An island particle Markov chain Monte Carlo algorithm for safety analysis

Christelle Vergé, Jérôme Morio, Pierre Del Moral

An island particle Markov chain Monte Carlo algorithm for safety analysis

Christelle Vergé [1,2,3] Jérôme Morio [1,*] Pierre Del Moral [2]

Abstract

Estimating rare event probability with accuracy is of great interest for safety and reliability applications. Nevertheless, some simulation parameters Θ such as the input density parameters in the case of input-output functions, are often set for simplification reasons. A bad estimation of Θ can strongly modify rare event probability estimations. In the present article, we design a new island particle Markov chain Monte Carlo algorithm to determine the parameters Θ that, in case of bad estimation, tend to increase the rare event probability value. This algorithm also gives an estimate of the rare event probability maximum taking into account the likelihood of Θ. The principles of this statistical technique are described throughout this article and its results are analyzed on different test cases.

Key words: rare event, sequential Monte Carlo, particle filtering, sensitivity analysis

1. Introduction

Numerous statistical and simulation techniques have been proposed to estimate rare event probabilities. Indeed, importance sampling [1], [2], [3], [4], Markov chain Monte Carlo (abbreviated MCMC) and subset simulation [5], [6], [7], [8], [9], First and Second Order Reliability Methods (FORM/SORM) [10], [11], [12], or extreme value theory [13], [14], [15] are notably well known simulation algorithms to estimate rare event probabilities on input-output black-box functions, like an output threshold exceedance. Their principles and advantages/drawbacks have also been deeply studied. Nevertheless, some parameters in black-box functions are implicitly set, such as the different parameters of the input parametric model density, and influence the rare event probability estimation. These hypotheses are assumed for simplification and computational reasons. However, important questions in safety are: what happen to the rare event probability if these parameters are not well tuned? Is it possible to determine the parameters that have to be the most carefully set because of their significant influence on the rare event probability ? The case where a bad tuning of Θ implies a decrease of the critical rare event probability can be troublesome for safety since the associated rare event risk is underestimated.

There are not many articles on this issue in the scientific literature. Nevertheless, some studies on close topics have been proposed recently, mostly based on sensitivity analysis methods. Monte Carlo filtering [16] consists in determining the differences between a

* corresponding author

Email addresses: christelle.verge@onera.fr (Christelle Vergé), jerome.morio@onera.fr (Jérôme Morio).

[1] Onera - The French Aerospace Lab, F-91761 PALAISEAU Tel.: +33 1 80 38 66 54

[2] INRIA Bordeaux 351, cours de la Libération 33405 Talence cedex, France

[3] CNES, 18 avenue Edouard Belin 31401 Toulouse France

"safe" sample and a "faulty" sample via standard statistical tests. The reliability index resulting from FORM/SORM [17] can also be used to analyze the influence of input parameters on the failure probability. Stratified sampling [18] and importance sampling [19] have been adapted with the same purpose.

The method proposed in this article is rather different. Indeed, in sensitivity analysis, a parameter is influent if its variation tends to increase or decrease the rare event probability. It is not the case here since one focuses more precisely on the Θ settings that increase the rare event probability, which is of major interest for safety and reliability. The proposed approach is to estimate and analyze the distribution of these input parameters conditionally to the rare event of interest. For that purpose, we adapt an universal Sequential Monte Carlo (abbreviated SMC) algorithm for computing a conditional distribution related to some parameter. This algorithm consists in the simultaneous run of two sets of mean field type interacting particles systems (abbreviated IPS) [7,8]. The first IPS allows to estimate the distribution of the input parameters conditionally to the rare event of interest. At each iteration of this IPS, the second IPS allows to estimated a required probability. Island particle Markov chain Monte Carlo (i-PMCMC) algorithm is thus defined by an IPS where, for each of its particle, another IPS has to be run. The validity of this algorithm is proved in [20] since it can be cast in the Feynman-Kac framework. To our knowledge, the development of this method for rare event case and safety application has not been proposed. In the context of filtering and hidden Markov chains, a related algorithm called SMC^2 (Sequential Monte Carlo square) has been also developed in [21]. Some general island particle algorithms are described in [22].

In this connection, the i-PMCMC algorithm presented in this work can be interpreted as an SMC version of the Particle MCMC (PMCMC) model developed in [23]. In particular, for a single island the i-PMCMC coincide with the PMCMC algorithm. The main advantage of the island particle version is to bypass the long time convergence issue of the MCMC algorithm using interacting island Monte Carlo spatial integrations. On the other hand, when there is a single individual in each island, the i-PMCMC coincide with the conventional SMC algorithm. In the following, we firstly raise the fundamental question of this article: how to estimate the law of Θ conditionally to a rare event? The i-PMCMC algorithm is thus proposed to solve this issue. Its principle and algorithmic implementation are notably described. The final section is devoted to the application of i-PMCMC on two different test cases.

2. General problem

Let us consider a d-dimensional random variable X with probability density f_X with respect to some dominating measure λ_X (like the Lebesgue measure), ϕ a continuous positive scalar function $\phi : \mathbb{R}^d \to \mathbb{R}$ and S a given critical threshold. The function ϕ is static, *i.e.* does not depend on time, and represents for instance an input-output model. This kind of model is notably used in various applications [24], [25], [26], [27] or [28]. We assume that the output $Y = \phi(X)$ is also a random variable of dimension 1 and with probability density f_Y w.r.t. λ_Y. The quantity of interest on the output Y is the probability of exceedance $\mathbb{P}(Y > S) = \mathbb{P}(\phi(X) > S)$. When $\mathbb{P}(Y > S)$ is rare relatively to the available simulation budget N and it is often the case in safety and reliability issues, different algorithms described in [29], [30], [31], [32], [11], [33] or [9] have notably been proposed to estimate accurately this probability.

In this article, one focuses on the case where f_X is uncertain. We assume that X is distributed according to a well known parametric model but its parameters, denoted by Θ, are random variables with probability density f_Θ w.r.t. λ_Θ. In the further development we let $\nu(d\theta) = f_\Theta(\theta)\ \lambda_\Theta(d\theta) = \mathbb{P}(\Theta \in d\theta)$ the distribution of Θ. For instance, if X is a random vector with a multivariate normal distribution, then Θ may describe the mean and the covariance matrix of X. It corresponds notably to realistic applications where it is not always possible to evaluate accurately the density of input parameters. This formalism enables thus to consider a large range of input PDF.

The probability of interest $\mathbb{P}(Y > S)$ depends of course on Θ and thus on f_Θ. In safety applications, it is important to estimate a superior bound of the rare event probability $\mathbb{P}(Y > S)$ taking also into account the likelihood of Θ. The likelihood of Θ is important since unrealistic bad tuning values of Θ which lead to high probabilities $\mathbb{P}(Y > S)$ are not relevant. The idea of this article is thus to determine the distribution of Θ conditionally to the fact that Y exceeds the threshold S, denoted by $\nu_{|Y>S}$. When Θ follows $\nu_{|Y>S}$, the probability $\mathbb{P}(Y > S|\Theta = \theta)$ can also be evaluated at the same time. To estimate $\nu_{|Y>S}$, one proposes to use an IPS which evolves according to iterative selection and mutation steps.

In the further development, when there is no confusion, we slightly abuse notation and sometimes write $\mathbb{P}(Y > S|\theta)$ instead of $\mathbb{P}(Y > S|\Theta = \theta)$ and $\mathbb{P}(Y > S_1|Y > S_2, \theta)$ instead of $\mathbb{P}(Y > S_1|Y > S_2, \Theta = \theta)$.

3. i-PMCMC algorithm

3.1. *Principle*

i-PMCMC algorithm is based on the use of two sets of particles to iteratively approach $\nu_{|Y>S}$. The first set of particles is defined on the parameter Θ and the second set of particles is useful to estimate the probabilities $\mathbb{P}(Y > S|\theta)$. The complete demonstration of IPS convergence and the link with Feynman-Kac framework is given in [20].

Define $S_1, S_2, ..., S_n = S$ a serie of increasing thresholds and denote $\pi_i = \nu_{|Y>S_i}$. The target law is of course $\pi_n = \nu_{|Y>S_n} = \nu_{|Y>S}$. Thanks to the Bayes theorem, one can rewrite π_n such as :

$$\pi_n(d\theta) = \frac{1}{\mathbb{P}(Y > S_n)} \mathbb{P}(Y > S_n|\theta)\ \nu(d\theta)\ . \tag{1}$$

The probability law π_n is thus proportional to

$$\pi_n(d\theta) \propto \mathbb{P}(Y > S_n|\theta)\ \nu(d\theta)$$

$$\pi_n(d\theta) \propto H_n(\theta)\ \nu(d\theta)\ ,$$

where $H_n(\theta) = \mathbb{P}(Y > S_n|\theta)$. The term $H_n(\theta)$ can be expressed as a product of conditional probabilities

$$H_n(\theta) = \left[\prod_{p=1}^{n-1} \mathbb{P}(Y > S_{p+1}|Y > S_p, \theta)\right] \times \mathbb{P}(Y > S_1|\theta) = \prod_{p=0}^{n-1} h_p(\theta)\ , \tag{2}$$

with

$$\begin{cases} h_p(\theta) = \mathbb{P}(Y > S_{p+1}|Y > S_p, \theta) \\ \\ h_0(\theta) = \mathbb{P}(Y > S_1|\theta)\ . \end{cases} \tag{3}$$

In this notation, we have

$$\pi_n(d\theta) \propto \left[\prod_{p=0}^{n-1} h_p(\theta)\right] \nu(d\theta) \ . \tag{4}$$

One can also remark that $H_p = H_{p-1} \times h_{p-1}$ and consequently the link between π_{p+1} and π_p can be written in the following way

$$\pi_{p+1} = \psi_{h_p}(\pi_p), \tag{5}$$

where ψ_{h_p} is the so-called Boltzmann-Gibbs transformation. Let $\mathcal{P}(\mathbb{E})$ be the set of probability measures on $\mathbb{E}$. For all positive bounded function G, the Boltzmann-Gibbs transformation $\Psi_G : \mathcal{P}(\mathbb{E}) \to \mathcal{P}(\mathbb{E})$ is defined for all $\mu \in \mathcal{P}(\mathbb{E})$ s.t. $\mu(G) = \int G(x)\,\mu(dx) > 0$ by

$$\Psi_G(\mu)(dx) := \frac{1}{\mu(G)}\, G(x)\,\mu(dx).$$

If one assumes that it is possible to determine a Markovian kernel M_p that let π_p invariant (which is not restrictive using, for example, an acceptance/rejection stage of Metropolis-Hastings algorithm) we have

$$\pi_p(d\theta) = (\pi_p M_p)(d\theta) := \int \pi_p(d\theta')M_p(\theta', d\theta). \tag{6}$$

This yields the evolution equation

$$\pi_{p+1} = \psi_{h_p}(\pi_p)M_{p+1} \ . \tag{7}$$

Equation 7 may be cast in the Feynman-Kac framework and then, each measure π_p can be approximated by an IPS which evolves in accordance with selection steps related to the so-called potential functions h_p and mutation steps related to the Markov kernel M_p. Denote by $\{(\theta_p^1, ..., \theta_p^{N_1})\}_{p\geq 0}$ such a system of N_1 particles.

$$\begin{bmatrix} \theta_0^1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \theta_0^{N_1} \end{bmatrix} \xrightarrow{\textbf{Selection}} \begin{bmatrix} \widehat{\theta}_0^1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \widehat{\theta}_0^{N_1} \end{bmatrix} \xrightarrow{\textbf{Mutation}} \begin{bmatrix} \theta_1^1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \theta_1^{N_1} \end{bmatrix} \xrightarrow{\textbf{Selection}} \begin{bmatrix} \widehat{\theta}_1^1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \widehat{\theta}_1^{N_1} \end{bmatrix} \xrightarrow{\textbf{Mutation}}$$

The selection stage consists in sampling $\{\widehat{\theta}_p^i\}_{i=1}^{N_1}$ independently according to the probability measure ψ_{h_p}, i.e. selecting the particles $\{\theta_p^i\}_{i=1}^{N_1}$ with probabilities proportional to their weights $\{h_p(\theta_p^i)\}_{i=1}^{N_1}$. The mutation stage consists in updating the selected particles conditionally independently using the Markov kernel M_{p+1} that let π_{p+1} invariant. This step enables to increase the diversity of $\widehat{\theta}_p$ without changing its probability law, that is already close to π_{p+1}. The Feynman-Kac theory [20] ensures that at each transition stage p:

$$\frac{1}{N_1} \sum_{i=1}^{N_1} \delta_{\theta_p^i} \xrightarrow[N_1 \to +\infty]{} \pi_p \ .$$

Thus, at the end of the n^{th} transition stage, the set of particles converges to the target law π_n so that

$$\frac{1}{N_1} \sum_{i=1}^{N_1} \delta_{\theta_n^i} \xrightarrow[N_1 \to +\infty]{} \pi_n \ .$$

Nevertheless, the knowledge of h_p is required to apply the different selection/mutation stages. In practice, for all $i = 1, ..., N_1$, $h_p(\theta_p^i)$ is not always analytically computable but can be estimated by defining a new set of particles $\{\xi_p^{i,j}\}_{j=1}^{N_2}$ on the random variable X conditionally to the different thresholds S_p and associated to each θ_p^i.

Consider $\{\theta_0^i\}_{i=1}^{N_1}$ generated with probability law ν. At iteration k of the algorithm, with $k \geq 1$, we assume that particles $\{\theta_k^i\}_{i=1}^{N_1}$ are available and then, the i-PMCMC algorithm consists in two iterative genetic type stages:

- **Selection stage:** The selection stage consists in choosing randomly and independently N_1 particles amongst $\{\theta_k^i\}_{i=1}^{N_1}$ with probabilities proportional to their weights $\{h_k(\theta_k^i)\}_{i=1}^{N_1}$. Thus, the particles with low weights are killed whereas those with high weights are multiplied. The number of particles is kept constant in this stage and a new set of particles $\{\widehat{\theta}_k^i\}_{i=1}^{N_1}$ can be defined.

  Remind that the potential functions h_k are defined by:

$$\begin{cases} h_k(\theta_k^i) = \mathbb{P}(Y > S_{k+1} | Y > S_k, \Theta = \theta_k^i), k \geq 1 \\ \\ h_0(\theta_0^i) = \mathbb{P}(Y > S_1 | \Theta = \theta_0^i) \end{cases}$$

  These quantities have to be computed.

- **Mutation stage:** Even if the number of particles is still equal to N_1, some particles have been duplicated, so we apply a Markov kernel to increase the diversity of the particles. Building a π_{k+1}-reversible transition kernel that thus let π_{k+1} invariant is the aim of mutation stage. For that purpose, the acceptance/rejection step of the Metropolis-Hastings algorithm [34] is useful. This approach results in the exploration of Θ space set without changing the $\left\{\widehat{\theta}_k^i\right\}_{i=1}^{N_1}$ distribution and the increase of the particle diversity. A new particle $\widehat{\theta}_k^{i\,\prime}$ is proposed with a ν-reversible kernel Q. The acceptation rate of this new proposal is consequently $1 \wedge \frac{H_{k+1}(\widehat{\theta}_k^{i\,\prime})}{H_{k+1}(\widehat{\theta}_k^i)}$. If $H_{k+1}(\widehat{\theta}_k^{i\,\prime}) > H_{k+1}(\widehat{\theta}_k^i)$, the proposal $\widehat{\theta}_k^{i\,\prime}$ is automatically accepted and replaces $\widehat{\theta}_k^i$ in the set of current particles. Otherwise, the proposal $\widehat{\theta}_k^{i}{}'$ is accepted with probability $\frac{H_{k+1}(\widehat{\theta}_k^{i}{}')}{H_{k+1}(\widehat{\theta}_k^{i})}$. This acceptance/rejection procedure is repeated N_{app} times to decrease the correlation between the particles. At the end of this stage, a new set of particles $\{\theta_{k+1}^i\}_{i=1}^{N_1}$ can be defined.

Mutation and selection stages are applied n times until reaching the target threshold S_n. At the end of the algorithm, the particles $\{\theta_n^i\}_{i=1}^{N_1}$ provides an estimate of π_n :

$$\widehat{\pi}_n^{N_1} = \frac{1}{N_1}\sum_{i=1}^{N_1} \delta_{\theta_n^i} \ .$$

For $i \in [\![1; N_1]\!]$ and $k \in [\![0; n]\!]$, the point is to compute each probability $\{h_l(\theta_k^i)\}_{l=1}^{k}$. If they cannot be analytically calculated, they can be estimated with another interacting particle system (also called, in that case, sequential Monte Carlo, importance splitting, subset simulation or subset sampling). It is a rare event estimation technique which consists in estimating several conditional probabilities that are easier to evaluate than estimating only one probability through a very tough simulation. Its principle is also based on selection and mutation stages. For each $\Theta = \theta_k^i$, let us define $\{\xi_0^{i,j}\}_{j=1}^{N_2}$ with probability density $f_{X|\theta_0^i}$ w.r.t. λ_X, *i.e.* the conditional density of X given $\Theta = \theta_k^i$. At iteration l of the algorithm, we assume that particles $\{\xi_l^{i,j}\}_{j=1}^{N_2}$ are available and then IPS consists in two iterative stages:

- **Selection stage:** The selection stage consists in choosing randomly and independently N_2 particles amongst the particles $\{\xi_l^{i,j}\}_{j=1}^{N_2}$ which are above S_l. The particles which have not reached the threashold S_l are thus killed. The number of particles is kept constant, and a new set of particles $\{\widehat{\xi}_l^{i,j}\}_{j=1}^{N_2}$ can be defined.
- **Mutation stage:** The mutation stage is patterned on the acceptance/rejection principle using the Metropolis-Hastings algorithm [34]. A new particle $\widehat{\xi}_l^{i,j}{}'$ is then proposed

with a Markov kernel $\tilde{Q}$. If $\phi(\widehat{\xi}_l^{i,j}{}') \geq S_{l+1}$, then the proposal is accepted with probability $1 \wedge \frac{f_{X|\theta_k^i}(\widehat{\xi}_l^{i,j}{}')\tilde{Q}(\widehat{\xi}_l^{i,j}{}',\widehat{\xi}_l^{i,j})}{f_{X|\theta_k^i}(\widehat{\xi}_l^{i,j})\tilde{Q}(\widehat{\xi}_l^{i,j},\widehat{\xi}_l^{i,j}{}')}$ and $\widehat{\xi}_l^{i,j}{}'$ replaces $\widehat{\xi}_l^{i,j}$ in the set of current particles. If $\phi(\widehat{\xi}_l^{i,j}{}') < S_{l+1}$, the proposal is automatically rejected and the particle $\widehat{\xi}_l^{i,j}$ is remained. This acceptance/rejection procedure is repeated N_{app2} times to decrease the correlation between the particles. At the end of this stage, a new set of particles $\{\xi_{l+1}^{i,j}\}_{j=1}^{N_2}$ can be defined. An estimate $\widehat{h}_l(\theta_k^i)$ of $h_l(\theta_k^i) = \mathbb{P}(Y > S_{l+1}|Y > S_l, \Theta = \theta_k^i)$ is given by the ratio between the number of $\{\xi_l^{i,j}\}_{j=1}^{N_2}$ particles such that $\phi(\xi_l^{i,j}) > S_{l+1}$ and the total number of particles N_2.

Mutation and selection stages are applied k times until reaching the target threshold S_k. At the end of the algorithm, $H_{k+1}(\theta_k^i) = \mathbb{P}(Y > S_{k+1}|\Theta = \theta_k^i)$ is estimated by

$$\widehat{H}_{k+1}(\theta_k^i) = \prod_{l=0}^{k} \widehat{h}_l(\theta_k^i)$$

For a given particle θ_k^i, a complete set of particles $\{\xi_l^{i,j}\}_{1\leq l\leq k}^{1\leq j\leq N_2}$ is thus generated. An island is thus constituted of a single particle θ_k^i and its associated $\{\xi_l^{i,j}\}_{1\leq l\leq k}^{1\leq j\leq N_2}$ particle set.

The i-PMCMC algorithm is described more precisely in **Algorithm** 1. Interacting particle system for probability estimation required in **Algorithm** 1 is developed in **Algorithm** 2.

In most applications, the potential functions H_k are strictly positive, so that the interacting island model described in **Algorithm** 1 is well defined at any time. In the reverse angle, the evolution of the particles within each island described in **Algorithm** 2 is based on a selection mechanism associated with indicator functions. When all particles in a given island, cannot enter into the desired event, their evolution is stopped and the corresponding level crossing probability is estimated by 0. To bypass this technical

Algorithm 1 i-PMCMC algorithm

Setting definition:
Define the thresholds $S_1, ..., S_n$, the sample sizes N_1, N_2 and the number of applications N_{app} of Markov kernel Q.
Initialization:
Sample $\left\{\theta_0^i\right\}_{i=1}^{N_1}$ according to the probability law ν.
for i from 1 to N_1 **do**
    Sample $\left\{\xi_0^{i,j}\right\}_{j=1}^{N_2}$ according to the conditional density $f_{X|\theta_0^i}$.
end for
Transition:
for k from 0 to n **do**
    Associate a system of particles $\left\{\xi_l^{i,j}\right\}_{1\le l\le k}^{1\le j\le N_2}$ to each θ_k^i in order to estimate $h_k\left(\theta_k^i\right)$ and $H_{k+1}\left(\theta_k^i\right)$ with **Algorithm** 2.
    Selection of the θ-particles:
    Sample $I_k = \left(I_k^1, \dots, I_k^{N_1}\right)$ multinomially with probability proportional to $\{h_k\left(\theta_k^i\right)\}_{i=1}^{N_1}$.
    Set $\widehat{\theta}_k^i = \theta_k^{I_k^i}$.
    Mutation of the θ-particles:
    for m from 1 to N_{app} **do**
        for i from 1 to N_1 **do**
            Sample $\widehat{\theta}_k^{i\,\prime}$ with a ν reversible kernel Q.
            Sample u with a uniform random variable.
            if $\left(u < 1 \wedge \frac{H_{k+1}(\widehat{\theta}_k^{i\,\prime})}{H_{k+1}(\widehat{\theta}_k^i)}\right)$ **then** set $\theta_{k+1}^i = \widehat{\theta}_k^{i\,\prime}$.
            else set $\theta_{k+1}^i = \widehat{\theta}_k^i$.
            end if
        end for
        if $m < N_{app}$ **then** set $\widehat{\theta}_k^i = \theta_{k+1}^i$.
        end if
    end for
end for
Estimation:
Estimate π_n with $\widehat{\pi}_n^{N_1} = \frac{1}{N_1}\sum_{i=1}^{N_1}\delta_{\theta_n^i}$

difficulty, we can use adaptive multiple levels evolution strategies [9].

The determination of Q and $\tilde{Q}$, in the general case, implies the use of Metropolis-Hastings algorithm. Nevertheless, if μ is a standard normal distribution, a transition from x to z defined with the following expression

$$x \mapsto z = \sqrt{1-a}\; x + \sqrt{a}\; W, \tag{8}$$

Algorithm 2 Interacting particle system for probability estimation

For a given value θ_k^i, we build an IPS which allows to estimate both $h_k\left(\theta_k^i\right) = \mathbb{P}\left(Y > S_{k+1} | Y > S_k, \Theta = \theta_k^i\right)$ and $H_{k+1}(\theta_k^i) = \prod_{l=0}^{k} h_l\left(\theta_k^i\right)$.

<u>Setting definition:</u>

Define the number of applications N_{app2} of Markov kernel $\tilde{Q}$, recall the iteration parameter k and the particle value θ_k^i, the thresholds $S_1, ..., S_{k+1}$ and the sample size N_2, that have been defined or obtained in **Algorithm** 1.

<u>Initialisation:</u>

Sample $\left\{\xi_0^{i,j}\right\}_{j=1}^{N_2}$ according to the conditional density $f_{X|\theta_k^i}$.

<u>Transition:</u>

for l from 0 to $k-1$ **do**
    <u>Selection of the ξ particles:</u>
    for j from 1 to N_2 **do**
        if $\phi\left(\xi_l^{i,j}\right) \geq S_{l+1}$ **then** set $\widehat{\xi}_l^{i,j} = \xi_l^{i,j}$.
        else Sample $\widehat{\xi}_l^{i,j}$ randomly and uniformly among particles which have reached the threshold S_{l+1}.
        end if
    end for
    <u>Mutation of the ξ particles:</u>
    for r from 1 to N_{app2} **do**
        for j from 1 to N_2 **do**
            Sample $\widehat{\xi}_l^{i,j\prime}$ according to $\tilde{Q}\left(\widehat{\xi}_l^{i,j}, .\right)$.
            if $\phi\left(\widehat{\xi}_l^{i,j\prime}\right) < S_{l+1}$ **then** set $\xi_{l+1}^{i,j} = \widehat{\xi}_l^{i,j}$.
            else Sample u with a uniform random variable.
                if $\left(u < 1 \wedge \frac{f_{X|\theta_k^i}(\widehat{\xi}_l^{i,j\prime})\tilde{Q}(\widehat{\xi}_l^{i,j\prime}, \widehat{\xi}_l^{i,j})}{f_{X|\theta_k^i}(\widehat{\xi}_l^{i,j})\tilde{Q}(\widehat{\xi}_l^{i,j}, \widehat{\xi}_l^{i,j\prime})}\right)$ **then** set $\xi_{l+1}^{i,j} = \widehat{\xi}_l^{i,j\prime}$
                else set $\xi_{l+1}^{i,j} = \widehat{\xi}_l^{i,j}$
                end if
            end if
        end for
        if $r < N_{app2}$ **then** set $\widehat{\xi}_l^{i,j} = \xi_{l+1}^{i,j}$.
        end if
    end for
    Set $\widehat{h}_l(\theta_k^i) = \frac{1}{N_2}\sum_{j=1}^{N_2} \mathbb{1}_{\phi\left(\xi_l^{i,j}\right) \geq S_{l+1}}$
end for

Set $\widehat{h}_k(\theta_k^i) = \frac{1}{N_2}\sum_{j=1}^{N_2} \mathbb{1}_{\phi\left(\xi_k^{i,j}\right) \geq S_{k+1}}$

<u>Estimation:</u>

Estimate $h_k\left(\theta_k^i\right)$ with $\widehat{h}_k(\theta_k^i)$ and $H_{k+1}\left(\theta_k^i\right)$ with $\widehat{H}_{k+1}(\theta_k^i) = \prod_{l=0}^{k} \widehat{h}_l(\theta_k^i)$.

where $W \sim \mathcal{N}(0,1)$ and a is scalar parameter such as $a \in [0,1]$, is μ-reversible. In order to use equation 8 instead of Metropolis-Hastings algorithm, it is also possible to apply a transformation on the variables X or Θ so that they follow a standard normal PDF. Depending on the available information on the distribution of X, several transformations can be proposed [35], [36], [37], [38] and [39].

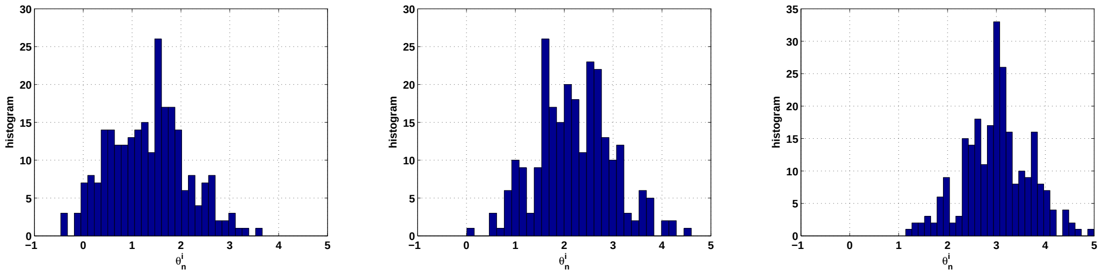
4. Application on different test cases

In this section, one applies i-PMCMC algorithm on a Gaussian toy case and on a realistic computer simulation of launch vehicle booster fallout.

4.1. *Gaussian toy case*

Let us consider the simple case where $Y = X$ with X following a Gaussian distribution $\mathcal{N}(\Theta, 1)$. The law of Θ is defined by a standard normal distribution $\mathcal{N}(0, 1)$. The probability of interest is $\mathbb{P}(Y > S)$ for different values of $S = \{2, 4, 6\}$. The i-PMCMC algorithm has been applied on this case with the following parameters : $N_1 = 250$, $N_2 = 500$, $N_{app} = 3$, $N_{app2} = 3$. The intermediate thresholds are set to $S_k = \{0, 0.5, 1, 0.5, 2, 2.5, 3, ..., 5.5, 6\}$. The histograms of $\widehat{\pi}_n$ obtained with i-PMCMC algorithm for the different values of S are given in figure 1. The corresponding probabilities $\left\{\widehat{\mathbb{P}}(Y > S|\Theta = \theta_n^i)\right\}_{i=1}^{N_1}$ available at the last iteration of **Algorithm** 1 are presented in figure 2. Computation times of the complete i-PMCMC algorithm for $S = \{2, 4, 6\}$ are respectively 230, 385 and 520 seconds.

The particles $\{\theta_n^i\}_{i=1}^{N_1}$ takes values that increase the probability $\widehat{\mathbb{P}}(Y > S|\Theta = \theta_n^i)$. In fact, the maximum of $\widehat{\mathbb{P}}(Y > S|\Theta = \theta_n^i)$ is reached when the value of Θ is higher than S, but it is not the case of the θ_n^i particles in i-PMCMC model. In fact, for $\mathbb{R}^d$-valued models, the i-PMCMC algorithm also takes into account the density f_Θ of Θ defined by



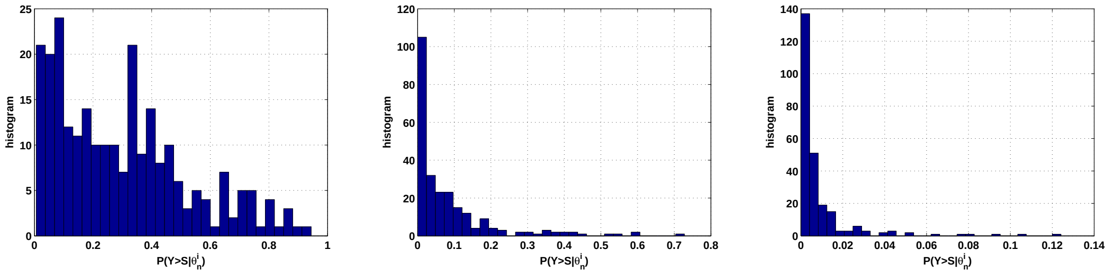
$S = 2$ $\qquad$ $S = 4$ $\qquad$ $S = 6$

Fig. 1. Estimate of π_n with particles $\left\{\theta_n^i\right\}_{i=1}^{N_1}$ for different values of S.

the probability law $\nu(d\theta) \propto f_\Theta(\theta)d\theta$. The density of high values of Θ is very low. Thus, the i-PMCMC model converges to a trade-off between maximization of $\widehat{\mathbb{P}}(Y > S|\Theta = \theta_n^i)$ and maximization of $f_\Theta(\theta_n^i)$. In this sense, f_Θ can be interpreted as a penalization for large values of Θ. Some estimates of the probability $\widehat{\mathbb{P}}(Y > S|\Theta = \theta)$ for different values of parameter Θ are presented in table 1. There is a significant difference between the probability estimated for $\theta = 0$, that is for the value of θ with the maximum likelihood, and the probability estimate for $\theta = \frac{1}{N_1}\sum_{i=1}^{N_1}\theta_n^i$, that is for the value of Θ that has the best trade-off between likelihood and probability. An overestimation of the parameter Θ can thus highly increase the probability $\widehat{\mathbb{P}}(Y > S|\Theta = \theta)$ and thus the tuning of Θ has to be carefully made (See table 1).

4.2. *Estimation of launch vehicle booster fallout zone*

Spatial launch vehicle fall-back safety zone estimation is a very important problem in space applications since the consequences of a mistake can be dramatic for the populations. We consider in this article a solid rocket booster that is the first stage of a launch vehicle. Its mass is about 35000 kilograms and the launch point is at 112 kilometer height



$S = 2$ $\quad$ $S = 4$ $\quad$ $S = 6$

Fig. 2. Estimate of $\mathbb{P}(Y > S|\Theta = \theta_n^i)$ with particles $\left\{\theta_n^i\right\}_{i=1}^{N_1}$ for different values of S.

S	$\widehat{\mathbb{P}}(Y > S\vert\Theta = 0)$	$\widehat{\mathbb{P}}(Y > S\vert\Theta = \frac{1}{N_1}\sum_{i=1}^{N_1}\theta_n^i)$
2	0.23	0.28
4	$3.2\ 10^{-5}$	0.036
6	$9.9\ 10^{-10}$	$1.9\ 10^{-3}$

Table 1

Estimates of the probability $\mathbb{P}(Y > S|\Theta = \theta)$ for different values of parameter θ.

with a slope of 15 degrees. At the end of its mission, the rocket booster falls into the sea at some distance of a predicted position. Similar models have already been analyzed in [40].

The launch vehicle stage fall-back is thus modeled as an input-output function ϕ with 4 Gaussian inputs X and one output $Y = \phi(X)$, representing the distance between the estimated launch stage fall-back position and the predicted one. In this study case, the aim is to estimate the probability that the distance to the predicted impact position exceeds 0.72 km: $\mathbb{P}(\phi(X) > 0.72)$. Several inputs of X can then influence its impact position:

- meteorological conditions (2 inputs). The wind variations during the fall-back can

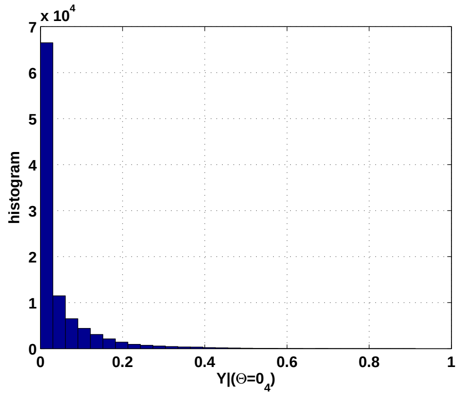


Fig. 3. Law of Y given $\Theta = 0_4$ estimated with 100000 runs of Monte Carlo simulations.

influence the impact position.

- launch vehicle mass (1 input). The mass of the different parts of the launch vehicle is also slightly random during the fall-back.
- the slope angle between the vertical axis and the velocity vector (1 input).

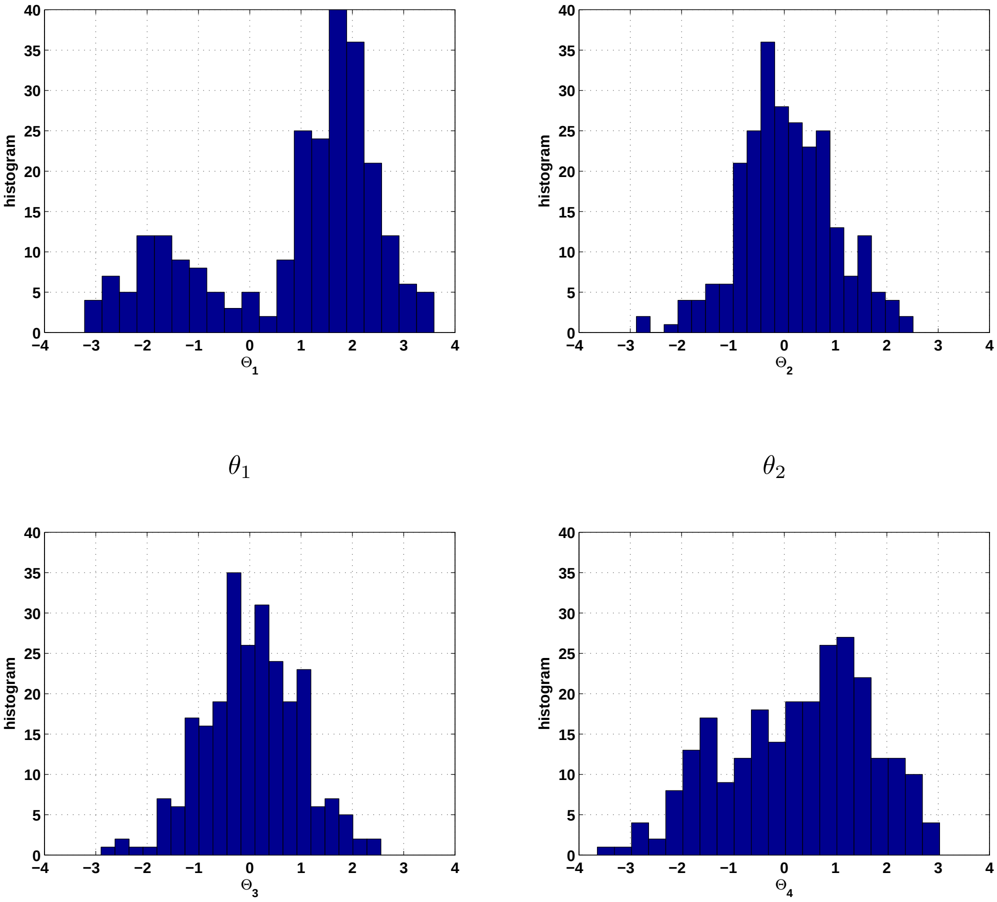
The uncertainty on X is modeled with a Gaussian random vector with mean $\Theta = (\Theta_1, \Theta_2, \Theta_3, \Theta_4)^t$ and a covariance matrix equal to the identity matrix. Its mean vector Θ is also uncertain since it is difficult to determine an accurate estimate of these quantities. We thus assume that Θ is a Gaussian random vector with mean 0_4 and a covariance matrix equal to the identity matrix.

The law of Y given $\Theta = 0_4$ is represented in figure 3 with 100000 runs of Monte-Carlo simulations. The probability $\mathbb{P}(\phi(X) > 0.72|\Theta = 0_4)$ is estimated to $8.5\ 10^{-4}$.

The i-PMCMC model has been applied on this case with the following parameters: $N_1 = 250$, $N_2 = 500$, $N_{app} = 3$, $N_{app2} = 3$. The intermediate thresholds are set to $S_k = \{0, 0.2, 0.4, 0.6, 0.72\}$. The estimate of $\widehat{\pi}_n$ obtained with i-PMCMC algorithm is given in figure 4 and 5. Table 2 also presents some estimations of probability $\mathbb{P}\left(Y > 0.72|\Theta = \theta\right)$ for different values of θ. Parameters Θ_1 and Θ_4 have to be tuned accurately in order to not underestimate $\mathbb{P}(\phi(X) > 0.72)$. Indeed, distributions of the θ-particles at the end of the i-PMCMC algorithm are very different from initial probability distribution ν (centered on 0) for parameters Θ_1 and Θ_4. It means that, if the assumption "Θ_1 and Θ_4 have a zero mean" is not valid, then, there is a high risk to underestimate $\mathbb{P}(\phi(X) > 0.72|\Theta = \theta)$ and it can be very problematic for obvious safety reasons. At the opposite, the initial assumption on parameters Θ_2 and Θ_3 already maximize the trade-off between likelihood of Θ and the probability $\mathbb{P}(\phi(X) > S)$. Thus, if Θ_2 and Θ_3 are not equal to 0, the probability $\mathbb{P}(\phi(X) > S)$ will not increase too much or even decrease which is positive for safety. The estimate probabilities $\widehat{\mathbb{P}}(Y > S|\Theta = \theta)$ with Θ following $\nu_{|Y>S}$ are given in figure 6. Their values can be very higher than $\widehat{\mathbb{P}}(Y > S|\Theta = 0_4)$, that is when Θ is equal to its mean value. When $\theta = \frac{1}{N_1}\sum_{i=1}^{N_1}\theta_n^i$, the probability $\widehat{\mathbb{P}}(Y > 0.72|\Theta = \frac{1}{N_1}\sum_{i=1}^{N_1}\theta_n^i)$ is equal to 0.034. This increase of the probability, compared to the initial situation where $\Theta = 0_4$, has to be taken into account in order to not underestimate the rare event probability. For large values of the threshold S the probabilities $\mathbb{P}(Y > S|\theta)$ can be extremely small and the computational cost rather high. For critical values of interest, which can be evaluated in a reasonable time, the i-PMCMC outperform the conventional Monte Carlo method, and as mentioned in the introduction, bypass the long time equilibrium convergence of the PMCMC.

5. Conclusion

In this article, we have proposed an original methodology to analyze the influence of the input PDF parameters on a rare failure probability. The proposed method based on i-PMCMC algorithm has been described in the case of a general problem where the



θ_1 θ_2

θ_3 θ_4

Fig. 4. Estimate of π_n w.r.t. the different components of parameter Θ.

model is a black-box system. The last part of this article deals with the application of this algorithm on the estimation of a launch vehicle booster fallout zone. We firstly show that it is important to estimate input PDF parameters since they influence strongly the value of the output model probability. The conditional law $\widehat{\pi}_n$ is estimated with success in that realistic case.

A possible perspective to this work could be the analysis of $\widehat{\pi}_n$ obtained by the i-PMCMC

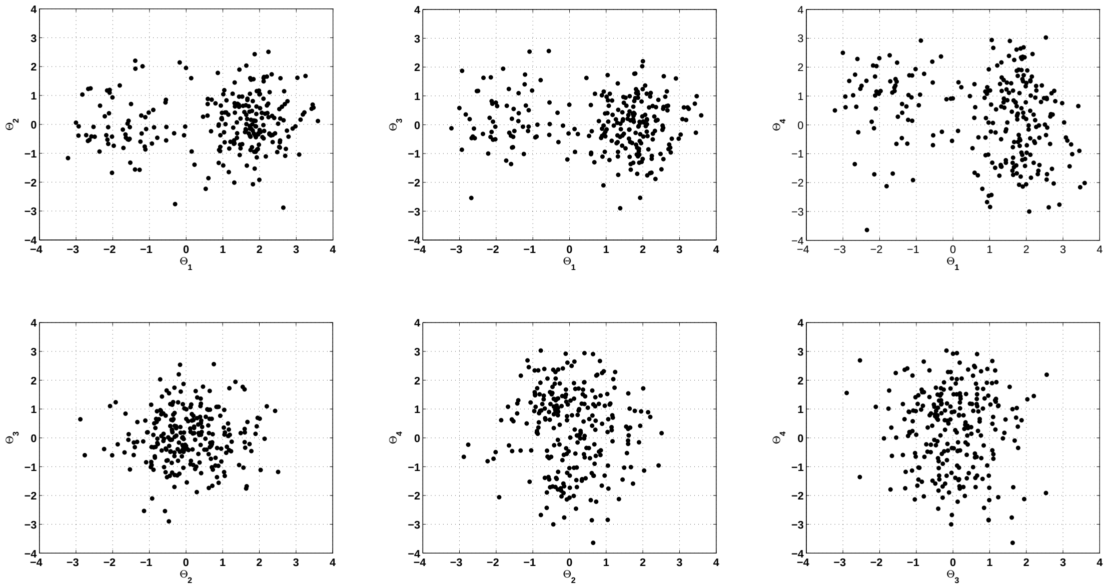


Fig. 5. Estimate of π_n w.r.t. the different components of parameter Θ.

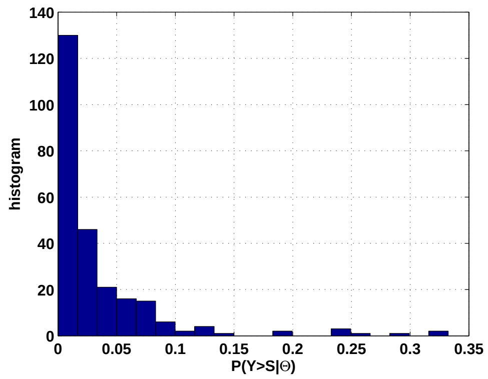


Fig. 6. Estimate of $\mathbb{P}(Y > S|\theta)$ with Θ following $\nu_{|Y>S}$.

algorithm with Sobol indices [41]. A possible ranking of the most influent input PDF parameters could then be derived.

θ_1	θ_2	θ_3	θ_4	$\hat{\mathbb{P}}(Y > 0.72\|\theta)$
0	0	0	0	$8.5\ 10^{-4}$
1	0	0	0	$3.5\ 10^{-3}$
1	0	0	1	$1.05\ 10^{-2}$
-1	0	0	1	$1.02\ 10^{-2}$
-1	0	0	-1	$1.14\ 10^{-2}$
0	-1	0	0	$9.4\ 10^{-4}$
0	-1	1	0	$9.8\ 10^{-4}$
0	1	1	0	$9.7\ 10^{-4}$
0	1	-1	0	$1.0\ 10^{-3}$
0	0	-1	0	$7.2\ 10^{-4}$

Table 2

Estimates of the probability $\mathbb{P}(Y > S|\theta)$ for different values of parameter θ.

6. Acknowledgement

The work of Christelle Vergé is financially supported by CNES (Centre National d'Etudes Spatiales) and Onera, The French Aerospace Lab.

References

[1] G. C. Orsak and B. Aazhang, "A class of optimum importance sampling strategies," *Information Sciences*, vol. 84, no. 1-2, pp. 139–160, 1995.

[2] D. P. Kroese and R. Y. Rubinstein, "Monte-Carlo methods," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 1, pp. 48–58, 2012.

[3] J.-F. Richard and W. Zhang, "Efficient high-dimensional importance sampling," *Journal of Econometrics*, vol. 141, no. 2, pp. 1385–1411, 2007.

[4] S. Engelund and R. Rackwitz, "A benchmark study on importance sampling techniques in structural reliability," *Structural Safety*, vol. 12, no. 4, pp. 255–276, 1993.

[5] H. Kahn and T. Harris, "Estimation of particle transmission by random sampling," *Appl. Math. Ser.*, vol. 12, pp. 27–30, 1951.

[6] M. Rosenbluth and A. Rosenbluth, "Monte-Carlo calculation of the average extension of molecular chains," *J. Chem. Phys*, vol. 356, pp. 356–359, 1955.

[7] P. Del Moral, *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications. Probability and its Applications.* New York: Springer, 2004.

[8] ——, *Mean field simulation for Monte Carlo integration.* Chapman and Hall/CRC Monographs on Statistics and Applied Probability, 2013.

[9] F. Cerou, P. Del Moral, T. Furon, and A. Guyader, "Sequential Monte-Carlo for rare event estimation," *Statistics and Computing*, vol. 22, pp. 795–808, 2012.

[10] Z. Yan-Gang and O. Tetsuro, "A general procedure for first/second-order reliabilitymethod (FORM/SORM)," *Structural Safety*, vol. 21, no. 2, pp. 95–112, 1999.

[11] R. Bjerager, *Methods for structural reliability computation.* New York: Springer Verlag, 1991, pp. 89–136.

[12] H. Madsen, S. Krenk, and N. C. Lind, *Methods of structural safety.* Englewood Cliffs: Springer-Verlag, 1986.

[13] R. Fisher and L. Tippett, "On the estimation of the frequency distributions of the largest or smallest member of a sample." Proceedings of the Cambridge Philosophical Society, 1928.

[14] M. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and related properties of random sequences and series.* New York, USA: Springer Verlag, 1983.

[15] M. Leadbetter and H. Rootzén, "Extremal theory for stochastic processes," *Annals of Probability*, vol. 16, pp. 431–478, 1988.

[16] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity analysis in practice: A guide to assessing scientific models.* Chichester, England: John Wiley & Sons, 2004.

[17] M. Lemaire, A. Chateauneuf, and J. Mitteau, *Structural reliability.* Wiley Online Library, 2009.

[18] M. Munoz Zuniga, J. Garnier, E. Remy, and E. de Rocquigny, "Adaptive directional stratification for controlled estimation of the probability of a rare event," *Reliability Engineering & System Safety*, vol. 96, no. 12, pp. 1691 – 1712, 2011.

[19] J. Morio, "Influence of input pdf parameters of a model on a failure probability estimation," *Simulation Modelling Practice and Theory*, vol. 19, no. 10, pp. 2244–2255, 2011.

[20] P. Del Moral, P. Hu, and L. Wu, "On the concentration properties of interacting particle processes," vol. 3, no. 3-4, pp. 225–389, 2012.

[21] N. Chopin, P. E. Jacob, and O. Papaspiliopoulos, "SMC^2: an efficient algorithm for sequential analysis of state space models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, pp. n/a–n/a, 2012.

[22] C. Vergé, C. Dubarry, P. Del Moral, and E. Moulines, "On parallel implementation of sequential monte carlo methods: the island particle model," 2013.

[23] C. Andrieu, A. Doucet, and R. Holenstein, "Particle markov chain monte carlo methods," *Journal of the Royal Statistical Society: Series B*, vol. 72, no. 3, pp. 269–342, 2010.

[24] A. J. Keane and P. B. Nair, *Computational Approaches for Aerospace Design.* John Wiley & Sons, Ltd, 2005.

[25] I. Banerjee, S. Pal, and S. Maiti, "Computationally efficient black-box modeling for feasibility analysis," *Computers & Chemical Engineering*, vol. 34, no. 9, pp. 1515 – 1521, 2010.

[26] A. Grancharova, J. Kocijan, and T. A. Johansen, "Explicit output-feedback nonlinear predictive control based on black-box models," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 2, pp. 388–397, 2011.

[27] V. Kreinovich and S. Ferson, "A new Cauchy-based black-box technique for uncertainty in risk analysis," *Reliability Engineering & System Safety*, vol. 85, no. 1-3, pp. 267–279, 2004.

[28] K. Worden, C. Wong, U. Parlitz, A. Hornstein, D. Engster, T. Tjahjowidodo, F. Al-Bender, D. Rizos, and S. Fassois, "Identification of pre-sliding and sliding friction dynamics: Grey box and black-box

models," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 514–534, 2007.

[29] I. M. Sobol, *A Primer for the Monte Carlo Method.* Boca Raton, Fl.: CRC Press, 1994.

[30] J. A. Bucklew, *Introduction to Rare Event Simulation.* Springer, 2004.

[31] R. Rubinstein and D. Kroese, *The Cross-Entropy Method : A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning (Information Science and Statistics).* Springer, 2004.

[32] P. Zhang, "Nonparametric importance sampling," *Journal of the American Statistical Association*, vol. 91, no. 434, pp. 1245–1253, Sept. 1996.

[33] Z. I. Botev and D. P. Kroese, "Efficient Monte-Carlo simulation via the generalized splitting method." *Statistics and Computing*, vol. 22, no. 1, pp. 1–16, 2012.

[34] L. Tierney, "Markov chains for exploring posterior distributions," *Annals of Statistics*, vol. 22, pp. 1701–1762, 1994.

[35] A. Nataf, "Distribution des distributions dont les marges sont données," *Comptes rendus de l'Académie des Sciences*, vol. 225, pp. 42–43, 1962.

[36] L. Pei-Ling and A. D. Kiureghian, "Optimization algorithms for structural reliability," *Structural Safety*, vol. 9, no. 3, pp. 161–177, 1991.

[37] R. Lebrun and A. Dutfoy, "An innovating analysis of the Nataf transformation from the copula viewpoint," *Probabilistic Engineering Mechanics*, vol. 24, no. 3, pp. 312–320, 2009.

[38] M. Rosenblatt, "Remarks on a multivariate transformation," *Annals of Mathematical Statistics*, vol. 23, pp. 470–472, 1952.

[39] R. Lebrun and A. Dutfoy, "A generalization of the Nataf transformation to distributions with elliptical copula," *Probabilistic Engineering Mechanics*, vol. 24, no. 2, pp. 172–178, 2009.

[40] J. Morio, "Non-parametric adaptive importance sampling for the probability estimation of a launcher impact position," *Reliability Engineering and System Safety*, vol. 96, no. 1, pp. 178 – 183, 2011.

[41] I. Sobol and S. Kuchereko, "Sensitivity estimates for non linear mathematical models," *Mathematical Modelling and Computationnal Experiments*, vol. 1, pp. 407–414, 1993.