Imprecise Random Variables, Random Sets, and Monte Carlo Simulation

Thomas Fetz
University of Innsbruck, Austria
thomas.fetz@uibk.ac.at

Michael Oberguggenberger
University of Innsbruck, Austria
michael.oberguggenberger@uibk.ac.at

Abstract

The paper addresses the evaluation of upper and lower probabilities induced by functions of an imprecise random variable. Given a function g and a family X_λ of random variables, where the parameter λ ranges in an index set Λ, one may ask for the upper/lower probability that $g(X_\lambda)$ belongs to some Borel set B. Two interpretations are investigated. In the first case, the upper probability is computed as the supremum of the probabilities that $g(X_\lambda)$ lies in B. In the second case, one considers the random set generated by all $g(X_\lambda)$, $\lambda \in \Lambda$, e.g. by transforming X_λ to standard normal as a common probability space, and computes the corresponding upper probability. The two results are different, in general. We analyze this situation and highlight the implications for Monte Carlo simulation. Attention is given to efficient simulation procedures and an engineering application is presented.

Keywords. Upper and lower probabilities, imprecise random variables, random sets, propagation of uncertainty through a function, Monte Carlo simulation.

1 Introduction

Methods of imprecise probability have increasingly attracted interest in the engineering community, see e.g. the recent survey article [3]. In most industrial applications, engineering structures are described by black box input-output models, given by large computer programs. To evaluate probabilities of output quantities, Monte Carlo simulation is the method of choice. Computational effort is aggravated, if the input variables are imprecise random variables (or even random fields) with set-valued parameters.

The present paper is motivated by recent papers on simulation of random sets [16, 17] and on simulation of upper and lower probabilities in engineering reliability [2]. These papers raise the question what is the appropriate model of imprecision and what are cost-saving ways of simulating the output quantities. The key issue is to keep the number of required evaluations of the expensive input-output map as low as possible. Suppose the imprecision of the input is described in terms of a family of random variables. Two interpretations of resulting lower and upper probabilities are at hand: the first one by taking infima and suprema over the probabilities generated by the individual members of the family, the second one by first forming random sets based on the family and then evaluating the lower and upper probability as belief and plausibility (see e.g. [4]). In general, the results differ. The second question is the appropriate simulation method in the two cases.

In civil engineering, the use of random sets on continuous probability spaces goes back at least to [20], see also the text book [4] and the survey in [15]. Various alternative approaches to simulation of random sets have been proposed in recent years [1, 14, 22, 21]. The employed notion of imprecise probability is of course crucial. This paper focusses on the two limiting cases given by a family of random variables versus the random set generated by it. Once the random set is given, other ways of generating upper and lower probabilities are known: by means of all measurable selections of the random set, by means of measurable selections of the measures living on the focal elements, or by means of its core (the set of probabilities dominated by the upper probability of the random set). For these notions we refer to [5, 11, 12] as well as the exhaustive comparison in [6] (see also the remarks in Section 3 of the paper).

The paper is organized as follows. In Section 2, the set-up is explained and sufficient conditions are given making the formation of a random set possible. Section 3 is devoted to comparing the two versions of lower and upper probabilities and to stating conditions under which they coincide. Section 4 presents an engineering example, which is used in the subsequent sections for the purpose of illustration. Section 5 addresses issues of simulating random sets and the corresponding lower/upper probabilities, while Section 6 addresses the issue of simulation for lower/upper probabilities derived from families of random variables. Section 7 contains a summary and conclusions.

2 The Set-Up

The issue of the paper is how to evaluate and to simulate a function of an imprecise random variable. To fix the notation, let Λ be an index set. Consider a family of random variables $\{X_\lambda : \lambda \in \Lambda\}$, defined on some probability space $(\Omega, \Sigma, \mathsf{m})$. At fixed $\lambda \in \Lambda$, the random variable X_λ may be univariate or multivariate, with values in some $\mathbb{R}^n$. Further, let $g : \mathbb{R}^n \to \mathbb{R}$ be a continuous function.

We wish to investigate lower and upper probabilities induced by the family of random variables $g(X_\lambda)$, $\lambda \in \Lambda$. From a probabilistic point of view, the function g could be suppressed, because $Y_\lambda = g(X_\lambda)$ is again a family of random variables. However, in the applications we have in mind, the function g will be an input-output map of a complex system, for which we want to compute the output distribution. Using Monte Carlo simulation, large samples of $g(X_\lambda)$ should be computed. While generating large samples of the input is computationally inexpensive, the evaluation of the function g might be very expensive. As mentioned, a focus of the paper will be on devising simulation methods of low computational cost. For this reason, the function g will play a role in Sections 4 to 6. One could also consider more generally parametrized functions of the form $g(\lambda, X_\lambda)$. For the initial probabilistic analysis, we drop mentioning the function g and consider families of real valued random variables.

Let $(\Omega, \Sigma, \mathsf{m})$ be the basic probability space, which we assume to be complete. Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of random variables $X_\lambda : \Omega \to \mathbb{R} : \omega \to X_\lambda(\omega)$. Then the probability of a Borel set $B \in \mathcal{B}(\mathbb{R})$ for a fixed random variable X_λ is

$$P(X_\lambda \in B) = \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B} \, \mathsf{dm}(\omega) \tag{1}$$

where

$$\mathbb{1}_E = \begin{cases} 1 & \text{if event } E \text{ occurs,} \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

denotes the indicator function of an event E. Modeling uncertainty by a family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables, it is natural to define lower and upper probabilities:

Definition 1 *For a family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables, the lower probability $\underline{P}$ and the upper probability $\overline{P}$ are given by*

$$\underline{P}(B) = \inf_{\lambda\in\Lambda} P(X_\lambda \in B) = \inf_{\lambda\in\Lambda} \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B} \, \mathsf{dm}(\omega), \tag{3}$$

$$\overline{P}(B) = \sup_{\lambda\in\Lambda} P(X_\lambda \in B) = \sup_{\lambda\in\Lambda} \int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B} \, \mathsf{dm}(\omega), \tag{4}$$

where $B \subset \mathbb{R}$ is a Borel set.

On the other hand, it is also natural to construct a random set based on the family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables. To this end, we define the set-valued map $\mathcal{X} : \Omega \to \mathbb{R}$ by

$$\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\}. \tag{5}$$

By definition, this set-valued map $\mathcal{X}$ is a random set, provided the upper inverses

$$\mathcal{X}^-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \cap B \neq \varnothing\} \tag{6}$$

are measurable subsets of Ω for every Borel set B. Then the same holds also for the lower inverses

$$\mathcal{X}_-(B) = \{\omega \in \Omega : \mathcal{X}(\omega) \subset B\}. \tag{7}$$

These lower and upper inverses lead to lower and upper probabilities based on the random set $\mathcal{X}$:

Definition 2 *For a random set $\mathcal{X}$, the lower probability $\underset{\sim}{P}$ and the upper probability $\widetilde{P}$ is defined by*

$$\underset{\sim}{P}(B) = \mathsf{m}(\mathcal{X}_-(B)) = \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\subseteq B} \, \mathsf{dm}(\omega), \tag{8}$$

$$\widetilde{P}(B) = \mathsf{m}(\mathcal{X}^-(B)) = \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\cap B\neq\varnothing} \, \mathsf{dm}(\omega). \tag{9}$$

We will show in the next section that

$$\underset{\sim}{P} \leq \underline{P} \leq \overline{P} \leq \widetilde{P} \tag{10}$$

with equality only in special cases, but first we give sufficient conditions so that $\mathcal{X}$ generated by the family $\{X_\lambda\}_{\lambda\in\Lambda}$ is a random set.

Theorem 1 *Let a family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables be given. Assume that Λ is a compact subset of a metric space and that the maps $\lambda \to X_\lambda(\omega)$ are continuous for each fixed $\omega \in \Omega$. Then*

(a) *the set-valued map $\mathcal{X}$ defined by*

$$\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\} \tag{11}$$

is a compact random set;

(b) *the map $\overline{\mathbb{1}}_B : \omega \to \sup_{\lambda\in\Lambda} \mathbb{1}_{X_\lambda(\omega)\in B}$ is measurable;*

(c) *the map $\underline{\mathbb{1}}_B : \omega \to \inf_{\lambda\in\Lambda} \mathbb{1}_{X_\lambda(\omega)\in B}$ is measurable;*

where B is a Borel set.

Proof. (a) As a compact subset of a metric space, Λ is separable. Take a countable dense subset of parameter values λ_k in Λ. From the continuity assumption, it follows that the sequence $\{X_{\lambda_k}(\omega) : k = 1, 2, 3, \ldots\}$ is dense in $\mathcal{X}(\omega)$ for

every fixed ω. (That is, it is a Castaing representation.) Let B be an open set. Then the set

$$\begin{aligned}\mathcal{X}^-(B) &= \{\omega\in\Omega : \mathcal{X}(\omega)\cap B\neq\varnothing\} \\ &= \{\omega\in\Omega : \text{ there is } k \text{ such that } X_{\lambda_k}(\omega)\in B\} \\ &= \bigcup_k \{\omega\in\Omega : X_{\lambda_k}(\omega)\in B\}\end{aligned} \tag{12}$$

is measurable as a countable union of measurable sets. (The individual sets are measurable, because each X_λ is a random variable.) The fundamental measurability theorem (see e.g. [13]) implies that $\mathcal{X}^-(B)$ is measurable for every Borel set B. In addition, $\mathcal{X}(\omega)$ is the continuous image of a compact set. Thus $\mathcal{X}$ is a compact random set.

(b) Considering $\overline{\mathbb{1}}_B$, we observe the equivalence

$$\begin{aligned}\overline{\mathbb{1}}_B(\omega) &= \sup_{\lambda\in\Lambda} \mathbb{1}_{X_\lambda(\omega)\in B} = 1 \\ &\iff \exists\lambda\in\Lambda : X_\lambda(\omega)\in B \iff \mathcal{X}(\omega)\cap B\neq\varnothing.\end{aligned} \tag{13}$$

Thus $\mathcal{X}^-(B) = \overline{\mathbb{1}}_B^{-1}(\{1\})$. Since $\overline{\mathbb{1}}_B$ takes only the two values 0 and 1, this proves that $\overline{\mathbb{1}}_B$ is measurable.

(c) For $\underline{\mathbb{1}}_B$ we have

$$\begin{aligned}\underline{\mathbb{1}}_B(\omega) &= \inf_{\lambda\in\Lambda} \mathbb{1}_{X_\lambda(\omega)\in B} = 1 \\ &\iff \forall\lambda\in\Lambda : X_\lambda(\omega)\in B \iff \mathcal{X}(\omega)\subseteq B\end{aligned} \tag{14}$$

which leads to $\mathcal{X}_-(B) = \underline{\mathbb{1}}_B^{-1}(\{1\})$. The same arguments as in (b) yield the measurability of $\underline{\mathbb{1}}_B$. □

3 Comparison Results

The purpose of this section is to prove the chain of inequalities formulated in Eq. (10), exhibit some circumstances when they are equal and illustrate the behavior by means of simple examples.

Theorem 2 *Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of random variables according to the assumptions of Theorem 1 and let $\mathcal{X}$ be the random set induced by this family together with the map $\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda\in\Lambda\}$. Then it holds*

$$\underset{\sim}{P} \leq \underline{P} \leq \overline{P} \leq \widetilde{P} \tag{15}$$

for the lower and upper probabilities in Def. 1 and 2.

Proof. For the upper probabilities of a Borel set B we have

$$\begin{aligned}\overline{P}(B) &= \sup_{\lambda\in\Lambda} P(X_\lambda\in B) = \sup_{\lambda\in\Lambda}\int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{dm}(\omega) \\ &\leq \int_\Omega \sup_{\lambda\in\Lambda}\mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{dm}(\omega) = \int_\Omega \overline{\mathbb{1}}_B\,\mathrm{dm}(\omega) \\ &= \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\cap B\neq\varnothing}\,\mathrm{dm}(\omega) = \widetilde{P}(B)\end{aligned} \tag{16}$$

using Eq. (13). Together with Eq. (14) we get

$$\begin{aligned}\underline{P}(B) &= \inf_{\lambda\in\Lambda} P(X_\lambda\in B) = \inf_{\lambda\in\Lambda}\int_\Omega \mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{dm}(\omega) \\ &\geq \int_\Omega \inf_{\lambda\in\Lambda}\mathbb{1}_{X_\lambda(\omega)\in B}\,\mathrm{dm}(\omega) = \int_\Omega \underline{\mathbb{1}}_B\,\mathrm{dm}(\omega) \\ &= \int_\Omega \mathbb{1}_{\mathcal{X}(\omega)\subseteq B}\,\mathrm{dm}(\omega) = \underset{\sim}{P}(B)\end{aligned} \tag{17}$$

for the lower probabilities. □

Remark. Let

$$\mathcal{M}(\widetilde{P}) = \{P : P(A)\leq\widetilde{P}(A), A\in\Sigma\} \tag{18}$$

be the set of all probability measures dominated by the upper probability $\widetilde{P}$ induced by the random set $\mathcal{X}$. Further let

$$P(\mathcal{X}) = \{P_X : X\in S(\mathcal{X})\} \tag{19}$$

be the set of all probability measures generated by the measurable selections

$$S(\mathcal{X}) = \{X : \Omega\to\mathbb{R} \text{ measurable} : X(\omega)\in\mathcal{X}(\omega)\} \tag{20}$$

of the random set $\mathcal{X}$. In [5, 11, 12] these two sets $P(\mathcal{X})$ and $\mathcal{M}(\widetilde{P})$ are investigated and it is proven that the relation $P(\mathcal{X})\subseteq\mathcal{M}(\widetilde{P})$ holds and that we have $P(\mathcal{X}) = \mathcal{M}(\widetilde{P})$ under certain conditions.

In our case the random variables $X\in\{X_\lambda\}_{\lambda\in\Lambda}$ are measurable selections of the random set $\mathcal{X}$ but in general not all of the selections in $S(\mathcal{X})$. That means we have $\{X_\lambda\}_{\lambda\in\Lambda}\subseteq S(\mathcal{X})$ and the following relations

$$\begin{aligned}\overline{P}(B) &= \underbrace{\sup_{\lambda\in\Lambda} P(X_\lambda\in B) = \sup_{X\in\{X_\lambda\}_{\lambda\in\Lambda}} P_X(B)}_{\text{family of random variables}} \leq \\ &\leq \underbrace{\sup_{X\in S(\mathcal{X})} P_X(B)}_{\text{all measurable selections}} \leq \underbrace{\sup_{P\in\mathcal{M}(\widetilde{P})} P(B)}_{\text{dominated probabilities}} = \widetilde{P}(B)\end{aligned}$$

for the upper probabilities and vice versa for the lower probabilities.

Example 1 Let $(\Omega,\Sigma,\mathsf{m}) = (\mathbb{R},\mathcal{B}(\mathbb{R}),\mathsf{m})$ be the probability space with probability measure

$$\mathsf{m}(B) = \int_{\mathbb{R}} \mathbb{1}_{\omega\in B}\,\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\omega^2/2}\,\mathrm{d}\omega, \quad B\in\mathcal{B}(\mathbb{R}). \tag{21}$$

On the one hand, the family $\{X_\lambda\}_{\lambda\in\Lambda}$ of random variables is given by

$$\{X_{(\mu,\sigma)}\}_{(\mu,\sigma)\in\Lambda}, \quad \Lambda = [\underline{\mu},\overline{\mu}]\times[\underline{\sigma},\overline{\sigma}], \quad \underline{\sigma}>0 \tag{22}$$

where

$$X_{(\mu,\sigma)}(\omega) = \sigma\omega+\mu. \tag{23}$$

This means that $X_{(\mu,\sigma)} \sim \mathcal{N}(\mu,\sigma^2)$ is a Gaussian random variable parameterized by (μ,σ). In particular, we have $X_{(0,1)} \sim \mathcal{N}(0,1)$ and $X_{(0,1)}(\omega) = \omega$.

On the other hand, the random set $\mathcal{X}$ is generated by the set-valued map

$$\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\}. \tag{24}$$

In this case, $\mathcal{X}(\omega)$ is an interval $[\underline{\mathcal{X}}(\omega), \overline{\mathcal{X}}(\omega)]$ with lower bound

$$\underline{\mathcal{X}}(\omega) = \inf_{\substack{\mu\in[\underline{\mu},\overline{\mu}]\\ \sigma\in[\underline{\sigma},\overline{\sigma}]}} X_{(\mu,\sigma)}(\omega) = \begin{cases} \overline{\sigma}\omega + \underline{\mu} & \omega < 0, \\ \underline{\sigma}\omega + \underline{\mu} & \omega \geq 0, \end{cases} \tag{25}$$

and upper bound

$$\overline{\mathcal{X}}(\omega) = \sup_{\substack{\mu\in[\underline{\mu},\overline{\mu}]\\ \sigma\in[\underline{\sigma},\overline{\sigma}]}} X_{(\mu,\sigma)}(\omega) = \begin{cases} \underline{\sigma}\omega + \overline{\mu} & \omega < 0, \\ \overline{\sigma}\omega + \overline{\mu} & \omega \geq 0. \end{cases} \tag{26}$$

As a specific example, we determine the lower and upper probabilities $\underline{P}(B)$ and $\overline{P}(B)$ given $\Lambda = [-0.5,2] \times [1,2]$ and $B = [1,2.5]$. For the family $\{X_{(\mu,\sigma)}\}_{(\mu,\sigma)\in\Lambda}$ of random variables we get

$$\begin{aligned} \underline{P}(B) &= \inf_{(\mu,\sigma)\in\Lambda} P(X_{(\mu,\sigma)} \in B) = P(X_{(-0.5,1)} \in B) \\ &= 0.065457, \\ \overline{P}(B) &= \sup_{(\mu,\sigma)\in\Lambda} P(X_{(\mu,\sigma)} \in B) = P(X_{(1.75,1)} \in B) \\ &= 0.546745, \end{aligned} \tag{27}$$

cf. Fig. 1 where the probability $P(X_{(\mu,\sigma)} \in [1,2])$ on Λ is visualized as a contour plot. The maximum probability is achieved at parameter values $(\mu,\sigma) = (1.75,1)$ ($\triangle$) and the minimum probability at $(-0.5,1)$ ($\triangledown$).

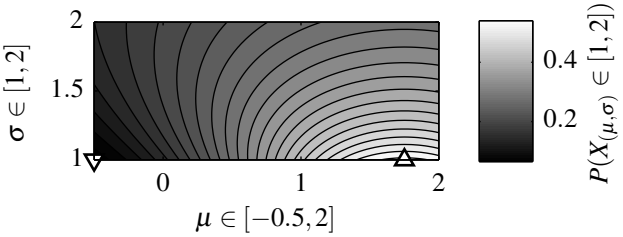

Figure 1: Set $\Lambda = [-0.5,2] \times [1,2]$ and probabilities $P(X_{(\mu,\sigma)} \in [1,2])$ visualized as a contour plot.

In Fig. 2 the random set $\mathcal{X}$ (gray area) corresponding to the above family of random variables, the bounds $\underline{\mathcal{X}}$, $\overline{\mathcal{X}}$, a single random variable $X_{(1.5,1.3)} \in \{X_{(\mu,\sigma)}\}_{(\mu,\sigma)\in\Lambda}$ and the focal set $\mathcal{X}(\omega)$ at $\omega = 1$ are depicted.

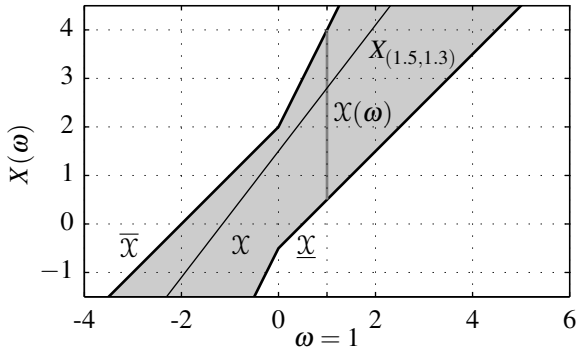

Figure 2: Random set $\mathcal{X}$, bounds $\underline{\mathcal{X}}$ and $\overline{\mathcal{X}}$, a single random variable $X_{(1.5,1.3)}$ and a focal set $\mathcal{X}(\omega)$.

To compute the lower and upper probabilities $\underset{\sim}{P}(B)$ and $\widetilde{P}(B)$ we need the lower and upper inverses. The lower inverse $\mathcal{X}_-(B)$ is the empty set, because there are no focal sets $\mathcal{X}(\omega)$ which are subsets of $B = [1,2.5]$. The upper inverse $\mathcal{X}^-(B)$ is the interval $[-1,3]$, cf. Fig. 3 where the random set $\mathcal{X}$ and the set B are depicted. In addition, some of the focal sets $\mathcal{X}(\omega)$, $\omega \in \mathcal{X}^-(B)$, with non-empty intersection with B are visualized as vertical lines. For comparison, the random variables $X_{(-0.5,1)}$ and $X_{(1.75,1)}$ resulting in $\underline{P}(B)$ and $\overline{P}(B)$ are plotted as well.

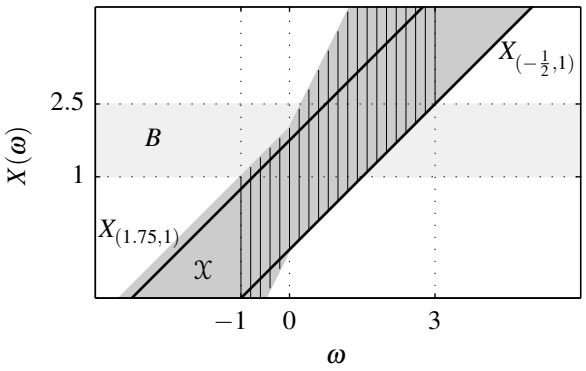

Figure 3: Set $B = [1,2.5]$, random set $\mathcal{X}$, random variables $X_{(1.75,1)}$ and $X_{(-\frac{1}{2},1)}$ resulting in $\underline{P}(B)$ and $\overline{P}(B)$. The vertical lines are focal sets with non-empty intersection with B leading to $\widetilde{P}(B)$.

Then the lower and upper probabilities are easily obtained:

$$\begin{aligned} \underset{\sim}{P}(B) &= \mathsf{m}(\mathcal{X}_-(B)) = \mathsf{m}(\varnothing) = 0 \\ &\leq 0.065457 = \underline{P}(B), \end{aligned} \tag{28}$$

$$\begin{aligned} \widetilde{P}(B) &= \mathsf{m}(\mathcal{X}^-(B)) = \mathsf{m}([-1,3]) = \Phi(3) - \Phi(-1) \\ &= 0.839994 \geq 0.546745 = \overline{P}(B) \end{aligned} \tag{29}$$

where $\Phi(\omega) = \mathsf{m}((-\infty,\omega])$ denotes the Gaussian cumulative distribution function.

In the following theorem we present special cases where $\underset{\sim}{P}(B) = \underline{P}(B)$ and/or $\widetilde{P}(B) = \overline{P}(B)$ holds for a fixed Borel set B. Since we are interested in probabilities $P(g(X) \le 0)$, we focus on sets B of type $(-\infty, b]$.

Theorem 3 *Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of random variables, Λ a compact subset of a metric space and assume that the maps $\lambda \to X_\lambda(\omega)$ are continuous for each fixed $\omega \in \Omega$. Further let $\mathcal{X}(\omega) = \{X_\lambda(\omega) : \lambda \in \Lambda\}$ and denote the lower and upper bounds by*

$$\underline{\mathcal{X}}(\omega) = \min \mathcal{X}(\omega) \quad \textit{and} \quad \overline{\mathcal{X}}(\omega) = \max \mathcal{X}(\omega). \tag{30}$$

(a) *If there is a $\lambda \in \Lambda$ such that $\mathbb{1}_{X_\lambda(\omega)\in B} = \underline{\mathbb{1}}_B(\omega)$ m-almost everywhere, then we have $\underline{P}(B) = \underset{\sim}{P}(B)$.*

If there is a $\lambda \in \Lambda$ such that $\mathbb{1}_{X_\lambda(\omega)\in B} = \overline{\mathbb{1}}_B$ m-almost everywhere, then we have $\overline{P}(B) = \widetilde{P}(B)$.

(b) *Let $B = (-\infty, b]$. If there is a $\lambda^* \in \Lambda$ such that $X_{\lambda^*}^{-1}(B) = \overline{\mathcal{X}}^{-1}(B)$, then we have $\underline{P}(B) = \underset{\sim}{P}(B)$.*

If there is a $\lambda_* \in \Lambda$ such that $X_{\lambda_*}^{-1}(B) = \underline{\mathcal{X}}^{-1}(B)$, then we have $\overline{P}(B) = \widetilde{P}(B)$.

(c) *Let $B = (-\infty, b]$. If there is a $\lambda^* \in \Lambda$ such that $X_{\lambda^*} = \overline{\mathcal{X}}$, then we have $\underline{P}(B) = \underset{\sim}{P}(B)$.*

If there is a $\lambda_* \in \Lambda$ such that $X_{\lambda_*} = \underline{\mathcal{X}}$, then we have $\overline{P}(B) = \widetilde{P}(B)$.

(d) *Let $B = (-\infty, b]$, $(\Omega, \Sigma, \mathsf{m}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathsf{m})$ and $\Lambda \times \mathbb{R} \to \mathbb{R} : (\lambda, \omega) \to X_\lambda(\omega)$ continuous and strictly monotonically increasing (decreasing) in the ω-direction, then $\underline{P}(B) = \underset{\sim}{P}(B)$ and $\overline{P}(B) = \widetilde{P}(B)$.*

Proof. (a) It follows directly from the Eqs. (16) and (17) in the proof of Theorem 2.

For the further proof we use that $\mathcal{X}_-(B) = \overline{\mathcal{X}}^{-1}(B)$ and $\mathcal{X}^-(B) = \underline{\mathcal{X}}^{-1}(B)$ holds for a set of the form $B = (\infty, b]$.

(b) Let $B = (-\infty, b]$. Then

$$\begin{aligned} \underset{\sim}{P}(B) &= \mathsf{m}(\mathcal{X}_-(B)) = \mathsf{m}(\overline{\mathcal{X}}^{-1}(B)) = \mathsf{m}(X_{\lambda^*}^{-1}(B)) \\ &= \int_\Omega \mathbb{1}_{X_{\lambda^*}(\omega)\in B}\, \mathsf{dm}(\omega) \ge \underline{P}(B) \end{aligned} \tag{31}$$

and

$$\begin{aligned} \widetilde{P}(B) &= \mathsf{m}(\mathcal{X}^-(B)) = \mathsf{m}(\underline{\mathcal{X}}^{-1}(B)) = \mathsf{m}(X_{\lambda_*}^{-1}(B)) \\ &= \int_\Omega \mathbb{1}_{X_{\lambda_*}(\omega)\in B}\, \mathsf{dm}(\omega) \le \overline{P}(B). \end{aligned} \tag{32}$$

(c) It is a special case of (b). E.g., if there is a $\lambda_* \in \Lambda$ such that $X_{\lambda_*} = \underline{\mathcal{X}}$, then it follows that $\underline{\mathcal{X}}^{-1}(B) = X_{\lambda_*}^{-1}(B)$.

(d) Let $B = (-\infty, b]$ and $(\lambda, \omega) \to X_\lambda(\omega)$ continuous and strictly monotonically increasing in the ω-direction. In this case the bound $\underline{\mathcal{X}}$ is continuous and strictly monotonically increasing in ω. (Assume $\underline{\mathcal{X}}(\omega_1) \ge \underline{\mathcal{X}}(\omega_2)$ for $\omega_1 < \omega_2$. Then it follows from Eq. (30) that there is a λ such that $\underline{\mathcal{X}}(\omega_2) = X_\lambda(\omega_2)$. This leads to the contradiction $X_\lambda(\omega_1) \ge \underline{\mathcal{X}}(\omega_1) \ge X_\lambda(\omega_2)$.)

There are three cases:

(i) Case $b < \underline{\mathcal{X}}$. Then $\underline{\mathcal{X}}^{-1}(B) = \varnothing$, $\widetilde{P}(B) = \mathsf{m}(\varnothing) = 0$ and $\overline{P}(B) = 0$.

(ii) Case $b > \underline{\mathcal{X}}$. Then $\underline{\mathcal{X}}^{-1}(B) = \mathbb{R}$, $\widetilde{P}(B) = \mathsf{m}(\mathbb{R}) = 1$. Further, $\overline{P}(B) = 1$. Indeed, take any $\overline{\omega} \in \mathbb{R}$. Then there is $\lambda \in \Lambda$ such that $\underline{\mathcal{X}}(\overline{\omega}) \le X_\lambda(\overline{\omega}) < b$ and by strict monotonicity $\underline{\mathcal{X}}(\omega) \le X_\lambda(\omega) < b$ for all $\omega \le \overline{\omega}$. This implies $\overline{P}(B) \ge \mathsf{m}((\infty, \overline{\omega}])$. Since $\overline{\omega}$ is arbitrary, we get $\overline{P}(B) = 1$ for $\overline{\omega} \to \infty$.

(iii) Case $b \in \underline{\mathcal{X}}(\mathbb{R})$. Then there is an $\omega^* \in \mathbb{R}$ such that $\underline{\mathcal{X}}(\omega^*) = b$.

In case (iii) we have on the one hand

$$\underline{\mathcal{X}}((-\infty, \omega^*]) \subseteq (-\infty, b] \tag{33}$$

because of the strict monotonicity of $\underline{\mathcal{X}}$ and on the other hand for the complement of $(-\infty, \omega^*]$

$$\underline{\mathcal{X}}((\omega^*, \infty)) \cap (-\infty, b] = \varnothing \tag{34}$$

which means that $\underline{\mathcal{X}}^{-1}((-\infty, b]) = (-\infty, \omega^*]$.

Further there is a $\lambda_* \in \Lambda$ such that $\underline{\mathcal{X}}(\omega^*) = b = X_{\lambda_*}(\omega^*)$ because of Eq. (30). This and the monotonicity of $X_{\lambda_*} \ge \underline{\mathcal{X}}$ leads to

$$\begin{aligned} &X_{\lambda_*}((-\infty, \omega^*]) \subseteq (-\infty, b], \\ &X_{\lambda_*}((\omega^*, \infty)) \cap (-\infty, b] = \varnothing, \\ &X_{\lambda_*}^{-1}((-\infty, b]) = (-\infty, \omega^*] \end{aligned} \tag{35}$$

and this in turn leads to the assumption

$$X_{\lambda_*}^{-1}((-\infty, b]) = \underline{\mathcal{X}}^{-1}((-\infty, b]) \tag{36}$$

of case (b). In particular, we get $\omega^* = \underline{\mathcal{X}}^{-1}(b) = X_{\lambda_*}^{-1}(b)$ and $\overline{P}(B) = \widetilde{P}(B) = \mathsf{m}((-\infty, \omega^*])$.

By the same arguments one can prove for the lower probabilities that $\underline{P}(B) = \underset{\sim}{P}(B) = \mathsf{m}((-\infty, \omega_*])$ with $\omega_* = \overline{\mathcal{X}}^{-1}(b) = X_{\lambda^*}^{-1}(b)$ and the results for decreasing functions $(\lambda, \omega) \to X_\lambda(\omega)$ as well. □

Remark on Theorem 3c. If there is a $\lambda_* \in \Lambda$ such that $X_{\lambda_*} = \underline{\mathcal{X}}$, then we have $F_{X_{\lambda_*}} = F_{\underline{\mathcal{X}}}$ for the cumulative distribution functions and therefore

$$\widetilde{P}((-\infty, b])) = F_{\underline{\mathcal{X}}}(b) = F_{X_{\lambda_*}}(b) = \overline{P}((-\infty, b]). \tag{37}$$

The relation $F_{X_{\lambda_*}} = F_{\underline{\mathcal{X}}}$ means in the notion of p-boxes [7], that one of the distribution functions F_{X_λ}, $\lambda \in \Lambda$, coincides with the upper envelope $\overline{F} = F_{\underline{\mathcal{X}}}$, cf. [7], of the p-box.

Example 2 We continue with Example 1. Here we want to compute the lower and upper probabilities for the set $B = (-\infty, b]$.

Since $(\mu, \sigma, \omega) \to X_{(\mu,\sigma)}(\omega) = \sigma\omega + \mu$, $\sigma > 0$, is continuous and a strictly monotonically increasing function in ω, we can apply Theorem 3 (d).

First we determine the inverses of $\underline{\mathcal{X}}$ and $\overline{\mathcal{X}}$ (cf. Eqs. (25) and (26)):

$$\underline{\mathcal{X}}^{-1}(x) = \begin{cases} (x-\underline{\mu})/\overline{\sigma} = X^{-1}_{(\underline{\mu},\overline{\sigma})}(x) & x < \underline{\mu}, \\ (x-\underline{\mu})/\underline{\sigma} = X^{-1}_{(\underline{\mu},\underline{\sigma})}(x) & x \geq \underline{\mu}, \end{cases} \tag{38}$$

and

$$\overline{\mathcal{X}}^{-1}(x) = \begin{cases} (x-\overline{\mu})/\underline{\sigma} = X^{-1}_{(\overline{\mu},\underline{\sigma})}(x) & x < \overline{\mu}, \\ (x-\overline{\mu})/\overline{\sigma} = X^{-1}_{(\overline{\mu},\overline{\sigma})}(x) & x \geq \overline{\mu}. \end{cases} \tag{39}$$

Now let again $\Lambda = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}] = [-0.5, 2] \times [1, 2]$ and $B = (-\infty, b] = (-\infty, 2.5]$. Then $2.5 \geq \overline{\mu} \geq \underline{\mu}$ which means that we have to take the second parts of Eqs. (38) and (39) to determine the lower and upper probabilities:

$$\begin{aligned} \underline{P}((-\infty, b]) &= \underline{P}((-\infty, b]) = \mathsf{m}((-\infty, \overline{\mathcal{X}}^{-1}(b)]) \qquad (40) \\ &= \mathsf{m}((-\infty, \omega_*]) = \Phi(X^{-1}_{(2,2)}(2.5)) = \Phi(0.25) \\ &= 0.598706, \\ \widetilde{P}((-\infty, b]) &= \overline{P}((-\infty, b]) = \mathsf{m}((-\infty, \underline{\mathcal{X}}^{-1}(b)]) \qquad (41) \\ &= \mathsf{m}((-\infty, \omega^*]) = \Phi(X^{-1}_{(-0.5,1)}(2.5)) = \Phi(3) \\ &= 0.998650 \end{aligned}$$

with $b = 2.5$, $\omega_* = \overline{\mathcal{X}}^{-1}(b) = 0.25$ and $\omega^* = \underline{\mathcal{X}}^{-1}(b) = 3$, see also Fig. 4.

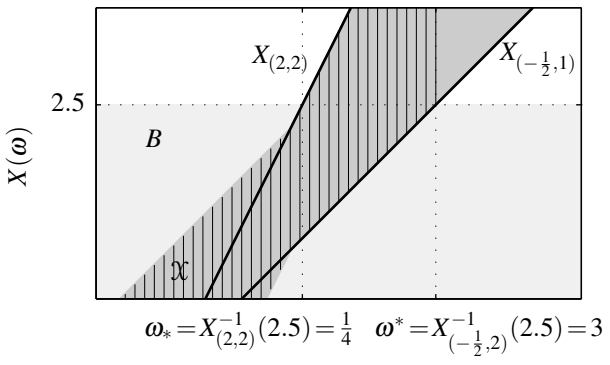

Figure 4: Set $B = (-\infty, b] = (-\infty, 2.5]$, random set $\mathcal{X}$, ω_*, ω^*, random variables $X_{(2,2)}$ and $X_{(-\frac{1}{2},1)}$ resulting in $\underline{P}(B)$ and $\overline{P}(B)$. The vertical lines are focal sets with non-empty intersection with B leading to $\widetilde{P}(B) = \overline{P}(B)$.

4 Numerical Example

As a simple engineering example we consider a beam of length $L = 3$ m supported on both ends and additionally bedded on a spring, cf. Fig. 5. The values of the beam rigidity $EI = 1\,\text{kNm}^2$, of the elastic limit moment $M_{\text{yield}} = 21$ kNm and of the load $f(\xi) = q = 100$ kN/m are deterministic, but the value of the spring constant x (in our notation for the variables of the function g) is assumed to be uncertain.

The beam will fail in the case where the value of limit state function g depending on the spring constant x is less or equal to 0. This means that we are interested in the failure probability $P(g(X) \leq 0)$ where the random variable X describes the uncertainty of the spring constant.

In this example the limit state function g is given as

$$\begin{aligned} \mathrm{g}(x) &= M_{\text{yield}} - \max_{\xi \in [0,3]} |M(\xi, x)| \qquad (42) \\ &= M_{\text{yield}} - \frac{qL^2}{4} \max\left(\frac{(1-c(x))^2}{2}, c(x) - \frac{1}{2} \right) \end{aligned}$$

with $c(x) = 5x/(384EI/L^3 + 8x)$, see Fig. 5 and [9]. Obviously this function is cheap to evaluate. Nevertheless we will apply the strategies for handling time consuming functions g as described in the following sections.

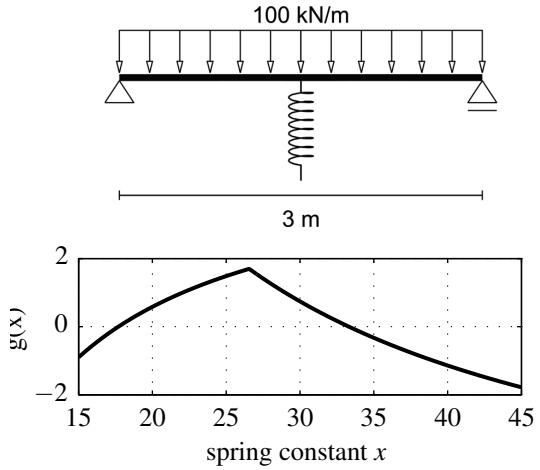

Figure 5: Beam bedded on a spring and deterministic limit state function g depending on the spring constant.

We model the uncertainty of the spring constant by a family $\{X_{(\mu,\sigma)}\}_{(\mu,\sigma)\in\Lambda}$ of random variables $X_{(\mu,\sigma)} \sim \mathcal{N}(\mu, \sigma^2)$ and, alternatively, by the induced random set $\mathcal{X}$. For this purpose, we simply continue with the set-up of Example 1. However, the set Λ is now given by

$$\Lambda = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}] = [20, 30] \times [0.5, 3]. \tag{43}$$

In the following sections the function g, the family of random variables as defined above and the corresponding random set will be used to exemplify the simulation techniques presented.

5 Simulation of Random Sets

First, let us recall how the propagation of random set data through a map is accomplished. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a continuous function and $\mathcal{X}(\omega)$, $\omega \in \Omega$ be a random set. The image set $\mathcal{G}(\omega) = g(\mathcal{X}(\omega))$, $\omega \in \Omega$, where $g(\mathcal{X}(\omega))$ is the set of values the function $y = g(x)$ attains when x ranges in $\mathcal{X}(\omega)$ is a random set. Computational aspects of random sets have been discussed at many places, see e.g. the references in [3]. In our context, the principles of Monte Carlo simulation of random sets of the form $g(\mathcal{X}(\omega)) = g(\{X_\lambda(\omega) : \lambda \in \Lambda\})$ have to be explained.

We assume that Λ is a compact subset of a metric space and that the maps $\lambda \to X_\lambda(\omega)$ are continuous for each fixed $\omega \in \Omega$. Then $\mathcal{G}(\omega) = g(\mathcal{X}(\omega))$ is contained in a random interval $[\underline{\mathcal{G}}(\omega), \overline{\mathcal{G}}(\omega)]$, where $\underline{\mathcal{G}}(\omega) = \min g(\mathcal{X}(\omega))$ and $\overline{\mathcal{G}}(\omega) = \max g(\mathcal{X}(\omega))$. Suppose we wish to compute its upper and lower distribution functions $\overline{F}(y)$ and $\underline{F}(y)$. Note that

$$\overline{F}(y) = P\big((-\infty, y] \cap [\underline{\mathcal{G}}, \overline{\mathcal{G}}] \neq \varnothing\big) = P\big(\underline{\mathcal{G}} \leq y\big) = F_{\underline{\mathcal{G}}}(y),$$

the cumulative distribution function of the random variable $\underline{\mathcal{G}}$. Similarly,

$$\underline{F}(y) = P\big([\underline{\mathcal{G}}, \overline{\mathcal{G}}] \subset (-\infty, y]\big) = P\big(\overline{\mathcal{G}} \leq y\big) = F_{\overline{\mathcal{G}}}(y).$$

Recall that $g(\mathcal{X}(\omega)) = \{g(X_\lambda(\omega))) : \lambda \in \Lambda\}$. Thus, in order to compute $\underline{\mathcal{G}}(\omega)$ and $\overline{\mathcal{G}}(\omega)$, an optimization problem has to be solved that determines the smallest and the largest value of the set $\{g(X_\lambda(\omega))) : \lambda \in \Lambda\}$. This leads to the following algorithm for computing the upper distribution function $\overline{F}(x)$.

- Generate a sample $\omega_1, \ldots, \omega_{N_{\text{samp}}}$ random elements of Ω, distributed according to m .
- For each ω_n, estimate $\underline{\mathcal{G}}(\omega_n) = \min g(X_\lambda(\omega_n))$ by minimization with respect to $\lambda \in \Lambda$.
- The empirical distribution function of the sample $\{\underline{\mathcal{G}}(\omega_n) : n = 1, \ldots, N_{\text{samp}}\}$ is an approximation to $\overline{F}(y)$.

In order to compute the respective minima and maxima, the parameter set should be discretized into $\lambda_1, \ldots, \lambda_{N_{\text{grid}}}$. The algorithm requires $N_{\text{grid}} \cdot N_{\text{samp}}$ evaluations of the function g. Generally, this is too expensive for large scale applications. Computational cost can be saved by suitably approximating the input-output function g by a surrogate model. For such an approximation, two levels are at hand:

$$\Omega \xrightarrow{X_\lambda} \mathbb{R}^n \xrightarrow{g} \mathbb{R}.$$

There are two possibilities to construct a surrogate model: either by a surrogate model $\widetilde{g}$ of the map $g : \mathbb{R}^n \to \mathbb{R}$ or by a family of stochastic surrogate models of the maps $\Omega \to g \circ X_\lambda$. Both approaches start with a set of collocation points x_j, $j = 1, \ldots, N_{\text{coll}}$ in $\mathbb{R}^n$, together with the corresponding function values $y_j = g(x_j)$. This requires N_{coll} evaluations of the input-output function g. In the first approach, the functions $\widetilde{g} \circ X_\lambda$ have to be simulated for λ belonging to a grid of representative parameter values λ_i, $i = 1, \ldots, N_{\text{grid}}$. Each X_{λ_i} has a different probability distribution. If a Monte Carlo sample of size N_{samp} is desired, this still requires $N_{\text{grid}} \cdot N_{\text{samp}}$ evaluations of the function $\widetilde{g}$. We will show below that reweighting technique can reduce the computational cost to N_{samp} evaluations.

We first discuss the second approach. Here the collocation points are pulled back to Ω as follows: For each λ_i and x_j, define a collocation point in Ω by $\omega_{ij} = X_{\lambda_i}^{-1}(x_j)$. Clearly, $y_j = g(X_{\lambda_i}(\omega_{ij}))$ for every i. Fitting a surrogate model $\widetilde{g}_i$ for each i, based on the data (ω_{ij}, g_j), $j = 1, \ldots, N_{\text{coll}}$ is computationally inexpensive. Typically, when $\Omega = \mathbb{R}^m$ and the measure $\mathsf{dm}(\omega)$ is absolutely continuous with respect to Lebesgue measure, one may use orthogonal polynomials with respect to the measure $\mathsf{dm}(\omega)$ (Hermite expansion in the Gaussian case) and then compute the coefficients by weighted regression through the data.

At fixed ω, the lower bound $\underline{\mathcal{G}}(\omega)$ of the focal set $\mathcal{G}(\omega)$ can simply be estimated by the smallest value among the $\widetilde{g}_i(\omega)$, $i = 1, \ldots, N_{\text{grid}}$. Repeating this procedure for a sample $\omega_1, \ldots, \omega_{N_{\text{samp}}}$ produces a Monte Carlo sample $\{\underline{\mathcal{G}}(\omega_n) : n = 1, \ldots, N_{\text{samp}}\}$ which can be used to estimate the desired upper distribution function $\overline{F}(y)$, and similarly for the lower distribution function $\underline{F}(y)$.

This approach requires N_{coll} evaluations of the expensive full model and N_{samp} evaluations of the inexpensive surrogate model. Details of the procedure can be found in [16]; further information on the construction of stochastic surrogate models is in [10].

Remarks. (a) In case X_λ are Gaussian variables with $\lambda = (\mu, \sigma)$ and $\Omega = \mathbb{R}$ with the standard Gaussian density, we simply have $X_\lambda(\omega) = \mu + \sigma\omega$ and $X_{\lambda_i}^{-1}(x) = (x - \mu)/\sigma$. The same procedure can be applied to a non-Gaussian family X_λ by transforming it to standard Gaussian space, i.e., $X_\lambda(\omega) = F_\lambda^{-1}(\Phi(\omega))$ where F_λ and Φ denote the cumulative distribution functions of X_λ and of a standard normal variable, respectively.

(b) Some indications on multivariate families X_λ are in order. We consider the case of an n-dimensional Gaussian variable $X_\lambda \sim \mathcal{N}\big(\mu(\lambda), \mathbf{S}(\lambda)\big)$ with mean $\mu(\lambda)$ and covariance $\mathbf{S}(\lambda)$, both assumed to depend continuously on a possibly multidimensional parameter λ. Performing the Cholesky factorization $\mathbf{S}(\lambda) = \mathbf{C}(\lambda)\mathbf{C}(\lambda)^\mathsf{T}$, the random variables X_λ can be realized as $X_\lambda = \mathbf{C}(\lambda)Y$ where Y is an n-dimensional standard Gaussian variable. The procedure outlined above can be applied in the same way, employing n-dimensional Hermite polynomials.

In this framework, finite dimensional discretizations of Gaussian random fields can be accommodated as well, either using the Cholesky factorization or – equivalently – a truncated Karhunen-Loève expansion.

Example 3 We continue with our engineering example. Due to its small size, we may use the full model without constructing a surrogate model.

Let the grid points (μ_i, σ_j) with

$$\mu_i = 20, 21, \ldots, 30 \quad \text{and} \quad \sigma_j = 0.5, 1, 1.5, \ldots, 3 \tag{44}$$

define a grid on $\Lambda = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}] = [20, 30] \times [0.5, 3]$. Then a focal set $[\underline{\mathcal{G}}(\omega), \overline{\underline{\mathcal{G}}}(\omega)]$ of the random set $\mathcal{G}$ at ω is approximated by

$$\begin{aligned} \underline{\mathcal{G}}(\omega) &\approx \min_{i,j} g \circ X_{(\mu_i,\sigma_j)}(\omega), \\ \overline{\mathcal{G}}(\omega) &\approx \max_{i,j} g \circ X_{(\mu_i,\sigma_j)}(\omega). \end{aligned} \tag{45}$$

We approximate the upper probability of failure of the beam by means of Monte Carlo simulation:

$$\begin{aligned} \widetilde{P}(g \le 0) &= \int_{\mathbb{R}} \mathbb{1}_{\mathcal{G}(\omega) \cap (-\infty, 0] \neq \varnothing} \, \mathsf{dm}(\omega) \\ &= \int_{\mathbb{R}} \mathbb{1}_{\underline{\mathcal{G}}(\omega) \le 0} \, \mathsf{dm}(\omega) \\ &\approx \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{\underline{\mathcal{G}}(\omega_k) \le 0} \cdot \frac{1}{N_{\text{samp}}} = 0.358. \end{aligned} \tag{46}$$

where $\omega_1, \ldots, \omega_{N_{\text{samp}}}$ is a standard normally distributed sample.

In Fig. 6 the random set $\mathcal{G}$ and one of the functions which are generating $\mathcal{G}$, namely $g \circ X_{(24,2)}$, are depicted. Further in Fig. 7 a sample of focal sets $\mathcal{G}(\omega_k)$, $k = 1, \ldots, N_{\text{samp}}$, is visualized. Counting the focal sets with non-empty intersection with $(-\infty, 0]$ leads to the upper probability $\widetilde{P}(g \le 0)$.

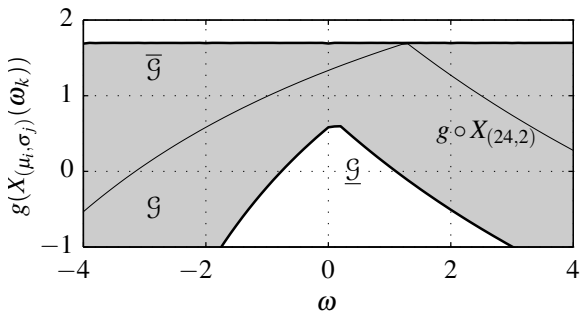

Figure 6: Random set $\mathcal{G}$, $g \circ X_{(24,2)}$ (one of the transformations of g), lower and upper bounds $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$.

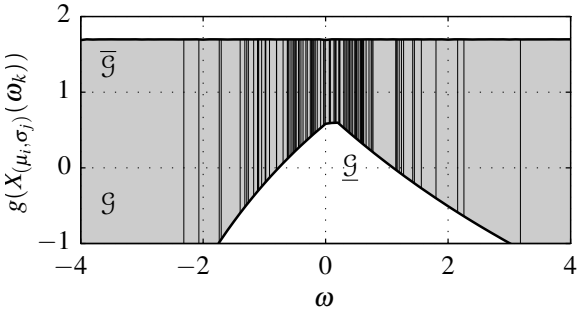

Figure 7: Random set $\mathcal{G}$ and focal sets $\mathcal{G}(\omega_k)$ for sample $\omega_1, \ldots, \omega_{N_{\text{samp}}}$ visualized as vertical lines.

6 Simulation of Families of Random Variables

In this section we show how to simulate families of random variables and how to save computational cost applying sample reweighting techniques.

We shortly present the ideas of resampling: Let v be a measurable function (later on an indicator function $\mathbb{1}_{g(x) \le y}$), $f_X > 0$ the strictly positive density function of a random variable X and $f_Y > 0$ the strictly positive density of a random variable Y. Then, for the random variable X we have the approximation

$$\int_{\mathbb{R}} v(x) f_X(x) \, \mathrm{d}x \approx \sum_{k=1}^{N_{\text{samp}}} v(x_k) \cdot \frac{1}{N_{\text{samp}}} \tag{47}$$

of the integral where $x_1, \ldots, x_{N_{\text{samp}}}$ is a sample distributed according to X together with weights $1/N_{\text{samp}}$. Similarly, for Y we get

$$\int_{\mathbb{R}} v(x) f_Y(x) \, \mathrm{d}x \approx \sum_{k=1}^{N_{\text{samp}}} v(y_k) \cdot \frac{1}{N_{\text{samp}}} \tag{48}$$

where now $y_1, \ldots, y_{N_{\text{samp}}}$ is a sample distributed as Y with weights $1/N_{\text{samp}}$. In this version, replacing the density function f_X by f_Y needs new samples and new function evaluations $v(y_k)$.

Alternatively, applying sample reweighting, we have

$$\begin{aligned} \int_{\mathbb{R}} v(x) f_Y(x) \, \mathrm{d}x &= \int_{\mathbb{R}} v(x) \cdot \frac{f_Y(x)}{f_X(x)} \cdot f_X(x) \, \mathrm{d}x \\ &\approx \sum_{k=1}^{N_{\text{samp}}} v(x_k) \cdot \frac{f_Y(x_k)}{f_X(x_k)} \frac{1}{N_{\text{samp}}} \end{aligned} \tag{49}$$

using the original sample $x_1, \ldots, x_{N_{\text{samp}}}$, but now with new weights

$$w_k := \frac{f_Y(x_k)}{f_X(x_k)} \frac{1}{N_{\text{samp}}} \tag{50}$$

instead of the uniform weights $1/N_{\text{samp}}$.

Our goal is to approximate probabilities $P(g(X_\lambda) \leq y)$ by means of Monte Carlo simulation using only one sample for all X_λ, $\lambda \in \Lambda$.

As a first step, we start with the generation of a sample $x_1, \ldots, x_{N_{\text{samp}}}$. This sample may be distributed as one of our random variables X_λ, $\lambda \in \Lambda$. But a better choice is a "basic" distribution covering a greater range than a distribution of a single X_λ, $\lambda \in \Lambda$, does. In our example with the family $\{X_{(\mu,\sigma)}\}_{(\mu,\sigma)\in\Lambda}$ of Gaussian random variables, this can be achieved by using an appropriate high variance $\sigma_*^2 > \overline{\sigma}^2$ and $\mu_* = (\underline{\mu} + \overline{\mu})/2$ for generating a Gaussian sample $x_1, \ldots, x_{N_{\text{samp}}} \sim \mathcal{N}(\mu_*, \sigma_*^2)$. In general we say that this "basic" sample is distributed as a random variable X_*.

As a second step, we compute all function values $g(x_k)$, $k = 1, \ldots, N_{\text{samp}}$, of the limit state function, either directly evaluating g or using a surrogate model $\tilde{g}$.

As a third step, we have to perform a reweighting of the sample generated above, since we need samples distributed according to certain random variables X_λ. These weights for a given X_λ are obtained by

$$w_k(\lambda) = \frac{f_{X_\lambda}(x_k)}{f_{X_*}(x_k)} \frac{1}{N_{\text{samp}}}. \tag{51}$$

Now we are able to approximately compute probabilities $P(g(X_\lambda) \leq y)$ for different random variables X_λ without additional function evaluations of g:

$$P(g(X_\lambda) \leq y) \approx \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(x_k)\leq y} \cdot w_k(\lambda). \tag{52}$$

For the computation of the upper/lower probabilities we use a grid of representative parameter values λ_i as mentioned in the previous section, estimate the probabilities $P(g(X_{\lambda_i}) \leq y)$ at the grid points λ_i by means of Monte Carlo simulation as in Eq. (52) and take the maximum/minimum value as an approximation:

$$\begin{aligned} \overline{P}(g \leq y) &= \sup_{\lambda\in\Lambda} P(g(X_\lambda) \leq y) \approx \max_{i=1,\ldots,N_{\text{grid}}} P(g(X_{\lambda_i}) \leq y) \\ &\approx \max_{i=1,\ldots,N_{\text{grid}}} \sum_{k=1}^{N} \mathbb{1}_{g(x_k)\leq y} \cdot w_k(\lambda_i), \end{aligned} \tag{53}$$

$$\underline{P}(g \leq y) \approx \min_{i=1,\ldots,N_{\text{grid}}} \sum_{k=1}^{N} \mathbb{1}_{g(x_k)\leq y} \cdot w_k(\lambda_i). \tag{54}$$

Example 4 Again, we continue with the engineering example.

We approximately compute the failure probability $P(g(X_{(\mu,\sigma)}) \leq 0)$ of the beam for a fixed pair $(\mu,\sigma) \in \Lambda$ using either Monte Carlo simulation in the space of the variables of the limit state function g, Eq. (57), or in the standard normal space, Eq. (56):

$$P(g(X_{(\mu,\sigma)}) \leq 0) = \int_{\mathbb{R}} \mathbb{1}_{g(X_{(\mu,\sigma)}(\omega))\leq 0} \, \mathrm{dm}(\omega) \tag{55}$$

$$\approx \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(X_{(\mu,\sigma)}(\omega_k))\leq 0} \cdot w_k(\mu,\sigma) \tag{56}$$

$$\approx \sum_{k=1}^{N_{\text{samp}}} \mathbb{1}_{g(x_k)\leq 0} \cdot w_k(\mu,\sigma) \tag{57}$$

where

$$X_{(\mu,\sigma)}(\omega_k) = \sigma\omega_k + \mu = x_k \tag{58}$$

and in the reverse direction

$$\omega_k = X_{(\mu,\sigma)}^{-1}(x_k) = (x_k - \mu)/\sigma. \tag{59}$$

The weights are given by

$$w_k(\mu,\sigma) = \frac{f_{X_{(\mu,\sigma)}}(x_k)}{f_{X_*}(x_k)} \frac{1}{N_{\text{samp}}} \tag{60}$$

where the sample $x_1, \ldots, x_{N_{\text{samp}}}$ is distributed according $X_* \sim \mathcal{N}(25, 6^2)$ for $N_{\text{samp}} = 100\,000$.

In the next step we have to compute

$$\overline{P}(g \leq 0) = \sup_{(\mu,\sigma)\in\Lambda} P(g(X_{(\mu,\sigma)}) \leq 0) \tag{61}$$

which is approximated using grid points (μ_i, σ_j) with

$$\mu_i = 20, 21, \ldots, 30 \quad \text{and} \quad \sigma_j = 0.5, 1, 1.5, \ldots, 3. \tag{62}$$

The probabilities $P(g(X_{(\mu_i,\sigma_j)}) \leq 0)$ at these grid points are computed as in Eq. (52) and depicted in Fig. 8. Then we simply take the maximum of all these probabilities similar to Eq. (53) and obtain the result

$$\overline{P}(g \leq 0) \approx \max_{i,j} P(g(X_{(\mu_i,\sigma_j)}) \leq 0) \approx 0.221 \tag{63}$$

which is the upper probability of failure of the beam considered.

Comparing this result with the result in Section 5 it holds that $0.221 \approx \overline{P}(g \leq 0) \leq \widetilde{P}(g \leq 0) \approx 0.358$.

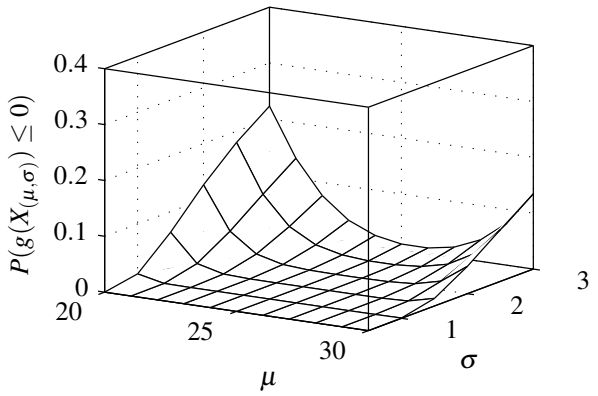

Figure 8: Failure probability $P(g(x_{(\mu,\sigma)}) \leq 0)$ computed at grid points (μ_i, σ_j).

7 Summary and Conclusions

We have discussed two interpretations of upper and lower probabilities, given a family of random variables. The random set approach is supported by the availability of a rich theory as well as various recent applications, e.g. [16, 17, 18, 19]. The approach based directly on the family of random variables has been favored e.g. in [2]. We have shown here that the latter approach gives tighter bounds, i.e., smaller probability intervals, in general. For both approaches, cost saving simulation methods have been presented. We hope that the paper stimulates further research into the computational aspects of imprecise probability.

References

[1] D. A. Alvarez. A Monte Carlo-based method for the estimation of lower and upper probabilities of events using infinite random sets of indexable type. *Fuzzy Sets and Systems*, 160:384–401, 2009.

[2] M. de Angelis, E. Patelli and M. Beer. Line Sampling for Assessing Structural Reliability with Imprecise Failure Probabilities. In: M. Beer, S.-K. Au, J.W. Hall (Eds.). *Vulnerability, Uncertainty, and Risk Quantification: Mitigation and Management*. CD-ROM Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, and 6th International Symposium on Uncertainty Modelling and Analysis. American Society of Civil Engineers, Reston VA 2014, pp. 915–924.

[3] M. Beer, S. Ferson and V. Kreinovich. Imprecise probabilities in engineering analysis. *Mechanical Systems and Signal Processing*, 37:4–29, 2013.

[4] A. Bernardini and F. Tonon. *Bounding Uncertainty in Civil Engineering - Theoretical Background*. Berlin: Springer-Verlag, 2010.

[5] A. Castaldo and M. Marinacci. Random correspondences as bundles of random variables. In: G. de Cooman, T. Fine, T. Seidenfeld (Eds.), *ISIPTA'01, Proceedings of the Second Symposium on Imprecise Probabilities and Their Applications*, Shaker Publ. BV, Maastricht 2001, pp. 77–82.

[6] S. Destercke, D. Dubois and E. Chojnacki. Relating practical representations of imprecise probabilities. In: G.de Cooman, J. Vejnarova and M. Zaffalon (Eds.), *ISIPTA '07, Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications.* Action M Agency, SIPTA, Prague 2007, pp. 155–164.

[7] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical report SAND2002-4015, Sandia National Laboratories, Albuquerque, NM, 2003.

[8] T. Fetz and M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering and System Safety*, 85:73–87, 2004.

[9] T. Fetz. Modelling uncertainties in limit state functions. *Int. J. of Approximate Reasoning*, 53(1):1–23, 2012.

[10] O. Le Maître and O. Knio. *Spectral methods for uncertainty quantification. With applications to computational fluid dynamics*. New York: Springer, 2010.

[11] E. Miranda, I. Couso and P. Gil. Random intervals as a model for imprecise information. *Fuzzy Sets and Systems*, 154(3):386–412, 2005.

[12] E. Miranda, I. Couso and P. Gil. Random sets as imprecise random variables. *Journal of Mathematical Analysis and Applications*, 307(1):32–47, 2005.

[13] I. Molchanov. *Theory of Random Sets*. Berlin: Springer-Verlag, 2005.

[14] R. I. Mullen and R. I. Muhanna. Probability bounds for the system response of non-linear structures. In G. Deodatis, B.R. Ellingwood and D.M. Frangopol (Eds.), *Safety, Reliability, Risk and Life-Cycle Performance of Structures & Infrastructures*. Taylor & Francis Group, London, 2013, pp. 499–502.

[15] M. Oberguggenberger. Engineering. In: T. Augustin, F. Coolen, G. de Cooman, M. Troffaes (Eds.), *Introduction to Imprecise Probabilities.* John Wiley & Sons Ltd, Chichester 2014, pp. 291–304.

[16] M. Oberguggenberger. Stochastic Response Surfaces, Interval Analysis, and the Reliability of Structures. In: M. Beer, S.-K. Au, J.W. Hall (Eds.). *Vulnerability, Uncertainty, and Risk Quantification: Mitigation and Management*. CD-ROM Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, and 6th International Symposium on Uncertainty Modelling and Analysis. American Society of Civil Engineers, Reston VA 2014, pp. 864–875.

[17] M. Oberguggenberger. Analysis and computation with hybrid random set stochastic models. *Structural Safety*, 52:233–243, 2015.

[18] B. Schmelzer. On solutions to stochastic differential equations with parameters modeled by random sets. *International Journal of Approximate Reasoning*, 51:1159–1171, 2010.

[19] B. Schmelzer. Set-valued assessments of solutions to stochastic differential equations with random set parameters. *Journal of Mathematical Analysis and Applications*, 400:425–438, 2013.

[20] F. Tonon, A. Bernardini and A. Mammino. Determination of parameters range in rock engineering by means of random set theory. *Reliability Engineering and System Safety*, 70:241–261, 2000.

[21] L. Utkin and S. Destercke. Computing expectations with continuous p-boxes: Univariate case. *Int. J. of Approximate Reasoning*, 50:778–798, 2009.

[22] R. Zhang. Structural reliability analysis with imprecise probability using interval importance sampling method. In G. Deodatis, B.R. Ellingwood and D.M. Frangopol (Eds.), *Safety, Reliability, Risk and Life-Cycle Performance of Structures & Infrastructures*. Taylor & Francis Group, London, 2013, pp. 563–570.