Enhancing Reliability Analysis with Limited Observations: A Statistical Framework for System Safety Margins

Guillaume Perrin, Julien Reygner, Vincent Chabridon

Enhancing Reliability Analysis with Limited Observations: A Statistical Framework for System Safety Margins

Guillaume Perrin[a], Julien Reygner[b], Vincent Chabridon[c]

[a] *Université Gustave Eiffel, COSYS, 14-20 Boulevard Newton, 77447 Marne-la-Vallée, France*
[b] *CERMICS, École des Ponts, Marne-la-Vallée, France*
[c] *EDF R&D, 6 quai Watier, 78401 Chatou, France*

Abstract

Reliability analysis of complex systems is essential for cost-effective evaluations. This study addresses scenarios where optimal system operation is analyzed using deterministic black-box models. The input vector, $\boldsymbol{x} \in \mathbb{X} \subset \mathbb{R}^d$, characterizes both the system and environmental conditions, while y represents the variable of interest associated to a "failure domain" expressed as $\mathcal{F} = \{\boldsymbol{x} \in \mathbb{X} \mid y(\boldsymbol{x}) \leqslant s^\star\}$, with $s^\star \in \mathbb{R}$ a given threshold. Given the uncertainties in $\boldsymbol{x}$, this deterministic model requires a probabilistic analysis for robust safety demonstrations. Such an analysis typically involves two key steps: first, estimating the system's probability of failure (denoted by p_f) which is the statistical quantity of interest, and then, evaluating it against safety standards or expert knowledge. While considerable effort has been invested in proposing efficient methods for estimating p_f, little attention has been paid to the decision phase, which should take into account the uncertainties (e.g., in the p_f estimate). This work focuses on the definition and use of safety margins in system reliability analysis with a final decision making purpose, especially when the knowledge of the input vectors $\boldsymbol{x}$ is limited to a finite set of n observations. An important distinction is made between cases where n is large or small relative to $1/p_\mathrm{f}$. The main contributions of the paper focus on the scenario where n is relatively small, and propose two dif-

Email addresses: guillaume.perrin@univ-eiffel.fr (Guillaume Perrin), julien.reygner@enpc.fr (Julien Reygner), vincent.chabridon@edf.fr (Vincent Chabridon)

September 24, 2024

ferent approaches, each of which allowing the definition of reasonable safety margins. The first method focuses on the estimation of the probability distribution of the code inputs $\boldsymbol{x}$, while the second method, closely related to extreme value theory, focuses on the influence of the direct estimation of the tail of the model output probability distribution. The developed framework is then numerically validated on two test cases.

Keywords: Reliability Analysis, Complex Systems, Probability of Failure, Safety Margins, Extreme Value Theory

1. Introduction

The reliability analysis of complex systems is increasingly based on simulation. In that context, this work is concerned with the case where the proper operation of these systems of interest can be analyzed using a deterministic black-box model. Let us note $\boldsymbol{x} \in \mathbb{X} \subset \mathbb{R}^d$ the input vector of this model, i.e., the vector characterizing the system of interest and the conditions under which it will evolve, and y the variable of interest (supposed to be scalar for the sake of simplicity) that is used to characterize the system's failure in the sense that one can define a so-called "failure domain" by $\mathcal{F} = \{\boldsymbol{x} \in \mathbb{X} \mid y(\boldsymbol{x}) \leqslant s^\star\}$, with $s^\star \in \mathbb{R}$ a prescribed threshold. By deterministic model, we mean that if we provide the same value for $\boldsymbol{x}$ to the code twice, we obtain exactly the same value for y twice. Here, the term "black-box" indicates that the equations linking the values of $\boldsymbol{x}$ to y are not explicit, and that only point estimates of y are available for the reliability assessment.

We assume that the value of $\boldsymbol{x}$ associated with a specific system is not perfectly known (due to manufacturing tolerances, for example), leading to consider it in a probabilistic setting as being a collection of random variables gathered in a random vector $\boldsymbol{X}$, whose probability distribution is denoted by $\mu_{\boldsymbol{X}}$. Given the random nature of the system characteristics, and denoting by $Y := y(\boldsymbol{X})$ the random variable with probability distribution ν_Y, the safety demonstration of the system now relies on two distinct steps:

1. The estimation of a quantity of interest which reflects the risk [26], typically, the system's probability of failure (supposed to be associated to a rare failure event), which is denoted by p_f, and which can be defined in several ways, whether at the level of the random inputs $\boldsymbol{X}$ or at the

level of the random output $Y = y(\boldsymbol{X})$:

$$p_{\mathrm{f}} := \mathbb{P}_{\mu_{\boldsymbol{X}}}(\boldsymbol{X} \in \mathcal{F}) = \int_{\mathcal{F}} \mathrm{d}\mu_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbb{P}_{\nu_Y}(Y \leqslant s^\star) = \int_{-\infty}^{s^\star} \mathrm{d}\nu_Y(y); \quad (1)$$

2. Its confrontation either to safety standards or expert knowledge, to decide whether this value is sufficiently small or not.

While much effort has been put to derive various efficient approximation or estimation techniques (possibly using advanced adaptive surrogate-based strategies) for static rare event estimation over the last decades (see, e.g., [18, 3] for comprehensive reviews), this decision phase does not seem to be really addressed in the literature. To be more precise, this type of problem has been seen in two paths of research:

- the first one focuses on the evaluation of the robustness of the failure probability estimation using several mathematical frameworks (see, e.g., [20, 7] in a Bayesian framework, [28] using "Optimal Uncertainty Quantification", [13] using an information-geometric framework and [1] using an info-gap approach);

- the second one focuses on the definition and evaluation of safety margins in the context of risk and reliability assessment (see, e.g., [10, 11, 9] and [29]).

However, the present work addresses this problem of safety margin and/or robustness evaluation from a slightly different perspective than the existing literature. Here, the focus is on a reliability-based decision making process while considering the constraint of limited statistical information about the input data.

Thus the definition of these safety margins, and their use in system reliability analysis, is therefore at the heart of this work. And of particular interest is the case where the maximum information about the input vector is limited to a set of n observations, which are denoted by $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ and are gathered in the set $\mathcal{D}_n = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$. We assume that these observations can be considered as n independent and identically distributed (i.i.d.) realizations of $\boldsymbol{X}$. In parallel, we define the realizations of Y associated with these realizations of $\boldsymbol{X}$ by:

$$Y_1 := y(\boldsymbol{X}_1), \ldots, Y_n := y(\boldsymbol{X}_n). \quad (2)$$

In this work, we focus on the case where n is **predetermined** (the realizations of $\boldsymbol{X}$ can thus correspond to the results of a completed measurement campaign for instance), and we distinguish two configurations: the case where n is high compared to $1/p_{\mathrm{f}}$, and the case where n is smaller than or of the same order as $1/p_{\mathrm{f}}$. The case where n is sufficiently large is mainly used to illustrate the proposed formalism on a simple case, and leads to relatively classical results. The main innovative contributions of this paper are then associated with the case where n is relatively small, for which it is highly likely that all the values of $Y_1, \ldots, Y_n$ will be greater than $s^\star$ (thus, not in the failure domain). Building an estimator of p_{f} from these values of Y alone is therefore unlikely to be very useful (in the sense that a large variance of the estimator is expected). In order to define more interesting estimators of p_{f}, we propose to add information on the structure of the probability distribution $\mu_{\boldsymbol{X}}$ of $\boldsymbol{X}$, or on the tail of the probability distribution ν_Y of Y. More precisely, these distributions are assumed to admit a probability density function (PDF), which can be searched among well-chosen sets of parametric PDFs [30]. The construction of the estimator of p_{f} is then based on two steps. The information gathered in $\mathcal{D}_n$ is first used to estimate the parameters that characterize $\mu_{\boldsymbol{X}}$ (or ν_Y). The estimator of p_{f} is then constructed from these estimated distributions, whether or not based on new evaluations of y. In that case, it is important to notice that two sources of uncertainty need to be considered in the reliability analysis and the final decision: the fact that $\mu_{\boldsymbol{X}}$ (or ν_Y) is estimated using a finite number of realizations of $\boldsymbol{X}$ (or Y), and can therefore be considered as random, and the fact that the estimator of p_{f} given the estimator of $\mu_{\boldsymbol{X}}$ (or ν_Y) may also be random (if we consider, for example, an estimator of p_{f} based on samples drawn according to the estimated measure, as will be the case in Section 4).

There are several papers in the literature that also looked at the consequences of this double source of uncertainty, sometimes called "bi-level" uncertainty (see for instance [6]). But as mentioned above, the positioning of this paper is different in that what interests us here is quantifying the influence of this double source of uncertainty on final decision-making, with a particular focus on the definition of safety margins in a statistical framework, which will no longer be as straightforward as in the case where n is large.

To meet these objectives, the outline of this paper is as follows. Section 2 presents the formalism to make reliability-based decision in a statistical

framework. Section 3 then applies this formalism to the case where a large number of observations of $\boldsymbol{X}$ is available, while Section 4 deals with the definition of safety margins in the context where n is relatively small, and brings together the paper's main original contributions. Section 5 illustrates numerically the theoretical developments of Section 4 on two analytical cases, and Section 6 concludes the paper while putting forward some directions for further work.

Notations. In the rest of this paper, the random quantities will be written in uppercase, and the deterministic quantities in lowercase. Thus, A will refer to a random variable that can take the value a, and similarly, $\boldsymbol{A}$ will refer to a random vector that can take the value $\boldsymbol{a}$. In addition, for any positive integer q, any real-valued measurable function h defined on $\mathbb{R}^q$, and any subset $\mathcal{A}$ of $\mathbb{R}$, the notation $\mathbb{P}_\mu(h(\boldsymbol{A}) \in \mathcal{A})$ designates the probability that $h(\boldsymbol{A})$ is in $\mathcal{A}$ under the assumption that $\boldsymbol{A}$ is of measure μ. And to simplify the reading, if $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n$ correspond to n independent copies of the same random vector $\boldsymbol{A}$ of measure μ, then we simply note $\mathbb{P}_\mu(\widetilde{h}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n) \in \mathcal{A})$ the probability that $\widetilde{h}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n)$ is in $\mathcal{A}$, with $\widetilde{h}$ a new deterministic measurable function. The typical situation in which we will use this notation is when we shall deal with an estimator $\widehat{P}_n$ of the probability of failure, which is a deterministic measurable function of the sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$. In Section 4.2, where the estimator $\widehat{P}_{n,m}$ will be constructed from m random variables $\boldsymbol{X}_1^{(n)}, \ldots, \boldsymbol{X}_m^{(n)}$ resampled according to some probability measure $\widehat{\mu}_n$ estimated from $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, we will still use the notation $\mathbb{P}_\mu(\widehat{P}_{n,m} \in \mathcal{A})$ to keep track of the fact that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are i.i.d. under μ, even though $\widehat{P}_{n,m}$ is no longer a deterministic function of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$.
Then, we denote by $\mathcal{M}(\mathbb{X})$ the set of probability distributions defined on $\mathbb{X}$, and by $\mathcal{P}$ the particular subset (to be specified on a case-by-case basis) of $\mathcal{M}(\mathbb{X})$ in which the (true but unknown) probability distribution of the system input vector $\boldsymbol{X}$ is assumed to belong. This probability distribution is noted $\mu_{\boldsymbol{X}}$, and the definition of $\mathcal{P}$ is thus based on a prior knowledge about $\mu_{\boldsymbol{X}}$. When there is no assumption about this probability distribution, as will be the case in Sections 2 and 3, then $\mathcal{P} = \mathcal{M}(\mathbb{X})$. For any $p \in [0, 1]$ and any $\alpha \in (0, 1)$, we finally denote by $\mathcal{P}_p$ and $\mathcal{P}_{\geqslant\alpha}$ the subsets of $\mathcal{P}$ such that:

$$\mathcal{P}_p := \{\mu \in \mathcal{P} : \mathbb{P}_\mu(\boldsymbol{X} \in \mathcal{F}) = p\}, \qquad \mathcal{P}_{\geqslant\alpha} := \bigcup_{p \geqslant \alpha} \mathcal{P}_p. \tag{3}$$

By definition, as $p_{\mathrm{f}} = \mathbb{P}_{\mu_{\boldsymbol{X}}}(\boldsymbol{X} \in \mathcal{F})$, $\mu_{\boldsymbol{X}}$ is a particular element of $\mathcal{P}_p$ for which $p = p_{\mathrm{f}}$, and if α is actually lower than p_{f}, it is also an element of $\mathcal{P}_{\geqslant\alpha}$.

2. Statistical test formalism for reliability-based decision making

Deciding whether a particular system is *sufficiently safe*, or not, amounts to comparing its failure probability p_{f} to a certain value which can be called the *acceptable risk*, and denoted by α in the following (the greater the consequences of a failure, the smaller α is). A first decision criterion can then be stated regarding the safety of the system.

Decision Criterion 1 (Theoretical case). *The system is considered as sufficiently safe if* $p_{\mathrm{f}} < \alpha$.

As the true value of p_{f} is (most of the time, for real applications) unknown, this criterion is difficult to use in practice. To circumvent this problem, let us suppose that one can build a statistical estimator $\widehat{P}_n$ of p_{f} relying on the n evaluations of y at the n i.i.d. realizations of $\boldsymbol{X}$ gathered in $\mathcal{D}_n$. Since $\widehat{P}_n$ is an estimator of p_{f}, it is however important to keep in mind that it is very likely it over- or underestimates the true p_{f} (since it is associated to a symmetric cost function, see [20]). As a consequence, ensuring that $\widehat{P}_n$ is smaller than α is likely to be not sufficient to guarantee that the system of interest is actually safe. To better control the risks associated with this replacement of p_{f} by $\widehat{P}_n$, the classical formalism of statistical tests [17] is particularly suited to this context of finite-size sample. Hence, one can take as null and alternative hypotheses

$$H_0 = \{p_{\mathrm{f}} \geqslant \alpha\}, \quad H_1 = \{p_{\mathrm{f}} < \alpha\}. \tag{4}$$

Since one desires to test whether the system is sufficiently safe or not, one chooses the null hypothesis H_0 as corresponding to an unsafe system, accordingly to standard formulations in statistical hypothesis testing. Then, if we denote by $\beta \in (0,1)$ the classical type I error, this formalism invites us to consider the following new decision criterion to assess the system safety.

Decision Criterion 2 (Estimator-based case). *The system is considered as sufficiently safe if* $\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha$, *where the constant* $c^*(n,\alpha,\beta)$ *is the solution of the following constrained optimization problem:*

$$c^*(n,\alpha,\beta) := \inf\left\{c \in [0,\alpha] : \left\{\sup_{\mu\in\mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right)\right\} \leqslant \beta\right\}. \tag{5}$$

In Decision Criterion 2, we recall that $\mathcal{P}_{\geqslant\alpha}$ corresponds to the set of input probability distributions μ such that $\mathbb{P}_\mu(\boldsymbol{X} \in \mathcal{F})$ is higher than α. The constant $c^*(n, \alpha, \beta)$ acts therefore as a *safety margin* as it is supposed to avoid false certification with (high) probability $1 - \beta$. Indeed, if the true (but unknown) failure probability $p_{\mathrm{f}} = \mathbb{P}_{\mu_{\boldsymbol{X}}}(\boldsymbol{X} \in \mathcal{F})$ is actually higher than α, that is to say if $\mu_{\boldsymbol{X}}$ is in $\mathcal{P}_{\geqslant\alpha}$, $\widehat{P}_n + c^*(n, \alpha, \beta)$ should also be higher than α with probability $1 - \beta$.

Note also that, for all $\mu \in \mathcal{P}_{\geqslant\alpha}$, $\mathbb{P}_\mu(\widehat{P}_n + c < \alpha)$ corresponds to the probability for $\widehat{P}_n + c$ to be strictly lower than α, under the assumption that $\widehat{P}_n$ relies on the evaluation of $\boldsymbol{x} \mapsto y(\boldsymbol{x})$ in n i.i.d. copies of a random vector with probability distribution μ (not $\mu_{\boldsymbol{X}}$). As a consequence, the computation of $c^*(n, \alpha, \beta)$ in Eq. (5) depends on a priori information about the true probability distribution of $\boldsymbol{X}$ (allowing the optimization problem to be solved on a search space $\mathcal{P}$ potentially smaller than $\mathcal{M}(\mathbb{X})$), but does not require its precise knowledge.

Under that formalism, the certification of a system by simulation refers to several constants, which have different meanings (see Figure 1 for a graphical illustration). First, the constant $s^\star$ is the *threshold* that y should not exceed, which is generally provided either by safety standards or by expert judgments. Then, we have what we have called the *acceptable risk* α, which clearly depends on $s^\star$, and β, which characterizes the *confidence level* associated with the replacement of p_{f} by its estimator $\widehat{P}_n$. At last, there is the *safety margin* $c^*(n, \alpha, \beta)$, which depends on n, α and β (the smaller $c^*(n, \alpha, \beta)$ is, the more usable the estimator $\widehat{P}_n$ is in practice). Like any statistical test, let us insist on the anti-symmetric character of this decision criterion, which minimizes in priority the risk of false certification of the system. This may result in a significant number of safe systems, but with a failure probability close to α, which may not be labeled as safe by Decision Criterion 2.

The following propositions then give conditions on $\widehat{P}_n$ to ensure that Decision Criterion 2 is consistent, i.e., that it gives the same answer as Decision Criterion 1 with probability 1 when n tends to infinity. In this statement, we recall that we say that the estimator $\widehat{P}_n$ is *consistent* if it converges to p_{f} in probability under $\mathbb{P}_{\mu_{\boldsymbol{X}}}$.

Proposition 1. *For $p_{\mathrm{f}} > \alpha$, if $\widehat{P}_n$ is a consistent estimator of p_{f},*

$$\lim_{n\to+\infty} \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right) = 0. \tag{6}$$

▫ **Proof:** Let $\varepsilon = p_{\mathrm{f}} - \alpha > 0$. As $c^*(n,\alpha,\beta) \geqslant 0$, it comes:

$$\mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right) \leqslant \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n < \alpha\right) = \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\varepsilon < p_{\mathrm{f}} - \widehat{P}_n\right) \leqslant \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\varepsilon < |\widehat{P}_n - p_{\mathrm{f}}|\right). \tag{7}$$

Hence, the consistency of $\widehat{P}_n$ directly implies that $\mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right)$ tends to 0 when n tends to infinity. ▫

By adding the following assumption of uniform consistency, we can go further and show the consistency of Decision Criterion 2 on H_1, that is to say for the cases where p_{f} is strictly lower than α.

Definition 1 (Uniform consistency). *The estimator $\widehat{P}_n$ is said to be uniformly consistent if, for any $\varepsilon > 0$,*

$$\lim_{n\to+\infty} \sup_{\mu\in\mathcal{P}} \mathbb{P}_{\mu}\left(|\widehat{P}_n - \mathbb{P}_{\mu}(\boldsymbol{X} \in \mathcal{F})| \geqslant \varepsilon\right) = 0. \tag{8}$$

Proposition 2 (see Appendix A.1 for the proof). *For $p_{\mathrm{f}} < \alpha$, if $\widehat{P}_n$ is uniformly consistent,*

$$\lim_{n\to+\infty} \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right) = 1. \tag{9}$$

Remark 2.0.1. *Note that Decision Criterion 2 is not consistent for the very specific and unlikely case $p_{\mathrm{f}} = \alpha$.*

Remark 2.0.2. *For any $\mu \in \mathcal{P}$, if $q_{n;\mu}(\beta)$ denotes the β-quantile of $\widehat{P}_n$, such that:*

$$q_{n;\mu}(\beta) := \inf\left\{q \in [0,1] : \beta \leqslant \mathbb{P}_{\mu}\left(\widehat{P}_n \leqslant q\right)\right\}, \tag{10}$$

and if

$$q_n^*(\beta) := \inf_{\mu\in\mathcal{P}_{\geqslant\alpha}} q_{n;\mu}(\beta), \tag{11}$$

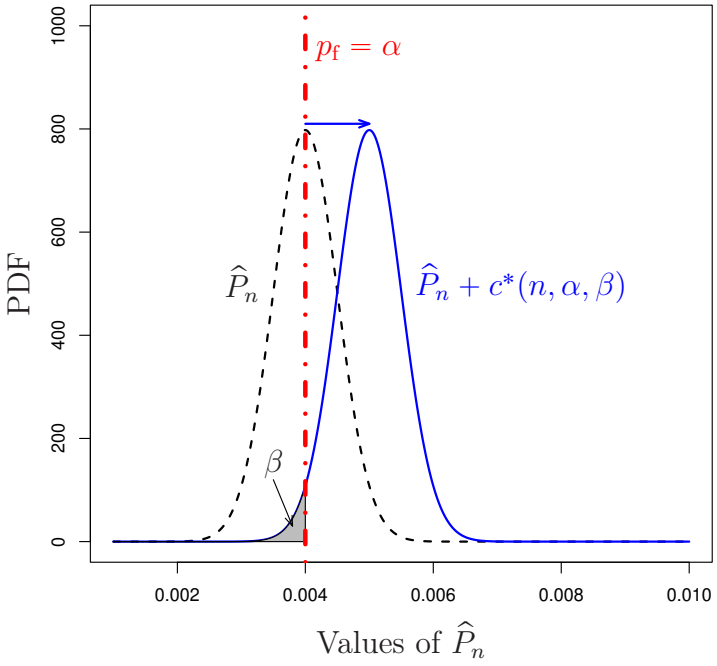


Figure 1: Graphical illustration of the meaning of the various constants leading to decision-making for the case $p_f = \alpha$. The vertical red dashed line indicates the value of α, the grey area corresponds to the probability β of false certification, the black dotted line characterizes the PDF of a particular unbiased estimator $\widehat{P}_n$ of p_f, while the blue solid line is associated with the PDF of $\widehat{P}_n + c^*(n, \alpha, \beta)$. As $p_f = \alpha$, the value of $c^*(n, \alpha, \beta)$ is supposed to be chosen such that $\mathbb{P}\left(\widehat{P}_n + c^*(n, \alpha, \beta) < \alpha\right) \leqslant \beta$.

then

$$
\begin{aligned}
c^*(n,\alpha,\beta) &= \inf\left\{c\in[0,\alpha] : \left\{\sup_{\mu\in\mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right)\right\} \leqslant \beta\right\} \\
&= \inf\left\{c\in[0,\alpha] : \forall\ \mu\in\mathcal{P}_{\geqslant\alpha},\ \mathbb{P}_\mu\left(\widehat{P}_n < \alpha - c\right) \leqslant \beta\right\} \\
&= \inf\left\{c\in[0,\alpha] : \forall\ \mu\in\mathcal{P}_{\geqslant\alpha},\ \alpha - c \leqslant q_{n;\mu}(\beta)\right\} \\
&= \inf\left\{c\in[0,\alpha] : \alpha - q_n^*(\beta) \leqslant c\right\} \\
&= \alpha - q_n^*(\beta).
\end{aligned}
\tag{12}
$$

Hence, the search for $c^*(n,\alpha,\beta)$ can be carried out in an equivalent way by solving the problem defined by Eq. (11), and Decision Criterion 2 can be equivalently reformulated by saying that the system can be considered sufficiently safe if $\widehat{P}_n < q_n^*(\beta)$. It is this second formulation that will be considered for the numerical estimation of the safety margins in Section 4.

3. Application to the case where n is sufficiently large

As a first application of the formalism presented in Section 2, we focus on the case where the size of $\mathcal{D}_n = \{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n\}$ is large. Given these i.i.d. realizations of $\boldsymbol{X}$ (and the associated realizations of $Y = y(\boldsymbol{X})$ denoted by $Y_1,\ldots,Y_n$), the estimator $\widehat{P}_n$ can be chosen as the usual Monte Carlo (MC) estimator of p_{f},

$$
\widehat{P}_n := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{y(\boldsymbol{X}_i)\leqslant s^\star\}} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{Y_i\leqslant s^\star\}},
\tag{13}
$$

where $\mathbb{1}_{\{a\}}$ is equal to 1 if a is true and to 0 otherwise. For this estimator, it is interesting to notice that for each subset $\mathcal{E}$ of $\mathbb{R}$, the value of $\mathbb{P}_\mu(\widehat{P}_n\in\mathcal{E})$ depends on μ only through the value of $p = \mathbb{P}_\mu(\boldsymbol{X}\in\mathcal{F})$. This means first, using the Chebyshev inequality, that for any $\varepsilon > 0$, $p\in[0,1]$ and $\mu\in\mathcal{P}_p$,

$$
\mathbb{P}_\mu\left(|\widehat{P}_n - p| \geqslant \varepsilon\right) \leqslant \frac{p(1-p)}{n\varepsilon^2} \leqslant \frac{1}{4n\varepsilon^2},
\tag{14}
$$

so that $\widehat{P}_n$ is uniformly consistent according to Definition 1. This also implies that if $0\leqslant\alpha\leqslant p\leqslant 1$, then for all $\mu_\alpha\in\mathcal{P}_\alpha$, $\mu\in\mathcal{P}_p$, and $c\in[0,\alpha]$,

$$
\mathbb{P}_{\mu_\alpha}\left(\widehat{P}_n + c < \alpha\right) \geqslant \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right),
\tag{15}
$$

so that:

$$\sup_{\mu\in\mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right) = \sup_{\mu\in\mathcal{P}_{\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right) = \mathbb{P}_{\mu_\alpha}\left(\widehat{P}_n + c < \alpha\right). \quad (16)$$

It is also well known, using the Central Limit Theorem (CLT), that for each $p \in [0,1]$, if $\widehat{P}_n$ is associated with n independent copies of a random vector with probability distribution $\mu \in \mathcal{P}_p$,

$$\frac{\sqrt{n}\left(\widehat{P}_n - p\right)}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \quad (17)$$

where $\xrightarrow{\mathcal{L}}$ is the convergence in law, and where for any $a \in \mathbb{R}$ and $b > 0$, $\mathcal{N}(a,b)$ denotes the set of Gaussian random variables whose mean and variance coefficients are equal to a and b respectively. In that case, for n sufficiently large and $G \sim \mathcal{N}(0,1)$ (i.e., distributed according to the standard Gaussian distribution), $\widehat{P}_n$ is close (in distribution) to $p + G\sqrt{p(1-p)/n}$, and therefore, for each $0 \leqslant c \leqslant \alpha$ and each $\mu \in \mathcal{P}_p$, $\mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right)$ is close to

$$\mathbb{P}_{\mathcal{N}(0,1)}(p + G\sqrt{p(1-p)/n} + c < \alpha) = \Phi\left((\alpha - c - p)\sqrt{n/(p(1-p))}\right), \quad (18)$$

where Φ is the cumulative distribution function (CDF) of G. The Gaussian assumption in the asymptotic case thus allows us to make $\mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right)$ depend explicitly on p. Hence, for sufficiently large n such that the Gaussian regime is reached, if $\phi_\beta = -\phi_{1-\beta}$ is the β-quantile of G, Eqs. (16) and (18) allow us to write

$$\begin{aligned} c^*(n,\alpha,\beta) &:= \inf\left\{c \in [0,\alpha] : \left\{\sup_{\mu\in\mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right)\right\} \leqslant \beta\right\} \\ &\approx \inf\left\{c \in [0,\alpha] : \Phi\left((\alpha - c - \alpha)\sqrt{n/(\alpha(1-\alpha))}\right) \leqslant \beta\right\} = \phi_{1-\beta}\sqrt{\frac{\alpha(1-\alpha)}{n}}. \end{aligned} \quad (19)$$

This leads to the following (asymptotic) Decision Criterion 3.

Decision Criterion 3 (Monte Carlo case). *If n is sufficiently large for the CLT to hold, the system is considered as sufficiently safe if $\widehat{P}_n + \phi_{1-\beta}\sqrt{\alpha(1-\alpha)/n} < \alpha$.*

It is important to note that Decision Criterion 3 only makes sense if $c^*(n,\alpha,\beta) < \alpha$, that is to say if n is greater than

$$n_{\min}^{\mathrm{MC}} := \phi_{1-\beta}^2(\alpha^{-1}-1). \tag{20}$$

For given values of p_{f} and α, and for n greater than $n_{\min}^{\mathrm{MC}}$ and sufficiently large for the Gaussian behavior of $\widehat{P}_n$ to be valid, the probability of certifying the system, which is defined by $\mathbb{P}_{\mu_{\boldsymbol{X}}}(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha)$, can then be approximated by:

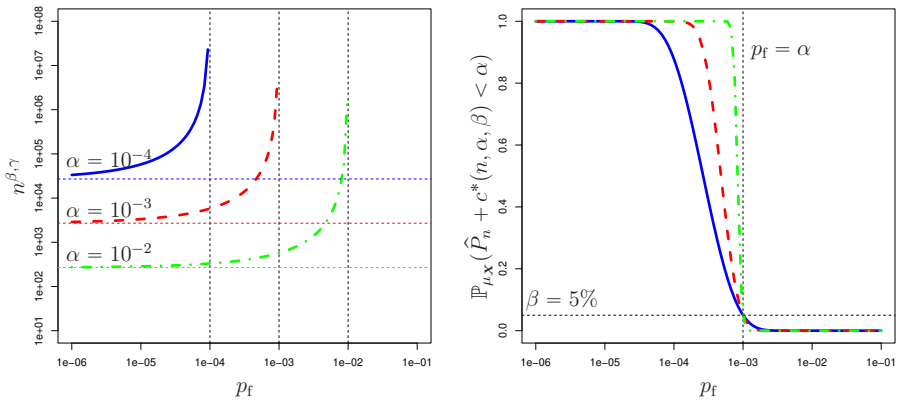
$$\Phi\left(\frac{\sqrt{n}(\alpha-p_{\mathrm{f}}) - \phi_{1-\beta}\sqrt{\alpha(1-\alpha)}}{\sqrt{p_{\mathrm{f}}(1-p_{\mathrm{f}})}}\right). \tag{21}$$

We recover the consistent character of this decision criterion, in the sense that when n tends to infinity this probability tends to 1 when $\alpha > p_{\mathrm{f}}$, and to 0 when $\alpha < p_{\mathrm{f}}$. We can also verify that for finite values of n, this probability is equal to β when $\alpha = p_{\mathrm{f}}$, which is coherent with the definition of β. In the context of the usual crude Monte Carlo reliability assessment, once n is sufficiently large, it is thus possible to specify a couple $(\widehat{P}_n, c^*(n,\alpha,\beta))$ allowing to evaluate, in a conservative manner, the proper operation of the system. Thus, three situations may occur:

- a favorable case, where $\widehat{P}_n + c^*(n,\alpha,\beta)$ is actually strictly smaller than α, which means that the risk of failure can reasonably be considered smaller than the acceptable risk;
- an unfavorable case, where $\widehat{P}_n \geqslant \alpha$ (necessarily implying $\widehat{P}_n + c^*(n,\alpha,\beta) \geqslant \alpha$), which shows a too important risk of failure for the system,
- an intermediate case, where $\widehat{P}_n < \alpha \leqslant \widehat{P}_n + c^*(n,\alpha,\beta)$, for which it is likely that the true failure probability p_{f} is smaller than α. But without sufficient statistical guarantees, it is preferable not to consider the system as sufficiently safe.

From Eq. (21), it is then possible to calculate, for different values of p_{f}, α and β (with $p_{\mathrm{f}} < \alpha$), the minimum value of n noted $n^{\beta,\gamma}$ allowing the *correct* labeling of the system with a probability greater than $1-\gamma$ (for $\gamma \in (0,1)$):

$$\begin{aligned} &\Phi\left(\frac{\sqrt{n^{\beta,\gamma}}(\alpha-p_{\mathrm{f}}) - \phi_{1-\beta}\sqrt{\alpha(1-\alpha)}}{\sqrt{p_{\mathrm{f}}(1-p_{\mathrm{f}})}}\right) = 1-\gamma \\ &\quad \Leftrightarrow \; n^{\beta,\gamma} := \frac{1}{(\alpha-p_{\mathrm{f}})^2}\left(\phi_{1-\gamma}\sqrt{p_{\mathrm{f}}(1-p_{\mathrm{f}})} + \phi_{1-\beta}\sqrt{\alpha(1-\alpha)}\right)^2. \end{aligned} \tag{22}$$



(a) $\beta = \gamma = 5\%$ (b) $\alpha = 10^{-3}$

Figure 2: Evolutions of $n^{\beta,\gamma}$ and $\mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right)$ with respect to p_f for different values of α and n, and for $\beta = \gamma = 5\%$. In the left figure, the blue solid line, the red dotted line and the green dashed line are respectively associated with $\alpha = 10^{-4}, 10^{-3}, 10^{-2}$, while the horizontal dotted lines indicate the lower bounds $n_{\min}^{\mathrm{MC}}$ in each case. In the right figure, α is chosen equal to 10^{-3}, and the blue solid line, the red dotted line and the green dashed line are associated with $N = 5 \times 10^3, 10^4, 10^5$ respectively, while the vertical and horizontal lines show the considered values of α and β respectively.

Figure 2 shows first the evolution of $n^{\beta,\gamma}$ and $\mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right)$ with respect to p_f for different values of n, α and $\beta = \gamma$. Focusing on Figure 2-(a), we find that, without surprise, the closer from above p_f is to α, the larger n must be to be able to say that p_f is smaller than α with sufficient confidence. A vertical asymptote in $\alpha = p_f$ can also be seen, which is due to the deliberately conservative nature of the reliability analysis. The case of p_f greater than α is of little interest, in the sense that $c^*(n,\alpha,\beta)$ is constructed to make, in that case and whatever the value of n, the probability of false labeling be lower than β. We also find an increase of n in $1/\alpha$ that is relatively classical for crude Monte Carlo approaches.

4. Application to the case where n is small

In this section, we now focus on the case where n is close to $1/\alpha$. Note that Decision Criterion 3 is no longer really useful in this case as $c^*(n,\alpha,\beta)$ will be

close to α and the probability of labeling the system as sufficiently safe will be close to 0 whatever the true value of p_{f}. This may also be explained by the fact that for such values of n and assuming that $p_{\mathrm{f}} < \alpha$, it is highly likely that there will be no failure points in the learning set $\mathcal{D}_n = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$.

To better support decision-making in such a context, it is then necessary to turn to alternative estimators of p_{f}, which will exploit additional assumptions about the statistical behavior of $\boldsymbol{X}$ and/or Y. This will result in the introduction of specific subsets $\mathcal{P}$ of $\mathcal{M}(\mathbb{X})$ for $\mu_{\boldsymbol{X}}$, whether working at the level of the output Y in Section 4.1, or at the level of the inputs $\boldsymbol{X}$ in Section 4.2.

4.1. Parametric approximation of the probability distribution of Y

4.1.1. Choice of the parametric family

From an output point of view, the objective of this section is to propose an estimator of p_{f} (different from the crude Monte Carlo estimator) based on the n i.i.d. realizations of $Y = y(\boldsymbol{X})$, which would allow the generalization of the developments presented in Section 3. If n is close to $1/\alpha$, we remind that the vast majority (if not all) of the values of $Y_1 = y(\boldsymbol{X}_1), \ldots, Y_n = y(\boldsymbol{X}_n)$ is likely to be above $s^\star$ (i.e., in the safe region). In this case, the Extreme Value Theory [15, 8] seems to be the most appropriate way of estimating the probability $\mathbb{P}_{\mu_{\boldsymbol{X}}}(y(\boldsymbol{X}) \leqslant s^\star) = \mathbb{P}_{\nu_Y}(Y \leqslant s^\star)$. According to this mathematical framework, the generalized Pareto distribution (GPD) is a good candidate [24] for modeling the left tail ("left" here, due to the formulation of the problem) of the CDF of Y. This CDF, which we denote by $F_{(\overline{s},\sigma,\xi)}$, is generally specified by three parameters (the location $\overline{s}$, the scale σ, and the shape ξ), such that

$$F_{(\overline{s},\sigma,\xi)}(s) = \begin{cases} \left(1 + \frac{\xi(\overline{s}-s)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ \exp\left(-\frac{\overline{s}-s}{\sigma}\right) & \text{for } \xi = 0. \end{cases} \tag{23}$$

By construction, this CDF is defined on $(-\infty, \overline{s}]$ if $\xi \geqslant 0$, and on $\left[\overline{s} + \frac{\sigma}{\xi}, \overline{s}\right]$ if $\xi < 0$. By introducing $\boldsymbol{w} := (\sigma, \xi, p_{\overline{s}})$, we then use the notation $Z \sim \mathrm{a-GPD}(\overline{s}, \boldsymbol{w})$ (for "adapted generalized Pareto distribution") to say that $\mathbb{P}(Z \leqslant \overline{s}) = p_{\overline{s}}$ and that $(Z \mid Z \leqslant \overline{s})$ has CDF $F_{(\overline{s},\sigma,\xi)}$. Notice that the statement that $Z \sim \mathrm{a-GPD}(\overline{s}, \boldsymbol{w})$ does not characterize the whole distribution of Z, but only describes its CDF $\mathbb{P}(Z \leqslant s)$ for values of s smaller than $\overline{s}$.

Hypothesis 4.1 (Asymptotic behavior of the PDF of Y). *In this section, we assume that there exist $\overline{s} \geqslant s^\star$ and $\boldsymbol{w}^\star := (\sigma^\star, \xi^\star, p^\star_{\overline{s}}) \in (0, +\infty) \times (-\infty, 1/2) \times (0, 1)$ such that*

$$Y \sim \mathrm{a} - \mathrm{GPD}(\overline{s}, \boldsymbol{w}^\star). \tag{24}$$

Hypothesis 4.1 tells us that $s \mapsto \mathbb{P}(Y \leqslant s \mid Y \leqslant \overline{s})$ coincides with a GPD with a shape parameter in $(-\infty, 1/2)$. The main reason for specifying that $\xi^\star$ is less than $1/2$ is to ensure that $(Y \mid Y \leqslant \overline{s})$ has a finite variance, which we accept here (other regimes could also be considered, in particular to ensure the uniform consistency of the estimator that will be proposed in Section 4.1.2, as explained in Proposition 3). We also assume that the value of the location parameter $\overline{s}$ is known (or empirically chosen as explained in Section 4.1.5).

4.1.2. Construction of a new estimator of p_{f}

Under Hypothesis 4.1, the probability of failure can be decomposed as

$$p_{\mathrm{f}} = \mathbb{P}_{\nu_Y}(Y \leqslant \overline{s}) \times \mathbb{P}_{\nu_Y}(Y \leqslant s^\star \mid Y \leqslant \overline{s}) = p^\star_{\overline{s}} \times F_{(\overline{s}, \sigma^\star, \xi^\star)}(s^\star). \tag{25}$$

In order to propose an alternative estimator of p_{f} to the crude Monte Carlo estimator, we therefore need to estimate $\boldsymbol{w}^\star$ from the available realizations of Y. Several methods can be found in the literature [24]. In this work, we limit ourselves to a method of moments, noting that if Hypothesis 4.1 is true,

$$\xi^\star = \frac{1}{2}\left(1 - \frac{\left(\overline{s} - \mathbb{E}_{\nu_Y}\left[Y \mid Y \leqslant \overline{s}\right]\right)^2}{\mathrm{Var}_{\nu_Y}(Y \mid Y \leqslant \overline{s})}\right), \quad \sigma^\star = \left(\overline{s} - \mathbb{E}_{\nu_Y}\left[Y \mid Y \leqslant \overline{s}\right]\right)(1 - \xi^\star). \tag{26}$$

Hence, denoting by

$$N_{\overline{s}} = \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \leqslant \overline{s}\}} \tag{27}$$

the number of realizations of $\boldsymbol{X}$ in $\mathcal{D}_n$ that are associated with values of Y that are lower than $\overline{s}$, and then introducing the estimators

$$\widehat{M}_n = \frac{1}{N_{\overline{s}}} \sum_{i=1}^{n} Y_i \, \mathbb{1}_{\{Y_i \leqslant \overline{s}\}}, \quad \widehat{S}^2_n = \frac{1}{N_{\overline{s}}} \sum_{i=1}^{n} Y_i^2 \, \mathbb{1}_{\{Y_i \leqslant \overline{s}\}} - \widehat{M}^2_n, \tag{28}$$

of $\mathbb{E}_{\nu_Y}[Y \mid Y \leqslant \overline{s}]$ and $\mathrm{Var}_{\nu_Y}[Y \mid Y \leqslant \overline{s}]$, the random variables

$$\widehat{\Xi}_n := \frac{1}{2}\left(1 - \frac{\left(\overline{s} - \widehat{M}_n\right)^2}{\widehat{S}_n^2}\right), \quad \widehat{\Sigma}_n := \frac{1}{2}\left(\overline{s} - \widehat{M}_n\right)\left(1 + \frac{\left(\overline{s} - \widehat{M}_n\right)^2}{\widehat{S}_n^2}\right), \tag{29}$$

are two natural moment estimators of $\xi^\star$ and $\sigma^\star$. As a consequence, and given Eq. (25), we introduce the following estimator of p_{f}:

$$\widehat{P}_n := \frac{N_{\overline{s}}}{n} \times F_{(\overline{s},\widehat{\Sigma}_n,\widehat{\Xi}_n)}(s^\star). \tag{30}$$

Remark 4.1.1. *Even if it is not possible to write the expression of the probability distribution of $\widehat{P}_n$ as an explicit function of $\boldsymbol{w}^\star$, it is relatively easy and quick to generate i.i.d. realizations of $\widehat{P}_n$ from i.i.d. realizations of uniform random variables, as explained in Appendix B.1.*

4.1.3. Uniform consistency of the proposed estimator

To prove the uniform consistency of the estimator $\widehat{P}_n$ that is introduced in Eq. (30), we need to reduce the possible ranges of variation of the components of $\boldsymbol{w}^\star$, and work with a priori bounds for $\sigma^\star$, $\xi^\star$ and $p^\star_{\overline{s}}$. We thus let $0 < \sigma_{\min} < \sigma_{\max} < +\infty$, $-\infty < \xi_{\min} < \xi_{\max} < 1/4$, and $0 < p_{\overline{s},\min} < p_{\overline{s},\max} \leqslant 1$, set

$$\mathcal{W} = [\sigma_{\min}, \sigma_{\max}] \times [\xi_{\min}, \xi_{\max}] \times [p_{\overline{s},\min}, p_{\overline{s},\max}],$$

and let $\mathcal{P}$ be the set of $\mu \in \mathcal{M}(\mathbb{X})$ such that, under $\mathbb{P}_\mu$, there exists $\boldsymbol{w} \in \mathcal{W}$ for which $y(\boldsymbol{X}) \sim \mathrm{a-GPD}(\overline{s}, \boldsymbol{w})$.

Proposition 3 (see Appendix A.2 for the proof). *The estimator $\widehat{P}_n$ defined in Eq. (30) is uniformly consistent for the parametric family $\mathcal{P}$.*

4.1.4. Practical solving of the constrained optimization problem

In order to apply Decision Criterion 2 in the former GPD context, one needs to solve the optimization problem

$$c^*(n,\alpha,\beta) := \inf\left\{c \in [0,\alpha] : \left\{\sup_{\mu \in \mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + c < \alpha\right)\right\} \leqslant \beta\right\}, \tag{31}$$

where $\widehat{P}_n$ is defined by Eq. (30), $\mathcal{P}_{\geqslant\alpha}$ is defined by Eq. (3), and where we recall that

$$\mathcal{P} := \left\{\mu \in \mathcal{M}(\mathbb{X}) : \exists\ \boldsymbol{w} \in \mathcal{W} \ /\ y(\boldsymbol{X}) \sim \mathrm{a-GPD}(\overline{s}, \boldsymbol{w})\right\}. \tag{32}$$

Let μ be an element of $\mathcal{P}$, and $\boldsymbol{w} = (\sigma, \xi, p_{\overline{s}}) \in \mathcal{W}$ be the vector that characterizes its left tail. Since $s^\star \leqslant \overline{s}$, $\mathbb{P}_\mu(y(\boldsymbol{X}) \leqslant s^\star)$ only depends on μ through $\boldsymbol{w}$, and there is no ambiguity in denoting it by $\mathbb{P}_{\boldsymbol{w}}(Y \leqslant s^\star)$. Similarly, the law of $\widehat{P}_n$ under $\mathbb{P}_\mu$ is completely characterized by $\boldsymbol{w}$, and for all $\mathcal{A} \subset \mathbb{R}$, we shall also use the notation $\mathbb{P}_{\boldsymbol{w}}\left(\widehat{P}_n \in \mathcal{A}\right)$ instead of $\mathbb{P}_\mu\left(\widehat{P}_n \in \mathcal{A}\right)$ to simplify reading. For any $p \geqslant \alpha$, we now introduce

$$\mathcal{W}_p := \{\boldsymbol{w} \in \mathcal{W} : \mathbb{P}_{\boldsymbol{w}}(Y \leqslant s^\star) = p\}, \qquad \mathcal{W}_{\geqslant\alpha} := \bigcup_{p \geqslant \alpha} \mathcal{W}_p, \tag{33}$$

so that the optimization problem in Eq. (31) rewrites in the form:

$$c^*(n, \alpha, \beta) := \inf\left\{c \in [0, \alpha] : \left\{\sup_{\boldsymbol{w} \in \mathcal{W}_{\geqslant\alpha}} \mathbb{P}_{\boldsymbol{w}}\left(\widehat{P}_n + c < \alpha\right)\right\} \leqslant \beta\right\}. \tag{34}$$

For any value of $c \in [0, \alpha]$, we are thus looking for the value of $\boldsymbol{w} = (\sigma, \xi, p_{\overline{s}}) \in \mathcal{W}$ that maximizes $\mathbb{P}_{\boldsymbol{w}}(\widehat{P}_n + c < \alpha)$ under the constraint that

$$p_{\overline{s}} \times F_{(\overline{s},\sigma,\xi)}(s^\star) \geqslant \alpha. \tag{35}$$

In the context of Section 3, Eq. (15) shows that $\widehat{P}_n$ is an increasing function of $p = \mathbb{P}_\mu(Y \leqslant s^\star)$, in the sense of stochastic ordering. In the computation of $c^*(n, \alpha, \beta)$, this allows one to reduce the search to measures $\mu \in \mathcal{P}_\alpha$ instead of $\mu \in \mathcal{P}_{\geqslant\alpha}$. In the present case, this stochastic monotonicity property may be tedious to prove, but it has been verified numerically by analyzing a very large number of configurations. It is also heuristically justified by Proposition 3, since the latter statement asserts that for n large enough, $\widehat{P}_n$ concentrates around p, uniformly in μ. So, in the present case we also reduce the search to $\mathcal{W}_\alpha$ instead of $\mathcal{W}_{\geqslant\alpha}$, which amounts to imposing the relation

$$p_{\overline{s}} \times F_{(\overline{s},\sigma,\xi)}(s^\star) = \alpha. \tag{36}$$

This first reduction allows us to express $p_{\overline{s}}$ as a function of (σ, ξ) and we now focus on maximizing $\mathbb{P}_{\boldsymbol{w}}\left(\widehat{P}_n + c < \alpha\right)$ as a function of (σ, ξ).

In addition, it is possible to restrict the domain of definition of (σ, ξ) by noting that if ξ is strictly negative, the definition domain of Y is necessarily $[\overline{s} + \sigma/\xi; \overline{s}]$. Assuming that the true probability of failure p_{f} is strictly positive (conservative assumption), and noting $Y_{\mathrm{mini}} = \min_{1 \leqslant i \leqslant n} Y_i$ the minimal observed value of Y when using the realizations of $\boldsymbol{X}$ that are in $\mathcal{D}_n$, this

also means that $\min(s^\star, Y_{\min}) \geqslant \overline{s} + \sigma/\xi$. This ensures that (σ, ξ) can be searched in

$$\mathcal{S}^{\text{cont}} := \left\{ (\sigma, \xi) \in [\sigma_{\min}, \sigma_{\max}] \times [\xi_{\min}, \xi_{\max}] : \ 1 + \frac{\xi(\overline{s} - \min(s^\star, Y_{\min}))}{\sigma} \geqslant 0 \right\}. \tag{37}$$

Note that this constraint on (σ, ξ) has been rewritten so that it always holds when ξ is positive, and only comes into play when ξ is negative. Finally, the safety margin $c^*(n, \alpha, \beta)$ can now be searched as the solution of the following simplified problem:

$$\inf \left\{ c \in [0, \alpha] : \left\{ \sup_{\boldsymbol{w} \in \mathcal{W}_\alpha^{\text{red}}} \mathbb{P}_{\boldsymbol{w}} \left(\widehat{P}_n + c < \alpha \right) \right\} \leqslant \beta \right\}, \tag{38}$$

where the reduced searching set is now defined by

$$\mathcal{W}_\alpha^{\text{red}} := \left\{ (\sigma, \xi, p_{\overline{s}}) \in \mathcal{W} : (\sigma, \xi) \in \mathcal{S}^{\text{cont}}, \ p_{\overline{s}} \times F_{(\overline{s}, \sigma, \xi)}(s^\star) = \alpha \right\}. \tag{39}$$

In order to solve this problem numerically, we propose to exploit the rewriting of $c^*(n, \alpha, \beta)$ proposed in Remark 2.0.2, saying that

$$c^*(n, \alpha, \beta) = \alpha - q_n^*(\beta), \tag{40}$$

where

$$q_n^*(\beta) = \inf_{\boldsymbol{w} \in \mathcal{W}_\alpha^{\text{red}}} q_{n;\boldsymbol{w}}(\beta), \quad \text{and} \quad q_{n;\boldsymbol{w}}(\beta) = \inf \left\{ q \in [0, 1] : \beta \leqslant \mathbb{P}_{\boldsymbol{w}} \left(\widehat{P}_n \leqslant q \right) \right\}. \tag{41}$$

We then proceed in three steps to find $q_n^*(\beta)$ (and $c^*(n, \alpha, \beta)$). First, we choose r_1 values of $\boldsymbol{w}$ in $\mathcal{W}_\alpha^{\text{red}}$ according to a space-filling design (see, e.g., [22] for more details about space-filling designs in such constrained spaces). Then, for each chosen value of $\boldsymbol{w}$, we then generate r_2 independent realizations of the associated estimator $\widehat{P}_n$ as it is explained in Appendix B.1. This allows us to estimate empirically its β-quantile $q_{n;\boldsymbol{w}}(\beta)$. Finally, a Bayesian optimization algorithm ([16, 14, 12, 23]) is used to construct a sequence of r_3 points of $\mathcal{W}_\alpha^{\text{red}}$ likely to converge towards $q_n^*(\beta)$ (in Section 5, we will take $r_1 = 40$, $r_2 = 10^4$ and $r_3 = 40$), from which we can deduce $c^*(n, \alpha, \beta)$ using Eq. (40).

4.1.5. Choice of the location parameter

The previous developments are conditional on the choice of $\overline{s}$. By construction, the higher $\overline{s}$, the more realizations of Y are available to estimate $\xi^\star$,

$\sigma^\star$ and $p_{\overline{s}}$, but the less likely it is that Hypothesis 4.1 will be appropriate, and the less likely it is that the tail approximation performed using data that are far from $s^\star$ will be able to approximate what is happening below $s^\star$. As a search for such a bias-variance compromise, we therefore propose to choose $\overline{s}$ such that the variance of $\widehat{P}_n$ is minimal, when relying to a bootstrap procedure over the n available realizations of Y to carry out the variance approximation.

4.2. Parametric approximation of the input probability distribution

4.2.1. Construction of a new estimator of p_{f}

As an alternative to the method proposed in Section 4.1, one can seek to further exploit the function y to explore the failure domain. We can then proceed in four steps: (1) select an adapted set for $\mu_{\boldsymbol{X}}$ noted $\mathcal{P}$; (2) estimate $\mu_{\boldsymbol{X}}$ in $\mathcal{P}$ from the n already available realizations of $\boldsymbol{X}$ (the approximate probability distribution is denoted by $\widehat{\mu}_n$, and the notation $\boldsymbol{X}^{(n)}$ is introduced to refer to random vectors with probability distribution $\widehat{\mu}_n$); (3) draw m supplementary i.i.d. realizations of $\boldsymbol{X}^{(n)}$, noted $\boldsymbol{X}_1^{(n)}, \ldots, \boldsymbol{X}_m^{(n)}$; and (4) estimate p_{f} by its crude Monte Carlo estimator from these m realizations:

$$\widehat{P}_{n,m} := \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{\left\{y(\boldsymbol{X}_j^{(n)}) \leqslant s^\star\right\}}. \tag{42}$$

It is important to note that this new estimator $\widehat{P}_{n,m}$ of p_{f} is affected by a double source of uncertainty (which explains the double index): the uncertainty related to the random nature of $\widehat{\mu}_n$ (due to the fact that the number of available realizations of $\boldsymbol{X}$ is finite) and the uncertainty related to the fact that the function y is evaluated only in a finite set of m points. The safety margin introduced in Decision Criterion 2 is now expected to depend on these values of n and m. It is then denoted $c^*(n, m, \alpha, \beta)$, and defined by:

$$c^*(n, m, \alpha, \beta) := \inf \left\{ c \in [0, \alpha] : \left\{ \sup_{\mu \in \mathcal{P}_{\geqslant \alpha}} \mathbb{P}_\mu \left(\widehat{P}_{n,m} + c < \alpha \right) \right\} \leqslant \beta \right\}. \tag{43}$$

For all $\mu \in \mathcal{P}_{\geqslant \alpha}$ and all $c \in [0, \alpha]$, $\mathbb{P}_\mu \left(\widehat{P}_{n,m} + c < \alpha \right)$ is the probability that $\widehat{P}_{n,m} + c$ is strictly smaller than α under the assumption that $\widehat{P}_{n,m}$ is constructed from m independent copies of a random vector with probability distribution $\widehat{\mu}_n$, where $\widehat{\mu}_n$ is itself constructed from n independent copies

of a random vector with probability distribution μ. For all $p \in [0,1]$, to emphasize the random nature of $\widehat{\mu}_n$, we choose to note the probability that $\widehat{P}_{n,m}$ is smaller than p conditionally on $\widehat{\mu}_n$ as $\mathbb{P}_\mu\left(\widehat{P}_{n,m} \leqslant p \mid \widehat{\mu}_n\right)$. Using the law of total probability, this allows $\mathbb{P}_\mu\left(\widehat{P}_{n,m} + c < p\right)$ to be decomposed into

$$\mathbb{P}_\mu\left(\widehat{P}_{n,m} + c < p\right) = \mathbb{E}_\mu\left[\mathbb{P}_\mu\left(\widehat{P}_{n,m} + c < p \mid \widehat{\mu}_n\right)\right]. \tag{44}$$

Remark 4.2.1. *In the context of this work, we would like to re-emphasize the important difference between n and m: while it is a priori not possible to modify the value of n, it is possible to adapt the value of m with respect to the targeted safety objective. The only potential constraint on m can be a maximum allowable computation budget. In particular, we assume that m can be chosen much larger than n and than $1/\alpha$.*

Remark 4.2.2. *Due to their common dependence on $\widehat{\mu}_n$ (and therefore on $\mathcal{D}_n$), the random vectors $\boldsymbol{X}_1^{(n)}, \ldots, \boldsymbol{X}_m^{(n)}$ are identically distributed but no longer statistically independent, which makes the analysis of the statistical properties of $\widehat{P}_{n,m}$ more complicated than in the crude Monte Carlo case.*

4.2.2. Uniform consistency of the proposed estimator

By construction, the convergence of $\widehat{P}_{n,m}$ to p_f relies heavily on the convergence of $\widehat{\mu}_n$ to $\mu_{\boldsymbol{X}}$. In this section, we place ourselves in the case where the process of constructing $\widehat{\mu}_n$ satisfies the following hypothesis, which will need to be verified on a case-by-case basis once the search set $\mathcal{P}$ has been further specified in numerical applications.

Hypothesis 1. *We assume that the process of constructing $\widehat{\mu}_n$ from n i.i.d. realizations of a vector $\boldsymbol{X}$, whose probability distribution is in $\mathcal{P}$, is uniformly consistent for the estimation of the failure probability, in the sense that*

$$\lim_{n \to +\infty} \sup_{\mu \in \mathcal{P}} \mathbb{E}_\mu\left[\left(\mathbb{P}_\mu(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n) - \mathbb{P}_\mu(\boldsymbol{X} \in \mathcal{F})\right)^2\right] = 0, \tag{45}$$

where extending the notation introduced in Eq. (44),

$$\mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right) := \int_{\mathcal{F}} \mathrm{d}\widehat{\mu}_n(\boldsymbol{x}) = \mathbb{E}_\mu\left[\widehat{P}_{n,m} \mid \widehat{\mu}_n\right] \tag{46}$$

is the (random) probability of failure associated with $\widehat{\mu}_n$.

Proposition 4 (see Appendix A.3 for the proof). *Under Hypothesis 1, $\widehat{P}_{n,m}$ is a uniformly consistent estimator of p_{f}, in the sense that for each $\varepsilon > 0$,*

$$\lim_{n\to+\infty} \lim_{m\to+\infty} \sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu \left(|\widehat{P}_{n,m} - \mathbb{P}_\mu(\boldsymbol{X} \in \mathcal{F})| \geqslant \varepsilon \right) = 0. \tag{47}$$

Remark 4.2.3. *From the proof of Proposition 4, we note that the limits in m and n could have been inverted to define the uniform consistency of $\widehat{P}_{n,m}$.*

4.2.3. Application to the case where the probability distribution admits a PDF
We focus in the rest of this section on the case where the probability distribution of $\boldsymbol{X}$ admits a continuous PDF $f_{\boldsymbol{X}}$ such that

$$\mathrm{d}\mu_{\boldsymbol{X}}(\boldsymbol{x}) = f_{\boldsymbol{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \boldsymbol{x} \in \mathbb{X}, \tag{48}$$

and we focus on the case where $\mu_{\boldsymbol{X}}$ belongs to a known parametric family $\mathcal{P}$ such that:

$$\mathcal{P} := \left\{ \begin{array}{c} \mu \in \mathcal{M}(\mathbb{X}) : \exists\ \boldsymbol{\theta} \in \mathbb{T} \subset \mathbb{R}^q,\ f(\cdot;\boldsymbol{\theta}) \in \mathbb{F}_1(\mathbb{X}, \mathbb{R}^+) \\ / \ \forall\ \boldsymbol{x} \in \mathbb{X},\ \mathrm{d}\mu(\boldsymbol{x}) = f(\boldsymbol{x};\boldsymbol{\theta})\mathrm{d}\boldsymbol{x} \end{array} \right\}, \tag{49}$$

where q is a positive integer characterizing the number of parameters that must be specified to define an element of $\mathcal{P}$, and $\mathbb{F}_1(\mathbb{X}, \mathbb{R}^+)$ is the family of PDFs defined on $\mathbb{X}$, that is to say the family of positive functions defined on $\mathbb{X}$, and whose integral on $\mathbb{X}$ is 1. For instance, we can think of the set of d-dimensional Gaussian PDFs, and in this case, $\boldsymbol{\theta}$ groups together all the parameters characterizing the mean vector and the covariance matrix. The set $\mathbb{T}$ in which the parameters $\boldsymbol{\theta}$ evolve characterizes the a priori information we may have about $\boldsymbol{\theta}$, positivity constraints for instance. As we shall see in the Application Section, constraints on $\mathbb{T}$ are often necessary to ensure that $\widehat{P}_{n,m}$ is uniformly consistent.

To simplify reading, for all $\boldsymbol{\theta} \in \mathbb{T}$, we note in the following $\mu_{\boldsymbol{\theta}}$ the probability distribution such that for each $\boldsymbol{x} \in \mathbb{X}$, $\mathrm{d}\mu_{\boldsymbol{\theta}}(\boldsymbol{x}) = f(\boldsymbol{x};\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}$. We next define, for any $\boldsymbol{\theta} \in \mathbb{T}$,

$$p_{\mathrm{f}}(\boldsymbol{\theta}) := \mathbb{P}_{\mu_{\boldsymbol{\theta}}}(\boldsymbol{X} \in \mathcal{F}) = \int_{\mathcal{F}} f(\boldsymbol{x};\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}. \tag{50}$$

Then, assuming that $\mu_{\boldsymbol{X}}$ is in $\mathcal{P}$ amounts to saying that there exists $\boldsymbol{\theta}^\star \in \mathbb{T}$ such that $\mu_{\boldsymbol{X}} = \mu_{\boldsymbol{\theta}^\star}$, and thus

$$p_{\mathrm{f}} = p_{\mathrm{f}}(\boldsymbol{\theta}^\star). \tag{51}$$

This value $\boldsymbol{\theta}^\star$ is however unknown, and must be estimated from the available information on $\boldsymbol{X}$, i.e., the n realizations of $\boldsymbol{X}$ gathered in $\mathcal{D}_n$. Let $\widehat{\boldsymbol{\theta}}_n$ be an estimator of $\boldsymbol{\theta}^\star$ based on $\mathcal{D}_n$. With the notation introduced above, Hypothesis 1 rewrites

$$\lim_{n\to+\infty} \sup_{\boldsymbol{\theta}\in\mathbb{T}} \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left(p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right) - p_{\mathrm{f}}(\boldsymbol{\theta})\right)^2\right] = 0. \tag{52}$$

As explained above, the estimator we are considering to measure to what extent p_{f} can be considered smaller than α is now

$$\widehat{P}_{n,m} := \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}_{\left\{y\left(\boldsymbol{X}_j^{\widehat{\boldsymbol{\theta}}_n}\right)\leqslant s^\star\right\}}, \tag{53}$$

where $\boldsymbol{X}_1^{\widehat{\boldsymbol{\theta}}_n}, \ldots, \boldsymbol{X}_m^{\widehat{\boldsymbol{\theta}}_n}$ correspond to m conditionally independent realizations of $\boldsymbol{X}^{\widehat{\boldsymbol{\theta}}_n}$, which is a random vector with PDF $f(\cdot;\widehat{\boldsymbol{\theta}}_n)$. As for Decision criterion 2, it finally amounts to considering the system as sufficiently safe if $\widehat{P}_{n,m} + c^*(n,m,\alpha,\beta) < \alpha$, where the constant $c^*(n,m,\alpha,\beta)$ is the minimal positive constant c such that

$$\sup_{\boldsymbol{\theta}\in\mathcal{S}_{\mathrm{INP}}(\alpha)} \mathbb{P}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m} + c < \alpha\right) \tag{54}$$

is lower than β, with $\mathcal{S}_{\mathrm{INP}}(\alpha) := \left\{\boldsymbol{\theta}\in\mathbb{T} \;\middle|\; p_{\mathrm{f}}(\boldsymbol{\theta}) \geqslant \alpha\right\}$ (INP stands for "input").

4.2.4. Practical solving of the optimization problem

In the same manner as the GPD case presented in Section 4.1.4, the method we propose for the optimization problem defined by Eq. (54) relies on the result of Remark 2.0.2. Hence, for each $\boldsymbol{\theta}\in\mathcal{S}_{\mathrm{INP}}(\alpha)$, we denote by $q_{n,m;\boldsymbol{\theta}}(\beta)$ the β-quantile of $\widehat{P}_{n,m}$ when assuming that $\boldsymbol{X}\sim f(\cdot;\boldsymbol{\theta})$, and by $q^*_{n,m}(\beta)$ the solution of

$$q^*_{n,m}(\beta) := \min_{\boldsymbol{\theta}\in\mathcal{S}_{\mathrm{INP}}(\alpha)} q_{n,m;\boldsymbol{\theta}}(\beta), \quad \text{with} \quad \mathbb{P}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m} \leqslant q_{n,m;\boldsymbol{\theta}}(\beta)\right) = \beta, \tag{55}$$

from which, using Eq. (12), we deduce

$$c^*(n,m,\alpha,\beta) = \alpha - q^*_{n,m}(\beta). \tag{56}$$

Using $q^*_{n,m}(\beta)$ to estimate $c^*(n,m,\alpha,\beta)$ is explained by the fact that once a value of $\boldsymbol{\theta}$ has been chosen, it is relatively simple to generate r independent

realizations of $\widehat{P}_{n,m}$, and then deduce from these values an estimation of $q_{n,m;\boldsymbol{\theta}}(\beta)$ (see Appendix B.2 for more details). However, the main difference with the GPD case is that each realization of $\widehat{P}_{n,m}$ is, this time, based on the evaluation of the underlying code y in m different points. Consequently, $r \times m$ evaluations of the code would be necessary for an estimate of $q_{n,m;\boldsymbol{\theta}}(\beta)$, which can quickly become prohibitive in terms of computational cost. The second major additional obstacle lies in the difficulty of selecting elements from $\mathcal{S}_{\mathrm{INP}}(\alpha)$. Indeed, as the failure domain $\mathcal{F}$ is a priori unknown, it is hard to know whether the constraint $p_{\mathrm{f}}(\boldsymbol{\theta}) \geqslant \alpha$ will be verified, or not, without a large number of supplementary calls to the code. To circumvent these issues and restrict the maximum number of code evaluations to be smaller than m, the strategy we propose is based on the coupling of an *augmented Lagrangian* (AL) algorithm with an *importance sampling* (IS) procedure.

The idea of this algorithm is to start by choosing an auxiliary PDF, called f_{IS}, such that the support of f_{IS} contains the support of all the PDFs associated with the elements of $\mathcal{P}$. Then, one needs to generate m i.i.d. realizations of a random vector $\boldsymbol{X}^{\mathrm{IS}}$ according to this PDF, denoted by $\boldsymbol{X}_1^{\mathrm{IS}}, \ldots, \boldsymbol{X}_m^{\mathrm{IS}}$, and then evaluate the numerical code at these m points. For a given $\boldsymbol{\theta} \in \mathbb{T}$, let $\widehat{\boldsymbol{\theta}}_n^{(1)}, \ldots, \widehat{\boldsymbol{\theta}}_n^{(r)}$ be r i.i.d. realizations of the estimator $\widehat{\boldsymbol{\theta}}_n$ under $\mathbb{P}_{\mu_{\boldsymbol{\theta}}}$. These values can be obtained by post-processing in a similar way r times a set of n realizations of a random vector with PDF $f(\cdot;\boldsymbol{\theta})$, as explained in Appendix B.2. For all $1 \leqslant \ell \leqslant r$, let $\widehat{p}_{\mathrm{IS}}\left(\widehat{\boldsymbol{\theta}}_n^{(\ell)}\right)$ and $\widehat{p}_{\mathrm{IS}}$ be the IS estimators of $p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n^{(\ell)}\right)$ and $p_{\mathrm{f}}(\boldsymbol{\theta})$ respectively,

$$\widehat{p}_{\mathrm{IS}}\left(\widehat{\boldsymbol{\theta}}_n^{(\ell)}\right) := \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}_{\{y(\boldsymbol{X}_j^{\mathrm{IS}}) \leqslant s^\star\}} \frac{f\left(\boldsymbol{X}_j^{\mathrm{IS}}; \widehat{\boldsymbol{\theta}}_n^{(\ell)}\right)}{f_{\mathrm{IS}}(\boldsymbol{X}_j^{\mathrm{IS}})}, \quad \widehat{p}_{\mathrm{IS}}(\boldsymbol{\theta}) := \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}_{\{y(\boldsymbol{X}_j^{\mathrm{IS}}) \leqslant s^\star\}} \frac{f(\boldsymbol{X}_j^{\mathrm{IS}}; \boldsymbol{\theta})}{f_{\mathrm{IS}}(\boldsymbol{X}_j^{\mathrm{IS}})}. \tag{57}$$

The AL function [25] can then be written as

$$L_A(\boldsymbol{\theta}; \rho, \lambda) = \widehat{q}^{(r)}_{\mathrm{IS};\boldsymbol{\theta}}(\beta) + \lambda\kappa(\boldsymbol{\theta}) + \frac{1}{2\rho}\max(0, \kappa(\boldsymbol{\theta})), \quad \kappa(\boldsymbol{\theta}) := \alpha - \widehat{p}_{\mathrm{IS}}(\boldsymbol{\theta}), \tag{58}$$

where $\rho > 0$ is a penalty parameter, $\lambda \geqslant 0$ serves as a Lagrange multiplier, and $\widehat{q}^{(r)}_{\mathrm{IS};\boldsymbol{\theta}}(\beta)$ is the empirical estimator of the β-quantile of $\widehat{P}_{n,m}$ using $\left\{\widehat{p}_{\mathrm{IS}}\left(\widehat{\boldsymbol{\theta}}_n^{(\ell)}\right)\right\}_{\ell=1}^{r}$. The AL method then transforms the constrained problem

into a sequence of unconstrained ones with automatic updates of the parameters (λ, ρ), as described in Algorithm 13.

Choose $\rho > 0, \lambda \geqslant 0, k_{\max} > 0, \mathrm{tol}, \boldsymbol{\theta}_{\min} \in \mathbb{T}$
for k *in* $1, \ldots, k_{\max}$ **do**
  Approximately solve the subproblem

$$\boldsymbol{\theta}_k \in \arg\min_{\boldsymbol{\theta} \in \mathbb{T}} L_A(\boldsymbol{\theta}; \rho, \lambda),$$

  Update $\lambda \leftarrow \max(0, \lambda + \kappa(\boldsymbol{\theta})/\rho)$. **if** $\kappa(\boldsymbol{\theta}) < 0$ **then**
    Set $\rho \leftarrow \rho/2$.
  end
  if $\kappa(\boldsymbol{\theta}) \geqslant 0$ *and* $\widehat{q}^{(r)}_{\mathrm{IS};\boldsymbol{\theta}_k}(\beta) < \widehat{q}^{(r)}_{\mathrm{IS};\boldsymbol{\theta}_{\min}}(\beta)$ **then**
    Set $\boldsymbol{\theta}_{\min} \leftarrow \boldsymbol{\theta}_k$.
  end
  if $\kappa(\boldsymbol{\theta}) \geqslant 0$ *and* $\|\boldsymbol{\theta}_k - \boldsymbol{\theta}_{\min}\| < \mathrm{tol}$ **then**
    Break the "for loop".
  end
end
Return $\widehat{q}^*_{n,m}(\beta) = \widehat{q}^{(r)}_{\mathrm{IS};\boldsymbol{\theta}_{\min}}(\beta)$ (as an approximation of $q_{n,m;\boldsymbol{\theta}_{\min}}(\beta)$).

Algorithm 1: Proposed algorithm for the approximation of $q^*_{n,m;\beta}$ using an augmented Lagrangian algorithm and an importance sampling strategy.

As with any IS method, the effectiveness of Algorithm 13 depends on a judicious choice of the auxiliary PDF f_{IS} [27], which in particular should allow a good exploration of the failure domain (i.e., leads to a sufficient number of failure points). There are many methods for optimizing the definition of such a PDF in a more or less sequential manner, which we will not discuss here for the sake of brevity. The interested reader can nevertheless turn to [4] and the associated references for a review of several adaptive IS algorithms. In our situation, as the aim is to find values of $\boldsymbol{\theta}$ such that $p_{\mathrm{f}}(\boldsymbol{\theta}) \geqslant \alpha$ but also such that $\widehat{P}_{n,m}$ is likely to be smaller than α, it seems a good idea to focus on PDFs f_{IS} such that

$$\int_{\mathbb{X}} \mathbb{1}_{\{y(\boldsymbol{x}) \leqslant s^\star)\}} f_{\mathrm{IS}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{59}$$

is close to α. Moreover, if $f(\cdot; \boldsymbol{\theta})$ deviates too far from f_{IS}, we can expect

to find potentially pathological values of the $f(\cdot;\boldsymbol{\theta})/f_{\mathrm{IS}}$ ratio both for the calculation of $p_{\mathrm{f}}(\boldsymbol{\theta})$ and for $\widehat{P}_{n,m}$. In this case, we can expect this value of $\boldsymbol{\theta}$ to be of little relevance for the minimization problem because it would unlikely lead to $p_{\mathrm{f}}(\boldsymbol{\theta}) \geqslant \alpha$ and $\widehat{P}_{n,m} < \alpha$ at the same time. Finally, we note that it is possible (as we did for the Monte Carlo and rare event cases) to add to Algorithm 13 the a priori information that the optimal value of $\boldsymbol{\theta}$ should be sought such that $p_{\mathrm{f}}(\boldsymbol{\theta}) = \alpha$ by adding the second inequality constraint $p_{\mathrm{f}}(\boldsymbol{\theta}) \leqslant \alpha$ in the AL function.

4.2.5. Regular case

In order to study in more detail the sensitivity of the $\widehat{P}_{n,m}$ estimator to the choices of n and m, we now focus on the case where the parametric class from which the PDF of $\boldsymbol{X}$ is chosen is regular, in the sense that for each $\mu \in \mathcal{P}$ so that $\mathrm{d}\mu(\boldsymbol{x}) = f(\boldsymbol{x};\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}$, $\partial f/\partial\boldsymbol{\theta}$ and $\partial^2 f/(\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top)$ exist and are integrable over the failure domain. For the sake of simplicity, we also assume that all the elements of this parametric class have the same support (i.e., the support of $f(\cdot;\boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta}$), that $\widehat{\boldsymbol{\theta}}_n$ is an unbiased estimator of $\boldsymbol{\theta}$ based on n realizations of $\boldsymbol{X}$ with distribution $\mu_{\boldsymbol{\theta}}$, and we denote by $\boldsymbol{C}_{\boldsymbol{\theta}}$ the asymptotic covariance matrix of $\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)$. Taking the maximum likelihood estimator for $\widehat{\boldsymbol{\theta}}_n$ is a way of placing oneself under this last hypothesis, at least asymptotically in n. Indeed, we can show [30] that in this case,

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{\theta}}), \tag{60}$$

where $\boldsymbol{C}_{\boldsymbol{\theta}}$ is a matrix that only depends on $\boldsymbol{\theta}$.

Under these assumptions, and taking n sufficiently large, we can then approximate $f(\boldsymbol{x};\widehat{\boldsymbol{\theta}}_n)$ by its second-order Taylor expansion:

$$\begin{aligned} f\left(\boldsymbol{x};\widehat{\boldsymbol{\theta}}_n\right) \approx f(\boldsymbol{x};\boldsymbol{\theta}) &+ \left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)^\top \frac{\partial f}{\partial\boldsymbol{\theta}}(\boldsymbol{x};\boldsymbol{\theta}) \\ &+ \frac{1}{2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)^\top \frac{\partial^2 f}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}(\boldsymbol{x};\boldsymbol{\theta})\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right). \end{aligned} \tag{61}$$

Integrating Eq. (61) over the failure domain $\mathcal{F}$, we obtain

$$p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right) \approx p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right) := p_{\mathrm{f}}(\boldsymbol{\theta}) \left(\begin{array}{l} 1 + \left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)^\top \partial p_{\mathrm{f}}(\boldsymbol{\theta}) \\ \quad + \dfrac{1}{2}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)^\top \partial^2 p_{\mathrm{f}}(\boldsymbol{\theta}) \left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \end{array} \right), \tag{62}$$

where for each $\boldsymbol{\theta} \in \mathbb{T}$,

$$\partial p_{\mathrm{f}}(\boldsymbol{\theta}) := \frac{1}{p_{\mathrm{f}}(\boldsymbol{\theta})} \int_{\mathcal{F}} \frac{\partial f}{\partial \boldsymbol{\theta}}(\boldsymbol{x}; \boldsymbol{\theta}) \mathrm{d}\boldsymbol{x} = \frac{1}{p_{\mathrm{f}}(\boldsymbol{\theta})} \frac{\partial p_{\mathrm{f}}}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}), \tag{63}$$

$$\partial^2 p_{\mathrm{f}}(\boldsymbol{\theta}) := \frac{1}{p_{\mathrm{f}}(\boldsymbol{\theta})} \int_{\mathcal{F}} \frac{\partial^2 f}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}(\boldsymbol{x}; \boldsymbol{\theta}) \mathrm{d}\boldsymbol{x} = \frac{1}{p_{\mathrm{f}}(\boldsymbol{\theta})} \frac{\partial^2 p_{\mathrm{f}}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}(\boldsymbol{\theta}). \tag{64}$$

This allows us to approximate the mean and variance of $\widehat{P}_{n,m}$ in explicit forms, as it is done in Proposition 1.

Proposition 1. *For n sufficiently large for Approximation (62) to be valid and any $\boldsymbol{\theta}$ in $\mathbb{T}$,*

$$\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\widehat{P}_{n,m}\right] \approx p_{\mathrm{f}}(\boldsymbol{\theta}) \left(1 + \frac{s_{\boldsymbol{\theta}}^2}{2n}\right), \tag{65}$$

$$\mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m}\right) \approx \frac{p_{\mathrm{f}}(\boldsymbol{\theta})}{m} \left(1 - p_{\mathrm{f}}(\boldsymbol{\theta}) + \frac{s_{\boldsymbol{\theta}}^2(1/2 - p_{\mathrm{f}}(\boldsymbol{\theta})) - p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2}{n}\right) + \frac{p_{\mathrm{f}}(\boldsymbol{\theta})^2 \sigma_{\boldsymbol{\theta}}^2}{n}, \tag{66}$$

where $s_{\boldsymbol{\theta}}^2$ and $\sigma_{\boldsymbol{\theta}}^2$ are two positive constants that depend on $\boldsymbol{C}_{\boldsymbol{\theta}}$ but not on n so that

$$\sigma_{\boldsymbol{\theta}}^2 := n \times \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left(\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)^\top \partial p_{\mathrm{f}}(\boldsymbol{\theta})\right)^2\right] = \partial p_{\mathrm{f}}(\boldsymbol{\theta})^\top \boldsymbol{C}_{\boldsymbol{\theta}}\, \partial p_{\mathrm{f}}(\boldsymbol{\theta}), \tag{67}$$

$$s_{\boldsymbol{\theta}}^2 := n \times \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)^\top \partial^2 p_{\mathrm{f}}(\boldsymbol{\theta}) \left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)\right] = \boldsymbol{C}_{\boldsymbol{\theta}} : \partial^2 p_{\mathrm{f}}(\boldsymbol{\theta}), \tag{68}$$

with ":" denotes the double contracted tensor product.

▫ **Proof:** Using Eq. (62), and remembering that $\widehat{\boldsymbol{\theta}}_n$ is an unbiased estimator of $\boldsymbol{\theta}$ with covariance $\boldsymbol{C}_{\boldsymbol{\theta}}/n$, we can calculate:

$$\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\widehat{P}_{n,m}\right] = \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right)\right] \approx p_{\mathrm{f}}(\boldsymbol{\theta}) \left(1 + \frac{s_{\boldsymbol{\theta}}^2}{2n}\right). \tag{69}$$

In the same manner, using the notations formerly introduced,

$$\begin{aligned}\mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m}\right) &= \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m} \mid \widehat{\boldsymbol{\theta}}_n\right)\right] + \mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\widehat{P}_{n,m} \mid \widehat{\boldsymbol{\theta}}_n\right]\right) \\ &= \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right)\left(1 - p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right)\right)/m\right] + \mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right)\right) \\ &\approx \frac{p_{\mathrm{f}}(\boldsymbol{\theta})}{m}\left(1 - p_{\mathrm{f}}(\boldsymbol{\theta}) + \frac{s_{\boldsymbol{\theta}}^2}{n}\left(\frac{1}{2} - p_{\mathrm{f}}(\boldsymbol{\theta})\right) - \frac{p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2}{n}\right) + \frac{p_{\mathrm{f}}(\boldsymbol{\theta})^2\sigma_{\boldsymbol{\theta}}^2}{n}.\end{aligned} \tag{70}$$

□

Looking at Proposition 1, we first note that even if $\widehat{\boldsymbol{\theta}}_n$ is an unbiased estimator of $\boldsymbol{\theta}$, the estimator $\widehat{P}_{n,m}$ is biased, and that its bias tends to 0 in $s_{\boldsymbol{\theta}}^2/n$. We then notice that the variance of $\widehat{P}_{n,m}$ does not converge to 0 when m tends to infinity, but remains greater than $p_{\mathrm{f}}(\boldsymbol{\theta})^2\sigma_{\boldsymbol{\theta}}^2/n$. Conversely, if n tends to infinity, that is to say if the PDF of $\boldsymbol{X}$ is perfectly known, then $\mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m}\right)$ tends to $p_{\mathrm{f}}(\boldsymbol{\theta})(1 - p_{\mathrm{f}}(\boldsymbol{\theta}))/m$, which corresponds to the usual Monte Carlo case. If we wanted to balance the contributions of m and n on the variance, we could choose $m = m^{\mathrm{bal}}(n, p_{\mathrm{f}}(\boldsymbol{\theta}))$, with

$$\begin{aligned}&\frac{1}{m^{\mathrm{bal}}(n, p_{\mathrm{f}}(\boldsymbol{\theta}))}\left(1 - p_{\mathrm{f}}(\boldsymbol{\theta}) + \frac{s_{\boldsymbol{\theta}}^2(1/2 - p_{\mathrm{f}}(\boldsymbol{\theta})) - p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2}{n}\right) = \frac{p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2}{n} \\ &\quad\Leftrightarrow\quad m^{\mathrm{bal}}(n, p_{\mathrm{f}}(\boldsymbol{\theta})) = n\frac{1 - p_{\mathrm{f}}(\boldsymbol{\theta})}{p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2} + \frac{s_{\boldsymbol{\theta}}^2(1/2 - p_{\mathrm{f}}(\boldsymbol{\theta}))}{p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2} - 1 \approx \frac{n + s_{\boldsymbol{\theta}}^2/2}{p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2}.\end{aligned} \tag{71}$$

Because of the multiplicative factor $1/(p_{\mathrm{f}}(\boldsymbol{\theta})\sigma_{\boldsymbol{\theta}}^2)$, we see that this required value of m can be much larger than n when the probability of failure is much lower than the first-order sensitivity of this probability of failure to the uncertainties about $\boldsymbol{\theta}$, i.e., when p_{f} is much smaller than $\sigma_{\boldsymbol{\theta}}^2$. In that balanced case, provided that $\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\widehat{P}_{n,m}\right]^2 > 0$, it is possible to approximate the square of the coefficient of variation of $\widehat{P}_{n,m}$, noted δ^2, under the form

$$\delta^2 := \frac{\mathrm{Var}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m}\right)}{\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\widehat{P}_{n,m}\right]^2} \approx \frac{2p_{\mathrm{f}}(\boldsymbol{\theta})^2\sigma_{\boldsymbol{\theta}}^2/n}{p_{\mathrm{f}}(\boldsymbol{\theta})^2(1 + s_{\boldsymbol{\theta}}^2/n)^2} = \frac{2\sigma_{\boldsymbol{\theta}}^2}{n\left(1 + s_{\boldsymbol{\theta}}^2/n\right)^2}. \tag{72}$$

The precision of the estimator is thus mainly controlled by $\sigma_{\boldsymbol{\theta}}^2/n$. Using the approximation $(1+s_{\boldsymbol{\theta}}^2/n)^2 \approx 1+2s_{\boldsymbol{\theta}}^2/n$, a link can eventually be made between

a targeted precision ε^2 and the minimum value n_ε of n for such precision in the form of

$$n \geqslant n_\varepsilon := \frac{2\sigma_{\boldsymbol{\theta}}^2}{\varepsilon^2} - 2s_{\boldsymbol{\theta}}^2. \tag{73}$$

The development presented in Eq. (61) can also be used to simplify the calculation of $c^*(n, m, \alpha, \beta)$ defined at the end of Section 4.2.3. Indeed, for each $\boldsymbol{\theta} \in \mathbb{T}$, if $\widehat{P}_{n,m}$ is an estimator of $p_{\mathrm{f}}(\boldsymbol{\theta})$ relying on n i.i.d. realizations of a random vector with PDF $f(\cdot; \boldsymbol{\theta})$, and if we approximate $p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right) = \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\widehat{P}_{n,m} \mid \widehat{\boldsymbol{\theta}}_n\right]$ by its Taylor expansion at $\boldsymbol{\theta}$, which was noted $p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)$ in Eq. (62), we can use the CLT to approximate the probability distribution of

$$\frac{\sqrt{m}\left(\widehat{P}_{n,m} - p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\right)}{\sqrt{p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\left(1 - p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\right)}} \;\Bigg|\; \widehat{\boldsymbol{\theta}}_n \tag{74}$$

by a standard Gaussian distribution, and therefore introduce the following approximation of $\mathbb{P}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m} + c < \alpha\right)$:

$$\begin{aligned}
\mathbb{P}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m} + c < \alpha\right) &= \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\mathbb{P}_{\mu_{\boldsymbol{\theta}}}\left(\widehat{P}_{n,m} < \alpha - c \mid \widehat{\boldsymbol{\theta}}_n\right)\right] \\
&\approx \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\Phi\left(\frac{\sqrt{m}\left(\alpha - c - p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\right)}{\sqrt{p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\left(1 - p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\right)}}\right)\right],
\end{aligned} \tag{75}$$

where it is recalled that Φ is the CDF of a standard Gaussian random variable. In that case, the safety margin $c^*(n, m, \alpha, \beta)$ can be searched as the minimal positive constant c such that

$$\sup_{\boldsymbol{\theta} \in \mathcal{S}_{\mathrm{INP}}(\alpha)} \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\Phi\left(\frac{\sqrt{m}\left(\alpha - c - p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\right)}{\sqrt{p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\left(1 - p_{\mathrm{f}}^{\mathrm{reg}}\left(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}\right)\right)}}\right)\right] \tag{76}$$

is lower than β. In Eq. (76), for each tested value of $\boldsymbol{\theta}$ in $\mathcal{S}_{\mathrm{INP}}(\alpha)$, rather than having to estimate the value of $p_{\mathrm{f}}(\widehat{\boldsymbol{\theta}}_n)$ for many realizations of $\widehat{\boldsymbol{\theta}}_n$, the

new problem defined by Eq. (76) only relies on the computation of $p_{\mathrm{f}}(\boldsymbol{\theta})$, $p_{\mathrm{f}}(\boldsymbol{\theta}) \times \partial p_{\mathrm{f}}(\boldsymbol{\theta})$ and $p_{\mathrm{f}}(\boldsymbol{\theta}) \times \partial^2 p_{\mathrm{f}}(\boldsymbol{\theta})$. And it is worth noting that these three quantities can again be estimated by $\widehat{p}_{\mathrm{IS}}(\boldsymbol{\theta})$, $\widehat{\kappa}_{\mathrm{IS}}^{(1)}(\boldsymbol{\theta})$ and $\widehat{\kappa}_{\mathrm{IS}}^{(2)}(\boldsymbol{\theta})$ using an IS strategy to limit the number of code evaluations, as it was done in Section 4.2.4,

$$\widehat{p}_{\mathrm{IS}}(\boldsymbol{\theta}) := \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{\left\{y(\boldsymbol{X}_j^{\mathrm{IS}}) \leqslant s^\star\right\}} \frac{f(\boldsymbol{X}_j^{\mathrm{IS}}; \boldsymbol{\theta})}{f_{\mathrm{IS}}(\boldsymbol{X}_j^{\mathrm{IS}})}, \tag{77}$$

$$\widehat{\kappa}_{\mathrm{IS}}^{(1)}(\boldsymbol{\theta}) := \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{\left\{y(\boldsymbol{X}_j^{\mathrm{IS}}) \leqslant s^\star\right\}} \frac{1}{f_{\mathrm{IS}}(\boldsymbol{X}_j^{\mathrm{IS}})} \frac{\partial f}{\partial \boldsymbol{\theta}}(\boldsymbol{X}_j^{\mathrm{IS}}; \boldsymbol{\theta}), \tag{78}$$

$$\widehat{\kappa}_{\mathrm{IS}}^{(2)}(\boldsymbol{\theta}) := \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{\left\{y(\boldsymbol{X}_j^{\mathrm{IS}}) \leqslant s^\star\right\}} \frac{1}{f_{\mathrm{IS}}(\boldsymbol{X}_j^{\mathrm{IS}})} \frac{\partial^2 f}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}(\boldsymbol{X}_j^{\mathrm{IS}}; \boldsymbol{\theta}), \tag{79}$$

where we recall that f_{IS} is the chosen auxiliary PDF, while $\boldsymbol{X}_1^{\mathrm{IS}}, \ldots, \boldsymbol{X}_m^{\mathrm{IS}}$ correspond to m i.i.d. realizations of a random vector with PDF f_{IS}.

Remark 4.2.4. *In the case where the function $\boldsymbol{\theta} \mapsto f(\boldsymbol{x}; \boldsymbol{\theta})$ is regular for each $\boldsymbol{x} \in \mathcal{F}$, the fact that $|p_{\mathrm{f}}(\widehat{\boldsymbol{\theta}}_n) - p_{\mathrm{f}}(\boldsymbol{\theta})| \leqslant 1$ for each $\boldsymbol{\theta} \in \mathbb{T}$ allows us to write the following inequalities*

$$\begin{aligned}
\left(p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right) - p_{\mathrm{f}}(\boldsymbol{\theta})\right)^2 &\leqslant \left|p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right) - p_{\mathrm{f}}(\boldsymbol{\theta})\right| \\
&\leqslant \int_{\mathcal{F}} \left|f\left(\boldsymbol{x}; \widehat{\boldsymbol{\theta}}_n\right) - f(\boldsymbol{x}; \boldsymbol{\theta})\right| \mathrm{d}\boldsymbol{x} \\
&\leqslant \left\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\| \times \int_{\mathcal{F}} \sup_{\boldsymbol{\theta} \in \mathbb{T}} \left\|\frac{\partial f}{\partial \boldsymbol{\theta}}(\boldsymbol{x}; \boldsymbol{\theta})\right\| \mathrm{d}\boldsymbol{x},
\end{aligned}$$

and therefore Eq. (52) *is satisfied as soon as the following two conditions hold:*

$$\int_{\mathcal{F}} \sup_{\boldsymbol{\theta} \in \mathbb{T}} \left\|\frac{\partial f}{\partial \boldsymbol{\theta}}(\boldsymbol{x}; \boldsymbol{\theta})\right\| \mathrm{d}\boldsymbol{x} < +\infty \qquad \textit{and} \qquad \lim_{n \to +\infty} \sup_{\boldsymbol{\theta} \in \mathbb{T}} \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\|\right] = 0. \tag{80}$$

5. Applications

We can list several objectives for this application section. On the one hand, when the uncertainty on the input or output probability distribution is not

Example	Function y	d	$s^\star$	p_{f} (IC 95%)
1	Waarts function	2	0	0.00223 [0.00220; 0.00226]
2	Borehole function	8	-225	0.00141 [0.00139; 0.00144]

Table 1: Characteristics of the two analyzed test functions.

taken into account, we highlight the significant risk of designating as safe a system that is not. By modifying the safety factors to include this new source of uncertainty, on the simulation inputs or outputs, we then show that the risk of designating a faulty system as safe is much better controlled. Although most of the results presented are based on the asymptotic behavior of estimators, we then empirically demonstrate the practical applicability of these results to finite regimes.

5.1. Presentation of the test functions

The developments presented in Sections 3 and 4 are illustrated on two analytical functions, whose characteristics are listed in Table 5.1. The reference values for the failure probability p_{f} are estimated using a crude Monte Carlo approach based on 10^7 function evaluations. To get sound comparisons between the different tests, the results presented in the next sections are averaged over 10^3 repetitions of the whole procedures.

Test function #1: the Waarts function

To begin, we focus on the Waarts function, which is a very common function in reliability analysis and has been studied in numerous papers in the literature, and which is given by the following expression:

$$y^{\mathrm{waarts}} : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ \boldsymbol{x} & \mapsto & \min\left(b_1(\boldsymbol{x}), b_2(\boldsymbol{x})\right) \end{array} \right. , \tag{81}$$

$$b_1(\boldsymbol{x}) := 3 + \frac{(x_1 - x_2)^2}{10} - \frac{\mathrm{sign}(x_1 + x_2)(x_1 + x_2)}{\sqrt{2}}, \;\; b_2(\boldsymbol{x}) := \mathrm{sign}(x_1 - x_2)(x_1 - x_2) + \frac{7}{\sqrt{2}}. \tag{82}$$

This function is in dimension $d = 2$, and its inputs are assumed to be independent and Gaussian, with unknown means μ_1, μ_2, and variances σ_1^2, σ_2^2:

$$\boldsymbol{X} \;\sim\; \mathcal{N}\left(\left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right), \left[\begin{array}{cc} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{array} \right] \right). \tag{83}$$

The number of parameters characterizing the distribution of the input vector $\boldsymbol{X}$ is therefore equal to four, and we can write $\boldsymbol{\theta} = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{T} \subset \mathbb{R} \times (0, +\infty) \times \mathbb{R} \times (0, +\infty)$, so that:

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \frac{1}{2\pi\sqrt{\theta_2\theta_4}} \exp\left(-\frac{1}{2}\left(\frac{(x_1-\theta_1)^2}{\theta_2} + \frac{(x_2-\theta_3)^2}{\theta_4}\right)\right), \quad \boldsymbol{x} \in \mathbb{X}. \tag{84}$$

Given n i.i.d. realizations of $\boldsymbol{X}$, which are noted $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, the vector $\widehat{\boldsymbol{\theta}}_n$ such that for $i \in \{1, 2\}$,

$$\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1} := \frac{1}{n}\sum_{i=1}^{n} (\boldsymbol{X}_n)_i\,, \quad \left(\widehat{\boldsymbol{\theta}}_n\right)_{2i} := \frac{1}{n-1}\sum_{i=1}^{n}\left((\boldsymbol{X}_n)_i - \left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1}\right)^2 \tag{85}$$

is the standard estimator of $\boldsymbol{\theta}$. It is unbiased, and it can be shown that

$$\mathrm{Cov}\left(\sqrt{n} \times \widehat{\boldsymbol{\theta}}_n\right) = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \frac{2n\sigma_1^4}{n-1} & 0 & 0 \\ 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 & \frac{2n\sigma_2^4}{n-1} \end{bmatrix} =: \boldsymbol{C}_{\boldsymbol{\theta}}. \tag{86}$$

With no real limitation on generality, if we limit the possible values for $\boldsymbol{\theta}$ (and therefore for $\widehat{\boldsymbol{\theta}}_n$), and choose for $\mathbb{T}$ the following closed set

$$\mathbb{T} = [-t_1, t_2] \times [t_3, t_4] \times [-t_1, t_2] \times [t_3, t_4], \tag{87}$$

with t_1, t_2, t_4 (resp. t_3) as large (resp. close to 0) as desired, we can verify that

$$\int_{\mathcal{F}} \sup_{\boldsymbol{\theta}\in\mathbb{T}} \left\| \frac{\partial f}{\partial \boldsymbol{\theta}}(\boldsymbol{x}; \boldsymbol{\theta}) \right\| \mathrm{d}\boldsymbol{x} < +\infty. \tag{88}$$

In addition, if χ is a random variable following a chi-squared distribution with $n-1$ degrees of freedom, we can show that

$$\frac{1}{\theta_2}\mathbb{E}\left[\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_1 - \theta_1\right)^2\right] = \frac{1}{\theta_4}\mathbb{E}\left[\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_3 - \theta_3\right)^2\right] = \frac{1}{n},$$

$$\frac{1}{\theta_2^2}\mathbb{E}\left[\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_2 - \theta_2\right)^2\right] = \frac{1}{\theta_4^2}\mathbb{E}\left[\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_3 - \theta_3\right)^2\right] = \frac{2}{n-1}.$$

As $\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\|\right] \leqslant \sqrt{\mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\|^2\right]}$, we then show that

$$\lim_{n\to+\infty} \sup_{\boldsymbol{\theta}\in\mathbb{T}} \mathbb{E}_{\mu_{\boldsymbol{\theta}}}\left[\left\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right\|\right] = 0, \tag{89}$$

which allows us to ensure that Hypothesis 1 is indeed verified by this estimator, using the results of Remark 4.2.4.

For this first test function, PDF $f(\cdot;\boldsymbol{\theta})$ is regular with respect to $\boldsymbol{\theta}$, and the constants $\partial p_{\mathrm{f}}(\boldsymbol{\theta})$ and $\partial^2 p_{\mathrm{f}}(\boldsymbol{\theta})$ can be estimated, using Eqs. (63) and (64), as functions of the statistical moments of the random vector $\boldsymbol{X}^{\mathcal{F}}$, whose PDF is $\boldsymbol{x} \mapsto \frac{1}{p_{\mathrm{f}}(\boldsymbol{\theta})} \mathbb{1}_{\{\boldsymbol{x}\in\mathcal{F}\}} f(\boldsymbol{x};\boldsymbol{\theta})$. In particular, if we choose $s^\star = 0$, $\boldsymbol{\theta}^\star = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) = (0, 1, 0, 1)$, as it will be done for generating the realizations of $\boldsymbol{X}$ in the following, we obtain:

$$s^2_{\boldsymbol{\theta}^\star} \approx 27.5, \quad \sigma^2_{\boldsymbol{\theta}^\star} \approx 23.9. \tag{90}$$

Using the result of Eq. (71), the value of m that balances the two variance terms of the estimator of p_{f}, which was noted $m^{\mathrm{bal}}(n, p_{\mathrm{f}})$, is then close to

$$m^{\mathrm{bal}}(n, p_{\mathrm{f}}) \approx 20 \times n. \tag{91}$$

Test function #2: the borehole function
The second function studied is the borehole function, which models the flow of water through a borehole [19], and which is given by the following expression:

$$y^{\mathrm{borehole}} : \left\{ \begin{array}{ccc} \mathbb{R}^8 & \to & \mathbb{R} \\ \boldsymbol{x} & \mapsto & -\frac{2\pi x_3 (x_4 - x_6)}{\log(x_2/x_1)\left(1 + \frac{2x_7 x_3}{\log(x_2/x_1) x_1^2 x_8} + \frac{x_3}{x_5}\right)} \end{array} \right. . \tag{92}$$

Once again, we have chosen this function for its simplicity and rapid evaluation, which have made it a benchmark function in recent years in computer experiments. This function is in dimension $d = 8$, and the components of its input vector $\boldsymbol{X}$ are assumed to be independent and uniformly distributed on

$$[\theta_1, \theta_2] \times [\theta_3, \theta_4] \times [\theta_5, \theta_6] \times [\theta_7, \theta_8] \times [\theta_9, \theta_{10}] \times [\theta_{11}, \theta_{12}] \times [\theta_{13}, \theta_{14}] \times [\theta_{15}, \theta_{16}]. \tag{93}$$

The number of parameters characterizing the PDF of $\boldsymbol{X}$ is thus equal to 16, and this PDF is this time not regular with respect to $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{16})$:

$$f(\boldsymbol{x};\boldsymbol{\theta}) = \frac{1}{\mathrm{Vol}(\Omega(\boldsymbol{\theta}))} \mathbb{1}_{\{\boldsymbol{x}\in\Omega(\boldsymbol{\theta})\}}, \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{94}$$

$$\Omega(\boldsymbol{\theta}) := \bigtimes_{i=1}^{d} [\theta_{2i}, \theta_{2i-1}], \quad \mathrm{Vol}(\Omega(\boldsymbol{\theta})) := \prod_{i=1}^{d} (\theta_{2i} - \theta_{2i-1}). \tag{95}$$

The uncertainty therefore relates to the support of the input PDF, which is a configuration that is often overlooked in the literature despite its great practical interest. Each component of $\boldsymbol{\theta}$ can a priori take any value in $\mathbb{R}$, provided that θ_{2i-1} is smaller than θ_{2i} for each $1 \leqslant i \leqslant d$. The definition domain for $\boldsymbol{\theta}$ can thus be written as

$$\mathbb{T} := \left\{ \boldsymbol{\theta} \in \mathbb{R}^{2d} : \ \forall \ i \in \{1, \dots, d\}, \ \theta_{2i} > \theta_{2i-1} \right\}. \tag{96}$$

Nevertheless, from the n available realizations of $\boldsymbol{X}$, it is possible to reduce $\mathbb{T}$ by integrating the fact that the lower bound θ_{2i-1} (resp. the upper bound θ_{2i}) of the domain of definition of X_i is necessarily smaller (resp. larger) than the smallest (resp. largest) observed value of X_i. For the numerical results, the true value of this vector, $\boldsymbol{\theta}^\star$, will be chosen equal to

$$(0.05, 0.15, 10^2, 5.10^4, 63070, 115600, 990, 1100, 63.1, 116, 700, 820, 1120, 1680, 9885, 12045),$$

as it was proposed in [32]. To estimate $\boldsymbol{\theta}^\star$, we focus on its maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_n$, so that:

$$\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1} = \min_{1 \leqslant i \leqslant n} \left(\boldsymbol{X}_n\right)_i, \quad \left(\widehat{\boldsymbol{\theta}}_n\right)_{2i} = \max_{1 \leqslant i \leqslant n} \left(\boldsymbol{X}_n\right)_i, \quad 1 \leqslant i \leqslant d. \tag{97}$$

For each $\boldsymbol{\theta} \in \mathbb{T}$, note that this estimator is biased,

$$\mathbb{E}\left[\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1}\right] - (\boldsymbol{\theta})_{2i-1} = (\boldsymbol{\theta})_{2i} - \mathbb{E}\left[\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i}\right] = \frac{(\boldsymbol{\theta})_{2i} - (\boldsymbol{\theta})_{2i-1}}{n+1}, \quad 1 \leqslant i \leqslant d, \tag{98}$$

but that its convergence to $\boldsymbol{\theta}$ is particularly quick as a function of n:

$$\mathrm{Var}\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1}\right) = \mathrm{Var}\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i}\right) = \frac{n}{(n+1)^2(n+2)} \left((\boldsymbol{\theta})_{2i} - (\boldsymbol{\theta})_{2i-1}\right)^2. \tag{99}$$

Using the notations of Eq. (95), we also see that $\Omega(\widehat{\boldsymbol{\theta}}_n) \subset \Omega(\boldsymbol{\theta})$ by construction, so that $\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n)) \leqslant \mathrm{Vol}(\Omega(\boldsymbol{\theta}))$ and $\mathrm{Vol}(\mathcal{F} \cap \Omega(\widehat{\boldsymbol{\theta}}_n)) \leqslant \mathrm{Vol}(\mathcal{F} \cap \Omega(\boldsymbol{\theta}))$.

For all $\boldsymbol{\theta}$ in $\mathbb{T}$, we can then calculate:

$$\begin{aligned}
\left|p_{\mathrm{f}}(\boldsymbol{\theta}) - p_{\mathrm{f}}\left(\widehat{\boldsymbol{\theta}}_n\right)\right| &= \left|\frac{\mathrm{Vol}(\mathcal{F}\cap\Omega(\boldsymbol{\theta}))}{\mathrm{Vol}(\Omega(\boldsymbol{\theta}))} - \frac{\mathrm{Vol}(\mathcal{F}\cap\Omega(\widehat{\boldsymbol{\theta}}_n))}{\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n))}\right| \\
&= \frac{\left|\mathrm{Vol}(\mathcal{F}\cap\Omega(\boldsymbol{\theta}))\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n)) - \mathrm{Vol}(\mathcal{F}\cap\Omega(\widehat{\boldsymbol{\theta}}_n))\mathrm{Vol}(\Omega(\boldsymbol{\theta}))\right|}{\mathrm{Vol}(\Omega(\boldsymbol{\theta}))\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n))} \\
&\leqslant \frac{\mathrm{Vol}(\mathcal{F}\cap\Omega(\boldsymbol{\theta}))\left|\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n)) - \mathrm{Vol}(\Omega(\boldsymbol{\theta}))\right| + \mathrm{Vol}(\Omega(\boldsymbol{\theta}))\left|\mathrm{Vol}(\mathcal{F}\cap\Omega(\boldsymbol{\theta})) - \mathrm{Vol}(\mathcal{F}\cap\Omega(\widehat{\boldsymbol{\theta}}_n))\right|}{\mathrm{Vol}(\Omega(\boldsymbol{\theta}))\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n))} \\
&\leqslant \frac{\left|\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n)) - \mathrm{Vol}(\Omega(\boldsymbol{\theta}))\right| + \left|\mathrm{Vol}(\mathcal{F}\cap\Omega(\boldsymbol{\theta})) - \mathrm{Vol}(\mathcal{F}\cap\Omega(\widehat{\boldsymbol{\theta}}_n))\right|}{\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n))} \\
&\leqslant \frac{2\left|\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n)) - \mathrm{Vol}(\Omega(\boldsymbol{\theta}))\right|}{\mathrm{Vol}(\Omega(\widehat{\boldsymbol{\theta}}_n))}.
\end{aligned}$$

As $\widehat{\boldsymbol{\theta}}_n$ tends to $\boldsymbol{\theta}$ in probability when n tends to infinity, and as $\boldsymbol{\theta} \mapsto \mathrm{Vol}(\Omega(\boldsymbol{\theta}))$ is a continuous function, we can then make Hypothesis 1 true by excluding the possibility that the components of $\boldsymbol{\theta}$ can take too large or too small values and that the upper bounds can be too close to the lower bounds. Without much limitation from an application point of view, we therefore ensure that Hypothesis 1 is verified by choosing

$$\mathbb{T} = \left\{\boldsymbol{\theta} \in [-t,t]^{2d} :\ \forall\ i \in \{1,\ldots,d\},\ \theta_{2i} - \theta_{2i-1} \geqslant \delta > 0\right\}, \tag{100}$$

with t and δ respectively as large and as small as desired.

Remark 5.1.1. *It is possible to consider an unbiased version of this estimator, without this having any real impact on the following.*

5.2. Comparing the performance of estimators

For each test function, we are interested in three values for the number of available realizations of $\boldsymbol{X}$, $n \in \{10^4, 10^3, 500\}$, and we take $m = 10^3/p_{\mathrm{f}}$ (we voluntarily place ourselves in the case $m \gg m^{\mathrm{bal}}(n, p_{\mathrm{f}})$ in order to focus mainly on the effects of a limited number of observations of $\boldsymbol{X}$). Five configurations are then compared:

- Case 1 (MC-m) corresponds to the case where the estimate of p_{f} relies on a MC procedure using a large number m of realizations of $\boldsymbol{X}$. As, in practice, only $n \ll m$ realizations of $\boldsymbol{X}$ are really available, this configuration is used as a reference for comparison.

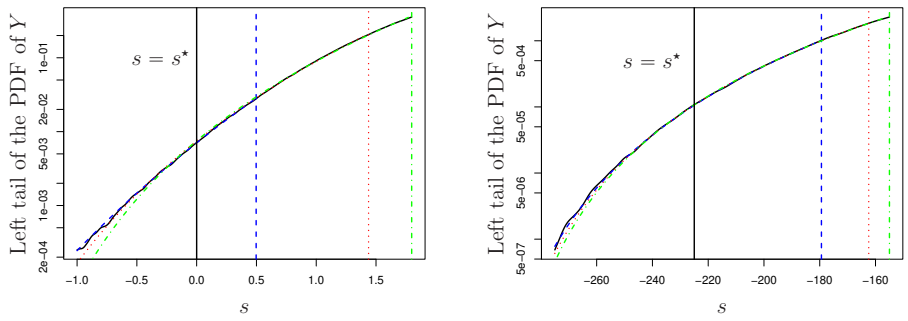
- Case 2 (MC-n) corresponds to the case where the estimate of p_{f} relies on a MC procedure using directly the n available realizations of $\boldsymbol{X}$. For values of n close to or smaller than $1/p_{\mathrm{f}}$, this method is not expected to be very relevant.

- Case 3 (EV-n) corresponds to the output parametric case relying on the extreme value theory presented in Section 4.1. This approach relies on the introduction of a location parameter $\overline{s}$. As explained in Section 4.1.5, this value may be adapted to n using a bootstrap strategy to minimize the variance of the estimator of p_{f}. For the Waarts function (respectively the borehole function), $\overline{s}$ is therefore chosen equal to 0.970 (resp. -179) when $n = 10^4$, to 1.33 (resp. -162) when $n = 10^3$, and to 1.64 (resp. -155) when $n = 500$. The relevance of the GPD to model $s \mapsto \mathbb{P}(Y \leqslant s^\star \mid Y \leqslant \overline{s})$ for the two studied functions and these values of the location parameter is illustrated graphically in Figure 3.

- Case 4 (InP-n) corresponds to the input parametric case presented in Section 4.2, assuming we know the parametric class to which belongs the PDF of $\boldsymbol{X}$. Whether or not we are in a regular case, the construction of the safety margins relies in this case on the introduction of an auxiliary PDF that we noted f_{IS}. In the case of the Waarts function, the Gaussian distribution with parameters $\widehat{\boldsymbol{\theta}}_n$ is taken for f_{IS}, while for the borehole function, a uniform distribution over the slightly increased domain

$$\prod_{i=1}^{d} \left[-\delta_i + \left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1}, \delta_i + \left(\widehat{\boldsymbol{\theta}}_n\right)_{2i} \right], \tag{101}$$

  is considered, with $\delta_i = 2\left(\left(\widehat{\boldsymbol{\theta}}_n\right)_{2i} - \left(\widehat{\boldsymbol{\theta}}_n\right)_{2i-1}\right)/n$ (i.e., an arbitrary shift of around 2 standard deviations on each bound). For these two PDFs, $\widehat{\boldsymbol{\theta}}_n$ refers to the estimator of the true value of the parameters obtained from the observed realizations of $\boldsymbol{X}$ contained in $\mathcal{D}_n$. For the second IS PDF, the reduced increase results from a necessary compromise: the greater the increase, the greater the exploration of the para-

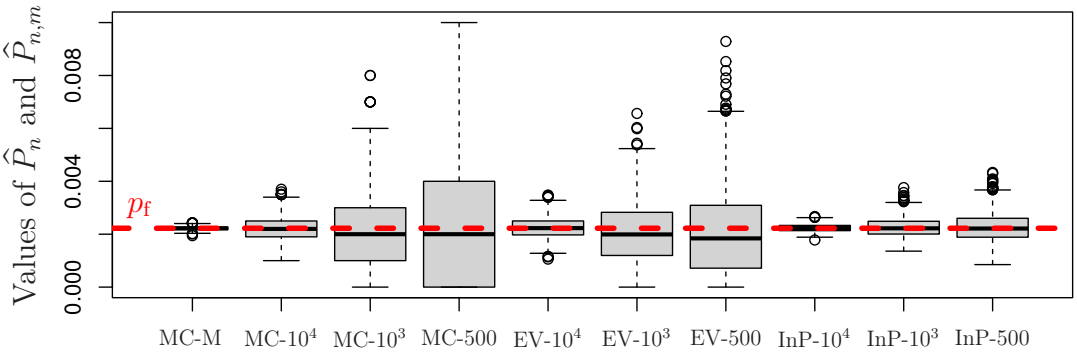


(a) Waarts function (b) Borehole function

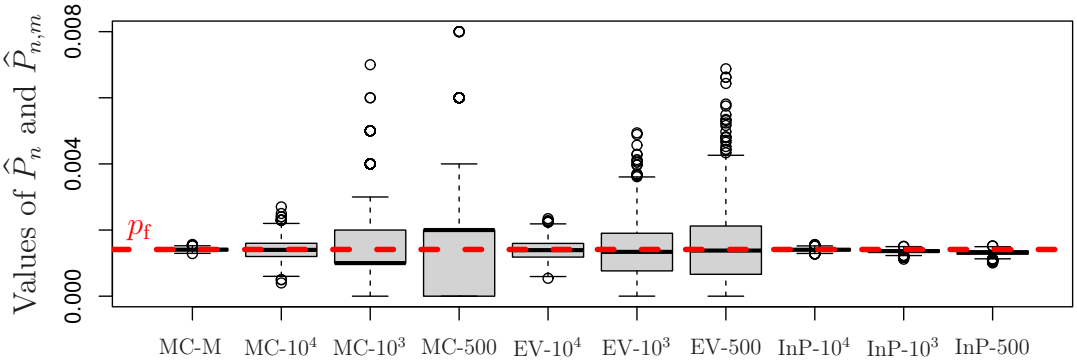
Figure 3: Assessment of the suitability of the GPD distribution for approximating the tail of the distribution of $Y = y(\boldsymbol{X})$. The black solid lines correspond to the empirical estimation of the PDF of Y using 10^7 i.i.d. realizations of $\boldsymbol{X}$. The blue dashed lines, the red dotted lines and the green dotted-dashed lines correspond to the approximation of this PDF using Hypothesis 4.1 when taking $\overline{s}$ equal to 0.970, 1.33 and 1.64 respectively for the Waarts case, and equal to -179, -162 and -155 for the borehole function. In each case, the vertical lines indicate the value of $s^\star$ and the three considered values of $\overline{s}$.

metric set $\mathcal{P}$, as the choice for f_{IS} necessarily implies new constraints for the search domain $\mathbb{T}$ in the optimization problems in Algorithm 13. But for a number of evaluations of y fixed at m, the greater the domain, the smaller the number of failure points in the smaller domains, and therefore the greater the risk of increasing the variance of the estimators based on this IS strategy. As it will be explained in conclusion, the introduction of more sophisticated and potentially more effective IS procedures is left as a working perspective.

For the values of n and m considered, Figure 4 then compares, in boxplot form, the dispersion of the values of the proposed estimators of p_{f}. For the two functions studied, we first observe, not surprisingly, that the greater the value of n, the more precise the estimate of p_{f}. Furthermore, we observe that switching from a crude Monte Carlo approach to another one based on the extreme value theory leads to a slight reduction in variance (reduction by a factor of 1.5 to 2), while knowing the parametric class in which the PDF of $\boldsymbol{X}$ evolves leads to an even stronger reduction in variance (reduction by a factor of around 15 for the Waarts function to more than 100 for the borehole

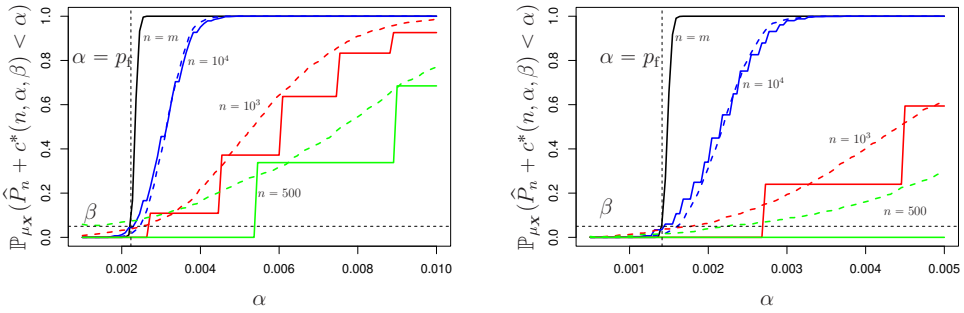


(a) Waarts function



(b) Borehole function

Figure 4: Comparison of the dispersions of the different proposed estimators of p_{f}. In each case, the boxplots group together 10^3 values of each estimator calculated from 10^3 independent drawns of $\mathcal{D}_n$, whose elements correspond to n i.i.d. realizations of $\boldsymbol{X} \sim f(\cdot;\boldsymbol{\theta}^\star)$.



(a) Waarts function (b) Borehole function

Figure 5: Comparison of the evolution of $\alpha \mapsto \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right)$ in the crude MC case (in solid lines) and in the extreme value case (in dashed lines). In both figures, the black lines correspond to the ideal case where $m \gg 1/p_{\mathrm{f}}$ i.i.d. realizations of $\boldsymbol{X}$ are available, when the curves in blue are associated to $n = 10^4$, the curves in red are associated to $n = 10^3$, and the curves in green are associated to $n = 500$.

function).

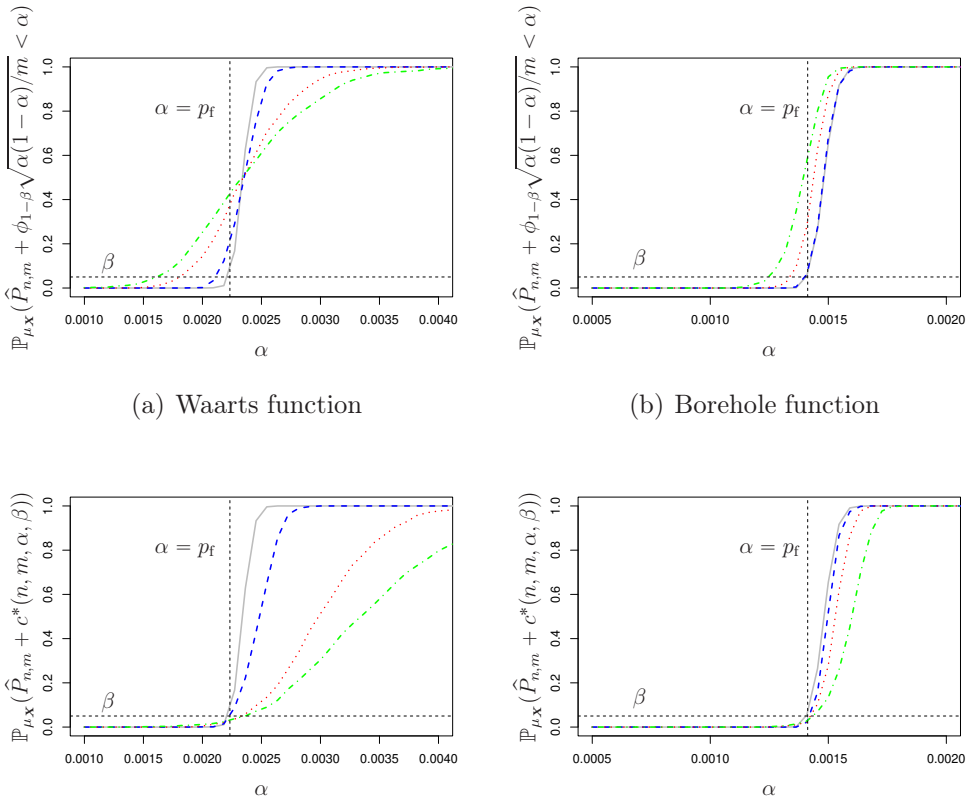
5.3. Construction of the safety margins

In this section, we aim to illustrate graphically the extent to which the decision criteria proposed in Sections 2, 3 and 4 can allow us to decide, with a chosen and controlled risk β, whether the system of interest can be considered sufficiently safe. In the following, we choose $\beta = 5\%$, which amounts to say that we accept being wrong when saying that a system is safe only 5 times over 100 in average. Depending on the system under study and the potential consequences of its failure, this value of β may of course be considered smaller or larger.

In order to numerically evaluate the relevance of the proposed decision criteria, Figure 5 first shows, for the chosen values of n and m, the evolution of $\alpha \mapsto \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_n + c^*(n,\alpha,\beta) < \alpha\right)$, that is to say the probability of labeling the true system as safe as a function of the acceptable risk α (remember that the system is considered as safe if $p_{\mathrm{f}} < \alpha$). These probabilities are estimated empirically from the 10^3 values of $\widehat{P}_n$ obtained in Section 5.2, the

values of the safety factors $c^*(n, \alpha, \beta)$ being estimated according to Eq. (19) for the crude MC case and as indicated in Section 4.1.4 for the extreme value case. In all these figures, the vertical lines correspond to the true values of p_{f}, while the horizontal lines are associated with the value of β. The safe regions are therefore to the right of these vertical lines (α is bigger than p_{f}), and the unsafe regions are to the left (α is in this case lower than p_{f}). By construction, we would like this probability to be as close to zero as possible for $\alpha < p_{\mathrm{f}}$, as close to 1 when $\alpha > p_{\mathrm{f}}$, and equal to β when $\alpha = p_{\mathrm{f}}$. And this is exactly what we obtain when looking at the black solid line, which is the reference assuming that we have access to $m \gg 1/p_{\mathrm{f}}$ i.i.d. realizations of $\boldsymbol{X}$. However, it becomes increasingly difficult for this probability to pass through the point (p_{f}, β) as n decreases. Indeed, in the MC case, this probability of labeling the system as safe is by construction a piecewise constant function, with a number of steps that corresponds to the number of times (out of the n values of $\boldsymbol{X}$ drawn) that a failure point is obtained. As n approaches $1/p_{\mathrm{f}}$, it becomes less and less likely to obtain even a single failure point, resulting in an increasingly long phase where this probability is equal to 0, i.e., a range of α values where, whatever the true probability of failure, the system will be considered unsafe. Unsurprisingly, we also observe that the lower n is, the more this probability decreases, which again reflects the increasingly difficult nature of being able to guarantee the reliability of the system as the level of information decreases. The results improve slightly when the extreme value theory is used rather than the MC approach. For the two functions studied, this approach based on the approximation of the tail of the distribution of $Y = y(\boldsymbol{X})$ by a GPD seems to make it possible to smooth the step functions of the Monte Carlo approach from above. Again, this is not necessarily surprising, as the MC and GPD approaches rely on the same information, namely n i.i.d. realizations of Y, very few of which being below $s^\star$.

Things improve greatly when we add as new information the knowledge of the parametric set to which the PDF of $\boldsymbol{X}$ belongs, as we can see in Figure 6. In these graphs, the values of $c^*(n, m, \alpha, \beta)$ are calculated as indicated in Section 4.2.4. Two types of evolution are thus compared in these graphs. In the top figures (a) and (b), we first plot the evolution of $\alpha \mapsto \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_{n,m} + \phi_{1-\beta}\sqrt{\alpha(1-\alpha)/m} < \alpha\right)$, which corresponds to the case where the decision would be based on the estimator $\widehat{P}_{n,m}$, but taking the safety factor provided by Eq. (19), i.e., not incorporating the double source



(a) Waarts function (b) Borehole function

(c) Waarts function (d) Borehole function

Figure 6: Comparison of the evolution of $\alpha \mapsto \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_{n,m} + \phi_{1-\beta}\sqrt{\alpha(1-\alpha)/m} < \alpha\right)$ (figures at the top) and $\alpha \mapsto \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_{n,m} + c^*(n,m,\alpha,\beta) < \alpha\right)$ (figures at the bottom). In each figure, the vertical line indicates the abscissa $\alpha = p_{\mathrm{f}}$, and the horizontal line corresponds to the chosen value of β. The grey lines correspond to the ideal MC case using $m \gg 1/p_{\mathrm{f}}$ i.i.d. realizations of $\boldsymbol{X}$. The curves in blue dashed line are associated to $n = 10^4$, the curves in red dotted line are associated to $n = 10^3$, and the curves in green dashed dotted line are associated to $n = 500$.

of uncertainty on which $\widehat{P}_{n,m}$ depends (crude MC approach). By neglecting the uncertainty associated to n, the risk of labeling the system as safe when it is not is not properly controlled, and this is all the more true as n decreases. In particular, if we focus on the case $\alpha = p_{\mathrm{f}}$, we can see that this probability can exceed 50%, which is much higher than the risk β we had chosen. On the contrary, if we take correctly into account the input uncertainty in the construction of the safety margin and now plot the evolution of $\alpha \mapsto \mathbb{P}_{\mu_{\boldsymbol{X}}}\left(\widehat{P}_{n,m} + c^*(n, m, \alpha, \beta) < \alpha\right)$, we find back curves that almost pass by the point (p_{f}, β), and we find again that the higher n, the quicker the increase of the probability in the safe region. In particular, with $n = 10^4$, we obtain results that are very similar to the reference case relying on $m = 1000/p_{\mathrm{f}}$ realizations of $\boldsymbol{X}$.
However, it is important to emphasize that this approach, which is based on an approximation of the PDF of the inputs, may generate a non-negligible additional numerical cost compared with the previous MC or extreme event approaches, in the sense that, this time, the code is always evaluated at m inputs points, whereas only $n \ll m$ evaluations of the code were previously necessary.

6. Conclusions and prospect

By focusing on statistical inference and uncertainty quantification, the main idea of this work was to provide a pathway to make informed decisions about system safety in the face of limited input information. Indeed, reliability analysis of systems traditionally relies on probabilistic models that require a precise understanding of input probability distributions. But in practical engineering settings, such distributions are often unknown or poorly characterized, which requires the adaptation of existing methods to determine whether the system is secure or not, while minimizing the risk of mislabeling due to uncertainty about input distributions. Several methods have thus been developed in this paper to enable a decision to be made on whether a system is safe or not in the case where the maximum information on the inputs is a set of n i.i.d. realizations. A distinction has been made between cases where n is large and cases where n is relatively small compared with the inverse of the probability of failure. And it is in the treatment of this case where n is relatively small, i.e., the cases where the estimation of the probability of failure must integrate this additional source of uncertainty linked to the lack of knowledge of the input probability distribution, that

the main contributions of this paper are to be found. Our findings demonstrate the potential of statistical frameworks to enhance reliability analysis by incorporating probabilistic assessments and robust decision-making tools. And by controlling the risk of misclassification, our approach offers practical solutions for assessing system safety based on limited input knowledge, paving the way for more effective and reliable engineering design and risk management practices.

In future work, further refinements and extensions of the proposed methods could explore alternative assumptions on the input and output probability distributions, and expand the applicability of the proposed framework to broader classes of estimators for the probability of failure. For this work, we limited ourselves to parametric approximations for the distributions of inputs and outputs, and we can wonder if it would not be possible to also obtain interesting decision criteria by working with non-parametric estimates. In addition, we have focused on estimates of the probability of failure relying on a Monte Carlo approach, without really worrying about the number of times we call the numerical code. If the calculation cost of the simulator is significant, it is nevertheless not reasonable to allow millions of calls to it, and we could seek to couple the work presented with estimators more suited to the estimation of a low probability, such as estimators from Multilevel Splitting / Subset Simulation (see, e.g., [2, 5]) or Moving Particles [31, 21] approaches. Finally, a number of numerical optimizations could be proposed to make the application of the introduced decision criteria more efficient, whether at the level of the choice of the auxiliary PDFs, or at the level of the constrained problems for the identification of safety margins.

[1] A. Ajenjo, E. Ardillon, V. Chabridon, B. Iooss, S. Cogan, and E. Sadoulet-Reboul, *An info-gap framework for robustness assessment of epistemic uncertainty models in hybrid structural reliability analysis*, Structural Safety, 96 (2022), p. 102196.

[2] S. Au and J. Beck, *Estimation of small failure probabilities in high dimensions by subset simulation*, Probabilistic Engineering Mechanics, 16 (4) (2001), pp. 263–277.

[3] J.-M. Bourinet, *Mechanical Engineering under Uncertainties: From Classical Approaches to Some Recent Developments*, ISTE - Wiley, 2021, ch. "5. Rare-event Probability Estimation", pp. 145–204.

[4] M. F. Bugallo, V. Elvira, L. Martino, D. Luengo, J. Miguez, and P. M. Djuric, *Adaptive importance sampling: The past, the present, and the future*, IEEE Signal Processing Magazine, 34 (2017), pp. 60–79.

[5] F. Cérou and A. Guyader, *Adaptive multilevel splitting for rare event analysis*, Stochastic Analysis and Applications, 25 (2) (2007), pp. 417–443.

[6] V. Chabridon, *Reliability-oriented sensitivity analysis under probabilistic model uncertainty–Application to aerospace systems*, PhD thesis, Université Clermont Auvergne [2017-2020], 2018.

[7] V. Chabridon, M. Balesdent, J.-M. Bourinet, J. Morio, and N. Gayton, *Evaluation of failure probability under parameter epistemic uncertainty: application to aerospace system reliability assessment*, Aerospace Science and Technology, 69 (2017), pp. 526–537.

[8] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An introduction to statistical modeling of extreme values*, vol. 208, Springer, 2001.

[9] Z. Del Rosario, *Precision Materials Indices: Materials Selection with Statistically Rigorous Reliability Analysis*, AIAA Journal, 60 (2022), pp. 578–586.

[10] Z. Del Rosario, R. W. Fenrich, and G. Iaccarino, *Fast Precision Margin with the First-Order Reliability Method*, AIAA Journal, (2019).

[11] Z. Del Rosario, R. W. Fenrich, and G. Iaccarino, *When Are Allowables Conservative?*, AIAA Journal, 59 (2021), pp. 1760–1772.

[12] P. I. Frazier, *A tutorial on Bayesian optimization*, arXiv preprint arXiv:1807.02811, (2018).

[13] C. Gauchy, J. Stenger, R. Sueur, and B. Iooss, *An Information Geometry Approach to Robustness Analysis for the Uncertainty Quantification of Computer Codes*, Technometrics, 64 (2022), pp. 80–91.

[14] R. B. Gramacy, *Surrogates: Gaussian process modeling, design, and optimization for the applied sciences*, Chapman and Hall/CRC, 2020.

[15] E. J. Gumbel, *Statistics of extremes*, Columbia university press, 1958.

[16] D. R. Jones, *A taxonomy of global optimization methods based on response surfaces*, Journal of Global Optimization, 21 (2001), pp. 345–383.

[17] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer Texts in Statistics, Springer New York, NY, Third ed. ed., 2005.

[18] J. Morio and M. Balesdent, *Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems: A Practical Approach*, Woodhead Publishing, Elsevier, 2015.

[19] M. D. Morris, T. J. Mitchell, and D. Ylvisaker, *Bayesian design and analysis of computer experiments: use of derivatives in surface prediction*, Technometrics, 35 (1993), pp. 243–255.

[20] A. Pasanisi, M. Keller, and E. Parent, *Estimation of a quantity of interest in uncertainty analysis: Some help from Bayesian decision theory*, Reliability Engineering and System Safety, 100 (2012), pp. 93–101.

[21] G. Perrin, *Point process-based approaches for the reliability analysis of systems modeled by costly simulators*, Reliability Engineering & System Safety, 214 (2021).

[22] G. Perrin and C. Cannamela, *A repulsion-based method for the definition and the enrichment of opotimized space filling designs in constrained input spaces*, Journal de la Société Française de Statistique, 158 (2017), pp. 37–67.

[23] G. Perrin and rodolphe le riche, *Bayesian optimization with derivatives acceleration*, Transactions on Machine Learning Research, (2024), `https://openreview.net/forum?id=JRjD0YF3Yd`.

[24] J. Pickands III, *Statistical inference using extreme order statistics*, the Annals of Statistics, (1975), pp. 119–131.

[25] R. T. Rockafellar, *Lagrange multipliers and optimality*, SIAM review, 35 (1993), pp. 183–238.

[26] R. T. Rockafellar and J. O. Royset, *Engineering Decisions under Risk Averseness*, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 1 (2015), pp. 1–12.

[27] R. T. Rubinstein and D. Kroese, *Simulation and the Monte Carlo method*, John Wiley and Sons, Inc., Hoboken, New Jersey, 2008.

[28] J. Stenger, F. Gamboa, M. Keller, and B. Iooss, *Optimal Uncertainty Quantification of a risk measurement from a thermal-hydraulic code using canonical moments*, International Journal for Uncertainty Quantification, 10 (2020), pp. 35–53.

[29] A. Touboul, J. Reygner, F. Mangeant, and P. Benjamin, *A formal framework to define margins in industrial processes*, HAL, (2019), pp. 1–27, `https://arxiv.org/abs/hal-02156493`.

[30] A. W. Van der Vaart, *Asymptotic statistics*, vol. 3, Cambridge university press, 2000.

[31] C. Walter, *Moving particles: A parallel optimal multilevel splitting method with application in quantiles estimation and meta-model based algorithms*, Structural Safety, 55 (2015), pp. 10–25.

[32] Q. Zhou, P. Z. Qian, and S. Zhou, *A simple approach to emulation for computer models with qualitative and quantitative factors*, Technometrics, 53 (2011), pp. 266–273.

A. Proofs

A.1. Proof of Proposition 2

In order to prove Proposition 2, we begin by introducing the following lemma.

Lemma A.1. *If $\widehat{P}_n$ is uniformly consistent, $c^*(n, \alpha, \beta)$ tends to 0 when n tends to infinity.*

▫ **Proof:** L ▫et $\varepsilon > 0$ and $\widehat{P}_n$ be a consistent estimator of p_{f}. We can then calculate:

$$\begin{aligned}\sup_{\mu\in\mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + \varepsilon < \alpha\right) &= \sup_{\alpha\leqslant p\leqslant 1}\left\{\sup_{\mu\in\mathcal{P}_p} \mathbb{P}_\mu\left(\widehat{P}_n + \varepsilon < \alpha\right)\right\} \\ &\leqslant \sup_{\alpha\leqslant p\leqslant 1}\left\{\sup_{\mu\in\mathcal{P}_p} \mathbb{P}_\mu\left(\widehat{P}_n + \varepsilon < p\right)\right\} \qquad \text{(A.1)} \\ &\leqslant \sup_{\alpha\leqslant p\leqslant 1}\left\{\sup_{\mu\in\mathcal{P}_p} \mathbb{P}_\mu\left(\varepsilon < |\widehat{P}_n - p|\right)\right\}.\end{aligned}$$

As $\widehat{P}_n$ is uniformly consistent, $\sup_{\alpha\leqslant p\leqslant 1}\left\{\sup_{\mu\in\mathcal{P}_p} \mathbb{P}_\mu\left(\varepsilon < |\widehat{P}_n - p|\right)\right\}$ tends to 0 when n tends to infinity. As long as the type I error β is strictly positive, which is the framework used in this paper, we deduce that there exists $\widetilde{n}$ such that for all $n \geqslant \widetilde{n}$,

$$\sup_{\mu\in\mathcal{P}_{\geqslant\alpha}} \mathbb{P}_\mu\left(\widehat{P}_n + \varepsilon < \alpha\right) \leqslant \sup_{\alpha\leqslant p\leqslant 1}\left\{\sup_{\mu\in\mathcal{P}_p} \mathbb{P}_\mu\left(\varepsilon < |\widehat{P}_n - p|\right)\right\} \leqslant \beta. \qquad \text{(A.2)}$$

Using Eq. (5), $c^*(n,\alpha,\beta)$ is then smaller than ε for all $n \geqslant \widetilde{n}$. This being true whatever the value of ε, we deduce that

$$\lim_{n\to+\infty} c^*(n,\alpha,\beta) = 0. \qquad \text{(A.3)}$$

Now, since $\widehat{P}_n$ is (uniformly) consistent, we have in particular that under $\mathbb{P}_{\mu_{\boldsymbol{X}}}$, $\widehat{P}_n$ converges to p_{f} in probability. By Lemma A.1, $c^*(n,\alpha,\beta)$ tends to 0, and therefore, since $p_{\mathrm{f}} < \alpha$, Eq. (9) holds.

A.2. Proof of Proposition 3

Let $\boldsymbol{X} \sim \mu \in \mathcal{P}$, and $\boldsymbol{w} = (\sigma, \xi, p_{\overline{s}}) \in \mathcal{W}$ be the vector of parameters characterizing the left tail of the CDF of $y(\boldsymbol{X})$, in the sense that $Y = y(\boldsymbol{X}) \sim \mathrm{a-GPD}(\overline{s}, \boldsymbol{w})$. For $k \in \{0,1,2\}$, if $Y_1, \ldots, Y_n$ are n i.i.d. copies of Y, we define

$$\tau_n^k = \frac{1}{n}\sum_{i=1}^{n} Y_i^k \, \mathbb{1}_{\{Y_i \leqslant \overline{s}\}}.$$

Recalling that $p_{\overline{s}} = \mathbb{P}_\mu(Y \leqslant \overline{s})$, it comes:

$$\begin{aligned}\mathbb{E}_\mu\left[\left(\tau_n^k - \mathbb{E}_\mu\left[Y^k\,\mathbb{1}_{\{Y\leqslant\overline{s}\}}\right]\right)^2\right] &= \frac{1}{n}\operatorname{Var}\left(Y^k\,\mathbb{1}_{\{Y\leqslant\overline{s}\}}\right)\\ &= \frac{1}{n}\left(p_{\overline{s}}\,\mathbb{E}_\mu\left[Y^{2k}\mid Y\leqslant\overline{s}\right] - p_{\overline{s}}^2\mathbb{E}_\mu\left[Y^k\mid Y\leqslant\overline{s}\right]^2\right).\end{aligned}$$

Using the properties of the GPD, for all $\ell \in \{0,\dots,4\}$, we have

$$\mathbb{E}_\mu\left[(Y-\overline{s})^\ell \mid Y\leqslant\overline{s}\right] = \sigma^\ell\varphi_\ell(\xi),$$

where φ_ℓ is a rational fraction which is well defined if $\xi < 1/\ell$. As a consequence, the assumptions that $-\infty < \xi_{\min} < \xi_{\max} < 1/4$ and that $\sigma_{\max} < +\infty$ yield

$$\forall k\in\{0,1,2\}, \qquad \lim_{n\to+\infty}\sup_{\mu\in\mathcal{P}}\mathbb{E}_\mu\left[\left(\tau_n^k - \mathbb{E}\left[Y^k\,\mathbb{1}_{\{Y\leqslant\overline{s}\}}\right]\right)^2\right] = 0,$$

and thus by the Chebychev inequality, for any $\delta > 0$,

$$\lim_{n\to+\infty}\sup_{\mu\in\mathcal{P}}\mathbb{P}_\mu\left(\|\boldsymbol{\tau}_n - \boldsymbol{\tau}(\mu)\|_1 \geqslant \delta\right) = 0, \tag{A.4}$$

with $\boldsymbol{\tau}_n := (\tau_n^0,\tau_n^1,\tau_n^2)$ and $\boldsymbol{\tau}(\mu) := \mathbb{E}_\mu[\boldsymbol{\tau}_n] = \left(\mathbb{P}_\mu(Y\leqslant\overline{s}),\ \mathbb{E}_\mu\left[Y\,\mathbb{1}_{\{Y\leqslant\overline{s}\}}\right],\ \mathbb{E}_\mu\left[Y^2\,\mathbb{1}_{\{Y\leqslant\overline{s}\}}\right]\right)$. We now introduce the function

$$\Psi(\tau^0,\tau^1,\tau^2) = \tau^0 \times F_{(\overline{s},\sigma,\xi)}(s^\star),$$

with

$$\xi = \frac{1}{2}\left(1 - \frac{(\overline{s}-\tau^1/\tau^0)^2}{\tau^2/\tau^0 - (\tau^1/\tau^0)^2}\right), \qquad \sigma = \frac{1}{2}(\overline{s}-\tau^1/\tau^0)\left(1 + \frac{(\overline{s}-\tau^1/\tau^0)^2}{\tau^2/\tau^0 - (\tau^1/\tau^0)^2}\right),$$

such that

$$\widehat{P}_n = \Psi(\boldsymbol{\tau}_n), \qquad \mathbb{P}_\mu(Y\leqslant s^\star) = \Psi(\boldsymbol{\tau}(\mu)).$$

The function Ψ is defined and continuous at all points $\boldsymbol{\tau} = (\tau^0,\tau^1,\tau^2) \in \mathbb{R}^3$ such that $\tau^0 > 0$ and $\tau^2/\tau^0 > (\tau^1/\tau^0)^2$. But the conditions that $p_{\overline{s},\min} > 0$ and $\sigma_{\min} > 0$ in the definition of $\mathcal{W}$ ensure that $\boldsymbol{\tau}(\mu)$ satisfies these conditions uniformly in $\mu\in\mathcal{P}$. Therefore, since the set $\{\boldsymbol{\tau}(\mu), \mu\in\mathcal{P}\} \subset \mathbb{R}^3$ is moreover bounded, we deduce that for any $\varepsilon > 0$, there exists $\delta > 0$ such that, for any

$\mu \in \mathcal{P}$ and for any $\boldsymbol{\tau} \in \mathbb{R}^3$, if $\|\boldsymbol{\tau} - \boldsymbol{\tau}(\mu)\|_1 < \delta$ then $|\Psi(\boldsymbol{\tau}) - \Psi(\boldsymbol{\tau}(\mu))| < \varepsilon$. As a consequence,

$$\begin{aligned}
\sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu\left(|\widehat{P}_n - \mathbb{P}_\mu(Y \leqslant s^\star)| \geqslant \varepsilon\right) &= \sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu\left(|\Psi(\boldsymbol{\tau}_n) - \Psi(\boldsymbol{\tau}(\mu))| \geqslant \varepsilon\right) \\
&\leqslant \sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu\left(|\Psi(\boldsymbol{\tau}_n) - \Psi(\boldsymbol{\tau}(\mu))| \geqslant \varepsilon, \|\boldsymbol{\tau}_n - \boldsymbol{\tau}(\mu)\|_1 \geqslant \delta\right) \\
&\quad + \sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu\left(|\Psi(\boldsymbol{\tau}_n) - \Psi(\boldsymbol{\tau}(\mu))| \geqslant \varepsilon, \|\boldsymbol{\tau}_n - \boldsymbol{\tau}(\mu)\|_1 < \delta\right) \\
&\leqslant \sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu\left(\|\boldsymbol{\tau}_n - \boldsymbol{\tau}(\mu)\|_1 \geqslant \delta\right),
\end{aligned}$$

and by Eq. (A.4), the right-hand side vanishes when $n \to +\infty$, which completes the argument.

A.3. Proof of Proposition 4

Let $\varepsilon > 0$. For each $\mu \in \mathcal{P}$, noting $p = \mathbb{P}_\mu(\boldsymbol{X} \in \mathcal{F})$, recalling that $\mathbb{E}_\mu\left[\widehat{P}_{n,m} \mid \widehat{\mu}_n\right] = \mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right)$, and using the Markov inequality, we have:

$$\begin{aligned}
\mathbb{P}_\mu\left(|\widehat{P}_{n,m} - p| \geqslant \varepsilon\right) &= \mathbb{E}_\mu\left[\mathbb{P}_\mu\left(|\widehat{P}_{n,m} - p| \geqslant \varepsilon \mid \widehat{\mu}_n\right)\right] \\
&\leqslant \mathbb{E}_\mu\left[\frac{\mathbb{E}_\mu\left[\left(\widehat{P}_{n,m} - p\right)^2 \mid \widehat{\mu}_n\right]}{\varepsilon^2}\right] \\
&\leqslant \mathbb{E}_\mu\left[\frac{\mathbb{E}_\mu\left[\left(\widehat{P}_{n,m} - \mathbb{E}_\mu\left[\widehat{P}_{n,m} \mid \widehat{\mu}_n\right] + \mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right) - p\right)^2 \mid \widehat{\mu}_n\right]}{\varepsilon^2}\right] \\
&\leqslant \frac{1}{\varepsilon^2}\mathbb{E}_\mu\left[\mathbb{E}_\mu\left[\left(\widehat{P}_{n,m} - \mathbb{E}_\mu\left[\widehat{P}_{n,m} \mid \widehat{\mu}_n\right]\right)^2 \mid \widehat{\mu}_n\right] + \left(\mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right) - p\right)^2\right] \\
&= \frac{1}{\varepsilon^2}\mathbb{E}_\mu\left[\frac{\mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right)\left(1 - \mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right)\right)}{m} + \left(\mathbb{P}_\mu\left(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n\right) - p\right)^2\right] \\
&\leqslant \frac{1}{4m\varepsilon^2} + \frac{1}{\varepsilon^2}\mathbb{E}_\mu\left[\left(\mathbb{P}_\mu(\boldsymbol{X}^{(n)} \in \mathcal{F} \mid \widehat{\mu}_n) - p\right)^2\right].
\end{aligned} \tag{A.5}$$

Using Hypothesis (1), we therefore see that $\sup_{\mu\in\mathcal{P}} \mathbb{P}_\mu(|\widehat{P}_{n,m} - p| \geqslant \varepsilon)$ tends to 0 when n and m tend to infinity, which proves the uniform consistency of $\widehat{P}_{n,m}$.

B. Generation of i.i.d. realizations of the failure probability estimator

B.1. The extreme-value case

Let $\overline{s} \geqslant s^\star$, $\boldsymbol{w} = (\sigma, \xi, p_{\overline{s}}) \in \mathcal{W}$, and $\widehat{P}_n$ be the estimator of p_{f} given by (30). Under $\mathbb{P}_{\boldsymbol{w}}$, this statistic is a function of a vector $(Z_1, \ldots, Z_n)$ of i.i.d. realizations of a variable $Z \sim \mathrm{a-GPD}(\overline{s}, \boldsymbol{w})$. With the notation of Section 4.1.2, $N_{\overline{s}}$ indicates the number of realizations of Z that are smaller than $\overline{s}$, and follows therefore a binomial distribution with parameters n and $p_{\overline{s}}$, while $\widehat{\Xi}_n$ and $\widehat{\Sigma}_n$ are constructed from these $N_{\overline{s}}$ realizations. By inversion of the CDF, we can then note that if U is a random variable uniformly distributed on $(0, 1]$, $(Z \mid Z \leqslant \overline{s})$ and

$$F^{-1}_{(\overline{s},\sigma,\xi)}(U) = \begin{cases} \overline{s} - \frac{\sigma(U^{-\xi}-1)}{\xi} & \text{if } \xi \neq 0, \\ \overline{s} + \sigma \log U & \text{if } \xi = 0, \end{cases}$$

have the same probability distributions. As a consequence, to simulate a realization of $\widehat{P}_n$ under $\mathbb{P}_{\boldsymbol{w}}$, one first draws a binomial random variable $N_{\overline{s}}$ with parameters n and $p_{\overline{s}}$, and then $N_{\overline{s}}$ independent uniform random variables $U_1, \ldots, U_{N_{\overline{s}}}$ on $(0, 1]$, sets $Z_i = F^{-1}_{(\overline{s},\sigma,\xi)}(U_i)$ for $1 \leqslant i \leqslant N_{\overline{s}}$ and returns the associated value of $\widehat{P}_n$.

B.2. The input case

Once a value of $\boldsymbol{\theta}$ has been chosen in $\mathcal{S}_{\mathrm{INP}}(\alpha)$ (assuming we are able to select such a vector, which is a priori not trivial), the PDF $f(\cdot; \boldsymbol{\theta})$ is completely defined. We can then generate n independent realizations of $\boldsymbol{X} \sim f(\cdot; \boldsymbol{\theta})$, and deduce a possible value for the estimator $\widehat{\boldsymbol{\theta}}_n$. We can then generate m new values of $\boldsymbol{X} \sim f(\cdot; \widehat{\boldsymbol{\theta}}_n)$ according to this new PDF. By evaluating the code at these m input points, we can generate a possible realization of $\widehat{P}_{n,m}$ that is associated with the chosen value of $\boldsymbol{\theta}$. By repeating this procedure independently r times, r independent realizations of $\widehat{P}_{n,m}$ are generated, which can eventually be post-processed to estimate the β-quantile of $\widehat{P}_{n,m}$.