comparison methods are faster than AK-ECO (in particular PMA and SORA). AK-ECO has been developed to solve problem where the execution time of one simulation is large. This is why the different approaches have been compared regarding the number of calls to the simulator in this thesis. However, even for relatively low simulation times as in the FOWT problem with 30 seconds per simulation, AK-ECO becomes faster than the comparison methods (remember that 7 processors are used with SORA and Stieng for this case). We can expect the gap between the execution times of the methods to widen as the simulation time increases since the total execution time of each approach will mainly depend on the number of calls to the simulator (which is much smaller with AK-ECO).

Appendix G

Results of the multistart resolutions

Contents

1 Results of the oscillator problem with multistart AK-ECO from different initial DoEs

In this section we display the result of the resolution of the oscillator problem described in section 3.3 with a multistart AK-ECO mentioned in section 3.3.4. For each initial design point d^0, a different initial DoE is used to calibrate the initial kriging models. We recall that the reference design point obtained with MC is $d^{min} = (5.0, 35.74)$ (see section 3.3.4). In Table G.1, we indicate for each initial design point d^0 of the multistart resolution, the design point d^{min} obtained and the value of the cost function at this point. For $i = 1, 2, 3$, the i-th reformulated failure probability at d^{min} is then estimated with a massive Monte Carlo of 30000 points. This estimation is denoted $p_i^{MC}(d^{min})$. The implementation of AK-ECO is described in section 3.3. The number of calls to the simulator for each resolution is denoted n_{call} and n_{cycles} corresponds to the number of cycle that AK-ECO required to converge.

run	d^0	d^{min}	$cost\left(d^{min}\right)$	$p_1^{MC}\left(d^{min}\right), p_2^{MC}\left(d^{min}\right), p_3^{MC}\left(d^{min}\right)$	n_{cycles}	n_{call}
1	$(2.17, 22.56)$	$(5, 35.70)$	-14.298	$0.81\times10^{-4}, 1.04\times10^{-4}, 0.15\times10^{-4}$	4	229
2	$(4.03, 31.12)$	$(5, 35.77)$	-14.225	$0.74\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	4	230
3	$(1.78, 44.36)$	$(5, 35.72)$	-14.284	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	6	232
4	$(4.26, 38.73)$	$(5, 35.71)$	-14.290	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	3	174
5	$(3.29, 48.59)$	$(5, 35.69)$	-14.313	$0.81\times10^{-4}, 1.03\times10^{-4}, 0.15\times10^{-4}$	4	181
6	$(3.80, 35.74)$	$(5, 35.76)$	-14.237	$0.75\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	4	220
7	$(1.51, 24.00)$	$(5, 35.72)$	-14.278	$0.80\times10^{-4}, 1.03\times10^{-4}, 0.15\times10^{-4}$	5	226
8	$(2.46, 37.99)$	$(5, 35.71)$	-14.287	$0.81\times10^{-4}, 1.03\times10^{-4}, 0.15\times10^{-4}$	5	210
9	$(3.46, 25.81)$	$(5, 35.75)$	-14.250	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	5	275
10	$(2.97, 29.54)$	$(5, 35.73)$	-14.268	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	5	254
11	$(3.77, 43.13)$	$(5, 35.72)$	-14.284	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	4	198
12	$(4.63, 47.07)$	$(5, 35.74)$	-14.259	$0.77\times10^{-4}, 1.00\times10^{-4}, 0.14\times10^{-4}$	4	193
13	$(4.43, 33.49)$	$(5, 35.76)$	-14.239	$0.75\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	3	185
14	$(3.03, 46.11)$	$(5, 35.73)$	-14.272	$0.77\times10^{-4}, 1.00\times10^{-4}, 0.14\times10^{-4}$	5	207
15	$(1.98, 40.32)$	$(5, 35.73)$	-14.274	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	5	230
16	$(2.34, 26.26)$	$(5, 35.69)$	-14.311	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	6	300
17	$(1.29, 33.93)$	$(5, 35.77)$	-14.227	$0.74\times10^{-4}, 0.98\times10^{-4}, 0.12\times10^{-4}$	9	416
18	$(1.04, 28.63)$	$(5, 35.73)$	-14.271	$0.76\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	4	193
19	$(4.82, 41.61)$	$(5, 35.72)$	-14.281	$0.81\times10^{-4}, 1.04\times10^{-4}, 0.15\times10^{-4}$	3	175
20	$(2.62, 21.29)$	$(5, 35.73)$	-14.268	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	5	254

Table G.1: Results of the oscillator problem with multistart AK-ECO from different initial DoEs

2 Results of the oscillator problem with multistart AK-ECO+GE

In this section we display the result of the resolution of the oscillator problem described in section 3.3 with a multistart AK-ECO+GE mentioned in section 4.3.3. The GE procedure requires 115 calls to the simulator (50 for the initial DoE and 65 for the enrichment of the metamodels). AK-ECO is then launched from 20 initial design points. For each resolution, the metamodels obtained with the GE are used during the first cycle of AK-ECO. We recall that the reference design point obtained with MC is $d^{min} = (5.0, 35.74)$ (see section 3.3.4). In Table G.2, we indicate for each initial design point d^0 of the multistart resolution, the design point d^{min} obtained and the value of the cost function at this point. For $i = 1, 2, 3$, the i-th reformulated failure probability at d^{min} is then estimated with a massive Monte

Carlo of 30000 points. This estimation is denoted $p_i^{MC}(d^{min})$. The implementation of AK-ECO is described in section 3.3. The number of additional calls to the simulator for each resolution is denoted n_{calls} and n_{cycles} corresponds to the number of cycle that AK-ECO required to converge. In Table G.3, the results of the same analysis but conducted without GE are displayed: the initial DoE used for each resolution is the same and it corresponds to the one used as initial DoE in the GE procedure.

run	d^0	d^{min}	$cost\left(d^{min}\right)$	$p_1^{MC}\left(d^{min}\right), p_2^{MC}\left(d^{min}\right), p_3^{MC}\left(d^{min}\right)$	n_{cycles}	n_{call}
1	$(2.17, 22.56)$	$(5, 35.78)$	-14.216	$0.74\times10^{-4}, 0.98\times10^{-4}, 0.13\times10^{-4}$	3	135
2	$(4.03, 31.12)$	$(5, 35.77)$	-14.227	$0.75\times10^{-4}, 0.99\times10^{-4}, 0.12\times10^{-4}$	3	135
3	$(1.78, 44.36)$	$(5, 35.77)$	-14.234	$0.76\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	3	107
4	$(4.26, 38.73)$	$(5, 35.74)$	-14.257	$0.77\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	3	114
5	$(3.29, 48.59)$	$(5, 35.71)$	-14.294	$0.79\times10^{-4}, 1.02\times10^{-4}, 0.14\times10^{-4}$	3	107
6	$(3.80, 35.74)$	$(5, 35.75)$	-14.245	$0.76\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	3	114
7	$(1.51, 24.00)$	$(5, 35.75)$	-14.252	$0.76\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	4	153
8	$(2.46, 37.99)$	$(5, 35.76)$	-14.239	$0.76\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	3	107
9	$(3.46, 25.81)$	$(5, 35.74)$	-14.261	$0.77\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	4	179
10	$(2.97, 29.54)$	$(5, 35.77)$	-14.234	$0.76\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	4	167
11	$(3.77, 43.13)$	$(5, 35.75)$	-14.246	$0.76\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	3	108
12	$(4.63, 47.07)$	$(5, 35.74)$	-14.261	$0.77\times10^{-4}, 1.00\times10^{-4}, 0.14\times10^{-4}$	3	109
13	$(4.43, 33.49)$	$(5, 35.83)$	-14.107	$0.72\times10^{-4}, 0.96\times10^{-4}, 0.12\times10^{-4}$	2	90
14	$(3.03, 46.11)$	$(5, 35.73)$	-14.272	$0.77\times10^{-4}, 1.00\times10^{-4}, 0.14\times10^{-4}$	3	107
15	$(1.98, 40.32)$	$(5, 35.77)$	-14.230	$0.75\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	4	152
16	$(2.34, 26.26)$	$(5, 35.76)$	-14.245	$0.76\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	4	163
17	$(1.29, 33.93)$	$(5, 35.76)$	-14.245	$0.76\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	4	152
18	$(1.04, 28.63)$	$(5, 35.77)$	-14.226	$0.75\times10^{-4}, 0.99\times10^{-4}, 0.13\times10^{-4}$	3	107
19	$(4.82, 41.61)$	$(5, 35.76)$	-14.235	$0.76\times10^{-4}, 1.00\times10^{-4}, 0.13\times10^{-4}$	3	114
20	$(2.62, 21.29)$	$(5, 35.78)$	-14.216	$0.75\times10^{-4}, 0.98\times10^{-4}, 0.13\times10^{-4}$	3	135

Table G.2: Results of the oscillator problem with multistart AK-ECO+GE

run	d^0	d^{min}	$cost\left(d^{min}\right)$	$p_1^{MC}\left(d^{min}\right), p_2^{MC}\left(d^{min}\right), p_3^{MC}\left(d^{min}\right)$	n_{cycles}	n_{call}
1	$(2.17, 22.56)$	$(5, 35.77)$	-14.232	$0.75 \times 10{-}4, 0.99 \times 10^{-4}, 0.13 \times 10^{-4}$	5	186
2	$(4.03, 31.12)$	$(4.40, 34.81)$	-9.204	$0.21 \times 10{-}4, 0.53 \times 10^{-4}, 0.02 \times 10^{-4}$	1	45
3	$(1.78, 44.36)$	$(5, 35.81)$	-14.190	$0.73 \times 10{-}4, 0.97 \times 10^{-4}, 0.12 \times 10^{-4}$	3	101
4	$(4.26, 38.73)$	$(5, 35.73)$	-14.268	$0.77 \times 10{-}4, 1.00 \times 10^{-4}, 0.14 \times 10^{-4}$	4	168
5	$(3.29, 48.59)$	$(5, 35.77)$	-14.229	$0.75 \times 10{-}4, 0.99 \times 10^{-4}, 0.13 \times 10^{-4}$	4	132
6	$(3.80, 35.74)$	$(5, 35.75)$	-14.255	$0.77 \times 10{-}4, 1.00 \times 10^{-4}, 0.13 \times 10^{-4}$	3	125
7	$(1.51, 24.00)$	$(5, 35.79)$	-14.208	$0.74 \times 10{-}4, 0.98 \times 10^{-4}, 0.13 \times 10^{-4}$	7	258
8	$(2.46, 37.99)$	$(5, 35.74)$	-14.257	$0.77 \times 10{-}4, 1.00 \times 10^{-4}, 0.13 \times 10^{-4}$	4	132
9	$(3.46, 25.81)$	$(5, 35.76)$	-14.243	$0.76 \times 10{-}4, 0.99 \times 10^{-4}, 0.13 \times 10^{-4}$	5	212
10	$(2.97, 29.54)$	$(5, 35.79)$	-14.210	$0.74 \times 10{-}4, 0.98 \times 10^{-4}, 0.13 \times 10^{-4}$	5	219
11	$(3.77, 43.13)$	$(5, 35.69)$	-14.307	$0.79 \times 10{-}4, 1.02 \times 10^{-4}, 0.14 \times 10^{-4}$	3	104
12	$(4.63, 47.07)$	$(5, 35.77)$	-14.233	$0.75 \times 10{-}4, 0.99 \times 10^{-4}, 0.13 \times 10^{-4}$	3	116
13	$(4.43, 33.49)$	$(5, 35.82)$	-14.182	$0.73 \times 10{-}4, 0.97 \times 10^{-4}, 0.12 \times 10^{-4}$	2	90
14	$(3.03, 46.11)$	$(5, 35.74)$	-14.261	$0.77 \times 10{-}4, 1.00 \times 10^{-4}, 0.13 \times 10^{-4}$	5	165
15	$(1.98, 40.32)$	$(5, 35.80)$	-14.204	$0.74 \times 10{-}4, 0.98 \times 10^{-4}, 0.12 \times 10^{-4}$	3	98
16	$(2.34, 26.26)$	$(5, 35.78)$	-14.222	$0.75 \times 10{-}4, 0.98 \times 10^{-4}, 0.13 \times 10^{-4}$	5	207
17	$(1.29, 33.93)$	$(5, 35.76)$	-14.238	$0.76 \times 10{-}4, 0.99 \times 10^{-4}, 0.13 \times 10^{-4}$	4	139
18	$(1.04, 28.63)$	$(5, 35.72)$	-14.284	$0.78 \times 10{-}4, 1.01 \times 10^{-4}, 0.14 \times 10^{-4}$	4	136
19	$(4.82, 41.61)$	$(5, 35.77)$	-14.229	$0.75 \times 10{-}4, 0.99 \times 10^{-4}, 0.13 \times 10^{-4}$	3	126
20	$(2.62, 21.29)$	$(5, 35.74)$	-14.263	$0.77 \times 10{-}4, 1.00 \times 10^{-4}, 0.14 \times 10^{-4}$	6	209

Table G.3: Results of the oscillator problem with multistart AK-ECO from the same initial DoE

In Table G.3, we observe that one resolution (run 2) does not converge towards the optimum. This is not due to the enrichment of AK-ECO but to the stopping condition of the optimization cycles: although the design point obtained at the end of the first cycle is far from the initial design point, the values of the cost function at the two points are equal (both points belong to the same level set). Therefore, the stopping condition is met.

3 Results of the FOWT problem with multistart AK-ECO from the same initial DoE

In this section we display the result of the resolution of the FOWT problem with the multistart AK-ECO mentioned in section 5.6.4. The same initial DoE of 60 points is used for every resolution. In Table G.4, we indicate for each initial design point d^0 of the multistart resolution, the design point d^{min} obtained and the value of the normalized cost function

at this point. The estimations of the reformulated failure probabilities $p_{\mathcal{S}}^{K1600}\left(d^{min}\right)$ and $p_{\mathcal{D}^3}^{K1600}\left(d^{min}\right)$ are indicated (see section 5.6.4 for more details). The estimations of the other reformulated failure probabilities at d^{min} ($p_{\mathcal{T}^l}^{K1600}\left(d^{min}\right)$ ($l = 1, 2, 3$), $p_{\mathcal{D}^1}^{K1600}\left(d^{min}\right)$, and $p_{\mathcal{D}^2}^{K1600}\left(d^{min}\right)$) obtained by each resolution are equal to zero. The implementation of AK-ECO is described in section 5.6. The number of additional calls to the simulator for each resolution is denoted n_{call} and n_{cycles} corresponds to the number of cycle that AK-ECO required to converge.

run	d^0	d^{min}	$\overline{cost\left(d^{min}\right)}$	$p_{\mathcal{S}}^{K1600}\left(d^{min}\right), p_{\mathcal{D}^3}^{K1600}\left(d^{min}\right)$	n_{cycles}	n_{call}
1	$(0.54, 116.13, 0.41)$	$(1.09, 110.55, 0)$	0.294	$1.01 \times 10^{-4}, 0.86 \times 10^{-4}$	8	280
2	$(0.30, 170.80, 0.51)$	$(0.95, 109.51, 0)$	0.287	$1.00 \times 10^{-4}, 0.95 \times 10^{-4}$	8	252
3	$(1.93, 75.12, 0.03)$	$(0.98, 109.75, 0)$	0.288	$0.93 \times 10^{-4}, 0.92 \times 10^{-4}$	11	371
4	$(1.29, 130.24, 0.82)$	$(1.13, 110.75, 0)$	0.296	$1.27 \times 10^{-4}, 0.84 \times 10^{-4}$	6	196
5	$(1.70, 85.56, 0.75)$	$(1.17, 111.13, 0)$	0.299	$1.02 \times 10^{-4}, 0.82 \times 10^{-4}$	8	252
6	$(1.04, 157.11, 0.24)$	$(0.91, 109.27, 0)$	0.285	$0.94 \times 10^{-4}, 0.96 \times 10^{-4}$	8	253
7	$(-0.45, 97.89, 0.57)$	$(1.1, 109.93, 0)$	0.290	$0.96 \times 10^{-4}, 0.91 \times 10^{-4}$	14	476
8	$(-0.17, 142.31, 0.16)$	$(0.99, 109.83, 0)$	0.289	$1.00 \times 10^{-4}, 0.92 \times 10^{-4}$	6	193
9	$(0.84, 154.63, 0.93)$	$(0.93, 109.39, 0)$	0.286	$0.97 \times 10^{-4}, 0.95 \times 10^{-4}$	11	378

Table G.4: Results of the FOWT problem with multistart AK-ECO from the same initial DoE

Appendix H

Comparison between time-domain and spectral approaches

Contents

In this appendix, we consider three models:

- in model 1, the equation of motion is not linearized and a turbulent wind corresponding to a class B wind turbine (IEC 61400-3 (2009)) is considered with an average wind speed equal to that considered for the other models;
- the equation of motion is not linearized in model 2 but constant wind forces are considered for a class B wind turbine;
- the non-linear external forces are linearized and constant wind forces are considered in model 3. It corresponds to the model used in this thesis and introduced in section 5.1.

For each model, we consider the nominal configuration of the structure.

In section H.1, we compare time series obtained with each model to examine the influence of the model choice.

In section H.2, the total damage is estimated for every sea state and for each model.

1 Comparison of the time series

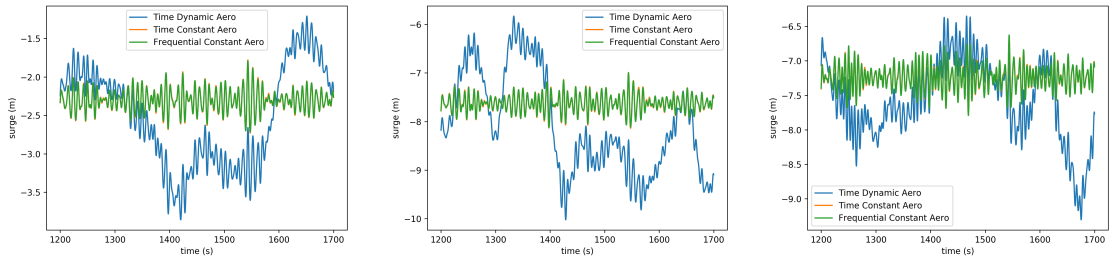
A realization of the sea elevation process is considered for each sea state s^j $(j = 1, \cdots, 7)$ introduced in section 5.2.5. For each sea state and for each model, the corresponding realizations of the surge and the tensions processes are evaluated. The time series considering model 1 and 2 are computed with a time-domain solver in Deelpines^{TM}. For the former, the wind loads are estimated using the Blade Element Momentum (Blondel et al. (2016)). For model 3, these time series are generated with a frequency-domain solver (Le Cunff et al. (2008)) in Deelpines^{TM}.

The resulting time series are displayed in Figures H.1, H.2, H.3, and H.4. Only the time series for the sea states s^1, s^4, and s^7 are indicated. Moreover, to observe more clearly the difference in the time series, only the time window $[1200\text{s}, 1700\text{s}]$ is visible. The power spectral densities (obtained as the Fourier transforms of the temporal autocorrelation functions) of the time series are evaluated with the psd function of the matplotlib Python package and a NFFT parameter set to 1024. The results are displayed as well in Figures H.5, H.6, H.7, and H.8.

In the legend of each figure, the results obtained with model 1, 2, and 3 are respectively labelled "Time Dynamic Aero", "Time Constant Aero", and "Frequential Constant Aero".

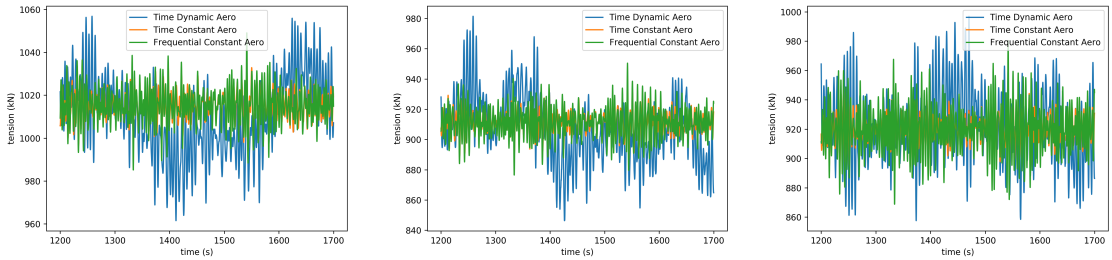
We notice from this analysis the following elements:

- the results obtained with model 2 and model 3 are close: the times series overlap for every figure. Therefore, the linearization of the external forces only slightly impacts the outputs of the model. However, considering a constant wind load significantly reduces the amplitude of the time series variations, as shown by the series obtained considering model 1;
- for model 1, we observe in the tension PSDs, peaks corresponding to the rotor speed (mean about 12 rpm thus 0.2Hz) harmonics. We observe peaks at t_p and close to the rotor frequency ;
- the difference between the PSDs obtained with model 1 and the ones considering a constant wind loading is clearly visible at high frequencies.



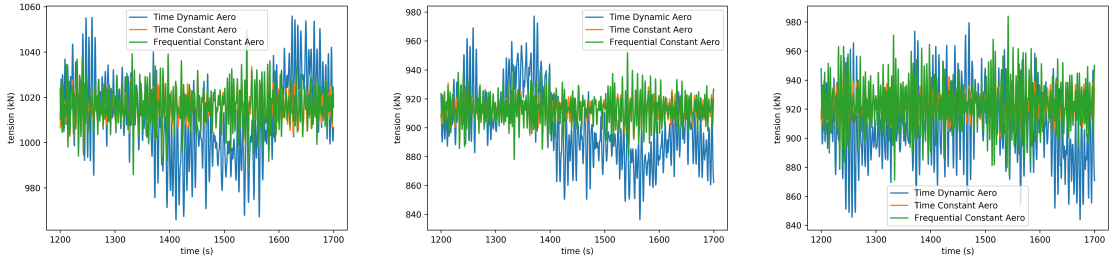
(a) Surge time series for s^1 (b) Surge time series for s^4 (c) Surge time series for s^7

Figure H.1: Time series of the surge



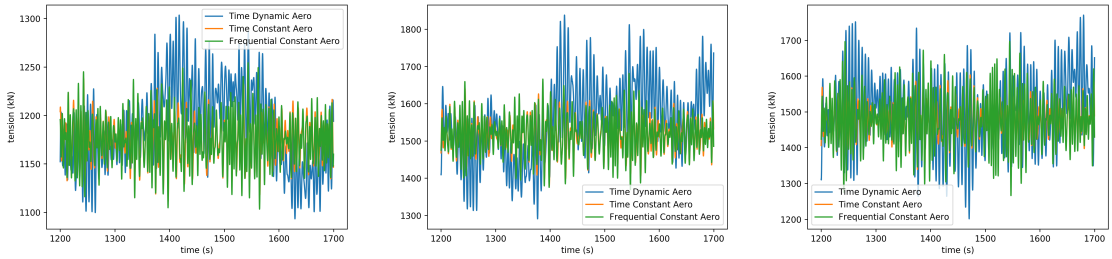
(a) Tension time series for s^1 (b) Tension time series for s^4 (c) Tension time series for s^7

Figure H.2: Time series of the tension at the top at the line 1



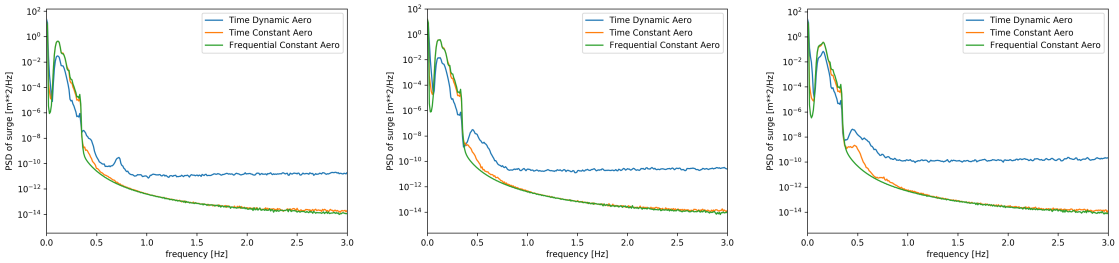
(a) Tension time series for s^1 (b) Tension time series for s^4 (c) Tension time series for s^7

Figure H.3: Time series of the tension at the top at the line 2



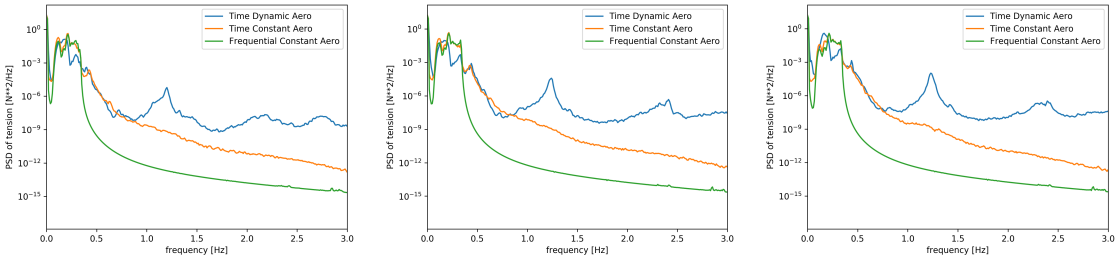
(a) Tension time series for s^1 (b) Tension time series for s^4 (c) Tension time series for s^7

Figure H.4: Time series of the tension at the top at the line 3



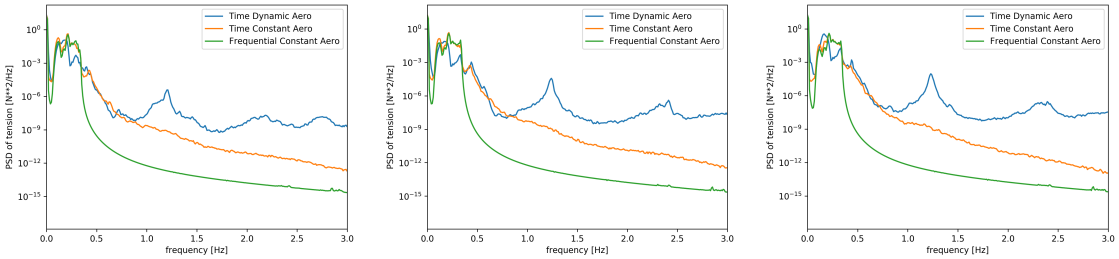
(a) Surge psd for s^1 (b) Surge psd for s^4 (c) Surge psd for s^7

Figure H.5: PSD of the centered and normalized surge



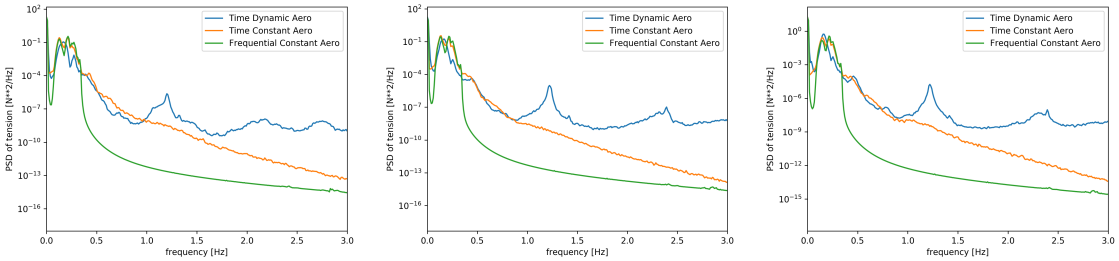
(a) Tension psd for s^1 (b) Tension psd for s^4 (c) Tension psd for s^7

Figure H.6: PSD of the centered and normalized tension at the top at the line 1



(a) Tension psd for s^1 (b) Tension psd for s^4 (c) Tension psd for s^7

Figure H.7: PSD of the centered and normalized tension at the top at the line 2



(a) Tension psd for s^1 (b) Tension psd for s^4 (c) Tension psd for s^7

Figure H.8: PSD of the centered and normalized tension at the top at the line 3

2 Comparison of the fatigue estimations

In this section, we estimate for each model and for each sea state s^j $(j = 1, \cdots, 7)$, the mean total damage caused by the tension at the top of each line during a time interval of one year. This quantity is denoted in section 5.3 as follows:

$$\mathbb{E}\left[\int_{I_i} \mathcal{D}_i^l\left(d, x_{d_2}, x_p, s_i; t\right) dt\right] = \Delta T \mathbb{E}\left[\mathcal{D}_i^l\left(d, x_{d_2}, x_p, s_i; 0\right)\right] = \mathbb{E}_{\eta_i}\left[D_{I_i}^{total,l}\left(d, x_{d_2}, x_p, s_i\right)\right]. \tag{H.1}$$

We point out that, in the thesis, $\Delta T = 3$ hours whereas, in the analyses performed in this appendix, we consider $\Delta T = 1$ year. Furthermore, the parameters d, x_{d_2}, and x_p are fixed at their nominal values.

The damage is first estimated with the time-domain approach and then with the spectral approach described in section 5.3. More precisely, in Figures H.9, H.10, and H.11, we indicate for each sea states s^j:

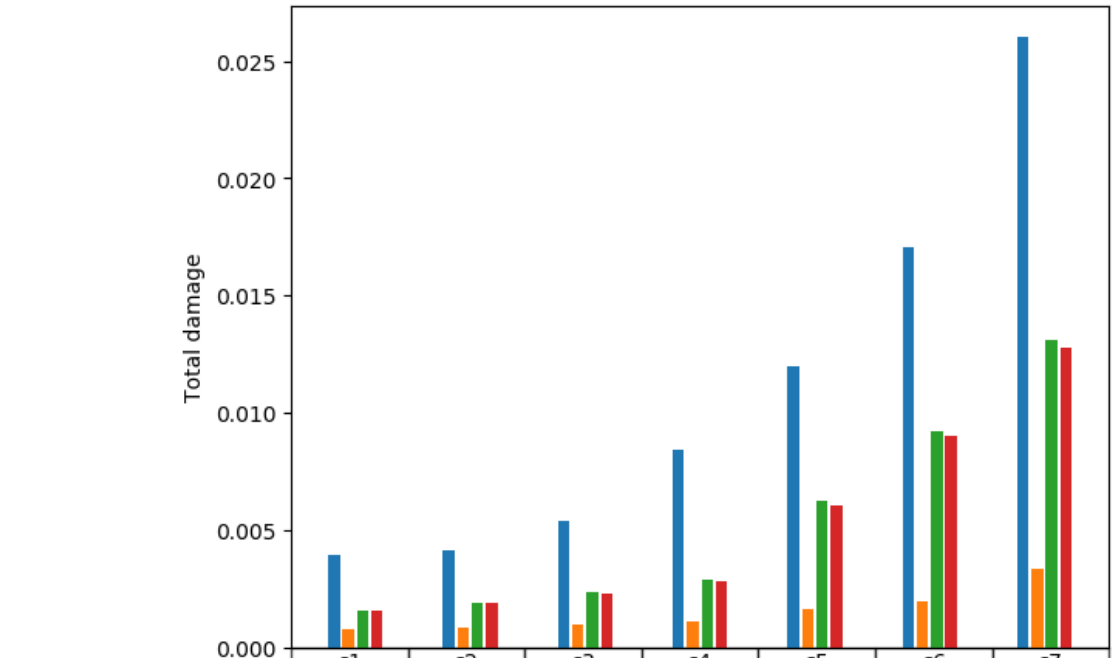
- the total damage estimated with the time-domain approach and considering successively model 1, 2, and 3. In the figures, the results are respectively labelled "Aero + Rainflow", " ConstWind + Rainflow", and "Freq + Rainflow";
- the total damage estimated with the spectral approach (i.e. the Dirlik method) and considering model 3. This estimation is labelled " Freq + Dirlik" in the figures.

For the time-domain approach, we use the Rainflow algorithm implemented by Jennifer Rinker of Duke University (the Python script is distributed with GNU GPL at

https://gist.github.com/jennirinker).

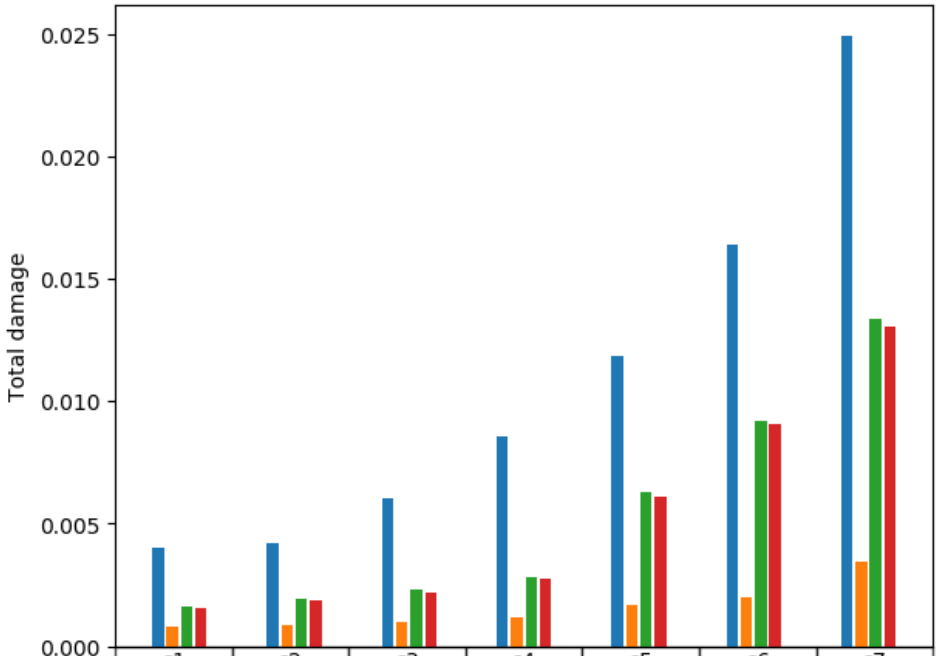
This analysis reveals that:

- the total damage considering model 1 and estimated with the time-domain approach is always greater that the estimations provided with the other models. We conclude that considering constant wind loads leads to underestimate the fatigue;
- the total damage estimations with the time-domain and the Dirlik method considering model 3 are close;
- a discrepancy persists between the estimated damage considering model 2 and model 3 but the gap decreases as the fatigue decreases. This point can be explained by slight differences in the areas of the PSDs peaks. It is assumed that this difference shows the limit of the quasi-linearization approach of Morison forces when the swell amplitude increases. This discrepancy could also be due to the fact that, to avoid non-linearities in the equation of motion in model 3, unlike in model 1 and 2, the part of the mooring lines in contact with the seabed is considered fixed. This effect is partially reduced by the small probability of occurrence of the most severe sea states.
- finally, line 3 (which corresponds to the line aligned with the wind direction) suffers more damage than the other two lines.



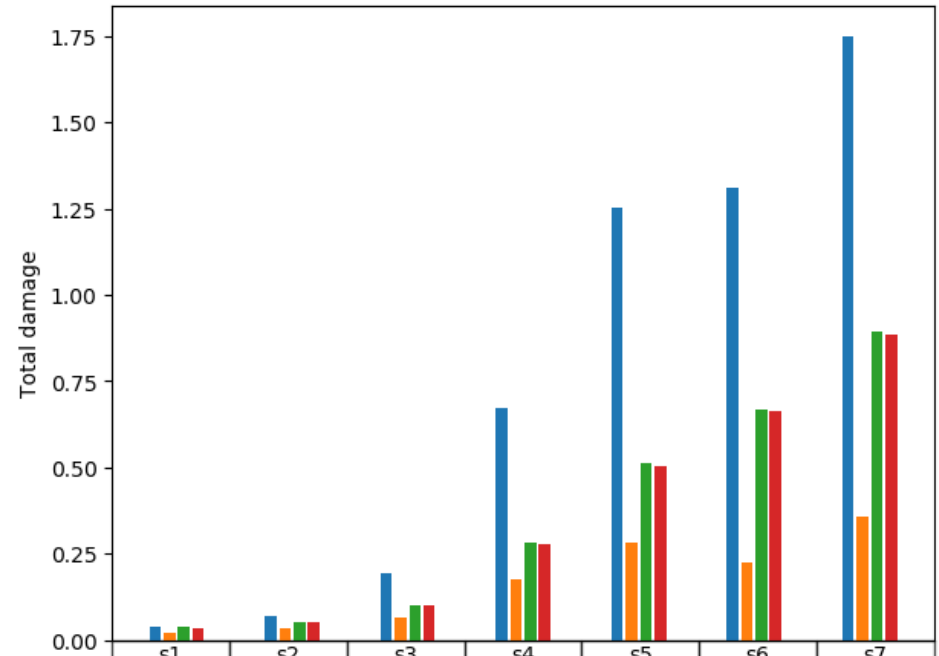
	s1	s2	s3	s4	s5	s6	s7
Aero + Rainflow	0.00396	0.00413	0.0054	0.00844	0.01199	0.01704	0.02603
ConstWind + Rainflow	0.00077	0.00086	0.00099	0.0011	0.00161	0.00194	0.00332
Freq + Rainflow	0.00157	0.00191	0.00233	0.00288	0.00622	0.00918	0.01307
Freq + Dirlik	0.00154	0.00186	0.00226	0.00281	0.00603	0.00902	0.01278

Figure H.9: Comparison of the one year total damage for line 1



	s1	s2	s3	s4	s5	s6	s7
Aero + Rainflow	0.00401	0.00423	0.00602	0.00859	0.01182	0.01641	0.02491
ConstWind + Rainflow	0.00077	0.00088	0.001	0.00115	0.0017	0.00202	0.00344
Freq + Rainflow	0.00159	0.00193	0.00228	0.00282	0.0063	0.00921	0.01335
Freq + Dirlik	0.00156	0.00188	0.00221	0.00276	0.00611	0.00905	0.01306

Figure H.10: Comparison of the one year total damage for line 2



	s1	s2	s3	s4	s5	s6	s7
Aero + Rainflow	0.03687	0.07004	0.19358	0.67314	1.25255	1.31149	1.7485
ConstWind + Rainflow	0.02153	0.03175	0.06668	0.1763	0.28107	0.22562	0.35844
Freq + Rainflow	0.03612	0.05151	0.10184	0.2811	0.51006	0.66729	0.89451
Freq + Dirlik	0.03549	0.05086	0.10034	0.27526	0.50312	0.66086	0.8864

Figure H.11: Comparison of the one year total damage for line 3

Appendix I

Reformulation of the surge constraint

The surge failure probability of the FOWT reformulated problem is expressed as:

$$\begin{aligned}\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]}\left|\mathcal{S}\left(d,X_p;t\right)\right|>\mathcal{S}_{\max}\right) &= \mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]}\mathcal{S}\left(d,X_p;t\right)>\mathcal{S}_{\max}\right)\\ &+\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]}-\mathcal{S}\left(d,X_p;t\right)>\mathcal{S}_{\max}\right)\\ &-\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]}\mathcal{S}\left(d,X_p;t\right)>\mathcal{S}_{\max},\max_{t\in[0,T]}-\mathcal{S}\left(d,X_p;t\right)>\mathcal{S}_{\max}\right).\end{aligned} \tag{I.1}$$

Since the surge process satisfies the sufficient conditions (see appendix J), we can reformulate the first term of the sum of the right-hand side of equation (I.1) as follows:

$$\begin{aligned}&\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]}\mathcal{S}\left(d,X_p;t\right)>\mathcal{S}_{\max}\right)\\ &\simeq\mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(a_{Tp^j}\left(d,X_p,s^j\right)^2-\frac{a_{Tp^j}\left(d,X_p,s^j\right)\left(\mathcal{S}_{\max}-\mu_{\mathcal{S}}\left(d,X_p,s^j\right)\right)}{\sqrt{m_{\mathcal{S},0}\left(d,X_p,s^j\right)}}\right)\right)\right]\end{aligned} \tag{I.2}$$

with $a_{Tp^j}\left(d,x_p,s^j\right)=\sqrt{2\log\left(\frac{Tp^j}{2\pi}\sqrt{\frac{m_{\mathcal{S},2}\left(d,x_p,s^j\right)}{m_{\mathcal{S},0}\left(d,x_p,s^j\right)}}\right)}$.

Let us denote

$$M_{\mathcal{S}}^{+}\left(d,x_p,s^j\right)=a_{Tp^j}\left(d,x_p,s^j\right)^2-\frac{a_{Tp^j}\left(d,x_p,s^j\right)\left(\mathcal{S}_{\max}-\mu_{\mathcal{S}}\left(d,x_p,s^j\right)\right)}{\sqrt{m_{\mathcal{S},0}\left(d,x_p,s^j\right)}}. \tag{I.3}$$

We have sampled 1600 points of the input space of $M_{\mathcal{S}}^{+}$ and the simulator has been called at each one to evaluate $M_{\mathcal{S}}^{+}$. This sample is the one used to calibrate the kriging models in the MC+K1600 method (see section 5.6.2 for more details). The maximum value of $M_{\mathcal{S}}^{+}$ obtained from this sample is −389.05 (denoted $M_{\mathcal{S},\max}^{+}$). Assuming that this is the maximum value of $M_{\mathcal{S}}^{+}$ over its input space, we have:

$$\mathbb{P}_{X_p,\eta}\left(\max_{t\in[0,T]}\mathcal{S}\left(d,X_p;t\right)>\mathcal{S}_{\max}\right)\simeq\mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(M_{\mathcal{S}}^{+}\left(d,X_p,s^j\right)\right)\right)\right] \tag{I.4}$$

$$\leq\mathbb{E}_{X_p}\left[F_\epsilon\left(\sum_{j=1}^{7}\exp\left(M_{\mathcal{S},\max}^{+}\right)\right)\right] \tag{I.5}$$

$$=F_\epsilon\left(7\exp\left(M_{\mathcal{S},\max}^{+}\right)\right) \tag{I.6}$$

$$=F_\epsilon\left(7\times(-389.05)\right) \tag{I.7}$$

and $F_\epsilon\left(7\times(-389.05)\right)$ is numerically null. Therefore, it is reasonable to consider that the probability in equation (I.2) is negligible in the estimation of the failure probability of the surge constraint for every $d\in\Omega_d$ which justifies the approximation of equation (5.50).

Appendix J

Proof of the sufficient conditions to reformulate the surge and tension constraints

To apply the reformulation procedure of extreme-based constraints involving a piece-wise stationary process described in section 2.3.1 to the surge and the tension constraints, for all d, x_p, and for all states s^j, the processes $\mathcal{S}_1\left(d, x_p, s^j; .\right)$ and $\mathcal{T}_1^l\left(d, x_p, s^j; .\right)$ $(l = 1, 2, 3)$ must meet the conditions of theorem 2.1. Thus, it is sufficient to prove that these processes satisfy the conditions (2.11) and (2.12) presented in section 2.3.1.

For the surge process, it is sufficient to show that for all d, x_p, and s^j:

$$m_{\mathcal{S},0}\left(d, x_p, s^j\right) < \infty, m_{\mathcal{S},2}\left(d, x_p, s^j\right) < \infty, \tag{J.1}$$

$$K_{\mathcal{S}}\left(d, x_p, s^j; .\right) \in C^1, K_{\mathcal{S}}\left(d, x_p, s^j; .\right) \text{ and } K'_{\mathcal{S}}\left(d, x_p, s^j; .\right) \text{ are integrable.} \tag{J.2}$$

where $\omega \to K'_{\mathcal{S}}\left(d, x_p, s^j; \omega\right)$ is the derivative of $\omega \to K_{\mathcal{S}}\left(d, x_p, s^j; \omega\right)$.

We recall that

$$K_{\mathcal{S}}\left(d, x_p, s^j; \omega\right) = \left|H_{\mathcal{S}}\left(d, x_p, s^j; \omega\right)\right|^2 K_\eta(s^j; \omega), \ \forall \omega \tag{J.3}$$

where $K_\eta(s^j; .)$ and $H_{\mathcal{S}}\left(d, x_p, s^j; .\right)$ are defined in section 5.2.5. Assuming that there is no resonance phenomena and that $\forall \omega$, $\left|H_{\mathcal{S}}\left(d, x_p, s^j; \omega\right)\right| < c_1$ with $c_1 > 0$, it follows, for $n = 0$ and 2:

$$\left|\int_{\mathbb{R}} \omega^n K_{\mathcal{S}}\left(d, x_p, s^j; \omega\right) d\omega\right| \leq c_1 \int_{\mathbb{R}} \left|\omega^n K_\eta\left(s^j; \omega\right)\right| d\omega \tag{J.4}$$

$$\leq c_1 \int_{\mathbb{R}^+} \left|\omega^n 2K_\eta^+\left(s^j; \omega\right)\right| d\omega \tag{J.5}$$

$$\leq c_1\, c_2 \int_{\mathbb{R}^+} \left|\omega^{-5+n} \exp\left(-c_3\omega^{-4}\right)\right| d\omega \tag{J.6}$$

$$\leq c_1\, c_2 \int_{\mathbb{R}^+} \left|\omega^{-5+n} \exp\left(-c_3\omega^{-4}\right)\right| d\omega \tag{J.7}$$

$$\leq c_1\, c_2 \left(\int_0^1 \left|\omega^{-5+n} \exp\left(-c_3\omega^{-4}\right)\right| d\omega + \int_1^\infty \left|\omega^{-5+n}\right| d\omega\right) \tag{J.8}$$

with $c_2 > 0$, $c_3 > 0$ and $K_\eta^+\left(s^j; .\right)$ is defined in section 5.2.5. In equation (J.8), the integrand of the first integral is a bounded function over $[0, 1]$ and the integrand of the second

integral is integrable over $[1, \infty]$ since $-5 + n < -1$. Therefore, condition (J.1) is met and $K_{\mathcal{S}}\left(d, x_p, s^j; .\right)$ is integrable.

Moreover, using the same arguments and assuming that $H_{\mathcal{S}}\left(d, x_p, s^j; .\right)$ is differentiable and that its derivative w.r.t to ω is bounded, we can show that $K'_{\mathcal{S}}\left(d, x_p, s^j; .\right)$ is integrable. Therefore condition (J.2) is satisfied as well.

Similar calculations and assumptions enable us to show that the tension processes meet the sufficient conditions for the reformulation of the tension constraints.

Bibliography

Abdo, T. and Rackwitz, R. (1991). A New Beta-Point Algorithm for Large Time-Invariant and Time-Variant Reliability Problems. In *Reliability and Optimization of Structural Systems '90*, pages 1–12, Berlin, Heidelberg. Springer Berlin Heidelberg.

Ahmadivala, M., Mattrand, C., Gayton, N., Dumas, A., Yalamas, T., and Orcesi, A. (2019). Application of AK-SYS method for time-dependent reliability analysis. In *Congrès Français de Mécanique*.

Alizadeh, R., Allen, J. K., and Mistree, F. (2020). Managing computational complexity using surrogate models: a critical review. *Research in Engineering Design*, 31(3):275–298.

Andrieu-Renaud, C., Sudret, B., and Lemaire, M. (2004). The PHI2 method: a way to compute time-variant reliability. *Reliability Engineering & System Safety*, 84(1):75–86.

Aoues, Y. and Chateauneuf, A. (2010). Benchmark study of numerical methods for reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 41(2):277–294.

Asmussen, S., Jensen, J. L., and Rojas-Nandayapa, L. (2011). A Literature Review on Log-Normal Sums. Technical report, University of Queensland.

Au, S. K. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277.

Azaïs, J. M. and Wschebor, M. (2009). *Level sets and extrema of random processes and fields*. John Wiley & Sons.

Barrera, C., Battistella, T., Guanche, R., and Losada, I. J. (2020). Mooring system fatigue analysis of a floating offshore wind turbine. *Ocean Engineering*, 195:106670.

Barrera, J., Homem-de-Mello, T., Moreno, E., Pagnoncelli, B. K., and Canessa, G. (2016). Chance-constrained problems and rare events: an importance sampling approach. *Mathematical Programming*, 157(1):153–189.

Baudin, M., Dutfoy, A., Iooss, B., and Popelin, A. L. (2016). *OpenTURNS: An Industrial Software for Uncertainty Quantification in Simulation*, pages 1–38. Springer International Publishing, Cham.

Benasciutti, D. and Tovo, R. (2004). Rainflow cycle distribution and fatigue damage in Gaussian random loadings. *Internal Report No. 129, Department of Engineering, University of Ferrara, Italy*.

Berman, S. M. (1991). Spectral conditions for sojourn and extreme value limit theorems for Gaussian processes. *Stochastic Processes and their Applications*, 39(2):201–220.

Bhattacharyya, B. (2020). Global sensitivity analysis: A Bayesian learning based polynomial chaos approach. *Journal of Computational Physics*, 415:109539.

Bichon, B. J. (2010). *Efficient surrogate modeling for reliability analysis and design*. PhD thesis, Vanderbilt University.

Biswas, R. and Sharma, D. (2021). A single-loop shifting vector method with conjugate gradient search for reliability-based design optimization. *Engineering Optimization*, 53(6):1044–1063.

Blondel, F., Boisard, R., Milekovic, M., Ferrer, G., Lienard, C., and Teixeira, D. (2016). Validation and comparison of aerodynamic modelling approaches for wind turbines. *Journal of Physics: Conference Series*, 753:022029.

Borgman, L. E. (1967). Random Hydrodynamic Forces on Objects. *The Annals of Mathematical Statistics*, 38(1):37 – 51.

Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784.

Bouhlel, M. A., Bartoli, N., Otsmane, A., and Morlier, J. (2016). Improving kriging surrogates of high-dimensional design models by partial least squares dimension reduction. *Structural and Multidisciplinary Optimization*, 53(5):935–952.

Bourinet, J. M. (2018). *Reliability analysis and optimal design under uncertainty-Focus on adaptive surrogate-based approaches*. HDR, Université Clermont Auvergne.

Breitung, K. (1984). Asymptotic approximations for multinormal integrals. *Journal of Engineering Mechanics*, 110(3):357–366.

Breitung, K. (1989). Asymptotic approximations for probability integrals. *Probabilistic Engineering Mechanics*, 4(4):187–190.

Buhmann, M. D. (2003). *Radial Basis Functions: Theory and Implementations*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press.

Burmester, S., Vaz, G., Gueydon, S., and el Moctar, O. (2020). Investigation of a semi-submersible floating wind turbine in surge decay using CFD. *Ship Technology Research*, 67(1):2–14.

Camus, P., Mendez, F. J., Medina, R., and Cofiño, A. S. (2011). Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coastal Engineering*, 58(6):453–462.

Carbon Trust (2015). *Floating Offshore Wind: Market and Technology Review*.

Chabridon, V. (2018). *Reliability-oriented sensitivity analysis under probabilistic model uncertainty–Application to aerospace systems*. PhD thesis, Université Clermont Auvergne.

Chakrabarti, S. K. (1987). *Hydrodynamics of offshore structures*. WIT press.

Chaudhuri, A., Kramer, B., and Willcox, K. E. (2020). Information Reuse for Importance Sampling in Reliability-Based Design Optimization. *Reliability Engineering & System Safety*, 201:106853.

Cheng, G., Xu, L., and Jiang, L. (2006). A sequential approximate programming strategy for reliability-based structural optimization. *Computers & Structures*, 84(21):1353–1367.

Cheng, K. and Lu, Z. (2020). Active learning polynomial chaos expansion for reliability analysis by maximizing expected indicator function prediction error. *International Journal for Numerical Methods in Engineering*, 121(14):3159–3177.

Chevreuil, M., Lebrun, R., Nouy, A., and Rai, P. (2015). A Least-Squares Method for Sparse Low Rank Approximation of Multivariate Functions. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):897–921.

Cui, L., Lü, Z., and Zhao, X. (2010). Moment-independent importance measure of basic random variable and its probability density evolution solution. *Science China Technological Sciences*, 53(4):1138–1145.

Cukier, R. I., Levine, H. B., and Shuler, K. E. (1978). Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26(1):1–42.

Da Veiga, S. (2015). Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305.

Det Norske Veritas (2005). Fatigue design of offshore steel structures. *Offshore Standard DNV-RP-C203*.

Det Norske Veritas (2013). Design of floating wind turbine structures. *Offshore Standard DNV-OS-J103*.

Dirlik, T. (1985). *Application of computers in fatigue analysis*. PhD thesis, University of Warwick.

Downing, S. D. and Socie, D. F. (1982). Simple rainflow counting algorithms. *International Journal of Fatigue*, 4(1):31–40.

Du, W., Wang, Y., and Luo, Y. (2020). A reliability-based fatigue design for mechanical components under material variability. *Quality and Reliability Engineering International*, 36(1):388–402.

Du, X. and Chen, W. (2004). Sequential Optimization and Reliability Assessment Method for Efficient Probabilistic Design. *Journal of Mechanical Design*, 126(2):225–233.

Du, X., Sudjianto, A., and Chen, W. (2004). An Integrated Framework for Optimization Under Uncertainty Using Inverse Reliability Strategy . *Journal of Mechanical Design*, 126(4):562–570.

Duarte, T. M., Sarmento, A. J., and Jonkman, J. (2014). Effects of Second-Order Hydrodynamic Forces on Floating Offshore Wind Turbines. In *32nd ASME Wind Energy Symposium*.

Dubourg, V. (2011). *Adaptive surrogate models for reliability analysis and reliability-based design optimization*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II.

Echard, B. (2012). *Assessment by kriging of the reliability of structures subjected to fatigue stress*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II.

Echard, B., Gayton, N., and Lemaire, M. (2011). AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. *Structural Safety*, 33(2):145–154.

Efron, B. and Stein, C. (1981). The Jackknife Estimate of Variance. *The Annals of Statistics*, 9(3):586 – 596.

Ehre, M., Papaioannou, I., and Straub, D. (2020). Global sensitivity analysis in high dimensions with PLS-PCE. *Reliability Engineering & System Safety*, 198:106861.

Eskandari, H. and Kim, H. S. (2017). *A Theory for Mathematical Framework and Fatigue Damage Function for the S-N Plane*, pages 299–336. ASTM International, West Conshohocken, PA.

Fauriat, W. and Gayton, N. (2014). AK-SYS: An adaptation of the AK-MCS method for system reliability. *Reliability Engineering & System Safety*, 123:137–144.

Fort, J. C., Klein, T., and Rachdi, N. (2016). New sensitivity analysis subordinated to a contrast. *Communications in Statistics - Theory and Methods*, 45(15):4349–4364.

Goswami, S., Chakraborty, S., Chowdhury, R., and Rabczuk, T. (2019). Threshold shift method for reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 60(5):2053–2072.

Gretton, A., Bousquet, O., Smola, A., and Schölkopf, B. (2005). Measuring statistical dependence with hilbert-schmidt norms. In *Algorithmic Learning Theory*, pages 63–77, Berlin, Heidelberg. Springer Berlin Heidelberg.

Guo, X. and Dias, D. (2020). Kriging based reliability and sensitivity analysis – Application to the stability of an earth dam. *Computers and Geotechnics*, 120:103411.

Hall, M. T. J. (2013). *Mooring line modelling and design optimization of floating offshore wind turbines*. PhD thesis, University of Victoria.

Hasofer, A. M. and Lind, N. (1974). Exact and invariant second-moment code format. *Journal of Engineering Mechanics*, 100:111–121.

Hasselmann, D. E., Dunckel, M., and Ewing, J. A. (1980). Directional wave spectra observed during JONSWAP 1973. *Journal of Physical Oceanography*, 10(8):1264–1280.

Hawchar, L. (2017). *Développement de méthodes fiabilistes dépendant du temps pour l'analyse de durabilité des structures: application au problème de conception fiabiliste dépendant du temps*. PhD thesis, Nantes.

Hawchar, L., El Soueidy, C. P., and Schoefs, F. (2018). Global kriging surrogate modeling for general time-variant reliability-based design optimization problems. *Structural and Multidisciplinary Optimization*, 58(3):955–968.

He, L., Lu, Z., and Feng, K. (2020). A novel estimation method for failure-probability-based-sensitivity by conditional probability theorem. *Structural and Multidisciplinary Optimization*, 61(4):1589–1602.

Helbert, C., Dupuy, D., and Carraro, L. (2009). Assessment of uncertainty in computer experiments from Universal to Bayesian Kriging. *Applied Stochastic Models in Business and Industry*, 25(2):99–113.

Hoeffding, W. (1992). *A Class of Statistics with Asymptotically Normal Distribution*, pages 308–334. Springer New York, New York, NY.

Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17.

Hou, G. J. W. (2004). A Most Probable Point-Based Method for Reliability Analysis, Sensitivity Analysis and Design Optimization. Technical report, NASA.

Hu, Y., Lu, Z., Wei, N., and Zhou, C. (2020). A single-loop Kriging surrogate model method by considering the first failure instant for time-dependent reliability analysis and safety lifetime analysis. *Mechanical Systems and Signal Processing*, 145:106963.

Hu, Z. and Du, X. (2015). Mixed Efficient Global Optimization for Time-Dependent Reliability Analysis. *Journal of Mechanical Design*, 137(5). 051401.

Hu, Z. and Mahadevan, S. (2016). A Single-Loop Kriging Surrogate Modeling for Time-Dependent Reliability Analysis. *Journal of Mechanical Design*, 138(6). 061406.

Huchet, Q., Mattrand, C., Beaurepaire, P., Relun, N., and Gayton, N. (2019). AK-DA: An efficient method for the fatigue assessment of wind turbine structures. *Wind Energy*, 22(5):638–652.

IEC 61400-3 (2009). Wind turbines - Part 3 : Design requirements for offshore wind turbines. International standard.

Iooss, B. and Lemaître, P. (2015). *A Review on Global Sensitivity Analysis Methods*, pages 101–122. Springer US, Boston, MA.

Janon, A., Klein, T., Lagnoux, A., Nodet, M., and Prieur, C. (2014). Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: PS*, 18:342–364.

Jiang, C., Qiu, H., Li, X., Chen, Z., Gao, L., and Li, P. (2020). Iterative reliable design space approach for efficient reliability-based design optimization. *Engineering with Computers*, 36(1):151–169.

Jiang, C., Wang, D., Qiu, H., Gao, L., Chen, L., and Yang, Z. (2019). An active failure-pursuing Kriging modeling method for time-dependent reliability analysis. *Mechanical Systems and Signal Processing*, 129:112–129.

Johnson, M. E., Moore, L. M., and Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, 26(2):131–148.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4):455–492.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G. (2009). Definition of a 5-MW Reference Wind Turbine for Offshore System Development. Technical report.

Jung, Y., Cho, H., and Lee, I. (2020). Intelligent initial point selection for MPP search in reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 62(4):1809–1820.

Kala, Z. (2020a). From Probabilistic to Quantile-Oriented Sensitivity Analysis: New Indices of Design Quantiles. *Symmetry*, 12(10).

Kala, Z. (2020b). Sensitivity Analysis in Probabilistic Structural Design: A Comparison of Selected Techniques. *Sustainability*, 12(11).

Keshtegar, B., Meng, D., Ben Seghier, M. E. A., Xiao, M., Trung, N. T., and Bui, D. T. (2021). A hybrid sufficient performance measure approach to improve robustness and efficiency of reliability-based design optimization. *Engineering with Computers*, 37(3):1695–1708.

Kraft, D. (1988). A software package for sequential quadratic programming. Technical report, Institut für Dynamik der Flugsysteme.

Kratz, M. F. and Rootzén, H. (1997). On the Rate of Convergence for Extremes of Mean Square Differentiable Stationary Normal Processes. *Journal of Applied Probability*, 34(4):908–923.

Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119–139.

Lataniotis, C., Marelli, S., and Sudret, B. (2017). UQLab user manual–Kriging (Gaussian process modelling). Technical report.

Le Cunff, C., Ryu, S., Heurtier, J. M., and Duggal, A. S. (2008). Frequency-Domain Calculations of Moored Vessel Motion Including Low Frequency Effect. volume Volume 1: Offshore Technology of *International Conference on Offshore Mechanics and Arctic Engineering*, pages 689–696.

Le Gratiet, L. and Cannamela, C. (2015). Cokriging-based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes. *Technometrics*, 57(3):418–427.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer series in statistics. Springer, New York, NY.

Lebrun, R. (2013). *Contributions à la modélisation de la dépendance stochastique*. PhD thesis, Université Paris-Diderot-Paris VII.

Lee, I., Choi, K. K., and Gorsich, D. (2010a). Sensitivity analyses of FORM-based and DRM-based performance measure approach (PMA) for reliability-based design optimization (RBDO). *International Journal for Numerical Methods in Engineering*, 82(1):26–46.

Lee, I., Choi, K. K., Noh, Y., Zhao, L., and Gorsich, D. (2010b). Sampling-Based Stochastic Sensitivity Analysis Using Score Functions for RBDO Problems With Correlated Random Variables. volume Volume 1: 36th Design Automation Conference, Parts A and B of *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 1055–1064.

Lee, I., Choi, K. K., and Zhao, L. (2011). Sampling-based RBDO using the stochastic sensitivity analysis and Dynamic Kriging method. *Structural and Multidisciplinary Optimization*, 44(3):299–317.

Lei, J., Lu, Z., and Wang, L. (2021). An efficient method by nesting adaptive Kriging into Importance Sampling for failure-probability-based global sensitivity analysis. *Engineering with Computers*.

Leira, B. J., Igland, R. T., Baarholm, G. S., Farnes, K. A., and Percy, D. (2005). Fatigue Safety Factors for Flexible Risers Based on Case Specific Reliability Analysis. volume 24th International Conference on Offshore Mechanics and Arctic Engineering: Volume 2 of *International Conference on Offshore Mechanics and Arctic Engineering*, pages 211–217.

Lelièvre, N., Beaurepaire, P., Mattrand, C., and Gayton, N. (2018). AK-MCSi: A Kriging-based method to deal with small failure probabilities and time-consuming models. *Structural Safety*, 73:1–11.

Lemaire, M. (2013). *Structural reliability*. John Wiley & Sons.

Li, C. C. and Der Kiureghian, A. (1993). Optimal Discretization of Random Fields. *Journal of Engineering Mechanics*, 119(6):1136–1154.

Li, J. and Chen, J. (2019). Solving time-variant reliability-based design optimization by PSO-t-IRS: A methodology incorporating a particle swarm optimization algorithm and an enhanced instantaneous response surface. *Reliability Engineering & System Safety*, 191:106580.

Li, X., Han, X., Chen, Z., Ming, W., Cao, Y., and Ma, J. (2020). A multi-constraint failure-pursuing sampling method for reliability-based design optimization using adaptive Kriging. *Engineering with Computers*.

Li, X., Meng, Z., Chen, G., and Yang, D. (2019). A hybrid self-adjusted single-loop approach for reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 60(5):1867–1885.

Liang, J., Mourelatos, Z. P., and Tu, J. (2008). A single-loop method for reliability-based design optimisation. *International Journal of Product Development*, 5(1-2):76–92.

Lindgren, G. (2010). A second course on stationary stochastic processes. *Center for Mathematical Sciences, Lund University*.

Ling, C. and Lu, Z. (2021). Support vector machine-based importance sampling for rare event estimation. *Structural and Multidisciplinary Optimization*, 63(4):1609–1631.

Ling, C., Lu, Z., and Zhu, X. (2019). Efficient methods by active learning Kriging coupled with variance reduction based sampling methods for time-dependent failure probability. *Reliability Engineering & System Safety*, 188:23–35.

Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2):4–22.

Liu, Y. and Li, L. (2020). Global reliability sensitivity analysis based on state dependent parameter method and efficient sampling techniques. *Aerospace Science and Technology*, 99:105740.

Loeve, M. (1977). *Probability Theory I*. Springer.

Lopez, R. H. and Beck, A. T. (2012). Reliability-based design optimization strategies based on FORM: a review. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 34(4):506–514.

Luyi, L., Zhenzhou, L., Jun, F., and Bintuan, W. (2012). Moment-independent importance measure of basic variable and its state dependent parameter solution. *Structural Safety*, 38:40–47.

Lv, Z., Lu, Z., and Wang, P. (2015). A new learning function for Kriging and its applications to solve reliability problems in engineering. *Computers & Mathematics with Applications*, 70(5):1182–1197.

Madsen, H. O., Krenk, S., and Lind, N. C. (2006). *Methods of structural safety*. Courier Corporation.

Marelli, S. and Sudret, B. (2018). An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis. *Structural Safety*, 75:67–74.

Matsuishi, M. and Endo, T. (1968). Fatigue of metals subjected to varying stress. *Japan Society of Mechanical Engineers, Fukuoka, Japan*, 68(2):37–40.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

Mehta, N. B., Wu, J., Molisch, A. F., and Zhang, J. (2007). Approximating a Sum of Random Variables with a Lognormal. *IEEE Transactions on Wireless Communications*, 6(7):2690–2699.

Melchers, R. E. (1989). Importance sampling in structural systems. *Structural Safety*, 6(1):3–10.

Miner, M. A. (2021). Cumulative Damage in Fatigue. *Journal of Applied Mechanics*, 12(3):A159–A164.

Molin, B. (2002). *Hydrodynamique des structures offshore*. Editions Technip.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174.

Moustapha, M. (2016). *Adaptive surrogate models for the reliable lightweight design of automotive body structures*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II.

Moustapha, M. and Sudret, B. (2019a). Surrogate-assisted reliability-based design optimization: a survey and a unified modular framework. *Structural and Multidisciplinary Optimization*, 60(5):2157–2176.

Moustapha, M. and Sudret, B. (2019b). A two-stage surrogate modelling approach for the approximation of models with non-smooth outputs. 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP 2019).

Moustapha, M., Sudret, B., Bourinet, J. M., and Guillaume, B. (2015). Adaptive kriging reliability-based design optimization of an automotive body structure under crashworthiness constraints. In *Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12), Vancouver, Canada*. University of British Columbia.

Müller, K. and Cheng, P. W. (2018). Application of a Monte Carlo procedure for probabilistic fatigue design of floating offshore wind turbines. *Wind Energy Science*, 3(1):149–162.

Munoz Zuniga, M. (2011). *Méthodes stochastiques pour l'estimation contrôlée de faibles probabilités sur des modèles physiques complexes: application au domaine nucléaire*. PhD thesis, Université Paris 7.

Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. M. (2016). *Response surface methodology: process and product optimization using designed experiments*. John Wiley & Sons.

Owen, A. B. (2013). Monte carlo theory, methods and examples.

Papaioannou, I., Geyer, S., and Straub, D. (2019). Improved cross entropy-based importance sampling with a flexible mixture model. *Reliability Engineering & System Safety*, 191:106564.

Papaioannou, I. and Straub, D. (2021). Combination line sampling for structural reliability analysis. *Structural Safety*, 88:102025.

Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes*. Communications and signal processing. McGraw-Hill.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Pegalajar-Jurado, A., Borg, M., and Bredmose, H. (2018). An efficient frequency-domain model for quick load analysis of floating offshore wind turbines. *Wind Energy Science*, 3(2):693–712.

Picheny, V. (2014). A Stepwise uncertainty reduction approach to constrained global optimization. In Kaski, S. and Corander, J., editors, *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, volume 33 of *Proceedings of Machine Learning Research*, pages 787–795, Reykjavik, Iceland. PMLR.

Powell, M. J. D. (1994). *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*, pages 51–67. Springer Netherlands, Dordrecht.

Rashki, M. (2021). SESC: A new subset simulation method for rare-events estimation. *Mechanical Systems and Signal Processing*, 150:107139.

Rasmussen, C. E. and Williams, C. K. (2006). *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Risser, M. D. and Turek, D. (2020). Bayesian inference for high-dimensional nonstationary Gaussian processes. *Journal of Statistical Computation and Simulation*, 90(16):2902–2928.

Robertson, A., Gueydon, S., Bachynski, E., Wang, L., Jonkman, J., Alarcón, D., Amet, E., Beardsell, A., Bonnet, P., Boudet, B., Brun, C., Chen, Z., Féron, M., Forbush, D., Galinos, C., Galvan, J., Gilbert, P., Gómez, J., Harnois, V., Haudin, F., Hu, Z., Dreff, J. L., Leimeister, M., Lemmer, F., Li, H., Mckinnon, G., Mendikoa, I., Moghtadaei, A., Netzband, S., Oh, S., Pegalajar-Jurado, A., Nguyen, M. Q., Ruehl, K., Schünemann, P., Shi, W., Shin, H., Si, Y., Surmont, F., Trubat, P., Qwist, J., and Wohlfahrt-Laymann, S. (2020). OC6 Phase I: Investigating the underprediction of low-frequency hydrodynamic loads and responses of a floating wind turbine. *Journal of Physics: Conference Series*, 1618:032033.

Robertson, A., Jonkman, J., Masciola, M., Song, H., Goupee, A., Coulling, A., and Luan, C. (2014a). Definition of the Semisubmersible Floating System for Phase II of OC4. Technical report NREL/TP-5000-60601.

Robertson, A., Jonkman, J., Vorpahl, F., Popko, W., Qvist, J., Frøyd, L., Chen, X., Azcona, J., Uzunoglu, E., Guedes Soares, C., Luan, C., Yutong, H., Pengcheng, F., Yde, A., Larsen, T., Nichols, J., Buils, R., Lei, L., Nygaard, T. A., Manolas, D., Heege, A., Vatne, S. R., Ormberg, H., Duarte, T., Godreau, C., Hansen, H. F., Nielsen, A. W., Riber, H., Le Cunff, C., Beyer, F., Yamaguchi, A., Jung, K. J., Shin, H., Shi, W., Park, H., Alves, M., and Guérinel, M. (2014b). Offshore Code Comparison Collaboration Continuation Within IEA Wind Task 30: Phase II Results Regarding a Floating Semisubmersible Wind System. volume Volume 9B: Ocean Renewable Energy of *International Conference on Offshore Mechanics and Arctic Engineering*. V09BT09A012.

Rossi, R. R. (2005). A Review of Fatigue Curves for Mooring Lines. volume 24th International Conference on Offshore Mechanics and Arctic Engineering: Volume 1, Parts A and B of *International Conference on Offshore Mechanics and Arctic Engineering*, pages 1097–1104.

Roustant, O., Ginsbourger, D., and Deville, Y. (2012). DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *Journal of Statistical Software, Articles*, 51(1):1–55.

Roy, A. and Chakraborty, S. (2020). Support vector regression based metamodel by sequential adaptive sampling for reliability analysis of structures. *Reliability Engineering & System Safety*, 200:106948.

Rychlik, I. (1996). Extremes, rainflow cycles and damage functionals in continuous random processes. *Stochastic Processes and their Applications*, 63(1):97–116.

Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global Sensitivity Analysis. The Primer*. John Wiley & Sons.

Saltelli, A., Tarantola, S., and Chan, K. P. S. (1999). A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, 41(1):39–56.

Santner, T. J., Williams, B. J., Notz, W. I., and Williams, B. J. (2018). *The design and analysis of computer experiments*. Springer.

Sclavounos, P. D. (2012). Karhunen–loeve representation of stochastic ocean waves. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2145):2574–2594.

Shang, X., Chao, T., Ma, P., and Yang, M. (2020). Derivative-based global sensitivity measure using radial basis function. *Structural and Multidisciplinary Optimization*, 62(1):107–129.

Shang, X., Ma, P., Yang, M., and Chao, T. (2021). An efficient polynomial chaos-enhanced radial basis function approach for reliability-based design optimization. *Structural and Multidisciplinary Optimization*, 63(2):789–805.

Shi, Y., Lu, Z., Huang, Z., Xu, L., and He, R. (2020a). Advanced solution strategies for time-dependent reliability based design optimization. *Computer Methods in Applied Mechanics and Engineering*, 364:112916.

Shi, Y., Lu, Z., Zhou, J., and Zio, E. (2020b). A novel time-dependent system constraint boundary sampling technique for solving time-dependent reliability-based design optimization problems. *Computer Methods in Applied Mechanics and Engineering*, 372:113342.

Shinozuka, M. (1964). Probability of Structural Failure Under Random Loading. *Journal of the Engineering Mechanics Division*, 90(5):147–170.

Shinozuka, M. and Deodatis, G. (1991). Simulation of Stochastic Processes by Spectral Representation. *Applied Mechanics Reviews*, 44(4):191–204.

Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86–112.

Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1):271–280.

Song, K., Zhang, Y., Zhuang, X., Yu, X., and Song, B. (2021). Reliability-based design optimization using adaptive surrogate model and importance sampling-based modified SORA method. *Engineering with Computers*, 37(2):1295–1314.

Spagnol, A. (2020). *Kernel-based sensitivity indices for high-dimensional optimization problems*. PhD thesis, Université de Lyon.

Stewart, G. M., Robertson, A., Jonkman, J., and Lackner, M. A. (2016). The creation of a comprehensive metocean data set for offshore wind turbine simulations. *Wind Energy*, 19(6):1151–1159.

Stieng, L. E. S. and Muskulus, M. (2020). Reliability-based design optimization of offshore wind turbine support structures using analytical sensitivities and factorized uncertainty modeling. *Wind Energy Science*, 5(1):171–198.

Sudret, B. (2007). *Uncertainty propagation and sensitivity analysis in mechanical models– Contributions to structural reliability and stochastic spectral methods*. HDR, Université Blaise Pascal, Clermont-Ferrand, France.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979. Bayesian Networks in Dependability.

Sun, Z., Wang, J., Li, R., and Tong, C. (2017). LIF: A new Kriging based learning function and its application to structural reliability analysis. *Reliability Engineering & System Safety*, 157:152–165.

Teixeira, R., Nogal, M., and O'Connor, A. (2021). Adaptive approaches in metamodel-based reliability analysis: A review. *Structural Safety*, 89:102019.

Teixeira, R., Nogal, M., O'Connor, A., Nichols, J., and Dumas, A. (2019). Stress-cycle fatigue design with Kriging applied to offshore wind turbines. *International Journal of Fatigue*, 125:454–467.

Toal, D. J. J. and Keane, A. J. (2012). Non-stationary kriging for design optimization. *Engineering Optimization*, 44(6):741–765.

Torii, A. J., Lopez, R. H., and Miguel, L. F. F. (2016). A general RBDO decoupling approach for different reliability analysis methods. *Structural and Multidisciplinary Optimization*, 54(2):317–332.

Tu, J., Choi, K. K., and Park, Y. H. (1999). A New Study on Reliability-Based Design Optimization. *Journal of Mechanical Design*, 121(4):557–564.

Vorpahl, F., Schwarze, H., Fischer, T., Seidel, M., and Jonkman, J. (2013). Offshore wind turbine environment, loads, simulation, and design. *WIREs Energy and Environment*, 2(5):548–570.

Wang, Z. and Chen, W. (2016). Time-variant reliability assessment through equivalent stochastic process transformation. *Reliability Engineering & System Safety*, 152:166–175.

Wang, Z., Li, H., Chen, Z., Li, L., and Hong, H. (2020). Sequential optimization and moment-based method for efficient probabilistic design. *Structural and Multidisciplinary Optimization*, 62:387–404.

Wang, Z. and Wang, P. (2012). A Nested Extreme Response Surface Approach for Time-Dependent Reliability-Based Design Optimization. *Journal of Mechanical Design*, 134(12). 121007.

Wang, Z. and Zhang, Y.and Song, Y. (2020). A Modified Conjugate Gradient Approach for Reliability-Based Design Optimization. *IEEE Access*, 8:16742–16749.

Wei, P., Lu, Z., Hao, W., Feng, J., and Wang, B. (2012). Efficient sampling methods for global reliability sensitivity analysis. *Computer Physics Communications*, 183(8):1728–1743.

Wiener, N. (1938). The Homogeneous Chaos. *American Journal of Mathematics*, 60(4):897–936.

WindEurope (2017). Floating Offshore Wind Vision Statement. WindEurope report.

WindEurope (2020). Wind energy in Europe. 2020 Statistics and the outlook for 2021-2025. WindEurope report.

Xiang, Z., Chen, J., Bao, Y., and Li, H. (2020). An active learning method combining deep neural network and weighted sampling for structural reliability analysis. *Mechanical Systems and Signal Processing*, 140:106684.

Xiong, Y., Chen, W., Apley, D., and Ding, X. (2007). A non-stationary covariance-based Kriging method for metamodelling in engineering design. *International Journal for Numerical Methods in Engineering*, 71(6):733–756.

Yang, M., Zhang, D., and Han, X. (2020a). Enriched single-loop approach for reliability-based design optimization of complex nonlinear problems. *Engineering with Computers*, pages 1–19.

Yang, M., Zhang, D., and Han, X. (2020b). New efficient and robust method for structural reliability analysis and its application in reliability-based design optimization. *Computer Methods in Applied Mechanics and Engineering*, 366:113018.

Youn, B. D., Choi, K. K., and Park, Y. H. (2003). Hybrid Analysis Method for Reliability-Based Design Optimization. *Journal of Mechanical Design*, 125(2):221–232.

Yu, W., Müller, K., and Lemmer, F. (2018). Qualification of innovative floating substructures for 10MW wind turbines and water depths greater than 50 m. Technical report, LIFES50+ project.

Yun, W., Lu, Z., and Jiang, X. (2018). A modified importance sampling method for structural reliability and its global reliability sensitivity analysis. *Structural and Multidisciplinary Optimization*, 57(4):1625–1641.

Zentner, I. (2010). Numerical computation of fragility curves for NPP equipment. *Nuclear Engineering and Design*, 240(6):1614–1621.

Zhang, H., Aoues, Y., Lemosse, D., and De Cursi, E. S. (2020a). A single-loop approach with adaptive sampling and surrogate kriging for reliability-based design optimization. *Engineering Optimization*, pages 1–17.

Zhang, J. and Ellingwood, B. (1994). Orthogonal Series Expansions of Random Fields in Reliability Analysis. *Journal of Engineering Mechanics*, 120(12):2660–2677.

Zhang, J., Xiao, M., and Gao, L. (2020b). A new local update-based method for reliability-based design optimization. *Engineering with Computers*.

Zhang, X., Lu, Z., Yun, W., Feng, K., and Wang, Y. (2020c). Line sampling-based local and global reliability sensitivity analysis. *Structural and Multidisciplinary Optimization*, 61(1):267–281.

Zhang, X., Wang, L., and Sørensen, J. D. (2019). REIF: A novel active-learning function toward adaptive Kriging surrogate models for structural reliability analysis. *Reliability Engineering & System Safety*, 185:440–454.

Zhang, Y., Gong, C., and Li, C. (2021a). Efficient time-variant reliability analysis through approximating the most probable point trajectory. *Structural and Multidisciplinary Optimization*, 63(1):289–309.

Zhang, Y., Tao, S., Chen, W., and Apley, D. W. (2020d). A Latent Variable Approach to Gaussian Process Modeling with Qualitative and Quantitative Factors. *Technometrics*, 62(3):291–302.

Zhang, Z., Deng, W., and Jiang, C. (2021b). A PDF-based performance shift approach for reliability-based design optimization. *Computer Methods in Applied Mechanics and Engineering*, 374:113610.

Zhao, L., Wang, P., Song, B., Wang, X., and Dong, H. (2020). An efficient kriging modeling method for high-dimensional design problems based on maximal information coefficient. *Structural and Multidisciplinary Optimization*, 61(1):39–57.

Zhou, T. and Peng, Y. (2020). Kernel principal component analysis-based gaussian process regression modelling for high-dimensional reliability analysis. *Computers & Structures*, 241:106358.

Zhou, Y. and Lu, Z. (2020). An enhanced Kriging surrogate modeling technique for high-dimensional problems. *Mechanical Systems and Signal Processing*, 140:106687.

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : Optimisation sous contraintes probabilistes d'un système complexe – Application au dimensionnement d'une éolienne offshore flottante

Mots clés : RBDO, théorie des valeurs extrêmes, krigeage adaptatif, Monte Carlo, éolienne offshore flottante, fatigue

Résumé : Nous proposons dans cette thèse une approche permettant d'optimiser la configuration des lignes d'ancrage d'une éolienne flottante offshore, en minimisant le coût des matériaux tout en respectant des contraintes d'un état limite de fatigue. Ces contraintes héritent du caractère aléatoire des conditions environnementales ainsi que d'incertitudes sur des paramètres du modèle. Par conséquent, nous sommes confrontés à un problème d'optimisation avec une fonction de coût déterministe et des contraintes impliquant des probabilités de dépassement de seuil du maximum et de l'intégrale de processus aléatoires dépendant du temps, évalués sur une période $[0, T]$.
La principale difficulté est de devoir évaluer ces probabilités à chaque boucle de l'algorithme d'optimisation. En effet, les méthodes de fiabilité nécessitent de nombreux appels à un code de calcul coûteux. L'estimation de ces probabilités est d'autant plus difficile que nous sommes confrontés à des événements rares. Pour résoudre efficacement ce problème, nous proposons une méthodologie en deux étapes. Premièrement, en considérant que T est suffisamment grand, nous utilisons les propriétés des contraintes et des théorèmes limite de la théorie des valeurs extrêmes et de la théorie ergodique pour reformuler les contraintes initiales en contraintes indépendantes du temps. Nous obtenons ainsi un problème équivalent pour lequel les algorithmes classiques sont peu performants. La deuxième étape de notre procédure consiste à résoudre le problème reformulé avec une nouvelle méthode basée sur une stratégie de krigeage adaptative. Cette méthode est appelée AK-ECO pour Adaptive Kriging for Expectation Constraints Optimization.
Le cas académique d'un oscillateur harmonique présentant toutes les caractéristiques du problème industriel est introduit afin d'illustrer notre méthodologie. La procédure est ensuite appliquée avec succès au problème de l'éolienne flottante. Les deux étapes qui composent cette méthodologie sont décrites dans un cadre général afin de pouvoir être appliquées à d'autres problèmes d'optimisation impliquant des contraintes probabilistes dépendant du maximum ou de l'intégrale de processus aléatoires.

Title : Chance constraint optimization of a complex system - Application to the fatigue design of a floating offshore wind turbine mooring system

Keywords : RBDO, extreme value theory, adaptive Kriging, Monte Carlo, floating offshore wind turbine, fatigue

Abstract : In this thesis, we propose a methodology to optimize the configuration of the mooring lines by minimizing the material cost while satisfying Fatigue Limit State (FLS) constraints. These constraints inherit the randomness of the marine environment as well as uncertainties on material properties and model parameters. Therefore, we face an optimization problem with a deterministic cost function and constraints involving probabilities of threshold exceedance of the maximum and the integral over a period $[0, T]$ of time-dependent random processes.
Having to evaluate these failure probabilities at each loop of the optimization algorithm is the main difficulty. Indeed, reliability methods require many time-consuming simulations. The estimation of these probabilities is all the more challenging as we are dealing with rare events. To solve this problem efficiently, we propose a two-step methodology. First, considering T sufficiently large, we use the properties of the constraints and limit theorems of the extreme value theory and the ergodic theory to reformulate the original constraints into time-independent ones. We thus obtain an equivalent problem for which classical algorithms perform poorly. The second step of the procedure consists in solving the reformulated problem with a new method based on an adaptive kriging strategy well suited to the reformulated constraints. This method is called AK-ECO for Adaptive Kriging for Expectation Constraints Optimization.
An academic case of a harmonic oscillator presenting all the characteristics of the industrial problem is introduced to illustrate the methodology. The procedure is then applied with success to the FOWT problem. The two steps composing this methodology are described in a general framework so that they can be applied to other optimization problems involving probabilistic constraints depending on the maximum or the integral of random processes.

Institut Polytechnique de Paris
91120 Palaiseau, France