Potential based clouds in robust design optimization

Martin Fuchs, *University of Vienna, Faculty of Mathematics,*
Nordbergstr. 15, 1090 Wien, Austria.
Email: martin.fuchs@univie.ac.at
Arnold Neumaier, *University of Vienna, Faculty of Mathematics,*
Nordbergstr. 15, 1090 Wien, Austria.
Email: arnold.neumaier@univie.ac.at

Received: *Revised:*

Abstract

Robust design optimization methods applied to real life problems face some major difficulties: how to deal with the estimation of probability densities when data are sparse, how to cope with high dimensional problems and how to use valuable information in the form of unformalized expert knowledge. In this paper we introduce in detail the *clouds* formalism as a means to process available uncertainty information reliably, even if limited in amount and possibly lacking a formal description. This enables a worst-case analysis with confidence regions of relevant scenarios which can be involved in an optimization problem formulation for robust design.

Keywords: clouds; potential clouds; robust design; design optimization; confidence regions; uncertainty modeling.

---

1 Introduction

Robust design optimization is the art of safeguarding reliably against uncertain perturbations while seeking an optimal design point. In every design process an engineer faces the task to qualify the object he has designed to be robust. That means the design should not only satisfy given requirements on functionalities, but should also work under uncertain, adverse conditions that may show up during employment of the designed object.

Hence the process of robust design optimization demands both the search of an optimal design with respect to a given design objective, and an appropriate method of handling uncertainties. In particular for early design phases, it is frequent engineering practice that experts first assign bounds in terms of intervals or safety margins, and then

decide iteratively to refine or coarsen the bounds on the basis of their judgments. These intervals or safety margins are propagated within the whole optimization process. Thus the design arising from this process is supposed to include robustness intrinsically. Note that the assessment of robustness is exclusively based on expert knowledge of the engineers who assign and refine the intervals. There is no quantification of reliability, no formal worst-case analysis involved.

Several methods exist to approach reliability quantification from a rigorous mathematical background, originating from classical probability theory, statistics, or fuzzy theory. However, real life applications of many methods disclose various problems. One of the most prominent is probably the fact that the dimension of many uncertain real life scenarios is very high. This can cause severe computational effort, even given the complete knowledge about the multivariate probability distributions of the uncertainties, also famous as the curse of dimensionality, Koch et al. (1999). High dimensional problems approached from the field of imprecise probabilities are investigated, e.g., by Oberguggenberger et al. (2007) with sensitivity analysis methods. Moreover, if the amount of available uncertainty information is very limited, well-known current methods using classical probabilities either do not apply at all, or are endangered to critically underestimate error probabilities, Ferson (1996). A simplification of the uncertainty model, e.g., a reduction of the problem to an interval analysis after assigning intervals to the uncertainties as described before (e.g., so called 3 σ boxes), entails a loss of valuable information which would actually be available, maybe only unformalized, but not at all considered in the uncertainty model. We find a fuzzy theory approach for engineering applications, e.g., by Ross (1995), simulation techniques, e.g., by McCormick and Olds (2002). A method based on possibility theory has been developed by Dubois and Prade (1986), and recently Dubois and Prade (2005) have given formal links and comparisons of various methods. Some works go into the arithmetic operations on random variables to bound the probability distributions of functions of random variables, cf. Berleant and Cheng (1998), Ferson (2002), Williamson (1989), recently reviewed by Regan et al. (2004). There are also attempts to generalize aspects of the different uncertainty approaches and put them into one framework, e.g., with random sets, cf. Kreinovich (1997). Different approaches to design optimization can be studied, e.g., in works of Alexandrov and Hussaini (1997), Lewis and Mistree (1997).

In this paper, we start from the theoretical basis of clouds to give practical means for multi-dimensional uncertainty modeling. We shortly introduce how to involve the model in an optimization problem formulation. It will be shown that clouds can process limited amounts of stochastic information in an understandable and computationally attractive way, even in higher dimensions, in order to perform a reliable worst-case analysis, reasonably safeguarded against perturbations that result from unmodeled and/or unavailable information. In higher dimensions full probabilistic models need to estimate high dimensional distributions for which rarely sufficient data are available. However, estimation of distributions in one dimension is simple and well-understood. We will transfer the higher dimensional case to the univariate case via the concept of potential clouds. Since the strength of our new methodology lies especially in the application to real life problems with a very limited amount of information available, we focus in particular on problem statements arising in early design phases where today's

methods of handling the limited information are very immature, i.e., no tractable solutions have been proposed for such problems so far. On the one hand, the information is considered to be available as bounds or marginal probability distributions on the uncertain variables, without any formal correlation information, as in many applications. On the other hand, unformalized expert knowledge will be captured to improve the uncertainty model adaptively by adding dependency constraints to exclude scenarios deemed irrelevant. Unformalized knowledge is available, e.g., if an expert does not know correlations exactly, but can formulate a statement like 'if variable a is high then variable b cannot be low'. Thus he is able to exclude irrelevant scenarios, although he is unable to give a formal description. The information can also be provided as real sample data set, if available.

This paper is organized as follows. First, we follow the clouds formalism from Neumaier (2004) in Section 2, picking those ideas that are most attractive for the handling of high dimensional uncertainties and investigating the special case of potential clouds. We will see how they help to cope with dimensionality issues. A short introduction how clouds can be involved in an optimization problem formulation for robust design is given in Section 3. In Section 4 we will learn how to interpret approximations and bounds on cumulative distribution functions in terms of clouds. Some considerations about suitable potential functions can be found in Section 5. Section 6 concludes our studies.

We thank the anonymous reviewers for their very valuable comments.

2 The potential clouds approach

The clouds formalism serves as the central theoretical background for our uncertainty handling:

2.1 General framework

We start with the formal definition of clouds and introduce the notation. Afterwards we restate the central results from Neumaier (2004), Sections 3 and 4, that will be relevant for our studies later on.

Let $\varepsilon \in \mathbb{M} \subseteq \mathbb{R}^n$ be an n-dimensional random vector. A **cloud** is a mapping $\chi(\varepsilon) = [\underline{\chi}(\varepsilon), \overline{\chi}(\varepsilon)]$, where $\chi(\varepsilon)$ is a nonempty, closed and bounded interval $\in [0,1]$ for all $\varepsilon \in \mathbb{M}$, and

$$]0,1[\subseteq \bigcup_{\varepsilon \in \mathbb{M}} \chi(\varepsilon) \subseteq [0,1]. \tag{2.1}$$

We call $\underline{\chi}(\varepsilon)$ the **lower level** and $\overline{\chi}(\varepsilon)$ the **upper level** of the cloud, and denote $\overline{\chi}(\varepsilon) - \underline{\chi}(\varepsilon)$ as the **width** of the cloud χ. A cloud is called **thin** if it has width 0.

2.1 Remark. Clouds where $\underline{\chi}(\varepsilon) \equiv 0$ are equivalent to fuzzy sets, cf. Neumaier (2004). They can also be considered as a special case of random sets, cf. Destercke et al. (2007).

We say that a random vector ε **belongs to** a cloud χ over $\mathbb{M}$, if

$$\Pr(\overline{\chi}(\varepsilon) \leq y) \leq y \leq \Pr(\underline{\chi}(\varepsilon) < y) \tag{2.2}$$

for all $y \in [0,1]$.

A cloud is called **discrete** if $\underline{\chi}$ and $\overline{\chi}$ only take finitely many different values. It can be shown that discrete clouds can be constructed from samples of discrete or discretized continuous probability distributions in low dimensions via histograms, cf. Neumaier (2004) Theorem 3.1, and in principle can approximate arbitrary distributions arbitrarily well. As histograms and the related discrete clouds are having computational problems in higher dimensions, we define continuous clouds that will be much more important for our purposes: A cloud is called **continuous** if the lower level $\underline{\chi}$ and the upper level $\overline{\chi}$ are continuous functions.

There exists a close relationship between thin continuous 1-dimensional clouds and CDFs of real univariate random variables, cf. Neumaier (2004) Proposition 4.1.

2.2 Proposition. *Let $F_\varepsilon(x) = \Pr(\varepsilon \leq x)$ be the CDF of the 1-dimensional random variable ε, then $\chi(x) := F_\varepsilon(x)$ defines a thin cloud and ε belongs to χ, i.e., for the thin case $\Pr(\chi(x) \leq y) = y, y \in [0,1]$, if x has the same distribution as ε.*

Proof. From the definition of a CDF it follows that $\chi := F_\varepsilon$ satisfies the conditions that define a cloud. Due to the fact that $F_\varepsilon(x)$ is uniformly distributed if x has the same CDF as ε we have $\Pr(\chi(x) \leq y) = \Pr(F_\varepsilon(x) \leq y) = y$. □

2.2 Potential clouds

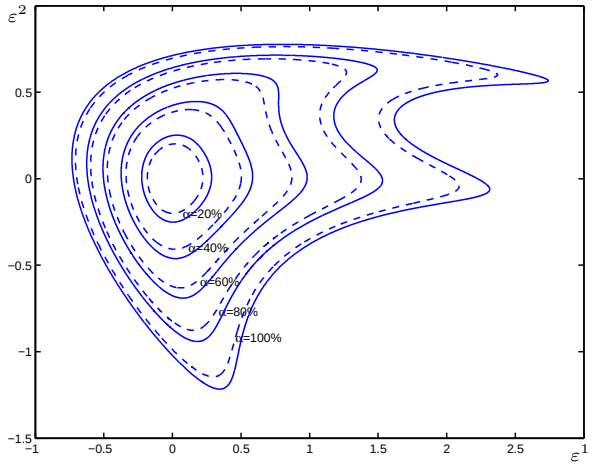
CDFs are well known from probability theory. In particular the univariate case is very handy, computationally appealing and intuitively understandable. However, we want to deal with significantly higher dimensions than 1. This leads to the idea to construct continuous clouds from user-defined potential functions $V : \mathbb{M} \to \mathbb{R}$. Let $\varepsilon \in \mathbb{R}^n$. Though the CDF of ε often cannot be estimated for high n due to a lack of available data, the random variable $V := V(\varepsilon)$ is univariate and its CDF can be easily approximated by an empirical CDF. The idea is to construct a **potential cloud** from an interval-valued function χ of the potential function V, i.e., $\chi \circ V : \mathbb{M} \to [a,b]$, where $[a,b]$ is an interval in $[0,1]$. How this potential can be constructed in practice will be discussed in Section 5.

2.3 Proposition. *The mapping χ defined by*

$$\chi(x) := [\underline{\alpha}(V(x)), \overline{\alpha}(V(x))], \tag{2.3}$$

defines a cloud and ε belongs to χ, suppose that $x \in \mathbb{M}$, $\underline{\alpha}(a) \leq \Pr(V(\varepsilon) < a)$, $\overline{\alpha}(a) \geq \Pr(V(\varepsilon) \leq a)$, $\underline{\alpha}$, $\overline{\alpha}$ continuous from the right and monotone, $\varepsilon \in \mathbb{M}$ a random vector, $a \in \mathbb{R}$, V bounded below.

Figure 2.1: Nested confidence regions in two dimensions for confidence levels $\alpha = 0.2, 0.4, 0.6, 0.8, 1$. The lower confidence regions $\underline{C}_\alpha$ plotted with dashed lines, the upper confidence regions $\overline{C}_\alpha$ with solid lines.



Proof. Obviously (2.1) holds for χ. One has to show (2.2). Let $y \in [0,1]$. Assume $\exists x \in \mathbb{R}$ with $\overline{\alpha}(x) = y$. By construction of $\overline{\alpha}$ and x we get $\Pr(\overline{\chi}(\varepsilon) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq \overline{\alpha}(x)) = \Pr(V(\varepsilon) \leq x) \leq \overline{\alpha}(x) = y$. If $\nexists x \in \mathbb{R}$ with $\overline{\alpha}(x) = y$ and $\overline{\alpha}(x) > y$ for all x. Then $\Pr(\overline{\chi}(\varepsilon) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq y) = 0 \leq y$. Otherwise if $\nexists x \in \mathbb{R}$ with $\overline{\alpha}(x) = y$, then $\exists y' \in [0,1], x, h \in \mathbb{R}$ with $\overline{\alpha}(x+h) = y' > y$, and $\overline{\alpha}(x) \leq y$ as $\overline{\alpha}$ is continuous from the right. This yields $\Pr(\overline{\chi}(\varepsilon) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq \overline{\alpha}(x+h)) = \Pr(V(\varepsilon) \leq x+h) = \overline{\alpha}(x+h) \leq y$ if $h \to 0$. Hence $\Pr(\overline{\chi}(\varepsilon) \leq y) \leq y$. We get $\Pr(\underline{\chi}(\varepsilon) < y) \geq y$ analogously, so χ fulfills (2.2). □

Proposition 2.3 leads us to an important interpretation in terms of confidence regions for ε, as it tells us that it is sufficient to find an appropriate bounding $\underline{\alpha}$, $\overline{\alpha}$ on the CDF of $V(\varepsilon)$ to construct a potential cloud. This can be achieved, e.g., by KS statistics Kolmogoroff (1941), cf. Figure 4.1.

If we have a look at Figure 2.1, we see fuzzy confidence levels on a two dimensional random vector ε. The curves displayed are level sets of the potential $V(\varepsilon) : \mathbb{R}^2 \to \mathbb{R}$. They represent nested confidence regions defined as follows: $\underline{C}_\alpha := \{\varepsilon \in \mathbb{M} \mid V(\varepsilon) \leq \underline{V}_\alpha\}$ if $\underline{V}_\alpha := \min\{V_\alpha \in V(\mathbb{M}) \mid \overline{\alpha}(V_\alpha) = \alpha\}$ exists, and $\underline{C}_\alpha := \emptyset$ otherwise; analogously $\overline{C}_\alpha := \{\varepsilon \in \mathbb{M} \mid V(\varepsilon) \leq \overline{V}_\alpha\}$ if $\overline{V}_\alpha := \max\{V_\alpha \in V(\mathbb{M}) \mid \underline{\alpha}(V_\alpha) = \alpha\}$ exists, and $\overline{C}_\alpha := \mathbb{M}$ otherwise. The region $\underline{C}_\alpha$ contains at most a fraction of α of all possible values of ε in $\mathbb{M}$, since $\Pr(\varepsilon \in \underline{C}_\alpha) \leq \Pr(\overline{\alpha}(V(\varepsilon)) \leq \alpha) \leq \Pr(F(V(\varepsilon)) \leq \alpha) = \alpha$; analogously $\overline{C}_\alpha$ contains at least a fraction of α of all possible values of ε in $\mathbb{M}$. Generally holds $\underline{C}_\alpha \subseteq \overline{C}_\alpha$.

Let us summarize what is needed to generate a potential cloud: a potential function V has to be chosen, then appropriate bounds on the CDF F of $V(\mathbb{M})$ must be found. How to find these bounds will be described in Section 4. But how to choose the

potential function? There are endless possibilities (see, e.g., Figure 2.1) to make the choice. A variation of the shape of the potential to improve the uncertainty model will be considered in Section 5.

3 Clouds in robust design optimization

In this section we shortly introduce how clouds can be involved in the formulation of a robust design optimization problem.

Provided an underlying model of a given structure to be designed, with several inputs and outputs, we denote as x the vector containing all output variables, and as z the vector containing all input variables. Let θ be a design point, i.e., it fully defines the design. Let T be the set of all possible designs. The input variables z consist of design variables which depend on the design θ, e.g., the thrust of a thruster, and external inputs with a nominal value that cannot be controlled for the underlying model, e.g., a specific temperature.

The input variables are affected by uncertainties. Let ε denote the related random vector of uncertain errors. One can formulate the optimization problem as a mixed-integer, bi-level problem of the following form:

$$\begin{array}{lll} \min\limits_{\theta} \ \max\limits_{x,z,\varepsilon} & g(x) & \text{(objective functions)} \\ \text{s.t.} & G(x,z) = 0 & \text{(functional constraints)} \\ & z = Z(\theta) + \varepsilon & \text{(input constraints)} \\ & \theta \in T & \text{(selection constraints)} \\ & V(\varepsilon) \leq \underline{V}_{\alpha} & \text{(cloud constraint)} \end{array} \tag{3.1}$$

where the design objective $g(x)$ is a function of the output variables of the underlying model. The functional constraints express the functional relationships defined in the underlying model. The input constraints assign to each design θ a vector z of input variables whose value is the nominal entry from $Z(\theta)$ plus its error ε with uncertainty specified by the cloud. The selection constraints specify which design points are allowed for θ, θ typically consists of integer and continuous components. The cloud constraint involves the potential function V as described in the Section 5 and models the worst-case relevant region $\underline{C}_{\alpha}$.

If the functional constraints can be solved numerically for x given z then z and x are determined by ε so that effectively only an optimization over ε is needed. The confidence level α should be chosen to reflect the seriousness of consequences of the worst case event. In our applications from spacecraft system design we used $\alpha = 0.95$, cf. Fuchs et al. (2008), Neumaier et al. (2007).

In the remainder of the paper, we concentrate on the cloud generation aspect, which is only loosely related to the optimization phase, cf. Remark 5.1.

4 Generation of potential clouds

This section will investigate how to find appropriate bounds on the CDF F of the potential $V(\varepsilon)$. As we do not have the knowledge of $F(V(\varepsilon))$ we have to approximate

it before we can assign bounds on it. To this end we will make use of the well-known KS statistics, cf. Kolmogoroff (1941), as suggested in the last section. That means we approximate F by an empirical distribution $\widetilde{F}$. The generation of an empirical distribution requires the existence of a sample S representing our uncertainties.

It depends on the given **uncertainty information** whether a sample already exists. We assume that a priori uncertainty information consists of given samples, boxes (i.e., lower and upper bounds on the variables), or continuous marginal CDFs F_i, $i \in I \subseteq \{1, 2, \ldots, n\}$, on the n-dimensional random vector ε, without any formal knowledge about correlations or joint distributions. In case there is no sample provided or the given sample is very small, a sample has to be generated. For these cases we first use a **Latin hypercube sampling** (**LHS**, cf. McKay et al. (1979)) inspired method to generate S. This introduces some preference for a simple structure. The effect of this preference will be later diminished by weighting of the sample points.

To generate N_S sample points we start with the creation of a $N_S \times \cdots \times N_S = N_S^n$ grid. In case of a given interval on ε^i the marginal grid points are chosen equidistantly in the interval. In case of a given F_i, $i \in I$, the marginal grid is transformed with respect to the marginal CDF F_i to ensure that each grid interval has the same marginal probability. Let $\alpha_S \in (0, 1]$, a confidence level for the sample generation, $p_S = 1 - \sqrt[n]{\alpha_S}$, $t_1 = \frac{p_S}{2}$, $t_2 = t_1 + 1 \cdot \frac{1-p_S}{N_S-1}$, $t_3 = t_1 + 2 \cdot \frac{1-p_S}{N_S-1}, \ldots, t_{N_S} = t_1 + (N_S - 1) \cdot \frac{1-p_S}{N_S-1} = 1 - \frac{p_S}{2}$, then the marginal grid points are chosen as $g_{i1} = F_i^{-1}(t_1)$, $g_{i2} = F_i^{-1}(t_2), \ldots, g_{iN_S} = F_i^{-1}(t_{N_S})$. From this grid the sample points $x_1, x_2, \ldots, x_{N_S}$ are chosen to satisfy the Latin hypercube condition, i.e.,

$$\forall i, k \in \{1, 2, \ldots, N_S\}, \forall j \in \{1, 2, \ldots, n\} : x_i^j \neq x_k^j \qquad \text{if } k \neq i, \tag{4.1}$$

where x_i^j is the projection of x_i to the j^{th} coordinate.

Note that given $I = \{1, 2, \ldots, n\}$, i.e., a marginal CDF is given for each coordinate of ε, then α_S is a confidence level for ε, supposed that ε^i, $i \in I$, are independent: $\Pr(\varepsilon \in [g_{11}, g_{1N_S}] \times [g_{21}, g_{2N_S}] \times \cdots \times [g_{n1}, g_{nN_S}]) = \prod_{i \in I} \Pr(\varepsilon^i \in [g_i 1, g_{iN_S}])$ $= \prod_{i=1}^n (F_i(g_{iN_S}) - F_i(g_{i1})) = \prod_{i=1}^n (t_{N_S} - t_1) = (1 - p_S)^n = \alpha_S$.

4.1 Proposition. *The marginal empirical distribution $\widetilde{F}_i(\xi) = \sum_{\{j | x_j^i \leq \xi\}} \frac{1}{N_S}$, $i \in I$, of our sample approximates F_i for $N_S \to \infty$ and $\alpha_S \to 1$.*

Proof. The LHS condition (4.1) implies that $x_i^j = g_{jk}$ for some k. One can write $k = q \cdot N_S$ with some rational number $q \in [\frac{1}{N_S}, 1]$. The ratio $q = \frac{k}{N_S}$ is constant for all N_S as to the construction of the points t_i, $i \in \{1, 2, \ldots, N_S\}$ equidistantly in direct dependence from N_S. Let $s \in [g_{jk}, g_{jk+1}]$, $k \in \{1, 2, \ldots, N_S - 1\}$, then $\widetilde{F}_i(s) = \frac{k}{N_S}$. For $N_S \to \infty$ and $\alpha_S \to 1$ there exists $k' : s = g_{jk'}$, so $\lim_{N_S \to \infty} F_i(s) = \lim_{N_S \to \infty} t_{k'} = \lim_{N_S \to \infty} t_1 + (k' - 1) \cdot \frac{1-p_S}{N_S-1} = q'$, with $k' = q' \cdot N_S$, and we get $\lim_{N_S \to \infty} F_i(s) - \widetilde{F}_i(s) = q' - q' = 0$. □

4.2 Remark. α_S can be freely chosen from $(0, 1]$. A choice of $\alpha_S = 1$ allows $g_{i1} = -\infty$, $g_{iN_S} = \infty$. In practice one chooses α_S smaller than but close to 1. We use

$\alpha_S = 0.998$ in our applications. Also note that by the use of $\alpha_S < 1$ the sample generation can distinguish between mere interval information and a marginal uniform distribution: the information that $\varepsilon^i \in [a, b]$ will result in a different sample than the information that ε^i is uniformly distributed in $[a, b]$.

Thus the generated sample $S := \{x_1, x_2, \ldots, x_{N_S}\}$ represents the marginal CDFs arbitrarily well. However after a modification of S, e.g., by cutting off sample points as we will do later, the marginal empirical distributions will no longer approximate the marginal CDFs. An assignment of weights to the sample points is necessary to preserve the marginal CDFs. In order to do so the weights $\omega_1, \omega_2, \ldots, \omega_{N_S} \in [0, 1]$, corresponding to the sample points $x_1, x_2, \ldots, x_{N_S}$, are required to satisfy the following conditions:

Let π_j be a sorting permutation of $\{1, 2, \ldots, N_S\}$, such that $x^j_{\pi_k(1)} \leq x^j_{\pi_k(2)} \leq \cdots \leq x^j_{\pi_k(N_S)}$. Let again I be the index set of those entries of the random vector ε where marginal distributions F_i are given on each component $i \in I$. Then the weights should satisfy the following equality for all components $i \in I$ and $k \in \{1, 2, \ldots, N_S\}$:

$$\sum_{j=1}^{k} \omega_{\pi_i(j)} \in [F_i(x^i_{\pi_i(k)}) - d, F_i(x^i_{\pi_i(k)}) + d], \qquad \sum_{k=1}^{N_S} \omega_k = 1. \tag{4.2}$$

The function

$$\widetilde{F}_i(\xi) := \sum_{\{j | x^i_j \leq \xi\}} \omega_j \tag{4.3}$$

is a weighted marginal empirical distribution. For trivial weights, $\omega_1 = \omega_2 = \cdots = \omega_{N_S} = \frac{1}{N_S}$, $\widetilde{F}_i$ is a standard empirical distribution as in Proposition 4.1. The constraints (4.2) require the weights to represent the marginal CDFs with some reasonable margin d. In other words, the weighted marginal empirical distributions $\widetilde{F}_i, i \in I$ should not differ from the given marginal CDF F_i by more than d. In practice, one chooses $d = d_{\mathrm{KS},1}$ with KS statistics, i.e.,

$$d_{\mathrm{KS},1} = \frac{\phi^{-1}(a)}{\sqrt{N_S} + 0.12 + \frac{0.11}{\sqrt{N_S}}}, \tag{4.4}$$

where ϕ is the Kolmogoroff function $\phi(\lambda) := \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2}$, a the confidence in the KS theorem, cf. Kolmogoroff (1941), Press et al. (1992). This choice is reasonable in practice although not precisely fitting the derivation of the KS theorem.

To achieve weights satisfying (4.2) we formulate the following linear program:

$$\begin{aligned} \min_{\omega_1, \omega_2, \ldots, \omega_{N_S}} \quad & e \\ \text{s.t.} \quad & \widetilde{F}_i(\xi_k) \in [F_i(\xi_k) - e \cdot d_{\mathrm{KS},1}, F_i(\xi_k) + e \cdot d_{\mathrm{KS},1}] \ \ \forall i, \xi_k \\ & \sum_{i=1}^{N_S} \omega_i = 1 \\ & \omega_i \geq 0 \ \ \forall i \\ & e \geq 0 \end{aligned} \tag{4.5}$$

where ξ_k are some given interpolation points on the related margin, $\widetilde{F}_i$ the weighted empirical marginal distributions as defined in (4.3). Then we compute d from

$$d = \max_{i \in I, j=1,2,\ldots,N_S} |\widetilde{F}_i(x_j^i) - F_i(x_j^i)| \tag{4.6}$$

the maximum deviation of $\widetilde{F}_i$ from F_i, $i \in I$.

It should be remarked that we make use of CVX, cf. Grant and Boyd (2007), to solve (4.5), and of the MATLAB Statistics Toolbox to evaluate probability distributions.

By the weight computation we get a weighted empirical distribution from

$$\widetilde{F}(\xi) := \sum_{\{j|V(x_j) \leq \xi\}} \omega_j \tag{4.7}$$

approximating the CDF of $V(\varepsilon)$. The achievement of weights satisfying (4.2) means, that all information of the marginal CDFs is reflected in the construction of $\widetilde{F}$. The information given as boxes or sample data is reflected anyway as this does not imply additional constraints to (4.2).

If weights satisfying (4.2) can only be achieved with $d > d_{\mathrm{KS},1}$, the relaxation d gives us an indicator for the quality of the approximation which will be useful to construct bounds on the CDF $F(V(\varepsilon))$. After the approximation of $F(V(\varepsilon))$ with $\widetilde{F}$, we are just one step away from generating a potential cloud. The last step is seeking an appropriate bounding on $F(V(\varepsilon))$.

From the knowledge of $d_{\mathrm{KS},1}$, and d we compute $d_{\mathrm{KS},2}$ similar to before (4.4), for a fixed confidence b:

$$d_{\mathrm{KS},2} = \frac{\phi^{-1}(b)}{\sqrt{N_S} + 0.12 + \frac{0.11}{\sqrt{N_S}}} \cdot \max(1, \frac{d}{d_{\mathrm{KS},1}}), \tag{4.8}$$

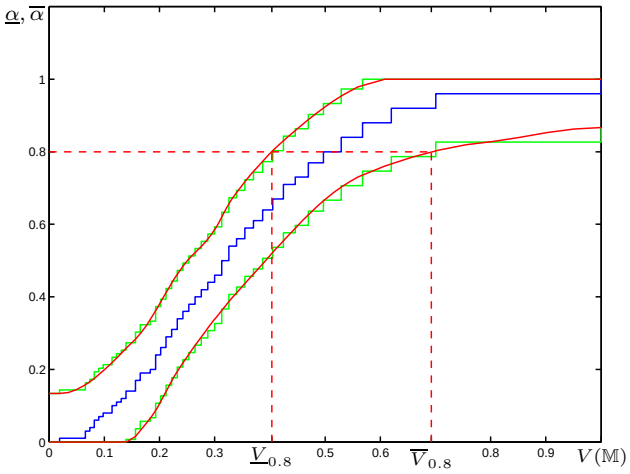
with the approximation quality factor $\frac{d}{d_{\mathrm{KS},1}}$ which corresponds to e in (4.5), if the interpolation and the sample points coincide. Now we define $\overline{F} := \min(\widetilde{F} + d_{\mathrm{KS},2}, 1)$ and $\underline{F} := \max(\widetilde{F} - d_{\mathrm{KS},2}, 0)$ and fit these two step functions to smooth, monotone lower bounds $\underline{\alpha}(V(\varepsilon))$ and upper bounds $\overline{\alpha}(V(\varepsilon))$, cf. Figure 4.1. Observe that if the quality of our approximation with $\widetilde{F}$ or the sample size N_S is decreased, the width of the bounds is increased correspondingly.

4.3 Remark. Figure 4.1 reminds of p-boxes. In fact a potential cloud can be considered as a p-box on the potential of a random vector. Clouds extend the p-box concept to the multivariate case without the exponential growth of work in the conventional p-box approach.

Thus we have found an **appropriate bounding** of the CDF $F(V(\varepsilon))$ and according to Proposition 2.3 we have generated a potential cloud via the mapping $\chi : \varepsilon \to [\underline{\alpha}(V(\varepsilon)), \overline{\alpha}(V(\varepsilon))]$.

Note that if a weighted sample is actually available (in the simplest case all weights are $\frac{1}{N_S}$), we can directly compute $\underline{F}$ and $\overline{F}$ from $\widetilde{F}$ setting $d = d_{\mathrm{KS},1}$.

Figure 4.1: The smoothed lower bounds $\underline{\alpha}(V(\varepsilon))$ and upper bounds $\overline{\alpha}(V(\varepsilon))$ enclosing the empirical distribution of $V(\varepsilon)$. The mapping $\varepsilon \to [\underline{\alpha}(V(\varepsilon)), \overline{\alpha}(V(\varepsilon))]$ is a potential cloud (cf. Section 2.2). The horizontal cut $\alpha = 0.8$ is illustrated and the corresponding $\underline{V}_\alpha$ and $\overline{V}_\alpha$, given by $\overline{\alpha}(\underline{V}_\alpha) = \alpha$ and $\underline{\alpha}(\overline{V}_\alpha) = \alpha$, respectively.



The cloud represents the given information and now enables us to interpret the potential level maps as confidence regions $\{\varepsilon \mid V(\varepsilon) \leq V_\alpha\}$ for the random vector ε: the **worst-case relevant** region is defined as $\underline{C}_\alpha$ (cf. Section 2.2), which is not empty except for the case α is very low and thus little interesting as a confidence level.

Thus the clouds give an intuition and guideline how to construct confidence regions for safety constraints. To this end we have combined several different theoretical means: potential functions, KS statistics to approximate CDFs with empirical distributions and estimate bounds, sample generation methods, and weighting techniques.

5 The choice of the potential

The choice of the potential function V that determines the shape of the corresponding potential cloud can be chosen freely before cloud generation. We will now investigate different choices of V and reflect on what characterizes a **good** choice of V.

Two special cases for choices of the potential function are

$$V(\varepsilon) := \max_k \frac{|\varepsilon^k - \mu^k|}{r^k}, \tag{5.1}$$

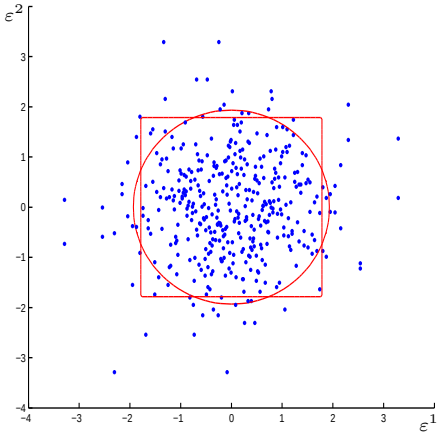
where $\varepsilon, \mu, r \in \mathbb{R}^n$, $\varepsilon^k, \mu^k, r^k$ are the k^{th} components of the vectors, defines a box-shaped potential.

$$V(\varepsilon) := \|A\varepsilon - b\|_2^2, \tag{5.2}$$

where $\varepsilon, b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, defines an ellipsoid-shaped potential.

Figure 5.1 visualizes two confidence regions for the same confidence level $\alpha = 0.95$, but different potentials V, and shows that uncertainties can be described by different clouds. We emphasize that a poor choice of the potential makes the worst-case analysis more pessimistic as the confidence regions are larger, but will still result in a valid robust uncertainty handling.

Figure 5.1: The regions $\underline{C}_{0.95}$ for two different choices of V, box- and circle-shaped, respectively. The 2-dimensional sample belongs to two independent $N(0,1)$-distributed random variables ε^1 and ε^2.



We are looking for a way to find a good choice of V that gives the possibility to improve the potential iteratively and allows for a simple computational realization of the confidence regions, e.g., by linear constraints. This leads us to the investigation of **polyhedron-shaped potentials** as a generalization of box-shaped potentials. A polyhedron potential centered at $m \in \mathbb{R}^n$ can be defined as:

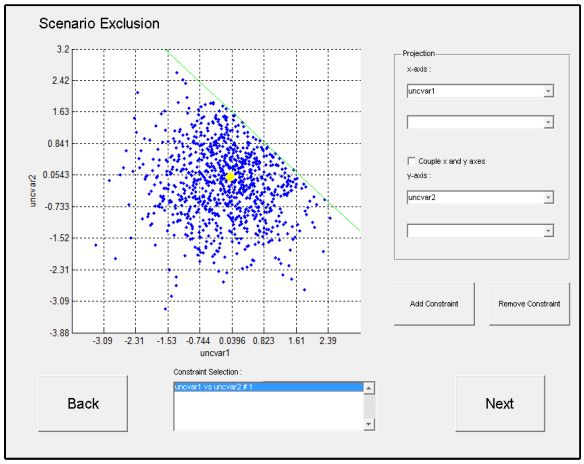
$$V_p(\varepsilon) := \max_k \frac{(A(\varepsilon - m))^k}{b^k}, \tag{5.3}$$

where $(A(\varepsilon - m))^k, b^k$ the k^{th} component of the vectors $A(\varepsilon - m)$ and b, respectively.

But how to achieve a polyhedron that reflects the given information in the best way? As mentioned previously we assume the information to consist of given samples, boxes or marginal distributions, and unformalized dependency constraints. After generation of a sample S as described in Section 4 we define a box b_0 containing 100% of the sample points by $b_0 := [g_{11}, g_{1N_S}] \times [g_{21}, g_{2N_S}] \times \cdots \times [g_{n1}, g_{nN_S}]$, and we define our potential $V_0(\varepsilon)$ box-shaped as in (5.1) taking the value 1 on the margin of b_0, i.e., $\mu^k = \frac{g_{k1}+g_{kN_S}}{2}$, $r^k = \frac{g_{kN_S}-g_{k1}}{2}$.

Based on expert knowledge, a user-defined variation of V_0 can be performed afterwards by cutting off sample points deemed irrelevant for the worst-case, cf. Figure 5.2: The optimization method provides a worst-case scenario which is highlighted in the graphical user interface. The expert can decide to exclude, e.g., the worst-case or different scenarios, based on his technical knowledge. Thus an expert can specify the

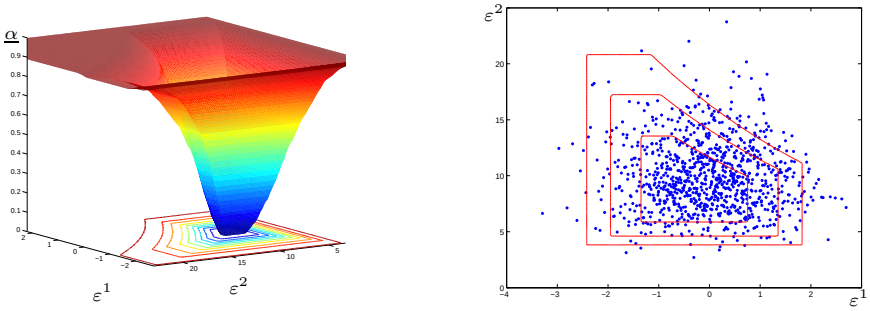
Figure 5.2: Graphical user interface for an interactive scenario exclusion. The exclusion is performed in 1 and 2 dimensional projections.



uncertainty information in the form of dependency constraints adaptively, even if the expert knowledge is only little formalized, resulting in a polyhedron shaped potential.

5.1 Remark. There is a close relationship between the dependency constraints and correlation information. Consider the 2 dimensional normally distributed case centered at 0, with non-zero correlation. Then the confidence regions are ellipse shaped $||A\varepsilon||_2^2 \leq \text{const}$. If one approximates the ellipse with a parallelepiped $||A\varepsilon||_\infty \leq \text{const}$ you end up in a polyhedron shaped confidence region.

Figure 5.3: On the left, the map $\varepsilon \to \underline{\alpha}(V(\varepsilon))$ ($\overline{\alpha}(V(\varepsilon))$ looks similar due to its construction) for a 2-dimensional potential cloud, on the right, the contour lines of $\underline{\alpha}(V(\varepsilon))$. The marginal distributions for $\varepsilon = (\varepsilon^1, \varepsilon^2)$ are a $N(0,1)$ and a $\Gamma(10,1)$ distribution, respectively.



Assume the linear constraints $A(\varepsilon - \mu) \leq b$ represent the exclusion of sample points and the box constraint from b_0, we define our polyhedron shaped potential as in

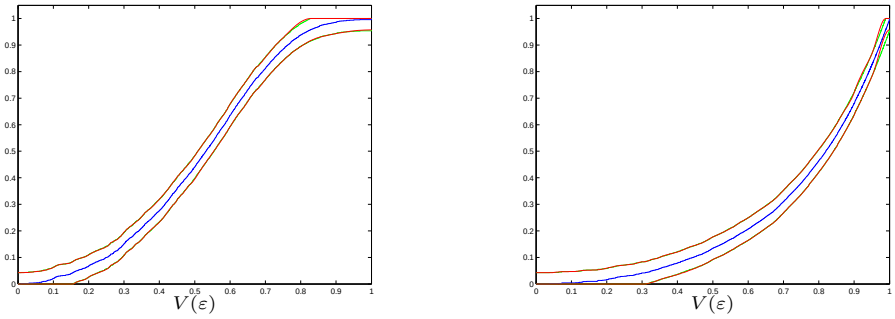
(5.3) with $m = \mu$.

This potential, originating from a box potential, is suitable for symmetric samples, but: If some uncertain variables are described by asymmetric marginal probability densities, a better choice V_t of the potential could be achieved by an appropriate **coordinate transformation** ψ, i.e.,

$$V_t(\varepsilon) := V_p(\psi(\varepsilon)). \tag{5.4}$$

An appropriate transformation would be, e.g., a logarithmic transformation of ε^i if $F_i : \mathbb{R}^+ \to [0,1]$. An example of a 2-dimensional potential cloud with $V = V_t$ is visualized in Figure 5.3.

Figure 5.4: The lower bounds $\underline{\alpha}(V(\varepsilon))$ and upper bounds $\overline{\alpha}(V(\varepsilon))$ for a potential cloud with transformation (left figure) and without transformation (right figure). The marginal distributions for $\varepsilon = (\varepsilon^1, \varepsilon^2)$ are a $N(0,1)$ and a $\Gamma(10,1)$ distribution, respectively.



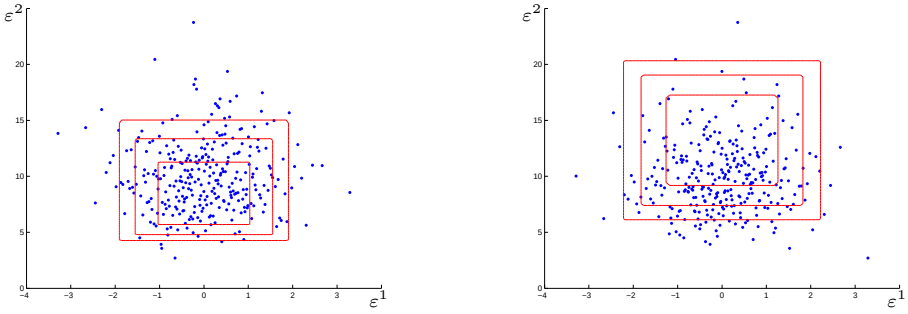
We can observe the advantage of a transformed potential in Figure 5.4. Without transformation the functions $\underline{\alpha}(V(\varepsilon))$ and $\overline{\alpha}(V(\varepsilon))$ are obviously steeper, and for α close to 1 the solution V_α of $\overline{\alpha}(V_\alpha) = \alpha$ is much closer to 1 than in the transformed case, which leads to larger confidence regions and a more pessimistic worst-case analysis. The reason for that becomes apparent looking at Figure 5.5. The confidence regions for the transformed box potential are obviously smaller than for the non-transformed potential.

6 Conclusions

In this paper we presented a new methodology to provide confidence regions for safety constraints in robust design optimization in terms of clouds. We can process the information from expert knowledge towards a reliable worst-case analysis, even if the information is limited in amount and high dimensional, formalized or not, we do not lose valuable unformalized information. The methods were successfully applied to real life problems from spacecraft system design.

The adaptive nature of our uncertainty model, i.e., manually adding dependency constraints, is one of the key features. The iteration steps significantly improve the

Figure 5.5: The figures show the regions $\underline{C}_\alpha$, $\alpha = 0.5, 0.8, 0.95$ for a box potential cloud with (left figure) and without (right figure) transformation. The two samples are generated for ε distributed as in Figure 5.4.



uncertainty information and we are able to process the new information to an improved uncertainty model.

All in all, the presented approach offers an attractive novel point of view on uncertainty handling and its involvement in robust design optimization.

References

Alexandrov N.M., Hussaini M.Y., 1997. Multidisciplinary design optimization: State of the art. In *Proceedings of the ICASE/NASA Langley Workshop on Multidisciplinary Design Optimization*.

Berleant D., Cheng H., 1998. A software tool for automatically verified operations on intervals and probability distributions. *Reliable Computing*, 4(1), 71–82.

Destercke S., Dubois D., Chojnacki E., 2007. Relating practical representations of imprecise probabilities. In *Proceedings of the* 5^{th} *International Symposium on Imprecise Probability: Theories and Applications*. Prague, Czech Republic.

Dubois D., Prade H., 1986. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press.

Dubois D., Prade H., 2005. Interval-valued fuzzy sets, possibility theory and imprecise probability. In *Proceedings of International Conference in Fuzzy Logic and Technology*.

Ferson S., 1996. What monte carlo methods cannot do. *Human and Ecological Risk Assessment*, 2, 990–1007.

Ferson S., 2002. *Ramas Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers*. Lewis Publishers, U.S.

Fuchs M., Girimonte D., Izzo D., Neumaier A., 2008. *Robust Intelligent Systems*, chapter Robust and Automated Space System Design, in press. Springer. Preprint available on-line at: `http://www.martin-fuchs.net/publications.php`.

Grant M.C., Boyd S.P., 2007. CVX: A system for disciplined convex programming. http://www.stanford.edu/~boyd/cvx/cvx_usrguide.pdf http://www.stanford.edu/~boyd/cvx/.

Koch P.N., Simpson T.W., Allen J.K., Mistree F., 1999. Statistical approximations for multidisciplinary optimization: The problem of size. *Special Issue on Multidisciplinary Design Optimization of Journal of Aircraft*, 36(1), 275–286.

Kolmogoroff A., 1941. Confidence limits for an unknown distribution function. *The Annals of Mathematical Statistics*, 12(4), 461–463.

Kreinovich V., 1997. *Random Sets: Theory and Applications*, chapter Random sets unify, explain, and aid known uncertainty methods in expert systems, 321–345. Springer-Verlag.

Lewis K., Mistree F., 1997. Modeling interactions in multidisciplinary design: A game theoretic approach. *AIAA Journal*, 35(8), 1387–1392.

McCormick D.J., Olds J.R., 2002. A distributed framework for probabilistic analysis. In *AIAA/ISSMO Symposium On Multidisciplinary Analysis And Design Optimization*.

McKay M., Conover W., Beckman R., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 221, 239–245.

Neumaier A., 2004. Clouds, fuzzy sets and probability intervals. *Reliable Computing 10*, 249–272. http://www.mat.univie.ac.at/~neum/ms/cloud.pdf.

Neumaier A., Fuchs M., Dolejsi E., Csendes T., Dombi J., Banhelyi B., Gera Z., 2007. Application of clouds for modeling uncertainties in robust space system design. ACT Ariadna Research ACT-RPT-05-5201, European Space Agency. Available on-line at http://www.esa.int/gsp/ACT/ariadna/completed.htm.

Oberguggenberger M., King J., Schmelzer B., 2007. Imprecise probability methods for sensitivity analysis in engineering. In *Proceedings of the* 5^{th} *International Symposium on Imprecise Probability: Theories and Applications*. Prague, Czech Republic.

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1992. *Numerical recipes in C*. Cambridge University Press, second edition.

Regan H., Ferson S., Berleant D., 2004. Equivalence of methods for uncertainty propagation of real-valued random variables. *International Journal of Approximate Reasoning*, 36(1), 1–30.

Ross T.J., 1995. *Fuzzy Logic with Engineering Applications*. New York, NY: McGraw-Hill.

Williamson R.C., 1989. *Probabilistic Arithmetic*. Ph.D. thesis, University of Queensland.