Potential based clouds in robust design optimization

Martin Fuchs and Arnold Neumaier

University of Vienna
Faculty of Mathematics
Nordbergstr. 15
1090 Wien
Austria
email: martin.fuchs@univie.ac.at

Abstract

Robust design optimization methods applied to real life problem 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 means to process available uncertainty information reliably, even if limited in amount and possibly lacking a formal description, to providing a worst-case analysis with confidence regions of relevant scenarios which can be involved in an optimization problem formulation for robust design.

Keywords. clouds, robust design, design optimization, confidence regions, uncertainty modeling

1   Background

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

1

of handling uncertainties. In particular for early design phases, it is frequent engineering practice to assign and refine intervals or safety margins to the uncertain variables. 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 rigorous worst-case analysis involved.

Several methods exist to approach reliability quantification from a rigorous mathematical background, originating from classical probability theory, statistics, fuzzy theory or based on simulation techniques. 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, also famous as the curse of dimensionality. Namely, even given the complete knowledge about the multivariate probability distributions of the uncertainties, the numerical computation of the error probabilities requires high dimensional integration and becomes very expensive. Moreover, if the amount of available uncertainty information is very limited, well-known current methods either do not apply at all, or are endangered to critically underestimate error probabilities. Also 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 $3\,\sigma$ boxes), entails a loss of valuable uncertainty information which would actually be available, maybe only unformalized, but not at all considered in the uncertainty model.

In the literature several discussions on the introduced problems and approaches to their solution can be found. Different approaches to design optimization can be studied in [1], [2], [14], [22]. An approach to achieve robustness in the special case of space shuttle design by means of probability risk analysis is given in [19]. A criticism on simulation techniques in case of limited uncertainty information can be found in [6], showing that the lack of information typically causes these techniques to underestimate the effects of the uncertain tails of the probability distribution. Fuzzy theory for engineering applications is investigated, e.g., in [21], simulation techniques, e.g., in [15]. A method based on possibility theory has been developed in [4], [5]. Some works go into the arithmetic operations on random variables to bound the probability distributions of functions of random variables, cf. [3], [7], [23]. There are also attempts to generalize aspects of the different uncertainty approaches and put them into one framework, cf. [13]. As mentioned many methods are limited to low dimensional problems. Problems that come with the curse of dimensionality are described, e.g., in [11].

2 Overview

In this paper we will present a detailed view on the theoretical basis of uncertainty handling with clouds which we developed. 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. Since the strength of our new methodology lies especially in the application to real life problems with a very limited amount of uncertainty 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. On the one hand, the information is usually available as bounds or marginal probability distributions on the uncertain variables, without any formal correlation information. On the other hand, unformalized expert knowledge will be captured to improve the uncertainty model adaptively by adding correlation constraints to exclude scenarios deemed irrelevant. The information can also be provided as real sample data, if available. We call the set of all uncertainty information available an uncertainty setting.
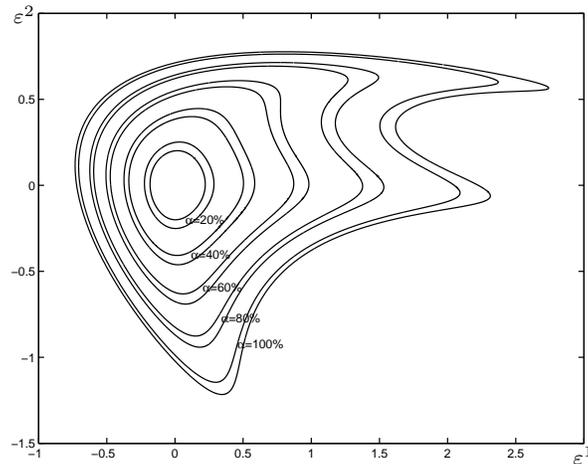

Figure 1: Nested confidence regions in two dimensions for confidence levels $\alpha = 0.2, 0.4, 0.6, 0.8, 1$.

If we have a look at Figure 1, we see confidence levels on some two dimensional random variable ε. The curves displayed can be considered to be level sets of a function $V(\varepsilon) : \mathbb{R}^2 \to \mathbb{R}$, called the potential. The potential characterizes confidence regions $C_\alpha := \{\varepsilon \in \mathbb{R}^2 \mid V(\varepsilon) \leq V_\alpha\}$, where V_α is determined by the

3

condition $\Pr(\varepsilon \in C_\alpha) = \alpha$. If the probability information is not precisely known nested regions

$$\underline{C}_\alpha := \{\varepsilon \in \mathbb{R}^2 \mid V(\varepsilon) \leq \underline{V}_\alpha\},$$

where $\underline{V}_\alpha$ is largest such that $\Pr(\varepsilon \in \underline{C}_\alpha) \leq \alpha$, and

$$\overline{C}_\alpha := \{\varepsilon \in \mathbb{R}^2 \mid V(\varepsilon) \leq \overline{V}_\alpha\},$$

where $\overline{V}_\alpha$ is smallest such that $\Pr(\varepsilon \in \overline{C}_\alpha) \geq \alpha$, are generated. The information in $\underline{C}_\alpha$ and $\overline{C}_\alpha$ is called a *potential cloud*. The $V_\alpha, \underline{V}_\alpha, \overline{V}_\alpha$ can easily be found from the **cumulative distribution function** (**CDF**) of $V(\varepsilon)$ and lower and upper bounds of it. These bounds in turn can be determined empirically using a **Kolmogoroff-Smirnov** (**KS**) distribution. Thus given a potential the corresponding potential cloud is easy to estimate, even for high dimensional data, and its interpretation is unambiguous in spite of the uncertainty about the full multidimensional probability distribution. The choice of the potential is dictated by the shape of the points set defined by the sample of available ε, and should allow for a simple computational realization of the confidence regions, e.g., by linear constraints. In terms of robust design the regions $\underline{C}_\alpha$ and $\overline{C}_\alpha$ yield safety constraints, as a design is called safe if all $\varepsilon \in \overline{C}_\alpha$ satisfy the design requirements, and it is called unsafe if one $\varepsilon \in \underline{C}_\alpha$ fails to satisfy these requirements. These safety constraints can easily be incorporated in an optimization problem formulation.

A first study on the application of the clouds theory in design optimization occurs in [18]. This study presented an initial step on what clouds are capable of, applied to the case of uncertainty handling in spacecraft design. A significant further step is given in [8] showing that clouds can be successfully applied to real life problems.

This paper is organized as follows. First, we follow the clouds formalism from [17] in Section 3, 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. 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. A short introduction how clouds can be involved in an optimization problem formulation for robust design is given in Section 6. Section 7 concludes our studies.

3 The theory of potential clouds

The clouds theory serves as the central theoretical background for our uncertainty handling. We start with the formal definition of clouds and introduce the notations.

Afterwards we restate the central results from [17], 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 vector of uncertain variables, we call ε an uncertain *scenario*. 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

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

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.

We introduce the term that a random variable ε *belongs to* a cloud χ over $\mathbb{M}$, if

$$\Pr(\overline{\chi}(\varepsilon) \leq y) \leq y \leq \Pr(\underline{\chi}(\varepsilon) < y) \tag{2}$$

for some value $y \in [0, 1]$.

3.1 Remark. To get an intuition about this definition, assume ε has the given CDF F, and F, $\underline{\chi}$ and $\overline{\chi}$ are continuous, invertible functions. Then it reads $F(\overline{\chi}^{-1}(y)) \leq y \leq F(\underline{\chi}^{-1}(y))$.

A cloud is called *discrete* if $\chi(\varepsilon)$, $\varepsilon \in \mathbb{M}$, only takes finitely many different values (interval valued). 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. Theorem 3.1 in [17], and in principle can approximate arbitrary distributions arbitrarily well. As histograms and the related discrete clouds are having problems in higher dimensions with the computational effort, 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, stated in [17] as Proposition 4.1.

3.2 Proposition. *Let $F_\varepsilon(x) = \Pr(\varepsilon \leq x)$ be the CDF of the random variable ε, then $\chi(x) := F_\varepsilon(x)$ defines a thin cloud and ε belongs to χ, i.e., for the thin case $\Pr(\chi(x) \leq y) = y, y \in \mathbb{M}$, 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 $\Pr(\chi(x) \leq y) = \Pr(F_\varepsilon(x) \leq y) = y$. $\square$

3.3 Remark. Comparing the notations we introduced in this section with those in [17] it should be remarked that ours correspond to those of the so called *mirror cloud* $\chi' = 1 - \chi$. This is unproblematic with respect to Corollary 2.2 in [17]. Observe that the same set of random variables can be characterized by different 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 $V : \mathbb{M} \to \mathbb{R}$. The idea is to construct a *potential cloud* from an interval-valued function χ of a user-defined potential function V, i.e. $\chi \circ V : \mathbb{M} \to [a, b]$, where $[a, b]$ is an interval in $[0, 1]$.

Define the mapping

$$\chi(x) := [\underline{\alpha}(V(x)), \overline{\alpha}(V(x))], \tag{3}$$

where $\underline{\alpha}(y) := \Pr(V(\varepsilon) < y), \overline{\alpha}(y) := \Pr(V(\varepsilon) \leq y), \varepsilon \in \mathbb{M}$ a random variable, $y \in \mathbb{R}$, V bounded below. Then we get the following result (cf. Theorem 4.3 in [17]).

3.4 Proposition. *The mapping χ as defined in* (3) *defines a cloud and ε belongs to χ.*

Proof. Obviously (1) holds for χ. One has to show (2). Assume $\exists x \in \mathbb{R}$ with $\overline{\alpha}(x) = y$. By construction of $\overline{\alpha}$ and x we get $\Pr(\overline{\chi}(\varepsilon) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq y) = \Pr(\overline{\alpha}(V(\varepsilon)) \leq \overline{\alpha}(x)) = \Pr(V(\varepsilon) \leq x) = \overline{\alpha}(x) = y$. 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). $\square$

3.5 Remark. The thus constructed cloud χ is thin if V has a continuous CDF: Let F be the CDF of $V(\varepsilon)$, then $\underline{\alpha}(y) = \Pr(V(\varepsilon) \leq y) = \Pr(V(\varepsilon) < y) = \overline{\alpha}(y) = F(y)$ and $\chi(x) = [F(V(x)), F(V(x))]$.

3.6 Remark. Usually the CDF of $V(\varepsilon)$ is unknown. However, if we find an lower bound $\underline{\alpha}(y) \leq \Pr(V(\varepsilon) < y)$ and upper bound $\overline{\alpha}(y) \geq \Pr(V(\varepsilon) \leq y)$, $\underline{\alpha}$, $\overline{\alpha}$ continuous from the right and monotone, then Proposition 3.4 is still valid.

The last remark leads us to an important interpretation in terms of confidence regions for the scenarios ε, as it tells us that it is sufficient to find an appropriate bounding $\underline{\alpha}$, $\overline{\alpha}$ on the CDF F of $V(\varepsilon)$. This can be achieved, e.g., by KS statistics [12]. We define $\underline{C}_\alpha := \{\varepsilon \mid V(\varepsilon) \leq \underline{V}_\alpha\}$ if a solution $\underline{V}_\alpha$ of $\overline{\alpha}(\underline{V}_\alpha) = \alpha$ exists and $\underline{C}_\alpha := \emptyset$ otherwise; analogously $\overline{C}_\alpha := \{\varepsilon \mid V(\varepsilon) \leq \overline{V}_\alpha\}$ if a solution $\overline{V}_\alpha$ of $\underline{\alpha}(\overline{V}_\alpha) = \alpha$ exists and $\overline{C}_\alpha := \mathbb{M}$ otherwise. The region $\underline{C}_\alpha$ contains at most a fraction of α of all scenarios 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 scenarios in $\mathbb{M}$. Generally holds $\underline{C}_\alpha \subseteq \overline{C}_\alpha$.

Let's 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 the following Section 4. But how to choose the potential function? There are endless possibilities (see, e.g., Figure 1) to make the choice. A variation of the shape of the potential to improve the uncertainty model will be considered in Section 5.

4 Generation of potential clouds

This section will investigate how to find appropriate bounds on $F(V(\varepsilon))$. 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 [12] 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 the uncertainty information consists of given samples, boxes, non-formalized correlation bounds or continuous marginal CDFs F_i, $i \in I \subseteq \{1, 2, \ldots, n\}$, on the n-dimensional vector of uncertainties 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. [16]) inspired method to generate S.

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.,

$$\forall i, k \in \{1, 2, \ldots, N_S\}, \forall j \in \{1, 2, \ldots, n\} : x_i^j \neq x_k^j \qquad \text{if } k \neq i, \quad (4)$$

where x_i^j is the projection of x_i to the jth coordinate. Consider the following simple facts about the sample generation:

4.1 Proposition. *The intervals between adjacent marginal grid points have the same marginal probability.*

Proof. Let $k \in \{1, 2, \ldots, N_S - 1\}$: $\Pr(x \in [g_{ik}, g_{ik+1}]) = F_i(g_{ik+1}) - F_i(g_{ik}) = t_{k+1} - t_k = \frac{1-p_S}{N_S-1}$, which is constant and independent from k. ☐

4.2 Proposition. *Assume $I = \{1, 2, \ldots, n\}$, i.e., a marginal CDF is given for each coordinate of the uncertain scenarios ε. Assume ε^i, $i \in I$, are independent, then α_S is a confidence level for ε.*

Proof. $\Pr(\varepsilon \in [g_{11}, g_{1N_S}] \times [g_{21}, g_{2N_S}] \times \cdots \times [g_{n1}, g_{nN_S}]) = \prod_{i \in I} \Pr(\varepsilon^i \in [g_i 1, g_{iN_S}]) = \prod_{i=1}^{n}(F_i(g_{iN_S}) - F_i(g_{i1})) = \prod_{i=1}^{n}(t_{N_S} - t_1) = (1 - p_S)^n = \alpha_S$.
☐

4.3 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 increasing N_S.*

Proof. The LHS condition (4) 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 $N_S \to \infty$ 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 $\lim_{N_S \to \infty} F_i(s) - \widetilde{F}_i(s) = q' - q' = 0$. ☐

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, 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 uncertainty vector ε where a marginal CDF $F_i, i \in I$ is given. Then the weights should satisfy (5) $\forall i \in I, k = 1, 2, \ldots, N_S$

$$\sum_{j=1}^{k} \omega_{\pi_i(j)} \in [F_i(x^i_{\pi_i(k)}) - d, F_i(x^i_{\pi_i(k)}) + d], \qquad \sum_{k=1}^{N_S} \omega_k = 1. \qquad (5)$$

The function

$$\widetilde{F}_i(\xi) := \sum_{\{j \mid x^i_j \leq \xi\}} \omega_j \qquad (6)$$

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 4.3. The constraints (5) 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 $d = d_{\text{KS},1}$ with KS statistics, i.e.,

$$d_{\text{KS},1} = \frac{\phi^{-1}(a)}{\sqrt{N_S} + 0.12 + \frac{0.11}{\sqrt{N_S}}}, \qquad (7)$$

where ϕ is the Kolmogoroff function $\phi(\lambda) := \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2}$, $a = \alpha_\omega$ the confidence in the KS theorem, cf. [12], [20].

To achieve weights satisfying (5) we formulate the following linear program:

$$\min_{\omega_1, \omega_2, \ldots, \omega_{N_S}} e$$

$$\text{s.t.} \quad \widetilde{F}_i(\xi_k) \in [F_i(\xi_k) - e \cdot d_{\text{KSD},1}, F_i(\xi_k) + e \cdot d_{\text{KSD},1}] \quad \forall i, \xi_k$$

$$\sum_{i=1}^{N_S} \omega_i = 1 \qquad (8)$$

$$\omega_i \geq 0 \quad \forall i$$

$$e \geq 0$$

where ξ_k are some given interpolation points on the related margin, $\widetilde{F}_i$ the weighted empirical marginal distributions as defined in (6). Then we compute d from

$$d = \max_{i \in I, j=1,2,\ldots,N_S} |\widetilde{F}_i(x^i_j) - F_i(x^i_j)| \qquad (9)$$

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

It should be remarked that we make use of CVX, cf. [10], to solve (8), and of the MATLAB Statistics Toolbox to evaluate probability distributions.

By the weight computation we get a weighted empirical distribution

$$\widetilde{F}(\xi) := \sum_{\{j|V(x_j)\leq\xi\}} \omega_j \tag{10}$$

approximating the CDF of $V(\varepsilon)$. The achievement of weights satisfying (5) means, that all uncertainty information of the marginal CDFs is reflected in the construction of $\widetilde{F}$. The uncertainty information given as boxes or sample data is reflected anyway as this does not imply additional constraints to (5).

If weights satisfying (5) 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 (7), for a fixed confidence $a = \alpha_b$:

$$d_{\mathrm{KS},2} = \frac{\phi^{-1}(a)}{\sqrt{N_S} + 0.12 + \frac{0.11}{\sqrt{N_S}}} \cdot \frac{d}{d_{\mathrm{KS},1}}, \tag{11}$$

with the approximation quality factor $\frac{d}{d_{\mathrm{KS},1}}$ which corresponds to e in (8), 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 2. Observe that if the the quality of our approximation with $\widetilde{F}$ or the sample size N_S is decreased, the width of the bounds is increased correspondingly.

Thus we have found an appropriate bounding of the CDF $F(V(\varepsilon))$ and according to remark 3.6 we have generated a potential cloud via the mapping $\chi : \varepsilon \to [\underline{\alpha}(V(\varepsilon)), \overline{\alpha}(V(\varepsilon))]$.

The cloud represents the given uncertainty information and now enables us to interpret the potential level maps as confidence regions $\{\varepsilon \mid V(\varepsilon) \leq V_\alpha\}$ for our uncertain vector ε: the *worst-case relevant* region is defined as $\underline{C}_\alpha := \{\varepsilon \mid \overline{\alpha}(V(\varepsilon)) \leq \alpha\}$ (cf. Section 3), i.e., $\underline{C}_\alpha = \{\varepsilon \mid V(\varepsilon) \leq V_\alpha\}$ if a solution V_α of $\overline{\alpha}(V_\alpha) = \alpha$ exists, and $\underline{C}_\alpha = \emptyset$ otherwise, which is only the case for very low α.

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 empir-
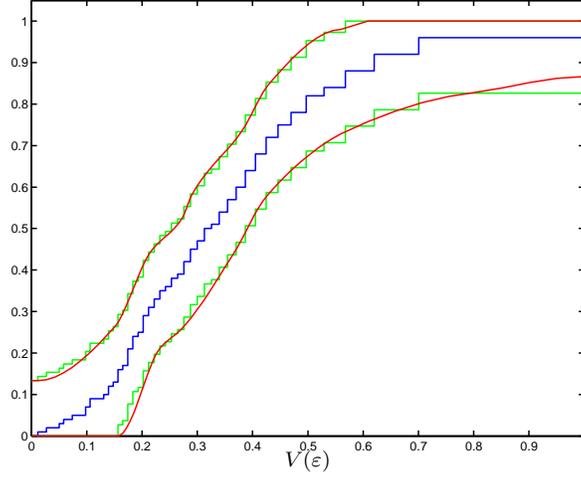
10

Figure 2: The smooth lower bounds $\underline{\alpha}(V(\varepsilon))$ and upper bounds $\overline{\alpha}(V(\varepsilon))$ for a potential cloud.

ical 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

$$V(\varepsilon) := \max_k \frac{|\varepsilon^k - \mu^k|}{r^k}, \tag{12}$$

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

$$V(\varepsilon) := \|A\varepsilon - b\|_2^2, \tag{13}$$

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

Figure 3 visualizes two confidence regions for the same confidence level $\alpha = 0.95$, but different potentials V, and shows that uncertainties can be described by

11

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.
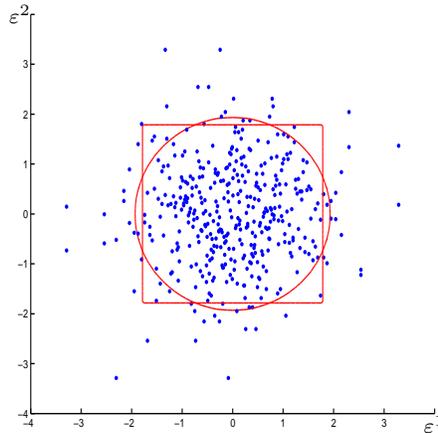

Figure 3: The regions $\underline{C}_{0.95}$ for two different choices of V, box- and circle-shaped, respectively. The 2-dimensional sample belongs to two independent $N(0,1)$-distributed random variables ε^1 and ε^2.

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 :

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

(14)

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

But how to achieve a polyhedron that reflects the given uncertainty information in the best way? As mentioned we assume the uncertainty information to consist of given samples, boxes or marginal distributions, and unformalized correlation 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 (12) 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 4. Thus an expert can specify the uncertainty information in the form of correlation bounds adaptively, even if the expert knowledge is only little formalized, resulting in a polyhedron shaped potential.
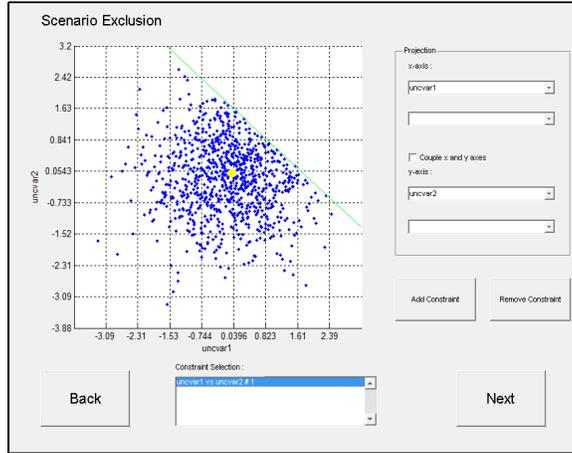

Figure 4: Graphical user interface for an interactive scenario exclusion.

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 (14) 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.,

$$V_{t,\mu}(\varepsilon) := V_\mu(\psi(\varepsilon)). \tag{15}$$

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 $V = V_{t,\mu}$ is visualized in Figure 5.

We can observe the advantage of a transformed potential in Figure 6. 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 7. The confidence regions for the transformed box potential are obviously smaller than for the non-transformed potential.
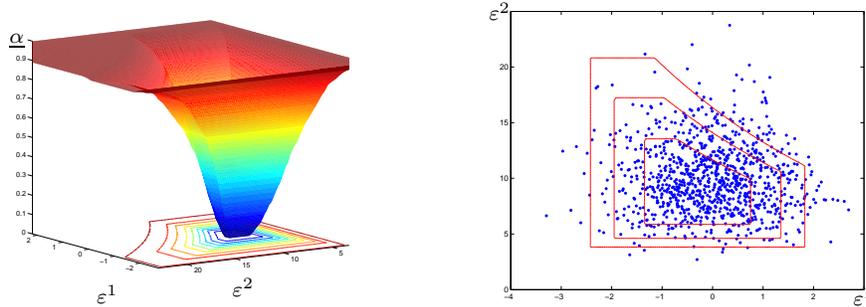
13

Figure 5: On the left, the map $\varepsilon \to \underline{\alpha}(V(\varepsilon))$ ($\underline{\alpha}(V(\varepsilon))$) 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.

6 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 vector of uncertain errors. One can formulate the optimization problem as a mixed-integer, bi-level problem of the following form:

$$
\begin{aligned}
\min_{\theta} \ \max_{x,z,\varepsilon} \ & g(x) && \text{(objective functions)} \\
\text{s.t.} \quad & G(x,z) = 0 && \text{(functional constraints)} \\
& z = Z(\theta) + \varepsilon && \text{(input constraints)} \\
& \theta \in T && \text{(selection constraints)} \\
& V_{t,0}(\varepsilon) \le \underline{V}_{\alpha} && \text{(cloud constraint)}
\end{aligned} \tag{16}
$$

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
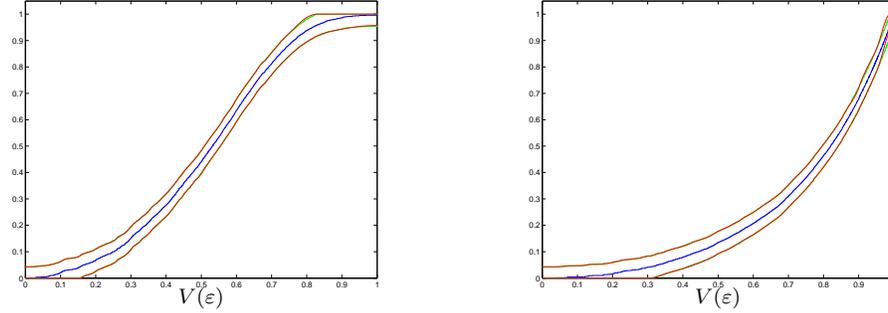
14

Figure 6: 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.
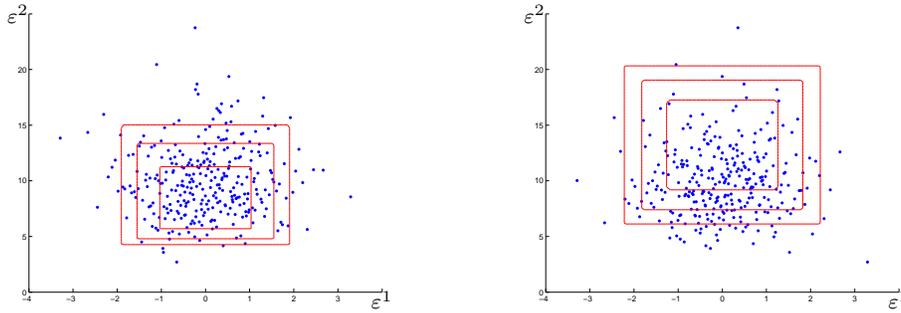

Figure 7: The figures show the regions $\underline{C}_\alpha$, $\alpha = 0.5, 0.8, 0.95$ for a box potential cloud with (left figure) and without (right figure) transformation. The two samples are generated for ε distributed as in Figure 6.

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. The cloud constraint involves the potential function $V = V_{t,0}(\varepsilon)$ as described in the Section 5 and models the worst-case relevant region $\{\varepsilon \mid V(\varepsilon) \le \underline{V}_\alpha\} = \underline{C}_\alpha$.

We already succeeded to apply this approach to robust optimization in real life problems arising from early phase spacecraft system design. For details see [8], [9], [18].

7　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 uncertainty 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 correlation bounds, is one of the key features. The iteration steps significantly improve the 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 to robust design optimization.

References

[1] Natalia M. Alexandrov and M. Yousuff Hussaini. Multidisciplinary design optimization: State of the art. In *Proceedings of the ICASE/NASA Langley Workshop on Multidisciplinary Design Optimization*, 1997.

[2] V. Belton and T. J. Stewart. *Multiple criteria decision analysis: an integrated approach*. Kluwer Academic Publishers, 2002.

[3] Daniel Berleant and Hang Cheng. A software tool for automatically verified operations on intervals and probability distributions. *Reliable Computing*, 4(1):71–82, 1998.

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

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

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

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

[8] Martin Fuchs, Daniela Girimonte, Dario Izzo, and Arnold Neumaier. Robust and autonomous space system design. submitted, 2007, Available on-line at: `http://www.mat.univie.ac.at/~mfuchs`.

[9] Martin Fuchs and Arnold Neumaier. Uncertainty modeling with clouds in autonomous robust design optimization. submitted, 2007, Available on-line at: `http://www.mat.univie.ac.at/~mfuchs`.

[10] Michael C. Grant and Stephen P. Boyd. CVX: A system for disciplined convex programming. 2007.
`http://www.stanford.edu/~boyd/cvx/cvx_usrguide.pdf`
`http://www.stanford.edu/~boyd/cvx/`.

[11] Patrick N. Koch, Timothy W. Simpson, Janet K. Allen, and Farrokh Mistree. Statistical approximations for multidisciplinary optimization: The problem of size. *Special Issue on Multidisciplinary Design Optimization of Journal of Aircraft*, 36(1):275–286, 1999.

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

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

[14] Kemper Lewis and Farrokh Mistree. Modeling interactions in multidisciplinary design: A game theoretic approach. *AIAA Journal*, 35(8):1387–1392, 1997.

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

[16] M.D. McKay, W.J. Conover, and R.J. Beckman. 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.

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

[18] A. Neumaier, M. Fuchs, E. Dolejsi, T. Csendes, J. Dombi, B. Banhelyi, and Z. Gera. Application of clouds for modeling uncertainties in robust space

system design. ACT Ariadna Research ACT-RPT-05-5201, European Space Agency, 2007. Available on-line at `http://www.esa.int/gsp/ACT/ariadna/completed.htm`.

[19] M. Pate-Cornell and P. Fischbeck. Probabilistic risk analysis and risk based priority scale for the tiles of the space shuttle. *Reliability Engineering and System Safety*, 40(3):221–238, 1993.

[20] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C*. Cambridge University Press, second edition, 1992.

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

[22] B. Roy. *Multicriteria methodology for decision aiding*. Kluwer Academic Publishers, 1996.

[23] R. C. Williamson. *Probabilistic Arithmetic*. PhD thesis, University of Queensland, 1989.