

# Clouds, $p$ -boxes, Fuzzy Sets, and Other Uncertainty Representations in Higher Dimensions

Martin Fuchs\*

## Abstract

Uncertainty modeling in real-life applications comprises some serious problems such as the curse of dimensionality and a lack of sufficient amount of statistical data.

In this paper we give a survey of methods for uncertainty handling and elaborate the latest progress towards real-life applications with respect to the problems that come with it. We compare different methods and highlight their relationships. We introduce intuitively the concept of potential clouds, our latest approach which successfully copes with both higher dimensions and incomplete information.

**Keywords:** uncertainty models, potential clouds, confidence regions, higher dimensions, incomplete information, reliability methods,  $p$ -boxes, Dempster-Shafer theory, fuzzy sets

## 1 Introduction

Among the major problems in real-life applications of uncertainty representations we have identified two particularly complicated ones. One concerns the dimensionality issue. High-dimensionality can cause computations to become very expensive, with an effort growing exponentially with the dimension in many cases. This phenomenon is famous as the curse of dimensionality [45]. Even given the full knowledge about a joint distribution the numerical computation of error probabilities may be very time consuming, if not impossible. Moreover, rigorous computation or (preferably tight) bounding of failure probabilities can only be done in very few cases because the space of possible scenarios is too large. In higher dimensions full probabilistic models need to estimate high-dimensional distributions for which rarely sufficient data are available. Frequently it is just the other way around, i.e., statistical data are scarce. This leads to the second issue which is incomplete, imprecise, or subjective information. Thus we can formulate our ultimate question for discussing an uncertainty method: How does the quality of the method respond

---

\*University of Vienna, Faculty of Mathematics, Nordbergstr. 15, 1090 Wien, Austria, E-mail: martin.fuchs@univie.ac.at, www.martin-fuchs.net

to a **lack of information** and to the **dimensionality** of the problem to solve? The dimension of a problem is determined by the number of uncertain variables involved. In some real-life design problems the dimension is low (say, smaller than about 5). In many problems, however, the dimension is significantly higher.

We will see that many methods exist for uncertainty modeling. Depending on the uncertainty information available one identifies the problem class and applies a suitable method for that class. It is just like choosing the appropriate tool from a toolbox. This point of view can be described as a **toolbox philosophy** [74]. This strategy to solve a problem is defined by the problem itself and the characteristics of the uncertainties involved. Thus different approaches to uncertainty modeling do not contradict each other, but rather constitute a mutually complementing framework.

The most general point of view describing scarce, vague, incomplete, or conflicting uncertainty information are **imprecise probabilities** [95]. This approach alludes to existing uncertainty models being not sufficiently general to handle all kinds of uncertainty, and it encourages to develop a **unified formulation** [97], opposed to the toolbox philosophy. Thus it rather aims at unification on a theoretical basis, whereas our focus is on applicability in real-life reliable engineering.

Uncertainty models in engineering applications are typically employed in the context of design optimization. An uncertainty method should enable to conduct a safety analysis for a given design and to weave this analysis into an optimization problem formulation as safety constraints towards finding a robust, optimal design point.

This paper presents a survey of conventional and modern approaches to uncertainty handling. For each method, the notation, the type of input information, and the basic concepts will be introduced. We will discuss the necessary assumptions, the rigor of results, and the sensitivity of the results to a lack of information. Typically, the more general a method is, the more expensive it becomes computationally, so we will also comment on computational effort, especially in higher dimensions. Eventually, we will highlight relationships between the presented methods and possible embedding in design optimization problems.

We start with a section on the basic principles used in uncertainty handling. Then we present the several different approaches to uncertainty handling: reliability methods,  $p$ -boxes, Bayesian methods, Dempster-Shafer theory, fuzzy sets, convex methods, and potential clouds. By means of the potential clouds formalism we present a reliable and tractable worst-case analysis for the given high-dimensional information. This approach enables to determine a nested collection of confidence regions parameterized by the confidence level  $\alpha$ , and has already been successfully applied to real-life engineering problems [31], [73] in 34 and 24 dimensions, respectively.

## 2 Basic principles

Throughout this study we assume familiarity with the basic principles of probability theory. In this section we introduce the notation and fundamental concepts which are the basis for classical methods of uncertainty modeling and of many modern methods as well.

We denote the **sample space** by  $\Omega$ , an  $n$ -dimensional **random vector** by  $\varepsilon : \Omega \rightarrow \mathbb{R}^n$ . A **random variable** is a 1-dimensional random vector which we denote by  $X$ . We denote the **probability** of the statement given as argument by  $\Pr$ . We denote the **expectation** of a random vector  $\varepsilon$  by  $\langle \varepsilon \rangle$ . We abbreviate the terms **probability density function** by PDF, and **cumulative distribution function** by CDF, respectively.

### 2.1 Reliability and failure

To employ uncertainty methods in design safety problems, we need to define failure probabilities  $p_f$  and reliability  $R$ . The **failure probability** of a fixed design is the probability that the random vector  $\varepsilon$  lies in a set  $\mathbb{F}$  of scenarios which lead to design failure. The **reliability** is the probability that the design will perform satisfactorily, i.e.,

$$R := \Pr(\varepsilon \notin \mathbb{F}) = 1 - \Pr(\varepsilon \in \mathbb{F}) = 1 - p_f, \quad (1)$$

so determining  $R$  and  $p_f$  are equivalent problems. A third important notion is that of a confidence region for  $\varepsilon$ . A set  $C_\alpha$  is a **confidence region** for the confidence level  $\alpha$  if

$$\Pr(\varepsilon \in C_\alpha) = \alpha. \quad (2)$$

The relation between confidence regions and failure probabilities can be seen as follows. Assume that we have determined a confidence region  $C_\alpha$  for the random vector  $\varepsilon$ , and  $C_\alpha$  does not contain a scenario which leads to design failure, i.e.,  $C_\alpha \cap \mathbb{F} = \emptyset$ . Then  $\Pr(C_\alpha \cup \mathbb{F}) = \Pr(C_\alpha) + \Pr(\mathbb{F}) \leq 1$ . Hence  $p_f = \Pr(\mathbb{F}) \leq 1 - \Pr(C_\alpha) = 1 - \alpha$ , the failure probability is at most  $1 - \alpha$ . For the reliability  $R = 1 - p_f$  we get  $R \geq \alpha$ .

### 2.2 Incomplete Information

To make use of probabilistic concepts one often assumes that for the random vector  $\varepsilon$  involved the joint distribution  $F$  is precisely known, provided by an objective source of information. In many design problems, the sources of information are merely subjective, provided by expert knowledge. Additionally, in higher dimensions joint distributions are rarely available, and the typically available distribution information consists of certain marginal distributions.

Often one simply fixes the CDF as normally distributed, arguing with the central limit theorem: a sufficiently large amount of statistical sample data justifies the

normal distribution assumption. The critical question is, what is sufficiently large in higher dimensions? The generalized Chebyshev inequality (3) gives rigorous bounds for the failure probability  $p_f = \Pr(\mathbb{F})$ , in case that  $\mathbb{F}(r) = \{\varepsilon \mid \|\varepsilon\|_2 \geq r\}$ ,  $r$  a constant radius. If the components of  $\varepsilon = (\varepsilon^1, \dots, \varepsilon^n)$  are uncorrelated, have mean 0 and variance 1, we get from [72]

$$p_f = \Pr(\mathbb{F}) \leq \min(1, \frac{n}{r^2}), \quad (3)$$

and this bound can be attained.

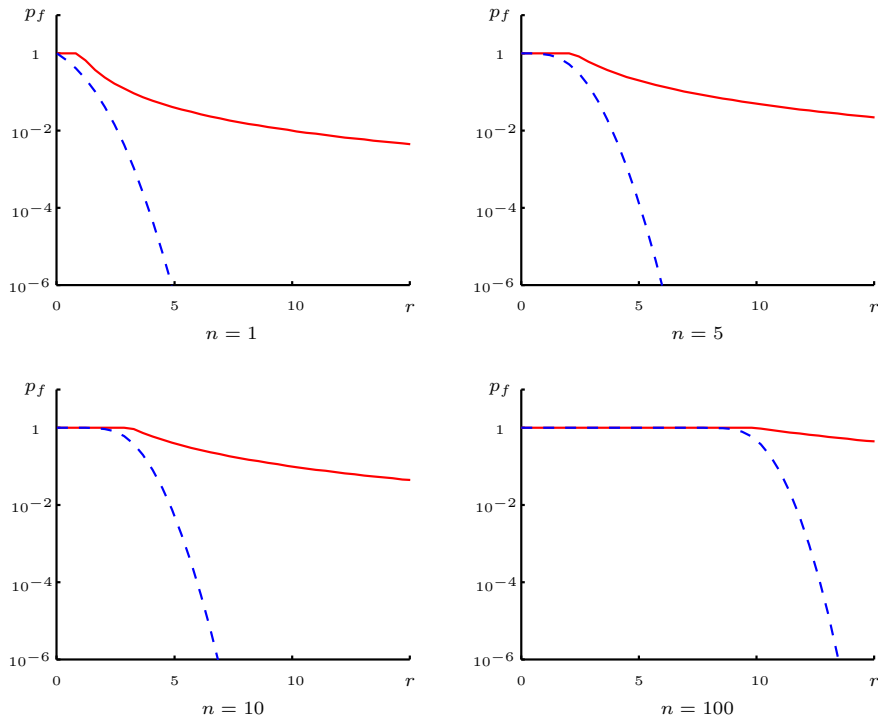


Figure 1: Failure probability  $p_f$  for the failure set  $\mathbb{F}(r)$  in different dimensions  $n$ , bounded from above by the Chebyshev inequality (solid line) and computed from a normal distribution (dashed line), respectively.

The failure probability bounds from (3) differ significantly from those of a normal distribution as shown in Figure 1. If we assume a multivariate normal distribution for  $\varepsilon$ , uncorrelated is equivalent to independent. The bounds for normal distribution assumption can then be computed from the  $\chi^2(n)$  distribution (i.e., a  $\chi^2$  distribution with  $n$  degrees of freedom). We see that the normal distribution assumption can be much too optimistic compared with the optimal worst-case bounds from (3).

An alternative justification of the normal distribution assumption is the **maximum entropy principle**, if the available information consists of mean and standard deviation only. The principle of maximum entropy originates from information theory [87], and is utilized in many fields of applications, cf., e.g., [34].

The intuitive meaning of entropy is: the larger the entropy the less information (relative to the uniformly distributed improper prior) is reflected in the probability measure with density  $\rho$ . In order to define a probability measure given incomplete information, the principle of maximum entropy consists in maximizing the entropy subject to constraints imposed by the information available. For example, in the case of given mean and standard deviation this ansatz leads to a normal distribution, in case of given interval information it leads to a uniform distribution assumption. Note that as soon as we employ the maximum entropy distribution as a probability measure we pretend to have more information than actually available. Hence critical underestimation of failure probabilities may show up.

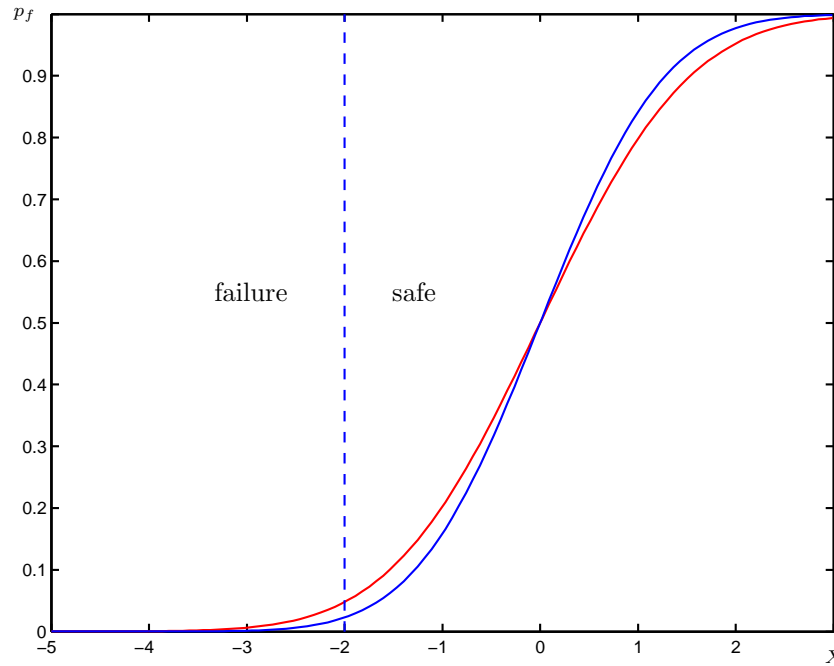


Figure 2: Small deviation in a distribution parameter (here 20% difference in the standard deviation of two normal distributions) can lead to critical underestimation of  $p_f$  for a random variable  $X$ . Here  $p_f$  is underestimated by the factor 2, if the design failure set was  $\mathbb{F} = \{X \mid X \leq -2\}$ .

In a nutshell, the concept of random variables and probability spaces enables one to derive rigorous statements about failure probabilities and reliability. But they require the probability measure to be precisely known. Otherwise, tails of CDFs

can be critically underestimated, so the estimation of failure probabilities becomes quite poor. In [18] one can find a demonstration that straightforward probabilistic computations are highly sensitive to imprecise information. Imagine a CDF is known almost precisely, but with a small deviation in some distribution parameter. This may easily lead to a situation as shown in Figure 2, which illustrates the fact that the failure probability is often underestimated (here by the factor 2).

In the univariate case it is simple to overcome problems with lack of information. One can apply Kolmogorov-Smirnov (KS) statistics as a powerful tool. Assume that the uncertainty information is given by empirical data on a random variable  $X$ , e.g., a small set of sample points  $x_1, \dots, x_N$ . The **empirical distribution**  $\tilde{F}$  is defined by

$$\tilde{F}(\xi) := \sum_{\{j|x_j \leq \xi\}} \frac{1}{N}. \quad (4)$$

The KS test uses  $D := \max|\tilde{F} - F|$ , the maximum deviation of  $\tilde{F}$  from  $F$ , as its test statistics, and it can be shown that  $\sqrt{N}D$  converges in distribution to the Kolmogorov function

$$\phi(\lambda) := \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2} \quad (5)$$

for  $N \rightarrow \infty$ , cf. [46]. Conversely, if we choose a fixed confidence level  $\alpha$ , we can compute  $D$  from

$$D = \frac{\phi^{-1}(\alpha)}{\sqrt{N} + 0.12 + \frac{0.11}{\sqrt{N}}}, \quad (6)$$

cf. [81], and thus find a maximum deviation of the unknown  $F$  from the known  $\tilde{F}$ . That means that we have non-parametric bounds  $[\tilde{F} - D, \tilde{F} + D]$  enclosing  $F$  with confidence  $\alpha$ , only given the knowledge of  $x_1, \dots, x_N$ .

In case of high-dimensional random vectors, classical probability theory has no means to cope with scarce data as in the univariate case with KS bounds. Although multivariate CDFs can be defined as in the 1D case using the componentwise partial order in  $\mathbb{R}^n$ , the computational effort for obtaining higher dimensional PDFs and their numerical integration prohibit the reliable use of standard probabilistic methods in higher dimensions.

### 2.3 Safety margins

A simple and widely spread non-probabilistic uncertainty model is based on so-called **safety margins**. This model is applied when very little information is available, in situations where most information is provided as interval information.

There are different kinds of sources for interval information, e.g., measurement accuracy. Safety margins are a special kind of interval information, namely one

which is provided subjectively by an expert designer, as typically the case in early design phases. If additional information is available, like marginal distributions or safety margins from further experts, the safety margins approach cannot handle it and thus loses some valuable information. Since safety margins are highly subjective information one cannot expect rigorous results for the safety analysis. However, engineers hope to achieve reasonably conservative bounds by using conservative margins.

The first approach – a tool to handle all kinds of interval information – is **interval analysis**, cf. [64], [70]. We write  $X \in [a, b]$  for  $a \leq X \leq b$  in the univariate case; in the higher dimensional case  $\varepsilon = (\varepsilon^1, \dots, \varepsilon^n)$ , we define interval information  $\varepsilon \in [a, b]$  pointwise via  $a^1 \leq \varepsilon^1 \leq b^1, \dots, a^n \leq \varepsilon^n \leq b^n$ , and call  $[a, b]$  a **box**. We always interpret equalities and inequalities of vectors componentwise. In the following we present two frequent approaches to handle the incoming interval information.

Assume that the cost or gain (or another assessment) function  $s : \mathbb{M} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ , with design space  $\mathbb{M}$ , models the response function of the design, and the information about the uncertain input vector  $\varepsilon$  is given by the bounds  $\varepsilon \in [a, b] \subseteq \mathbb{M}$ . By way of interval calculations one achieves bounds componentwise on  $s(\varepsilon)$  – also called an **interval extension** of  $s$ .

Computing an interval extension is often affected by overestimation. A variable  $\varepsilon^i \in [a^i, b^i]$  should take the same value from the interval  $[a^i, b^i]$  each time it occurs in an expression in the computation of  $s$ . However, this is not considered by straightforward interval calculations, so the range is computed as if each time the variable  $\varepsilon^i$  occurs it can take a different value from  $[a^i, b^i]$ , leading to an enclosure which may be much wider than the range for  $f(\varepsilon)$ . One possible way out is based on Taylor series, cf. [55]. Nonlinear interval computations in higher dimensions may become expensive, growing exponentially with  $n$ , but can often be done efficiently and complemented by simulation techniques or sensitivity analysis, as we will see later. Note that in case that  $s$  is given as a black box evaluation routine – as in many real-life applications – the interval extension cannot be determined rigorously anyway. Also interval methods are often not applicable as a toolbox, but require problem specific expert knowledge to overcome overestimation issues.

In the literature we find much utilization of interval computations in uncertainty modeling. Analyzing the statistics for interval valued samples one seeks to bound mean or variance, which are then also interval valued, cf. [49]: Finding an upper bound on the variance is NP-hard, a lower bound can be found in quadratic time. The field of applications of interval uncertainty for uncertainty handling is vast, e.g., [23], [53], [67], [68], [79].

Also probability theory proper makes use of non-probabilistic interval uncertainty models. For example, consider a Markov chain model with transition matrix  $(p_{ij})$ , where the transition probabilities  $p_{ij}$  are uncertain, and only given as intervals. Then one can build a generalized Markov chain model, cf. [88], [89]. In [30] one can find a study of imprecise transition probabilities in Markov decision processes.

The second approach to handle safety margin interval information is a simpli-

fication of the information by **fixing each uncertain variable**  $\varepsilon^i \in [a^i, b^i]$  at the value of one of the safety margins  $a^i$  or  $b^i$ , and simply insert this value, for instance  $a^i$  for all  $i$ , as worst-case scenario to compute the worst-case design response  $s(a)$ . The decision where to fix the worst-case scenario is taken merely subjectively or via a list of relevant cases. A designer may overestimate intentionally the subjective safety margin assignments, e.g., by adding  $20\% = 2 \cdot 0.1$  to the nominal interval bounds for a variable, i.e.,  $\varepsilon \in [a - 0.1(b - a), b + 0.1(b - a)]$ , in order to be suitably conservative in computing the worst-case design response.

The computational effort with this latter approach is not very high and also applies well in higher dimensions. Actually, there is no extra effort in addition to the cost for evaluating  $s$  involved in this simple uncertainty model.

A field where safety margins are very popular is multidisciplinary design optimization [1]. In many cases, in particular in early design phases, it is common engineering practice to combine the assignment of safety margins with an iterative process of refining or coarsening the margins, while converging to a robust optimal design. The refinement of the intervals is done by experts who assess whether the worst-case scenario determined for the design at the current stage of the iteration process is too pessimistic or too optimistic. The goal of the whole iteration includes both optimization of the design and safeguarding against uncertainties. The achieved design is thus supposed to be robust. This procedure enables a very dynamic design process and quick interaction between the involved disciplines.

All in all, safety margins allow for a simple, efficient handling of uncertainties, also in large-scale problems, if no information is available but an interval bounding from a single source. Otherwise, we have to look for improved methods, which can handle more uncertainty information. It should be remarked that in most cases, even in early design stages, there is more information available than assumed for the safety margin approach.

## 2.4 Safety factors

Remember the concept of failure probabilities as introduced in Section 2.1. The failure probability was defined as  $p_f = \Pr(\mathbb{F})$ , where  $\mathbb{F}$  was the set of events which lead to design failure. Let  $s : \mathbb{R}^n \rightarrow \mathbb{R}$  be the design response of a fixed design for uncertain inputs  $\varepsilon \in \mathbb{R}^n$ . Assume that there is a limit state  $\ell$  for which  $s(\varepsilon) < \ell$  means design failure, and  $s(\varepsilon) \geq \ell$  represents satisfying design performance, i.e., that  $\mathbb{F}$  is defined by

$$\mathbb{F} = \{\varepsilon \mid s(\varepsilon) < \ell\}. \quad (7)$$

The idea behind safety factors is to build the design in a way that the expected value of  $s(\varepsilon)$  is greater than the limit state  $\ell > 0$  multiplied by a factor  $k_{\text{safety}} > 1$ , called the **safety factor**. In other words, a design should fulfill

$$\langle s(\varepsilon) \rangle \geq k_{\text{safety}} \ell. \quad (8)$$



where the expectation has to be suitably estimated. For  $s(\varepsilon) \in \mathbb{R}^m$  we interpret this definition componentwise for each safety requirement on the design responses  $s^1, \dots, s^m$ , and  $\ell = (\ell^1, \dots, \ell^m)$ .

Conversely, suppose that we are given a fixed design and  $\langle s(\varepsilon) \rangle \geq \ell$ ,  $s(\varepsilon) \in \mathbb{R}$ ,  $s \in C^1$ . Define the maximal feasible safety factor as  $k := \frac{\langle s(\varepsilon) \rangle}{\ell}$ . To see the relation between  $k$  and the design failure probability  $p_f$  we assume that we have fixed an admissible  $p_f$  for the fixed design, then  $p_f = \Pr(s(\varepsilon) \leq \ell) = F_s(\ell)$ . Here  $F_s$  is the CDF of  $s(\varepsilon)$  with density  $\rho_s$  given by

$$\rho_s(x) = \rho(s^{-1}(x)) \cdot |\det s'(x)|^{-1}, \quad (9)$$

with  $|\det s'(x)|$  the absolute value of the determinant of the Jacobian of  $s$ . Hence we get  $\ell = F_s^{-1}(p_f)$ , assuming that  $F_s$  is invertible, and

$$k = \frac{\langle s(\varepsilon) \rangle}{F_s^{-1}(p_f)}. \quad (10)$$

As we are applying standard methods from probability theory to compute safety factors, precise knowledge of  $\rho$  and of the **limit state function**  $s(\varepsilon) - \ell$  is required to achieve rigorous probability statements.

In the lower dimensional case, if  $\rho$  is unknown, but a narrow bounding interval and certain expectations (e.g., means and covariances) for the random vector  $\varepsilon$  are known, safety factors can still be well described approximately. The expectation of smooth functions  $s$  of  $\varepsilon$  is then achievable from the Taylor series for  $s$ , cf., e.g., [5], since expectation is a linear operator on random variables – similar to Taylor models for interval computations. The problems mentioned in Section 2.2 concerning lack of information in the higher dimensional case remain.

Probabilistic computation of safety factors is not as much affected by subjective opinions of the designer as, for example, safety margins. Safety factors are directly associated with required reliability. One important subjective decision is how to fix the required reliability or the admissible failure probability, respectively. The decision can be based, e.g., on the assessment of failure cost or on regulations by legislation.

## 2.5 Simulation methods

Simulation methods are ubiquitous tools, and uncertainty handling is one of their application fields. Simulation means computational generation of sample points as realizations of a random vector  $\varepsilon$ , assuming that the joint CDF, marginal CDFs, or interval bounds are given. Thus the not necessarily probabilistic uncertainties involved are simulated, which gives rise to the term simulation methods. Simulation methods are also referred to via the terms **random sampling** or **Monte Carlo sampling**.

After sample generation, the design response  $s : \mathbb{M} \rightarrow \mathbb{R}^m$  is evaluated for each generated sample point. If all or at least a reasonable majority of the points meet the safety requirements  $s(\varepsilon) \geq \ell$ , the design is considered to be safe.

The core part of simulation is the sample generation. There is a large number of different techniques addressing it. Three classical variants are based on CDF inversion, Markov chains, and Latin hypercube sampling, respectively. CDF inversion requires the CDF to be invertible, and is particularly not applicable in higher dimensions. The Markov chain method constructs a reversible Markov chain using the Metropolis approach [58], or the more general Hastings method [36], by way of a rejection method. The rejection rule assures that it is not necessary to compute  $p_i$  (i.e., the stationary probability of the state  $i$ ), but only the easy-to-compute quotient  $\frac{p_i}{p_j}$  of two different states which is independent of the dimension of  $X_n$ . This makes the method highly attractive in higher dimensions. The **Latin hypercube sampling** (LHS) method [57], first determines a finite grid of size  $N^n$ , where  $N$  is the desired number of sample points and  $n$  is the dimension of the random vector for which we want to generate a sample. The grid is preferably constructed such that the intervals between adjacent marginal grid points have the same marginal probability. The  $N$  sample points  $x_1, x_2, \dots, x_N$ ,  $x_i = (x_i^1, \dots, x_i^n) \in \mathbb{R}^n$  are then placed to satisfy the Latin hypercube requirement,

$$\text{For } i, k \in \{1, \dots, N\}, j \in \{1, \dots, n\} : x_i^j \neq x_k^j \text{ if } k \neq i. \quad (11)$$

This procedure introduces some preference for a simple structure, i.e., we disregard correlations, tacitly assuming independence. The advantage of the method is that the full range of  $\varepsilon$  is much better covered than with a Markov chain based method, giving deeper insight to the distribution tails of  $\varepsilon$ . Hence failure probabilities can be better estimated. Moreover, one does not require more sample points for higher  $n$ , so the application of LHS in higher dimensions is still attractive.

Often **importance sampling** is used to speed up simulation techniques by a reduction of the number of required simulations, e.g., [37], [44], [92]. The sample points are generated from a different distribution than the actual distribution of the involved random variables. The sampling density is weighted by an importance density, e.g., a normal distribution with standard deviation  $\sigma$  depending on where the most probable failure points are expected, for instance, depending on the curvature of  $s$ . Thus the generated sample is more likely to cover the 'important' regions for the safety analysis.

Considering the rigor of the results one should be aware of the fact that no estimation of failure probabilities computed from a simulation technique is a rigorous bound. These methods are based on the law of large numbers, and their results are only valid for a sufficient amount of sample points. It is difficult to assess what 'sufficient amount' means in a higher dimensional space; one might need to generate an excessively large number of sample points for estimating very small failure probabilities. That is why simulation methods are endangered to critically underestimate CDF tails [20]. It gets particularly dangerous when the CDFs to sample from are unknown.

On the other hand, simulation methods are computationally very efficient, they can be parallelized [51], and also apply well in higher dimensions, where almost no alternatives exist at present.

Another important aspect comes with **black box response functions**  $s$ . They principally impose no additional difficulties applying simulation methods. However, if the computational cost for evaluating  $s$  is very high, problems will arise as simulation typically requires many evaluations, hence is limited to simple models for  $s$ , often surrogate functions (cf., e.g., [17], [39]) for more complex models.

As mentioned earlier simulation techniques have many applications, e.g., the computation of multi-dimensional integrals. They are related to many uncertainty methods, also non-probabilistic ones like interval uncertainty.

## 2.6 Sensitivity Analysis

Sensitivity analysis is actually not an independent uncertainty method itself, it rather applies in several different fields one of which is uncertainty handling. Sensitivity analysis investigates the variability of a model function output  $f(\varepsilon)$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\varepsilon = (\varepsilon^1, \varepsilon^2, \dots, \varepsilon^n)^T$ , with respect to changes in the input variables  $\varepsilon^i$ .

To this end one can follow different approaches, e.g., investigate the partial derivatives of  $f$  if they are available, using  $\frac{\partial f}{\partial \varepsilon^i}$  as an indicator for the influence of  $\varepsilon^i$  on  $f$ . One can also vary a subset of all single input variables  $\varepsilon^i$  of  $f$  while keeping all other inputs constant. Then one assesses the impact of this subset by the variability of the output  $f$  by means of some uncertainty methods introduced in this paper, e.g., fuzzy set or simulation methods. Thus one hopes to achieve a dimensionality reduction of  $f$  fixing those input variables which turn out to have little influence on  $f$ . Frequently, e.g., [50], one assumes monotonicity of  $f$  and interval uncertainty of  $\varepsilon$ , since this enables the use of very fast techniques in higher dimensions, the effort is then growing only linearly in the dimension  $n$ .

As a particular case of handling interval uncertainties in high dimensions with computationally expensive black box response functions  $s$  we mention the Cauchy distribution based simulation for interval uncertainty [51]: Assuming that the intervals are reasonably small, e.g., given as measurement errors, it is reasonable to assume that  $s$  is linear. Generate  $N$  independent sample points for the measurement errors from the scaled Cauchy(0, 1) distribution, which is easy as the inverse CDF of a Cauchy(0, 1) distribution is known explicitly in this case. Linear functions of Cauchy distributed variables are again Cauchy distributed [98], with an unknown parameter  $\Theta$ . Having estimated the parameter  $\Theta$ , e.g., by means of a maximum likelihood estimator, one can infer probabilistic statements about errors in  $s$  which are Cauchy(0,  $\Theta$ ) distributed. Thus this method exploits the characteristics of a Cauchy distribution to produce results the accuracy of which can be investigated statistically depending on  $N$ , also for low  $N$  in case of expensive  $s$ . No derivatives are required, only  $N$  black box evaluations of  $s$ .

Applications of sensitivity analysis can be found, e.g., in [77], [80].

### 3 Reliability methods

Reliability methods are a very popular approach based on the concepts of reliability and failure probability and transformation to standard normal space, cf. [83]. They represent a significant improvement in computational modeling of reliability compared to rather old-fashioned methods like safety factors.

In order to investigate the failure probability  $p_f$  given the joint CDF  $F$ , or at least marginal CDFs, of the involved random vector  $\varepsilon$ , one first applies a coordinate transformation  $u = \mathcal{T}(\varepsilon)$  to standard normal space, cf. [69], [84]. Then the failure surface  $\mathbb{F} = \{s(\varepsilon) < \ell\}$  is approximated and embedded in an optimization problem to estimate  $p_f$ .

Once a  $\mathcal{T}$  is found the new coordinates  $u$  live in standard normal space, that means the level sets of the density of  $u$  are  $\{u \mid \|u\|_2 = \text{const}\}$ , due to the shape of the multivariate normal distribution. Let  $s(u)$  be the design response in the transformed coordinates. Then the most probable failure point  $u^*$  from the failure set  $\mathbb{F} = \{u \mid s(u) < \ell\}$  is the solution of

$$\begin{aligned} \min_u \quad & \|u\|_2 \\ \text{s.t.} \quad & s(u) < \ell \end{aligned} \tag{12}$$

i.e., the point from  $\mathbb{F}$  with minimal 2-Norm. This critical point is called  $\beta$ -point, and

$$p_f \approx \Phi(-\beta) \tag{13}$$

approximates the failure probability, where  $\beta = \|u^*\|_2$  and  $\Phi$  denotes the CDF of the univariate  $N(0, 1)$  distribution.

Thus we have reduced the estimation of  $p_f$  to the standard problem of finding  $\mathcal{T}$  and the remaining problem of solving the optimization problem (12). The latter is a nonlinear optimization problem with all the problems that come with it. Even if the limit state function is convex, after transformation it may become a strongly non convex problem in case that the CDF  $F$  significantly differs from a normal distribution. Using a linear approximation of the limit state function in the computation of  $\beta$  is called **first order reliability method** (FORM), a quadratic approximation is called **second order reliability method** (SORM).

One hopes that a unique solution for  $\beta$  exists; however in general, there is usually no guaranteed global and unique solution for this optimization problem. Another problem about this approach is that the  $\beta$ -point found may not be representative for the failure probability. A discussion on the involved optimization problem can be studied in [14], investigating difficulties like multiple  $\beta$ -points. The entailed difficulties require some caveats assessing the results of reliability methods: the methods may fail to estimate  $p_f$  correctly without warning the user. Especially when additional problems appear – like higher dimensionality or black box response functions  $s$  – the reliability methods become less attractive in many large-scale real-life situations. It should be remarked that the search for  $u^*$  can be supported by

sampling and simulation techniques like importance sampling, cf. Section 2.5, as means for corrections and reduction of the computational effort, e.g., [56].

Reliability methods are associated with design optimization within the field of reliability based design optimization (RBDO). Instead of the often occurring bilevel problem formulation (i.e., design optimization in the outer level, worst-case scenario search in the inner level) one formulates a one level problem as follows, cf. [43], [65]. Let  $s_{\mathcal{T}} = s_{\mathcal{T}}(\theta, u) = s(\theta, \mathcal{T}(\varepsilon))$ , the design response in transformed coordinates, with the controllable design vector  $\theta$  which fully specifies the design. Let  $g(\theta)$  be the objective function, e.g., the cost of the design or the cost of failure. One seeks to minimize  $g$  subject to some reliability constraint  $p_f \leq p_a$  where  $p_f = \Pr(s(\theta, u) < \ell)$  is approximated by equation (13), and  $p_a$  the fixed admissible failure probability. We get

$$\begin{aligned} \min_{\theta} \quad & g(\theta) \\ \text{s.t.} \quad & \Phi(-\beta) \leq p_a \\ & \theta \in T \end{aligned} \tag{14}$$

where  $T$  is the set of possible design choices. For  $s(\varepsilon) \in \mathbb{R}^m$  we have several constraints  $\Phi(-\beta_i) \leq p_a$ ,  $i = 1, \dots, m$ .

Usually simulation techniques are employed to solve (14), e.g., [85]. In [78] it is suggested to use Monte Carlo methods to check the probabilistic constraints, and to train a neural network to check the deterministic constraints, or even both probabilistic and deterministic. This can be implemented as parallelized computations which improves computation time significantly. In any case, one should be aware that one uses a soft solution technique on top of a soft uncertainty model.

## 4 $p$ -boxes

A  $p$ -box – or  $p$ -bound, or probability bound – is an enclosure of the CDF of a univariate random variable  $X$ ,  $F_l \leq F \leq F_u$ , in case of partial ignorance about specifications of  $F$ . Such an enclosure enables, e.g., to compute lower and upper bounds on expectation values or failure probabilities.

There are different ways to construct a  $p$ -box depending on the available information about  $X$ , cf. [22]. For example, assume that we have empirical data for  $X$ . Then we can construct a  $p$ -box with KS statistics, cf. (6), after fixing a confidence level  $\alpha$ . In [25] we find an exhaustive description which construction techniques can be applied to construct a  $p$ -box, related to the type of available information. Moreover, it is illustrated how to construct  $p$ -boxes from different uncertainty models like Dempster-Shafer structures (cf. Section 6) or Bayesian update estimates (cf. Section 5). The studies on  $p$ -boxes have already lead to successful software implementations, cf. [6], [21].

Higher order moment information on  $X$  (e.g., correlation bounds) cannot be handled or processed yet. This is a current research field, cf., e.g., [24].

To compute functions  $f$  of  $p$ -boxes, that means we have a  $p$ -box for  $\varepsilon^1, \dots, \varepsilon^n$  and seek a  $p$ -box for  $f = f(\varepsilon^1, \dots, \varepsilon^n)$ , one first regards  $f$  consisting of elementary arithmetical operations and finds bounds for these expressions. To this end one discretizes the bounds for  $\varepsilon^1, \dots, \varepsilon^n$  towards a discretization of the bounds for  $f$ , and then finds an expression for the bound of  $f$  in terms of the bounds for  $\varepsilon^1, \dots, \varepsilon^n$ . This can be done for all elementary arithmetic operations, without independence assumption for  $\varepsilon^1, \dots, \varepsilon^n$ , cf. [99], [100]. Thus the research on arithmetics for random variables actually builds the foundation of  $p$ -boxes. The dependency problem is not trivial, assume that one has independent random variables  $X, Y, Z$ , then the variables  $S = X + Y$  and  $T = Y \cdot Z$  are not independent in general.

One learns that the problem of rigorously quantifying probabilities given incomplete information – as done with probability arithmetic and  $p$ -boxes – is highly complex, even for simple problems, e.g., [52]. Due to their constructions the methods are rather restricted to lower dimensions and non-complex models  $f$ . Black box functions  $f$  cannot be handled as one requires knowledge about the involved arithmetic operations. All in all, they often appear not to be reasonably applicable in many real-life situations. On the other hand, as soon as we can apply methods like  $p$ -boxes to calculate with bounds on probability distributions, we are not restricted to the use of selecting less rigorous single distribution assumptions (e.g., maximum entropy) anymore.

Two more remarks about  $p$ -boxes: First, the definition of  $p$ -boxes can be generalized to higher dimensions based on the definition of higher dimensional CDFs, cf. [15]. However, this has not lead to practical results yet. Second, probability arithmetic can be regarded as a generalization of interval arithmetic which would be the special case given only the information  $X \in [a, b]$ . It is also related to the world of imprecise probabilities via sets of measures. From a  $p$ -box  $[F_l, F_u]$  for  $X$  one can infer bounds on the expectation for  $f(X)$  by  $\langle f \rangle_l = \inf_{F_l \leq F \leq F_u} \int_{\Omega} f dF$ ,  $\langle f \rangle_u = \sup_{F_l \leq F \leq F_u} \int_{\Omega} f dF$ , regarding  $F_l \leq F \leq F_u$  as a set of measures, e.g., [47], [96]. The bounds can be computed numerically by discretization and formulation of a linear programming problem (LP), cf. [93].

## 5 Bayesian inference

As soon as incomplete information is based on subjective knowledge and can be updated iteratively by additional information, one can consider using Bayesian inference to handle uncertainties. Bayesian inference means reasoning on the basis of Bayes' rule, working with conditional probabilities.

Here we have the crucial problem, how to select a suitable prior distribution. For a reasonable choice real statistical data is needed in sufficient amount. Additionally, of course, incoming new observations are required for updating. Priors can be chosen, e.g., with the maximal entropy principle, cf. Section 2.2. In practice one often chooses a normal distribution to simplify calculations, or conjugate priors, i.e., a distribution where the posterior has a similar shape like the prior except from a change in some parameters. Actually, it is a well-known criticism that the

choice of the prior often seems to be quite arbitrary and merely in the will of the statistician.

A typically employed model in Bayesian inference is a so-called Bayesian or belief network (BN). A BN is a directed acyclic graph (DAG) between states of a system and observables. A node  $\mathcal{N}$  and its parent nodes in the DAG represent the input information of the network which consists of tables of conditional probabilities of  $\mathcal{N}$  conditional on its parent nodes. The whole DAG represents the joint distribution of all involved variables, even in higher dimensional situations. Computations using BNs can be done efficiently on the DAG structure, assumed that all conditional probabilities are precisely known.

What if the conditional PDFs of the tables of conditional probabilities in BNs are unknown or not precisely known? This happens frequently in practice, in particular for variables conditional on multiple further variables. The Bayesian approach appears to become useless in this case. A generalized approach to BNs with imprecise probabilities can be studied on the basis of so-called **credal networks**, e.g., [11], [35]. A credal network is a set of BNs with the same DAG structure, but imprecise values in the conditional probability tables. The probabilities can be given as intervals, or more generally described.

The Bayesian approach applies in design optimization, cf. [103]. Similar to RBDO (14) one minimizes a certain objective like design cost subject to probabilistic constraints involving the failure distribution. The associated joint distribution is estimated and updated from available data, starting with conjugate priors.

## 6 Dempster-Shafer theory

Dempster-Shafer theory enables to process incomplete uncertainty information allowing to compute bounds for failure probabilities and reliability.

We start with defining fuzzy measures, cf. [90]. A **fuzzy measure**  $\tilde{\mu} : 2^\Omega \rightarrow [0, 1]$ , fulfills

$$\tilde{\mu}(\emptyset) = 0, \tilde{\mu}(\Omega) = 1, \quad (15)$$

$$A \subseteq B \Rightarrow \tilde{\mu}(A) \leq \tilde{\mu}(B). \quad (16)$$

The main difference to a probability measure is the absence of additivity. Instead, fuzzy measures only satisfy monotonicity (16). To find lower and upper bounds for an unknown probability measure given incomplete information one seeks two fuzzy measures **belief** Bel and **plausibility** Pl, where Bel is a fuzzy measure with  $\text{Bel}(A \cup B) \geq \text{Bel}(A) + \text{Bel}(B) - \text{Bel}(A \cap B)$ , and Pl is a fuzzy measure with  $\text{Pl}(A \cup B) \leq \text{Pl}(A) + \text{Pl}(B) - \text{Pl}(A \cap B)$ .

To construct the measures Bel and Pl from the given uncertainty information one formalizes the information as a so-called **basic probability assignment**  $m : 2^\Omega \rightarrow [0, 1]$  on a finite set  $\mathcal{A} \subseteq 2^\Omega$  of non-empty subsets  $A$  of  $\Omega$ , such that

$$m(A) \begin{cases} > 0 & \text{if } A \in \mathcal{A}, \\ = 0 & \text{otherwise,} \end{cases} \quad (17)$$

and the normalization condition  $\sum_{A \in \mathcal{A}} m(A) = 1$ . Sometimes  $m$  is also called **basic belief assignment**.

The basic probability assignment  $m$  is interpreted as the exact belief focussed on  $A$ , and not in any strict subset of  $A$ . The sets  $A \in \mathcal{A}$  are called **focal sets**. The structure  $(m, \mathcal{A})$ , i.e., a basic probability assignment together with the related set of focal sets, is called a **Dempster-Shafer structure** (DS structure).

Given a DS structure  $(m, \mathcal{A})$  we can construct Bel and Pl by

$$\text{Bel}(B) = \sum_{\{A \in \mathcal{A} | A \subseteq B\}} m(A), \quad (18)$$

$$\text{Pl}(B) = \sum_{\{A \in \mathcal{A} | A \cap B \neq \emptyset\}} m(A) \quad (19)$$

for  $B \in 2^\Omega$ .

Thus Bel and Pl have the sought property  $\text{Bel} \leq \text{Pr} \leq \text{Pl}$  by construction and, moreover, satisfy  $\text{Bel}(B) = 1 - \text{Pl}(B^c)$ . The information contained in the two measures Bel and Pl induced by the DS structure is often called a **random set**.

In the classical case the additivity of non-fuzzy measures would yield  $\text{Pl}(B) = 1 - \text{Pl}(B^c) = \text{Bel}$ . Thus  $\text{Bel} = \text{Pr} = \text{Pl}$  and classical probability theory becomes a special case of DS theory. Also note that if we have a DS structure on the singletons of a finite  $\Omega$ , then we have full stochastic knowledge equivalent to a CDF.

DS structures can be obtained from expert knowledge or in lower dimensions from histograms, or from the Chebyshev inequality  $\Pr(|X - \mu| \leq r) > 1 - \frac{\sigma^2}{r^2}$  given expectation value  $\mu$  and variance  $\sigma^2$  of a random variable  $X$ , cf. [75], [76], [77]: Let  $r = \frac{\sigma}{\sqrt{1-\alpha}}$  for a fixed confidence level  $\alpha$ , then  $\Pr(\{|X - \mu| \leq \frac{\sigma}{\sqrt{1-\alpha}}\}) > \alpha$ . The sets  $C_\alpha := \{\omega \in \Omega \mid |X(\omega) - \mu| \leq \frac{\sigma}{\sqrt{1-\alpha}}\}$  for different values of  $\alpha$  define focal sets, and we get Belief and Plausibility measures by  $\text{Bel}(C_\alpha) = \alpha$  and  $\text{Pl}(C_\alpha^c) = 1 - \alpha$ , respectively.

To extend one-dimensional focal sets to the multi-dimensional case one can generate joint DS structures from the Cartesian product of marginal basic probability assignments assuming random set independence, cf. [10], or from weighting the 1-dimensional marginal focal sets, cf. [27]. In [103] we find the suggestion to employ Bayesian techniques to estimate and update DS structures from little amount of information.

To combine different, or even conflicting DS structures  $(m_1, \mathcal{A}_1)$ ,  $(m_2, \mathcal{A}_2)$  (in case of multiple bodies of evidence, e.g., several different expert opinions) to a new basic probability assignment  $m_{\text{new}}$  one uses Dempster's rule of combination [12], forming the basis of **Dempster-Shafer theory** or **evidence theory** [86],

$$m_{\text{new}}(B) = \sum_{\{A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2 | A_1 \cap A_2 = B\}} \frac{m_1(A_1)m_2(A_2)}{K} \quad (20)$$

with the normalization constant  $K = 1 - \sum_{\{A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2 | A_1 \cap A_2 = \emptyset\}} m_1(A_1)m_2(A_2)$  which is interpreted as the **conflict**.



The combination rule enables to compute a joint DS structure. Also note that the combination rule is a generalization of Bayes' rule, motivated by the criticism that a single probability assignment cannot model the amount of evidence one has.

The complexity of the rule is strongly increasing in higher dimensions, and in many cases requires independence assumptions for simplicity reasons avoiding problems with interacting variables. It is not yet understood how the dimensionality issue can be solved. Working towards more efficient computational implementations of evidence theory it can be attempted to decompose the high-dimensional case in lower dimensional components which leads to so-called compositional models, cf. [41].

The **extension** of a function  $f$  is based on the joint DS structure  $(m, \mathcal{A})$ . The new focal sets of the extension are  $B_i = f(A_i)$ ,  $A_i \in \mathcal{A}$ , the new basic probability assignment is  $m_{\text{new}}(B_i) = \sum_{\{A_i \in \mathcal{A} | f(A_i) = B_i\}} m(A_i)$ .

To embed DS theory in design optimization one formulates a constraint on the upper bound of the failure probability  $p_f$  which should be smaller than an admissible failure probability  $p_a$ , i.e.,  $\text{Pl}(\mathbb{F}) \leq p_a$ , for the failure set  $\mathbb{F}$ . This can be studied in [66] as evidence based design optimization (EBDO). One can also find further direct applications in engineering computing, e.g., in [29], [75].

DS structures enable to construct  $p$ -boxes [7], [15], [93], i.e., to determine lower bounds  $F_l$  and upper bounds  $F_u$  of the CDF of a random variable  $X$ ,

$$\begin{aligned} F_l(t) &= \text{Bel}(\{\omega \in \Omega \mid X(\omega) \leq t\}), \\ F_u(t) &= \text{Pl}(\{\omega \in \Omega \mid X(\omega) \leq t\}). \end{aligned}$$

Conversely it is possible to generate a DS structure that approximates a given  $p$ -box discretely, cf. [2], [13], [25]. Fix some levels  $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_N = 1$  of the  $p$ -box, then generate focal sets by

$$\begin{aligned} A_i &:= [\inf\{x \mid F_u(x) = \alpha_i\}, \inf\{x \mid F_l(x) = \alpha_i\}], \\ m(A_1) &= \alpha_1, m(A_i) = \alpha_i - \alpha_{i-1}, i = 2, \dots, N. \end{aligned} \tag{21}$$

Another relation to a different uncertainty representation concerns nested focal sets, i.e.,  $\mathcal{A} = \{A_1, A_2, \dots, A_m\}$ ,  $A_1 \subseteq A_2 \subseteq \dots \subseteq A_m$ . In this case

$$\text{Bel}(A \cap B) = \min(\text{Bel}(A), \text{Bel}(B)), \tag{22}$$

$$\text{Pl}(A \cup B) = \max(\text{Pl}(A), \text{Pl}(B)). \tag{23}$$

For nested focal sets the fuzzy measures Bel and Pl directly correspond to possibility and necessity measures, respectively, which appear in fuzzy set theory, cf. [16], as we will see in the next section.

We have learned that DS structures can unify several different uncertainty models, see, e.g., [48], but cannot overcome the curse of dimensionality being prohibitively expensive in higher dimensions.

## 7 Fuzzy sets

The development of fuzzy sets has started roughly in parallel to the development of DS theory with the goal to model vague verbal descriptions in absence of any statistical data. It is a generalization of conventional set theory redefining the characteristic function of a set  $A$  by a so-called **membership function**  $\mu_A$ . The value  $\mu_A(x)$  indicates the membership value of an uncertain variable  $x$  with respect to  $A$ . The value can be any real number between 0 and 1 as opposed to the characteristic function  $1_A(x)$  which only takes binary values. A **fuzzy set** is a set  $A$  together with its related membership function  $\mu_A$ .

This section will give a short overview on fuzzy sets, focussing on their application for uncertainty handling. The following terms play an important role in the theory of fuzzy sets. The **height**  $h$  of a fuzzy set is defined by  $h := \max_x \mu_A(x)$ . The **support** of a fuzzy set is the set  $\{x \mid \mu_A(x) \neq 0\}$ . The **core** or **modal values** of a fuzzy set is the set  $\{x \mid \mu_A(x) = 1\}$ . The  $\alpha$ -**cut**  $C_\alpha$  of a fuzzy set for a fixed value  $\alpha \in [0, 1]$  is the set

$$C_\alpha := \{x \mid \mu_A(x) \geq \alpha\}. \quad (24)$$

The  $\alpha$ -cut is determined by the values of the membership function. Conversely one can construct  $\mu_A$  from the knowledge of the  $\alpha$ -cuts, cf. [102], to achieve an  $\alpha$ -cut based representation of a fuzzy set:

$$\mu_A(x) = \sup_{\alpha} \min(\alpha, 1_{C_\alpha}(x)). \quad (25)$$

Note the relationship between BPA-structures on nested focal sets, cf. Section 6, and  $\alpha$ -cuts of a fuzzy set with non-empty core, which are nested by definition, i.e.,  $C_\alpha \subseteq C_\beta$  for  $\alpha \geq \beta$ . Let  $1 = \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_N = 0$  be  $\alpha$ -levels of a fuzzy set, then we can construct a BPA  $m$  on the  $\alpha$ -cuts  $C_{\alpha_i}$  by  $m(C_{\alpha_i}) = \alpha_i - \alpha_{i+1}$ ,  $i < N$ ,  $m(C_{\alpha_N}) = \alpha_N$ . Conversely a BPA-structure on nested focal sets  $A_1 \subseteq A_2 \subseteq \dots \subseteq A_N$  allows to construct a fuzzy set by  $\alpha_N = m(A_N)$ ,  $C_{\alpha_N} = A_N$ ,  $\alpha_{N-1} = m(A_N) + m(A_{N-1})$ ,  $C_{\alpha_{N-1}} = A_{N-1}$ ,  $\dots$ ,  $\alpha_1 = \sum_{i=1}^N m(A_i) = 1$ ,  $C_{\alpha_1} = A_1$ , and then applying (25). Thus it is possible to convert expert knowledge modeled by a fuzzy set into a DS structure. Using the Dempster's rule, however, to combine different bodies of evidence in general leads to non-nested focal sets, hence a conversion back to the fuzzy set formalism is not possible after applying a combination rule.

Some special cases of fuzzy sets motivated the notation of fuzzy intervals and fuzzy numbers, cf. [101]. A **fuzzy interval** or **convex fuzzy set** is a fuzzy set with  $\mu_A(x) \geq \min(\mu_A(a), \mu_A(b))$  for all  $a, b, x \in [a, b]$ . A **fuzzy number** is a fuzzy interval with closed  $\alpha$ -cuts, compact support, and a unique modal value.

The definition of a fuzzy set and its membership function in higher dimensions is a straightforward generalization of the one-dimensional case. The extension of a function  $f(x) = z$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , for a fuzzy set with membership function  $\mu$  is constructed by the **extension principle** for a new membership function

$$\mu_{\text{new}}(z) = \sup_{x \in f^{-1}(z)} \mu(x), \quad (26)$$

cf. [101]. The construction involves an optimization problem with rapidly increasing complexity in higher dimensions.

It can be attempted to solve this problem by reduction of the problem to the  $\alpha$ -cuts of the fuzzy set, cf. Section 7.1, or by sensitivity analysis, cf. Section 2.6.

Except from the dimensionality issue another criticism of fuzzy sets is the fact that the assignment of membership functions appears to be quite arbitrary, often defined by a single expert opinion. In lower dimensions membership functions can be estimated, e.g., from histograms, but there is no general, statistically well-grounded basis for the assignment of membership functions. Of course, if only vague verbal descriptions, i.e., highly informal uncertainty information, is available statistical properties are entirely absent. In this case, which represents the classical motivation of fuzzy sets, it can be argued that it is impossible to formulate a general recipe for processing the information. However, usually the information consists of a mixture of statistical and fuzzy descriptions, and conventional fuzzy methods cannot combine both. The concept of fuzzy randomness, cf. [54], [59], [63], [82], is one attempt of a combination.

The applications of fuzzy methods in engineering computing are vast. A famous application of fuzzy methods is fuzzy control, cf. [91]. Moreover, most design analyzing methods have their counterparts in the context of fuzzy sets, for instance, fuzzy reliability methods (e.g., [9], [63]), fuzzy differential equations (e.g., [28]), fuzzy finite element methods (e.g., [26], [60], [67]), fuzzy ARMA and other stochastic processes (e.g., [61]).

In fuzzy statistics, i.e., with sample points that are modeled as fuzzy numbers, one can apply statistical methods on non-precise data, cf. [94].

In design optimization fuzzy methods can be used to find clusters of permissible designs with fuzzy clustering methods, e.g., [38], [42]. Seeking the optimal design one can use fuzzy methods to compare different design points of different clusters with respect to some criterion, e.g., weighted distances from design constraints [3], [8], [40].

The following subsection presents a special fuzzy set based method which is highlighted because of its relationship to our approach based on clouds, cf. Section 9.

## 7.1 $\alpha$ -level optimization

The  $\alpha$ -level optimization approach [62] is the most relevant fuzzy set based method for our purposes as it applies also in higher dimensional real-life situations and uses similar techniques as we will use employing the clouds formalism.

The  $\alpha$ -level optimization method combines the  $\alpha$ -cut representation (25) and the extension principle to determine the membership function  $\mu_f$  of a function  $f(\varepsilon)$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , given the membership function  $\mu$  of the variable  $\varepsilon$ . This is achieved

by constructing the  $\alpha$ -cuts  $C_{f_{\alpha_i}}$  belonging to  $\mu_f$  from the  $\alpha$ -cuts  $C_{\alpha_i}$  belonging to  $\mu$ . To this end one solves the optimization problems

$$\min_{\varepsilon \in C_{\alpha_i}} f(\varepsilon), \quad (27)$$

$$\max_{\varepsilon \in C_{\alpha_i}} f(\varepsilon) \quad (28)$$

for different discrete values  $\alpha_i$ . Finally from the solution  $f_{i_*}$  of (27) and  $f_i^*$  of (28) one constructs the  $\alpha$ -cuts belonging to  $\mu_f$  by  $C_{f_{\alpha_i}} = [f_{i_*}, f_i^*]$ .

To simplify the optimization step one assumes sufficiently nice behaving functions  $f$  and computationally nice fuzzy sets, i.e., convex fuzzy sets, typically triangular shaped fuzzy numbers.

In  $n$  dimensions one optimizes over a hypercuboid, obtained by the Cartesian product of the  $\alpha$ -cuts  $C_{\alpha_i} = C_{\alpha_i}^1 \times C_{\alpha_i}^2 \times \dots \times C_{\alpha_i}^n$ , where  $C_{\alpha_i}^j := \{\varepsilon^j \mid \mu^j(\varepsilon^j) \geq \alpha_i\}$ ,  $\mu^j(\varepsilon^j) := \sup_{\varepsilon^k, k \neq j} \mu(\varepsilon)$ ,  $\varepsilon = (\varepsilon^1, \varepsilon^2, \dots, \varepsilon^n)$ . Here one has to assume non-interactivity of the uncertain variables  $\varepsilon^1, \dots, \varepsilon^n$ .

Using a discretization of the  $\alpha$ -levels by a finite choice of  $\alpha_i$  the computational effort for this methods becomes tractable. From (25) one gets a step function for  $\mu_f$  which is usually linearly approximated through the points  $f_{i_*}$  and  $f_i^*$  to generate a triangular fuzzy number.

## 8 Convex methods

Convex methods model uncertainty by so-called **anti-optimization** over convex sets, cf. [4], [19]. Assume that we wish to find the design point  $\theta = (\theta^1, \theta^2, \dots, \theta^{n_o})$  with the minimal design objective function value  $g(\theta, \varepsilon)$ ,  $g : \mathbb{R}^{n_o} \times \mathbb{R}^n \rightarrow \mathbb{R}$  under uncertainty of the vector of input variables  $\varepsilon$ . Also assume that the uncertainty of  $\varepsilon$  is described by a convex set  $\mathcal{C}$ . Anti-optimization means finding the worst-case scenario for a fixed design point  $\theta$  by the solution of an optimization problem of the type

$$\begin{aligned} \max_{\varepsilon} g(\theta, \varepsilon) \\ \text{s.t. } \varepsilon \in \mathcal{C} \end{aligned} \quad (29)$$

The corresponding design optimization problem would be

$$\begin{aligned} \min_{\theta} \max_{\varepsilon} g(\theta, \varepsilon) \\ \text{s.t. } \varepsilon \in \mathcal{C} \\ \theta \in T \end{aligned} \quad (30)$$

where  $T$  is the set of possible selections for the design  $\theta$ . As the inner level of problem (30), i.e., equation (29), maximizes the objective which is sought to be minimized for the design optimization in the outer level (i.e., one seeks the design

with minimal worst-case), the term *anti-optimization* has been proposed for this approach.

Investigating convex regions for the worst-case search is motivated by the fact that in many cases the level sets of probability densities are convex sets, e.g., ellipsoids for normal distributions. In this respect the term convex uncertainty for a random vector  $\varepsilon \in \mathbb{R}^n$  is characterized by a convex set  $\mathcal{C} = \{\varepsilon \mid Q(\varepsilon) \leq c\}$ , where  $Q$  is a quadratic form and  $\varepsilon$  is known to belong to  $\mathcal{C}$  with some confidence. The quadratic form could be, e.g.,  $Q(\varepsilon) = (\varepsilon - m)^T C^{-1}(\varepsilon - m)$  with a vector of nominal values  $m$  and an estimated covariance matrix  $C$ .

Once one has a description by convex uncertainty one can apply optimization methods which can make convex methods applicable even in higher dimensions.

It should be remarked that this particular idea is one of the inspirations for the potential clouds concept, see the next section, where the potential function will be constructed to have convex level sets.

## 9 Potential clouds

This section gives an intuitive introduction of uncertainty representation by means of clouds [71], inspired by and combining ideas from  $p$ -boxes, random and fuzzy sets, convex methods, interval and optimization methods. We will learn in this section about the natural approach that has led to use clouds for uncertainty modeling, handling incomplete information in higher dimensions, and to weave the methodology into an optimization problem formulation similar to (30).

The goal is to construct confidence regions in which we should be able to search for worst-case scenarios via optimization techniques. The construction should be possible on the basis of scarce, high-dimensional data, incomplete information, unformalized knowledge and information updates. As mentioned in previous sections, in lower dimensions and provided real empirical data one has powerful tools, like KS, e.g., to bound the CDF of a random variable  $X$ . What could one do to tackle the same problems for higher dimensional random vectors  $\varepsilon \in \mathbb{R}^n$  with little or no information available? To generate data we will first simulate a data set and modify it with respect to the available uncertainty information. To reduce the dimensionality of the problem we will use a potential function  $V : \mathbb{R}^n \rightarrow \mathbb{R}$ . We will bound the CDF of  $V(\varepsilon)$  using KS as in the one-dimensional case (like a  $p$ -box on  $V(\varepsilon)$ , cf. Section 4). From the bounds on the CDF of  $V(\varepsilon)$  we get lower and upper confidence regions for  $V(\varepsilon)$ , and finally lower and upper confidence regions for  $\varepsilon$  as level sets of  $V$ .

Assume that we have a lower bound  $\underline{\alpha}$  and an upper bound  $\bar{\alpha}$  for the CDF  $F$  of  $V(\varepsilon)$ ,  $\underline{\alpha}$  continuous from the left and monotone,  $\bar{\alpha}$  continuous from the right and monotone. Then we find nested lower and upper confidence regions for  $\varepsilon$  by:  $\underline{C}_\alpha := \{x \in \mathbb{R}^n \mid V(x) \leq \underline{V}_\alpha\}$  if  $\underline{V}_\alpha := \min\{V_\alpha \in \mathbb{R} \mid \bar{\alpha}(V_\alpha) = \alpha\}$  exists, and  $\underline{C}_\alpha := \emptyset$  otherwise; analogously  $\bar{C}_\alpha := \{x \in \mathbb{R}^n \mid V(x) \leq \bar{V}_\alpha\}$  if  $\bar{V}_\alpha := \max\{V_\alpha \in \mathbb{R} \mid \underline{\alpha}(V_\alpha) = \alpha\}$  exists, and  $\bar{C}_\alpha := \mathbb{R}^n$  otherwise.

The regions  $\underline{C}_\alpha$  and  $\bar{C}_\alpha$  are lower and upper confidence regions in the following

sense: the region  $\underline{C}_\alpha$  contains at most a fraction of  $\alpha$  of all possible values of  $\varepsilon$  in  $\mathbb{R}^n$ , since  $\Pr(\varepsilon \in \underline{C}_\alpha) \leq \Pr(\bar{\alpha}(V(\varepsilon)) \leq \alpha) \leq \Pr(F(V(\varepsilon)) \leq \alpha) = \alpha$ ; analogously  $\bar{C}_\alpha$  contains at least a fraction of  $\alpha$  of all possible values of  $\varepsilon$  in  $\mathbb{R}^n$ . Generally holds  $\underline{C}_\alpha \subseteq \bar{C}_\alpha$ .

The interval-valued mapping  $x \rightarrow [\underline{\alpha}(V(x)), \bar{\alpha}(V(x))]$  is called a **potential cloud**.

Note that potential clouds extend the  $p$ -box concept to the multivariate case without the exponential growth of work in the conventional  $p$ -box approach. From the fact that we construct a  $p$ -box on  $V(\varepsilon)$  one can also see the relation to DS structures generated from  $p$ -boxes as in (21), with  $A_i = \bar{C}_{\alpha_i} \setminus \underline{C}_{\alpha_i}$ . Thus the focal sets are determined by the level sets of  $V$ . To see an interpretation in terms of fuzzy sets one may consider  $\underline{C}_\alpha, \bar{C}_\alpha$  as  $\alpha$ -cuts of a multi-dimensional interval valued membership function defined by  $\underline{\alpha}$  and  $\bar{\alpha}$ . However, clouds allow for probabilistic statements, so they become a more powerful tool in the estimation of failure probabilities.

The potential clouds approach not only helps us to overcome the curse of dimensionality in real-life applications, but also it turns out to enable a flexible uncertainty representation. It can process incomplete knowledge of different kinds and allows for an adaptive interaction between the uncertainty elicitation and the optimization phase, reducing the incompleteness of epistemic information via information updating. The adaptive step is realized by a modification of the shape of  $V$  in a graphical user interface. This is a unique feature in higher dimensions.

To illustrate how unformalized knowledge can be provided by clouds assume that a set of data points is given, but no formal information about the probability distribution is available, in particular no correlation information. Frequently an expert still has some knowledge about the dependence between the uncertain variables involved, and he may be able to provide linear constraints as shown in Figure 3. The lower and upper confidence regions constructed with clouds then become polyhedra, cf. Figure 4. We also see that these regions reasonably approximate the confidence regions in case that we knew the correlations exactly.

The basic concept of embedding our approach in a design optimization problem is as follows. The designing expert provides an underlying system model – e.g., given as a black box model – and all currently available uncertainty information on the input variables of the model. The information is processed to generate a cloud that provides a nested collection of regions of relevant scenarios parameterized by a confidence level  $\alpha$ . Thus we produce safety constraints for the optimization. The optimization minimizes a certain objective function (e.g., cost, mass) subject to the safety constraints to account for the robustness of the design, and subject to the functional constraints which are represented by the system model. The results of the optimization, i.e., the automatically found optimal design point and the worst-case analysis, are returned to the expert, who is given an interactive possibility to provide additional uncertainty information afterwards and rerun the procedure. For further details on the construction of potential clouds and cloud based design optimization the interested reader is referred to [32] and [33].

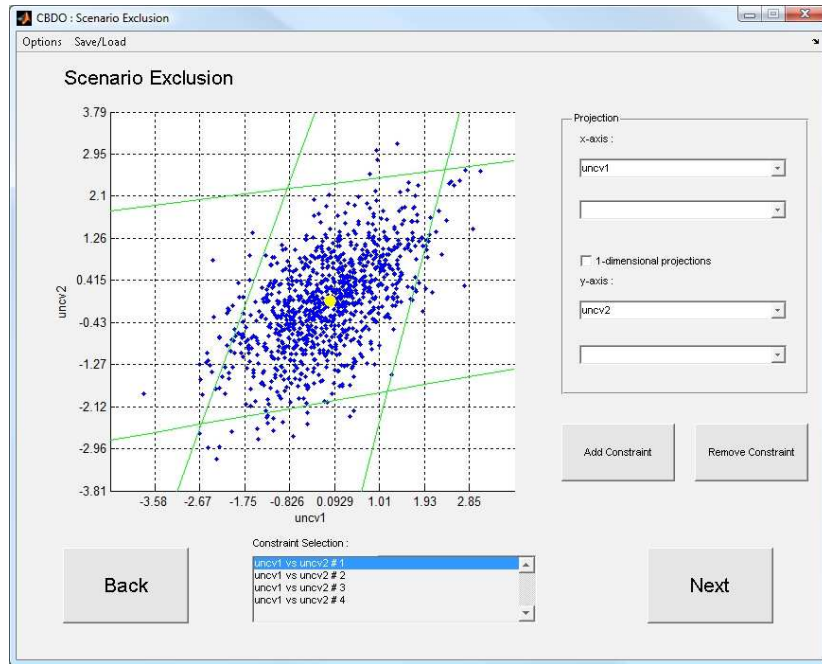


Figure 3: Two random variables  $\varepsilon^1, \varepsilon^2$  with non zero correlation. The linear constraints model the unformalized knowledge of the expert about the dependence of the variables.

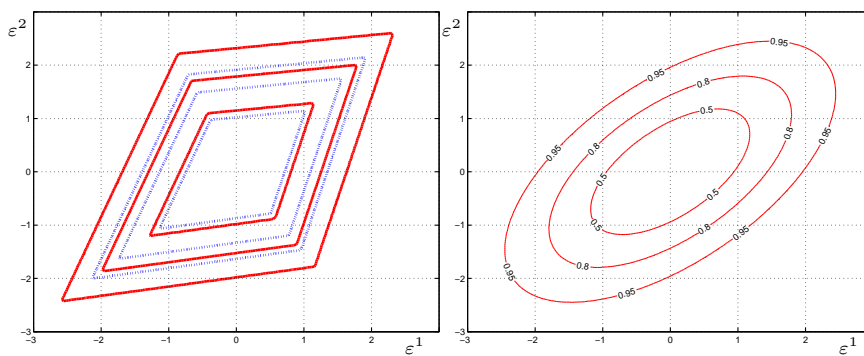


Figure 4: On the left, plotted with dotted and solid lines respectively, are the lower and upper confidence regions for  $\alpha = 50\%, 80\%, 95\%$  of a 2-dimensional random variable  $(\varepsilon^1, \varepsilon^2)$  belonging to a polyhedral potential cloud. On the right the corresponding confidence regions if the correlation was exactly known.

## Acknowledgements

I would like to thank Arnold Neumaier who has significantly contributed to the creation of this paper with various comments. Also I would like to thank the anonymous reviewers for their fruitful remarks.

## References

- [1] Alexandrov, N.M. and Hussaini, M.Y. Multidisciplinary design optimization: State of the art. In *Proceedings of the ICASE/NASA Langley Workshop on Multidisciplinary Design Optimization*, Hampton, Virginia, USA, 1997.
- [2] Aregui, A. and Denœux, T. Constructing predictive belief functions from continuous sample data using confidence bands. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 11–19, Prague, Czech Republic, 2007.
- [3] Beer, M., Liebscher, M., and Möller, B. Structural design under fuzzy randomness. In *Proceedings of the NSF Workshop on Reliable Engineering Computing*, pages 215–234, Savannah, Georgia, USA, 2004.
- [4] Ben-Haim, Y. and Elishakoff, I. *Convex Models of Uncertainty in Applied Mechanics*. Elsevier, 1990.
- [5] Berleant, D., Ferson, S., Kreinovich, V., and Lodwick, W.A. Combining interval and probabilistic uncertainty: Foundations, algorithms, challenges – an overview. In *Proceedings of the 4th International Symposium on Imprecise Probabilities and Their Applications*, Pittsburgh, Pennsylvania, USA, 2005.
- [6] Berleant, D. and Xie, L. An interval-based tool for verified arithmetic on random variables of unknown dependency. Manuscript, 2005.
- [7] Bernardini, A. *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization*, chapter What are random and fuzzy sets and how to use them for uncertainty modelling in engineering systems?, pages 63–125. Springer, 1999.
- [8] Chen, S.H. Ranking fuzzy numbers with maximizing set and minimizing set. *Fuzzy Sets and Systems*, 17:113–129, 1985.
- [9] Cheng, C.H. and Mon, D.L. Fuzzy system reliability analysis by interval of confidence. *Fuzzy Sets and Systems*, 56(1):29–35, 1993.
- [10] Couso, I., Moral, S., and Walley, P. Examples of independence for imprecise probabilities. In *Proceedings of the 1st International Symposium on Imprecise Probabilities and Their Applications*, pages 121–130, Ghent, Belgium, 1999.
- [11] Cozman, F.G. Credal networks. *Artificial Intelligence*, 120(2):199–233, 2000.



- [12] Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38(2):325–339, 1967.
- [13] Denceux, T. Constructing belief functions from sample data using multinomial confidence regions. *International Journal of Approximate Reasoning*, 42(3):228–252, 2006.
- [14] Der Kiureghian, A. and Dakessian, T. Multiple design points in first and second-order reliability. *Structural Safety*, 20(1):37–49, 1998.
- [15] Destercke, S., Dubois, D., and Chojnacki, E. Relating practical representations of imprecise probabilities. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 155–163, Prague, Czech Republic, 2007.
- [16] Dubois, D. and Prade, H. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press, 1986.
- [17] Eldred, M.S., Brown, S.L., Adams, B.M., Dunlavy, D.M., Gay, D.M., Swiler, L.P., Giunta, A.A., Hart, W.E., Watson, J.-P., Eddy, J.P., Griffin, J.D., Hough, P.D., Kolda, T.G., Martinez-Canales, M.L., and Williams, P.J. DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 4.0 Users Manual. Sand Report SAND2006-6337, Sandia National Laboratories, 2006.
- [18] Elishakoff, I. *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization*, chapter What may go wrong with probabilistic methods?, pages 265–283. Springer, 1999.
- [19] Elishakoff, I. *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization*, chapter Are probabilistic and anti-optimization approaches compatible?, pages 263–355. Springer, 1999.
- [20] Ferson, S. What Monte Carlo methods cannot do. *Human and Ecological Risk Assessment*, 2:990–1007, 1996.
- [21] Ferson, S. *Ramas Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Lewis Publishers, U.S., 2002.
- [22] Ferson, S., Ginzburg, L., and Akcakaya, R. Whereof one cannot speak: When input distributions are unknown. *Risk Analysis*, 1996. In press, available online at: <http://www.ramas.com/whereof.pdf>.
- [23] Ferson, S., Ginzburg, L., Kreinovich, V., Longpre, L., and Aviles, M. Exact bounds on finite populations of interval data. *Reliable Computing*, 11(3):207–233, 2005.

- [24] Ferson, S., Ginzburg, L., Kreinovich, V., and Lopez, J. Absolute bounds on the mean of sum, product, etc.: A probabilistic extension of interval arithmetic. In *Extended Abstracts of the 2002 SIAM Workshop on Validated Computing*, pages 70–72, Toronto, Canada, 2002.
- [25] Ferson, S., Kreinovich, V., Ginzburg, L., Myers, D.S., and Sentz, K. Constructing probability boxes and Dempster-Shafer structures. Sand Report SAND2002-4015, Sandia National Laboratories, 2003. Available on-line at <http://www.sandia.gov/epistemic/Reports/SAND2002-4015.pdf>.
- [26] Fetz, T. Finite element method with fuzzy parameters. In *Proceedings of the IMACS Symposium on Mathematical Modelling*, volume 11, pages 81–86, Vienna, Austria, 1997.
- [27] Fetz, T. Sets of joint probability measures generated by weighted marginal focal sets. In *Proceedings of the 2nd International Symposium on Imprecise Probabilities and Their Applications*, pages 171–178, Maastricht, The Netherlands, 2001.
- [28] Fetz, T., Oberguggenberger, M., Jager, J., Koll, D., Krenn, G., Lessmann, H., and Stark, R.F. Fuzzy models in geotechnical engineering and construction management. *Computer-Aided Civil and Infrastructure Engineering*, 14(2):93–106, 1999.
- [29] Fetz, T., Oberguggenberger, M., and Pittschmann, S. Applications of possibility and evidence theory in civil engineering. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(3):295–309, 2000.
- [30] Filho, R.S., Cozman, F.G., Trevizan, F.W., de Campos, C.P., and de Barros, L.N. Multilinear and integer programming for Markov decision processes with imprecise probabilities. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 395–403, Prague, Czech Republic, 2007.
- [31] Fuchs, M., Girimonte, D., Izzo, D., and Neumaier, A. *Robust intelligent systems*, chapter Robust and automated space system design, pages 251–272. Springer, 2008.
- [32] Fuchs, M. and Neumaier, A. Autonomous robust design optimization with potential clouds. *International Journal of Reliability and Safety, Special Issue on Reliable Engineering Computing*, 2008. Accepted, preprint available on-line at: <http://www.martin-fuchs.net/publications.php>.
- [33] Fuchs, M. and Neumaier, A. Potential based clouds in robust design optimization. *Journal of Statistical Theory and Practice, Special Issue on Imprecision*, 2008. Accepted, preprint available on-line at: <http://www.martin-fuchs.net/publications.php>.

- [34] Grandy, W.T. and Schick, L.H. *Maximum Entropy and Bayesian Methods*. Springer, 1990.
- [35] Haenni, R. Climbing the hills of compiled credal networks. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 213–221, Prague, Czech Republic, 2007.
- [36] Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [37] Hohenbichler, M. and Rackwitz, R. Improvement of second-order reliability estimates by importance sampling. *Journal of Engineering Mechanics*, 114(12):2195–2199, 1988.
- [38] Höppner, F., Klawonn, F., Kruse, R., and Runkler, T. *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*. Wiley, 1999.
- [39] Huyer, W. and Neumaier, A. SNOBFIT – Stable Noisy Optimization by Branch and Fit. *ACM Transactions on Mathematical Software*, 35(2), 2008. Article 9, 25 pages.
- [40] Jain, R. Decision making in the presence of fuzzy variables. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(10):698–703, 1976.
- [41] Jirousek, R., Vejnarova, J., and Daniel, M. Compositional models of belief functions. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 243–251, Prague, Czech Republic, 2007.
- [42] Kaufman, L. and Rousseeuw, P.J. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley, 1990.
- [43] Kaymaz, I. and Marti, K. Reliability-based design optimization for elastoplastic mechanical structures. *Computers & Structures*, 85(10):615–625, 2007.
- [44] Kijawatworawet, W., Pradlwarter, H.J., and Schueller, G.I. Structural reliability estimation by adaptive importance directional sampling. In *Structural Safety and Reliability: Proceedings of ICOSSAR'97*, pages 891–897, Balkema, 1998.
- [45] Koch, P.N., Simpson, T.W., Allen, J.K., and Mistree, F. Statistical approximations for multidisciplinary optimization: The problem of size. *Special Issue on Multidisciplinary Design Optimization of Journal of Aircraft*, 36(1):275–286, 1999.
- [46] Kolmogorov, A. Confidence limits for an unknown distribution function. *The Annals of Mathematical Statistics*, 12(4):461–463, 1941.

- [47] Kozine, I. and Krymsky, V. Enhancement of natural extension. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 253–261, Prague, Czech Republic, 2007.
- [48] Kreinovich, V. *Random Sets: Theory and Applications*, chapter Random sets unify, explain, and aid known uncertainty methods in expert systems, pages 321–345. Springer, 1997.
- [49] Kreinovich, V. Probabilities, intervals, what next? optimization problems related to extension of interval computations to situations with partial information about probabilities. *Journal of Global Optimization*, 29(3):265–280, 2004.
- [50] Kreinovich, V., Beck, J., Ferregut, C., Sanchez, A., Keller, G.R., Averill, M., and Starks, S.A. Monte-Carlo-type techniques for processing interval uncertainty, and their engineering applications. In *Proceedings of the NSF Workshop on Reliable Engineering Computing*, pages 139–160, Savannah, Georgia, USA, 2004.
- [51] Kreinovich, V. and Ferson, S. A new Cauchy-based black-box technique for uncertainty in risk analysis. *Reliability Engineering & System Safety*, 85(1–3):267–279, 2004.
- [52] Kreinovich, V., Ferson, S., and Ginzburg, L. Exact upper bound on the mean of the product of many random variables with known expectations. *Reliable Computing*, 9(6):441–463, 2003.
- [53] Kreinovich, V. and Trejo, R. *Handbook of Randomized Computing*, chapter Error estimations for indirect measurements: randomized vs. deterministic algorithms for 'black-box' programs, pages 673–729. Kluwer, 2001.
- [54] Kwakernaak, H. Fuzzy random variables – ii: Algorithms and examples for the discrete case. *Information Sciences*, 17:253–278, 1979.
- [55] Makino, K. and Berz, M. Efficient control of the dependency problem based on taylor model methods. *Reliable Computing*, 5(1):3–12, 1999.
- [56] Marti, K. and Kaymaz, I. Reliability analysis for elastoplastic mechanical structures under stochastic uncertainty. *Zeitschrift fr Angewandte Mathematik und Mechanik*, 86(5):358–384, 2006.
- [57] McKay, M.D., Conover, W.J., and Beckman, R.J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 221:239–245, 1979.
- [58] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.

- [59] Möller, B. and Beer, M. *Fuzzy Randomness: Uncertainty in Civil Engineering and Computational Mechanics*. Springer-Verlag Berlin Heidelberg, 2004.
- [60] Möller, B., Beer, M., Graf, W., and Sickert, J. U. Fuzzy finite element method and its application. In *Trends in Computational Structural Mechanics*, pages 529–538, Barcelona, Spain, 2001.
- [61] Möller, B., Beer, M., and Reuter, U. Theoretical basics of fuzzy-randomness – application to time series with fuzzy data. In *Safety and Reliability of Engineering Systems and Structures: Proceedings of the 9th International Conference on Structural Safety and Reliability*, Rome, Italy, 2005.
- [62] Möller, B., Graf, W., and Beer, M. Fuzzy structural analysis using  $\alpha$ -level optimization. *Computational Mechanics*, 26(6):547–565, 2000.
- [63] Möller, B., Graf, W., and Beer, M. Fuzzy probabilistic method and its application for the safety assessment of structures. In *Proceedings of the European Conference on Computational Mechanics*, Cracow, Poland, 2001.
- [64] Moore, R.E. *Methods and Applications of Interval Analysis*. Society for Industrial & Applied Mathematics, 1979.
- [65] Mourelatos, Z.P. and Liang, J. An efficient unified approach for reliability and robustness in engineering design. In *Proceedings of the NSF Workshop on Reliable Engineering Computing*, pages 119–138, Savannah, Georgia, USA, 2004.
- [66] Mourelatos, Z.P. and Zhou, J. A design optimization method using evidence theory. *Journal of Mechanical Design*, 128(4):901–908, 2006.
- [67] Muhanna, R.L. and Mullen, R.L. Formulation of fuzzy finite element methods for mechanics problems. *Computer-Aided Civil and Infrastructure Engineering*, 14(2):107–117, 1999.
- [68] Muhanna, R.L. and Mullen, R.L. Uncertainty in mechanics problems – interval-based approach. *Journal of Engineering Mechanics*, 127(6):557–566, 2001.
- [69] Nataf, A. Determination des distributions de probabilités dont les marges sont données. *Comptes Rendus de l'Académie des Sciences*, 225:42–43, 1962.
- [70] Neumaier, A. *Interval Methods for Systems of Equations*. Cambridge University Press, 1990.
- [71] Neumaier, A. Clouds, fuzzy sets and probability intervals. *Reliable Computing*, 10(4):249–272, 2004. Available on-line at:  
<http://www.mat.univie.ac.at/~neum/ms/cloud.pdf>.

- [72] Neumaier, A. Uncertainty modeling for robust verifiable design. Slides, 2004. Available on-line at: <http://www.mat.univie.ac.at/~neum/ms/unclides.pdf>.
- [73] Neumaier, A., Fuchs, M., Dolejsi, E., Csendes, T., Dombi, J., Banhelyi, B., and Gera, Z. Application of clouds for modeling uncertainties in robust space system design. ACT Ariadna Research ACT-RPT-05-5201, European Space Agency, 2007.
- [74] Nguyen, H. Fuzzy sets and probability. *Fuzzy Sets and Systems, Special Issue on Fuzzy Sets: Where do we stand? Where do we go?*, 90(2):129–132, 1997.
- [75] Oberguggenberger, M. and Fellin, W. Assessing the sensitivity of failure probabilities: a random set approach. In *Safety and Reliability of Engineering Systems and Structures: Proceedings of the 9th International Conference on Structural Safety and Reliability*, pages 1755–1760, Rome, Italy, 2005.
- [76] Oberguggenberger, M. and Fellin, W. Reliability bounds through random sets: Non-parametric methods and geotechnical applications. *Computers and Structures*, 86(10):1093–1101, 2008.
- [77] Oberguggenberger, M., King, J., and Schmelzer, B. Imprecise probability methods for sensitivity analysis in engineering. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 317–325, Prague, Czech Republic, 2007.
- [78] Papadrakakis, M. and Lagaros, N.D. Reliability-based structural optimization using neural networks and Monte Carlo simulation. *Computer Methods in Applied Mechanics and Engineering*, 191(32):3491–3507, 2002.
- [79] Pownuk, A. Calculation of displacement in elastic and elastic-plastic structures with interval parameters. In *Proceedings of the 33rd Solid Mechanics Conference*, pages 5–9, Zakopane, Poland, 2000.
- [80] Pownuk, A. General interval FEM program based on sensitivity analysis method. In *Proceedings of the 3rd International Workshop on Reliable Engineering Computing*, pages 397–428, Savannah, Georgia, USA, 2008.
- [81] Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. *Numerical Recipes in C*. Cambridge University Press, 2nd edition, 1992.
- [82] Puri, M. and Ralescu, D. *Madan Lal Puri Selected Collected Works, Volume 3: Time Series, Fuzzy Analysis and Miscellaneous Topics*, chapter Fuzzy Random Variables. Brill Academic Pub, 2003.
- [83] Rackwitz, R. Reliability analysis – a review and some perspectives. *Structural Safety*, 23(4):365–395, 2001.
- [84] Rosenblatt, M. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 23(3):470–472, 1952.

- [85] Saad, E. Structural optimization based on evolution strategy. In *Advanced Computational Methods in Structural Mechanics*, pages 266–280, Barcelona, Spain, 1996.
- [86] Shafer, G. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [87] Shannon, C.E. and Weaver, W. *The Mathematical Theory of Communication*. University of Illinois Press, 1949.
- [88] Skulj, D. *Soft Methods for Integrated Uncertainty Modelling*, chapter Finite Discrete Time Markov Chains with Interval Probabilities, pages 299–306. Springer, 2006.
- [89] Skulj, D. Regular finite Markov chains with interval probabilities. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 405–413, Prague, Czech Republic, 2007.
- [90] Sugeno, M. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.
- [91] Sugeno, M. An introductory survey of fuzzy control. *Information Science*, 36:59–83, 1985.
- [92] ter Marten, E.J.W., Doorn, T.S., Croon, J.A., Bargagli, A., di Bucchianico, A., and Wittich, O. Importance sampling for high speed statistical Monte-Carlo simulations. NXP Technical Note NXP-TN-2007-00238, NXP Semiconductors, 2007.
- [93] Utkin, L. and Destercke, S. Computing expectations with p-boxes: two views of the same problem. In *Proceedings of the 5th International Symposium on Imprecise Probability: Theories and Applications*, pages 435–443, Prague, Czech Republic, 2007.
- [94] Viertl, R. *Statistical Methods for Non-Precise Data*. CRC Press, 1996.
- [95] Walley, P. *Statistical Reasoning with Imprecise Probability*. Chapman and Hall, 1991.
- [96] Walley, P. Measures of uncertainty in expert systems. *Artificial Intelligence*, 83(1):1–58, 1996.
- [97] Walley, P. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2–3):125–148, 2000.
- [98] Weisstein, E.W. Cauchy distribution. MathWorld – A Wolfram Web Resource, 2008. Available on-line at: <http://mathworld.wolfram.com/CauchyDistribution.html>.

- [99] Williamson, R.C. *Probabilistic Arithmetic*. PhD thesis, University of Queensland, 1989.
- [100] Williamson, R.C. and Downs, T. Probabilistic arithmetic. I. numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning*, 4(2):89–158, 1990.
- [101] Zadeh, L.A. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.
- [102] Zadeh, L.A. Similarity relations and fuzzy orderings. *Information Sciences*, 3(2):177–200, 1971.
- [103] Zhou, J. and Mourelatos, Z.P. Design under uncertainty using a combination of evidence theory and a Bayesian approach. In *Proceedings of the 3rd International Workshop on Reliable Engineering Computing*, pages 171–198, Savannah, Georgia, USA, 2008.