

# Design Optimization with Imprecise Random Variables

Jeffrey W Herrmann

The  
Institute for  
**Systems**  
Research



**A. JAMES CLARK**  
SCHOOL OF ENGINEERING

ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.

ISR is a permanent institute of the University of Maryland, within the A. James Clark School of Engineering. It is a graduated National Science Foundation Engineering Research Center.

[www.isr.umd.edu](http://www.isr.umd.edu)

# Design Optimization with Imprecise Random Variables

**Jeffrey W. Herrmann**

Department of Mechanical Engineering  
and Institute for Systems Research  
University of Maryland  
College Park, MD 20742  
jwh2@umd.edu

## **ABSTRACT**

Design optimization is an important engineering design activity. Performing design optimization in the presence of uncertainty has been an active area of research. The approaches used require modeling the random variables using precise probability distributions or representing uncertain quantities as fuzzy sets. This work, however, considers problems in which the random variables are described with imprecise probability distributions, which are highly relevant when there is limited information about the distribution of a random variable. In particular, this paper formulates the imprecise probability design optimization problem and presents an approach for solving it. We present examples for illustrating the approach.

## **1. INTRODUCTION**

Design optimization is an important engineering design activity in automotive, aerospace, and other development processes. In general, design optimization determines values for design variables such that an objective function is optimized while performance and other constraints are satisfied (Papalambros and Wilde, 2000; Ravindran *et al.*, 2006; Arora, 2004). The use of design optimization in engineering design continues to increase, driven by more powerful software packages and the formulation of new design optimization problems motivated by the decision-based design (DBD) framework (Hazelrigg, 1998; Renaud and Gu, 2006).

Because many engineering problems must be solved in the presence of uncertainty, developing approaches for solving design optimization problems that have uncertain variables has been an active area of research. The approaches used require modeling the random variables using precise probability distributions or representing uncertain quantities as fuzzy sets. Haldar and Mahadevan (2000) give a

general introduction to reliability-based design optimization, and many different solution techniques have been developed. For examples, see Liang *et al.* (2007), Tu *et al.* (1999), Youn (2007), and the references therein. Other approaches include evidence-based design optimization (Zhou and Mourelatos, 2008b), possibility-based design optimization (Zhou and Mourelatos, 2008a), and approaches that combine possibilities and probabilities (Nikolaidis, 2007). Zhou and Mourelatos (2008b) discussed an evidence theory-based design optimization (EBDO) problem. They used a hybrid approach that first solves a RBDO to get close to the optimal solution and then generates response surfaces for the active constraints and uses a derivative-free optimizer to find a solution.

Different models of uncertainty are more or less appropriate given

The amount of information available and the outlook of the decision-maker (design engineer) determines the appropriateness of different models of uncertainty. No single model should be considered universally valid. In this paper, we consider situations in which there is insufficient information about the random variables to model them with precise probability distributions. Instead, imprecise probability distributions (described in more detail below) are used to capture the limited information or knowledge. In the extreme case, the imprecise probability distribution may be a simple interval. This paper presents an approach for solving design optimization problems in which the random variables are described with imprecise probability distributions because there exists limited information about the uncertainties.

## **2. IMPRECISE PROBABILITIES**

In traditional probability theory, the probability of an event is defined by a single number between in the range  $[0, 1]$ . However, because this may be inappropriate in cases of incomplete or conflicting information, researchers have proposed theories of imprecise probabilities. For these situations, probabilities can be intervals or sets, rather than precise numbers (Dempster, 1967; Walley, 1991; Weichselberger, 2000). The theory of imprecise probabilities, formalized by Walley (1991), uses the same fundamental notion of rationality as the work of de Finetti (1974, 1980). However, the theory allows a range of indeterminacy—prices at which a decision-maker will not enter a gamble as either a buyer or a seller. These in turn correspond to ranges of probabilities.

Imprecise probabilities have previously been considered in reliability analysis (Coolen, 2004; Utkin, 2004a, b) and engineering design (Aughenbaugh and Paredis, 2006; Ling *et al.*, 2006; Rekuc *et al.*, 2006).

Aughenbaugh and Herrmann (2007, 2008a, b) have compared techniques using imprecise probabilities to other statistical approaches for making reliability-based design decisions. The work described in this paper builds upon these previous results.

In many engineering applications, the relevant random variables (e.g., parameters or manufacturing errors) are continuous variables. One common way to represent the imprecision in the probability distribution of such a random variable is a probability box (“p-box”) that is a set of cumulative probability distributions bounded by an upper distribution  $\bar{F}$  and a lower distribution  $\underline{F}$ . These bounds model the epistemic uncertainty about the probability distribution for the random variable. Of course, a traditional precise probability distribution is a special case of a p-box, in which the upper and lower bounds are equal.

There are multiple ways to construct a p-box for a random variable (Ferson *et al.*, 2003). In some cases, the type of distribution is known (or assumed) but its parameters are imprecise (such as an interval for a mean). In other cases, the distribution is constructed from sample data. Additionally, one can create a p-box from a Dempster-Shafer structure, in which intervals (not points) within the range of the random variable are assigned probabilities. For more about p-boxes and the link between p-box representation and Dempster-Shafer structures, see Ferson *et al.* (2004).

Functions of random variables that have imprecise probability distributions also have imprecise probability distributions. Methods for calculating these convolutions are given by Yager (1986), Williamson and Downs (1990), and Berleant (1993, 1996). Wang (2008) proposes a new interval arithmetic that could be used as well.

Therefore, p-boxes are a very general way to represent uncertainty. For computational purposes, in the approach below, we will convert a p-box into a “canonical” Dempster-Shafer structure (Ferson *et al.* 2003), which will necessarily be bounded.

### **3. DESIGN OPTIMIZATION WITH IMPRECISE PROBABILITIES**

In the imprecise probability design optimization (IPDO) problem, there is a set of deterministic design variables for which the designer chooses values and a set of random variables, which may be manufacturing errors, uncertain engineering parameters, or other sources of uncertainty. Unlike other

work, this formulation does not include in the model “random design variables.” Such variables are typically those in which the designer chooses the mean, but the actual value is random. In the IPDO formulation presented here, each such quantity is modeled with two quantities: a deterministic design variable and a random parameter that represents the error of that variable. This does not limit the scope of the model. For instance, suppose we have a “random design variable”  $X$  that is a dimension of a part. The mean of  $X$ , denoted  $\mu_X$ , is chosen by the designer, but the dimension is a normally distributed random variable with a standard deviation of  $\sigma$ . Examples of this type of variable have been considered in Zhou and Mourelatos (2008a) and elsewhere. In this formulation, we replace the variable  $X$  with  $X = d_X + Z_X$ , where the first term, which corresponds to the mean, is a deterministic design variable, and the second term is a random variable that is normally distributed with a mean of 0 and a standard deviation of  $\sigma$ .

The general IPDO is formulated as follows:

$$\begin{aligned} & \min_{\mathbf{d}} V[f(\mathbf{d}, \mathbf{Z})] \\ \text{s.t. } & \bar{P}\{g_i(\mathbf{d}, \mathbf{Z}) \leq 0\} \leq p_i \quad i = 1, \dots, n \\ & \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U \end{aligned} \quad (1)$$

In this formulation,  $\mathbf{d} \in R^k$  is the vector of deterministic design variables, and  $\mathbf{Z} \in R^r$  is the vector of random variables that have imprecise probability distributions. The probabilistic constraints are functions of the deterministic design variables and the random variables. We want  $g_i(\mathbf{d}, \mathbf{Z}) \geq 0$  (which is the “safe region”) but will be satisfied if the upper bound on the failure probability is less than the target  $p_i$ . We choose the upper probability in order to be conservative.

The function  $f$  is the system performance, which may be random, in which case the function  $V$  is a moment of that random performance, such as the upper limit for the mean; thus  $V$  is a deterministic function of  $\mathbf{d}$ . In many cases, the objective is specified as a function of only the deterministic design variables, in which case we get the following formulation:

$$\begin{aligned} & \min_{\mathbf{d}} f(\mathbf{d}) \\ \text{s.t. } & \bar{P}\{g_i(\mathbf{d}, \mathbf{Z}) \leq 0\} \leq p_i \quad i = 1, \dots, n \\ & \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U \end{aligned} \quad (2)$$

First, this formulation has the usual difficulty of computing the failure probability for each constraint. Analytically evaluating the failure probability is possible only in special cases. An additional complication is the imprecision of the random variables, which makes applying standard RBDO techniques difficult.

Now,  $\mathbf{Z}$  has an imprecise joint probability distribution, which can be considered as a set  $H$  of precise joint probability distributions. For any precise joint probability distribution  $F_j \in H$ , let

$P_j \{g_i(\mathbf{d}, \mathbf{Z}) \leq 0\}$  be the probability of violating constraint  $i$  when  $\mathbf{Z}$  has that precise joint probability distribution. Then, we could reformulate the IPDO as the following RBDO:

$$\begin{aligned} & \min_{\mathbf{d}} f(\mathbf{d}) \\ \text{s.t. } & P_j \{g_i(\mathbf{d}, \mathbf{Z}) \leq 0\} \leq p_i \quad i=1, \dots, n, F_j \in H \\ & \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U \end{aligned} \quad (3)$$

Unfortunately, because of the large number of constraints, this reformulation is not helpful unless the set  $H$  is limited to a reasonable number of “extreme” distributions that can be used as surrogates for the entire set. Research on this topic is ongoing and may provide a way to increase the computational efficiency of IPDO in the future.

Due to these difficulties, we will pursue a numerical approach. To do this, we will first partition the constraints  $g_i(\mathbf{d}, \mathbf{Z}) \geq 0$  into two sets. Set  $S_1$  includes any constraint that can be rearranged so that  $g_i(\mathbf{d}, \mathbf{Z}) = a_i(h_i(\mathbf{Z}) - b_i(\mathbf{d}))$ , where  $a_i$  is a positive scalar. Note  $g_i(\mathbf{d}, \mathbf{Z}) \leq 0$  if and only if  $h_i(\mathbf{Z}) \leq b_i(\mathbf{d})$ . The constraints that cannot be rearranged in this way are placed in set  $S_2$ .

For each constraint in  $S_1$ , we will perform the convolution needed to get the imprecise distribution of  $h_i(\mathbf{Z})$  by combining the Dempster-Shafer structures for the relevant random variables. Because the upper cumulative probability distribution will be a discontinuous function, we will approximate it with  $F_i(x) \approx \bar{P}\{h_i(\mathbf{Z}) \leq x\}$ . Therefore, we can replace each of the constraints in  $S_1$  by  $F_i(b_i(\mathbf{d})) \leq p_i$ .

#### 4. SOLUTION APPROACH

To solve the IPDO, we will use a sequential approach similar to that of Du and Chen (2004) and Zhou and Mourelatos (2008a).

A key part of the approach is to solve the following deterministic optimization problem  $P$  given values for the random variables in each constraint in  $S_2$ . Let  $\mathbf{Z}^{i(k)}$  be specific values for the random variables in constraint  $i \in S_2$  in iteration  $k$ .

$$\begin{aligned}
 & \min_{\mathbf{d}} f(\mathbf{d}) \\
 \text{s.t. } & F_i(b_i(\mathbf{d})) \leq p_i \quad i \in S_1 \\
 & g_i(\mathbf{d}, \mathbf{Z}^{i(k)}) \geq 0 \quad i \in S_2 \\
 & \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
 \end{aligned} \tag{4}$$

In the space of the design variables, the  $g_i(\mathbf{d}, \mathbf{Z}^{i(k)}) \geq 0$  constraints move the boundaries of the “safe region” (by making it smaller) in order to reduce the probability of failure. However, it is still necessary to determine the probability of failure and compare it to the target. If it is too large, then we have to move that constraint some more.

Given these preliminaries, the complete approach follows:

1. Let  $k = 0$ . For  $i \in S_2$ , let each component of  $\mathbf{Z}^{i(k)}$  equal a value within the range of its expected value.
2. Solve  $P$  to get the solution  $\mathbf{d}^{(k+1)}$ .
3. For  $i \in S_2$ , evaluate  $\bar{P}\{g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}) \leq 0\}$ . If  $\bar{P}\{g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}) \leq 0\} \leq p_i$  for all  $i \in S_2$ , then the design point is feasible; stop. Otherwise, for all constraints  $i \in S_2$  where  $\bar{P}\{g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}) \leq 0\} \leq p_i$ , set  $\mathbf{Z}^{i(k+1)} = \mathbf{Z}^{i(k)}$ . For the others, find  $\mathbf{Z}^{i(k+1)}$  by solving the following problem:

$$\begin{aligned} & \min_{\mathbf{Z}^{i(k+1)}} g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}^{i(k+1)}) \\ \text{s.t. } & \bar{P}\left\{g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}) \leq g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}^{i(k+1)})\right\} = p_i \end{aligned} \quad (5)$$

This yields a “very bad” (but not worst-case) value of those random variables used in that constraint. (A technique for solving this problem is described below.)

4.  $k = k + 1$ . Repeat steps 2 and 3 until a feasible design point is found.

At this point we have no proof that the algorithm will converge, and the approach may fail on problems with irregular objective functions and constraints. Further analysis and experimentation is needed to study this aspect of the method.

We use the following reliability analysis technique to determine if  $\bar{P}\left\{g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}) \leq 0\right\} \leq p_i$  and to find  $\mathbf{Z}^{i(k+1)}$ . This reliability analysis technique corresponds roughly to solving the inverse “most probable point” problem discussed in Du and Chen (2004) or finding the “shifting vector” of Zhou and Mourelatos (2008a). Given  $\mathbf{d}^{(k+1)}$ , set  $p_f = 0$ . Without loss of generality, we assume that  $g_i(\mathbf{d}^{(k+1)}, \mathbf{Z})$  is a function of  $m$  random variables  $Z_1, \dots, Z_m$ . The Dempster-Shafer structure of  $Z_i$  is represented by  $n_i$  equally likely intervals. Let  $N = n_1 n_2 \dots n_m$ . Let  $N^* = \lfloor p_i N \rfloor$  be the number of values and intervals to save. Consider each of the  $N$  combinations of intervals for the random variables. For each combination, let each random variable range over its interval and find  $g_i^{\min}$ , the minimum of  $g_i(\mathbf{d}^{(k+1)}, \mathbf{Z})$  for that combination. If  $g_i^{\min} \leq 0$ , add  $\frac{1}{N}$  (the probability of that combination) to  $p_f$ . As the  $N$  combinations of intervals are checked, keep the  $N^*$  smallest minima found along with the values of  $Z_1, \dots, Z_m$  that yield those minima. The final value of  $p_f$  is used to estimate  $\bar{P}\left\{g_i(\mathbf{d}^{(k+1)}, \mathbf{Z}) \leq 0\right\}$ . We set  $\mathbf{Z}^{i(k+1)}$  equal to the values of  $Z_1, \dots, Z_m$  that yield the largest of the  $N^*$  smallest minima found.

## 5. EXAMPLES

This section presents three examples to illustrate the IPDO solution method. The first example has two design variables and three random variables:



$$\begin{aligned}
& \min_{\mathbf{d}} d_1 + d_2 \\
& \text{s.t. } \bar{P}\left\{z_3 - \frac{4}{d_1} \leq 0\right\} \leq 0.02 \\
& \bar{P}\left\{(d_1 + z_1)^2 (d_2 + z_2) - 20 \leq 0\right\} \leq 0.02 \\
& \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
\end{aligned} \tag{6}$$

The bounds  $\mathbf{d}^L = (0.01, 0.01)$  and  $\mathbf{d}^U = (10, 10)$ . Note that the first constraint is in set  $S_1$ , whereas the second one is in set  $S_2$ .

All three random variables have imprecise probability distributions. The random errors  $z_1$  and  $z_2$  have the same distribution, each characterized by the intervals  $[-1 + 1.5(k-1)/99, -0.5 + 1.5(k-1)/99]$ , for  $k = 1, \dots, 100$  (each interval has a probability of 0.01). Therefore, they can range from -1 to 1. The distribution of random parameter  $z_3$  is characterized by the intervals  $[1 + 0.5(k-1)/99, 1.5 + 0.5(k-1)/99]$ , for  $k = 1, \dots, 100$  (each interval has a probability of 0.01). This random parameter ranges from 1 to 2.

For  $z_3$ , we will approximate its upper cumulative probability as follows:

$$\bar{P}\{z_3 \leq x\} = \begin{cases} 0 & \text{if } x < 1 \\ 2(x-1) & \text{if } 1 \leq x \leq 1.5 \\ 1 & \text{if } x > 1.5 \end{cases} \tag{7}$$

The IPDO solution approach begins with  $z_1$  and  $z_2$  both set to zero; that is,  $\mathbf{Z}^{2(0)} = (0, 0)$ :

$$\begin{aligned}
& \min_{\mathbf{d}} d_1 + d_2 \\
& \text{s.t. } \bar{P}\left\{z_3 - \frac{4}{d_1} \leq 0\right\} \leq 0.02 \\
& d_1^2 d_2 \geq 20 \\
& \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
\end{aligned} \tag{8}$$

This yields  $\mathbf{d}^{*(1)} = (3.9604, 1.2751)$ , but the upper probability of violating the second constraint is too high. Our reliability analysis technique estimates that  $\bar{P}\left\{20 - (d_1^{*(1)} + z_1)^2 (d_2^{*(1)} + z_2) \leq 0\right\} = 0.7830$ .

Because there are  $10,000 = 100^2$  combinations of intervals, we keep the worst 200 interval lower bounds.  $(z_1^{2(1)}, z_2^{2(1)}) = (-0.5303, -0.9697)$  gives the best of these worst. (The first superscript refers to the second constraint, the second to the iteration number.)

Now we solve by adding the shifting vector to the second constraint:

$$\begin{aligned}
& \min_{\mathbf{d}} d_1 + d_2 \\
& \text{s.t. } \bar{P} \left\{ z_3 - \frac{4}{d_1} \leq 0 \right\} \leq 0.02 \\
& (d_1 - 0.5303)^2 (d_2 - 0.9697) \geq 20 \\
& \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
\end{aligned} \tag{9}$$

This yields  $\mathbf{d}^{*(2)} = (3.9604, 2.6696)$ . Our reliability analysis technique estimates that

$$\bar{P} \left\{ 20 - (d_1^{*(2)} + z_1)^2 (d_2^{*(2)} + z_2) \leq 0 \right\} = 0.0655.$$

Thus, the upper probability of violating the second constraint is lower but still too high. We also determine that  $(z_1^{2(2)}, z_2^{2(2)}) = (-0.7727, -0.9545)$  gives the best of the worst for this design point.

Now we solve with the new shifting vector:

$$\begin{aligned}
& \min_{\mathbf{d}} d_1 + d_2 \\
& \text{s.t. } \bar{P} \left\{ z_3 - \frac{4}{d_1} \leq 0 \right\} \leq 0.02 \\
& (d_1 - 0.7727)^2 (d_2 - 0.9545) \geq 20 \\
& \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U
\end{aligned} \tag{10}$$

This yields  $\mathbf{d}^{*(3)} = (4.1927, 2.6645)$ . The first constraint is not active, but the upper probability of violating the second constraint is now acceptable. Our reliability analysis technique estimates that

$$\bar{P} \left\{ 20 - (d_1^{*(3)} + z_1)^2 (d_2^{*(3)} + z_2) \leq 0 \right\} = 0.0194.$$

The second example is the mathematical example from Zhou and Mourelatos (2008a). In our version, the problem has two design variables  $\mathbf{d} = (d_1, d_2)$  and two random variables, the error for each one:

$\mathbf{Z} = (z_1, z_2)$ . The bounds are  $0 \leq d_i \leq 10$  for both design variables. The objective is to minimize the sum of the design variables. In the terms of our general IPDO, we have the following relationships:

$$\begin{aligned}
 f(\mathbf{d}) &= d_1 + d_2 \\
 g_1(\mathbf{d}, \mathbf{Z}) &= (d_1 + z_1)^2 (d_2 + z_2) - 20 \\
 g_2(\mathbf{d}, \mathbf{Z}) &= 4(d_1 + z_1 + d_2 + z_2 - 5)^2 + (d_1 + z_1 - d_2 - z_2 - 12)^2 - 120 \\
 g_3(\mathbf{d}, \mathbf{Z}) &= 75 - (d_1 + z_1)^2 - 8(d_2 + z_2)
 \end{aligned} \tag{11}$$

Both random variables have the same imprecise probability distribution, which is characterized by the intervals  $[-1 + 1.5(k-1)/99, -0.5 + 1.5(k-1)/99]$ , for  $k = 1, \dots, 100$  (each interval has a probability of 0.01). Therefore, they can range from -1 to 1.

We cannot separate any of the constraints, so  $S_1 = \emptyset$  and  $S_2 = \{1, 2, 3\}$ . The two random variables have the same imprecise probability distribution, which is approximately an imprecise uniform distribution with a lower bound in the range [-1, 0.5] and the upper bound in the range [0.5, 1]. Figure 1 shows the actual p-box.

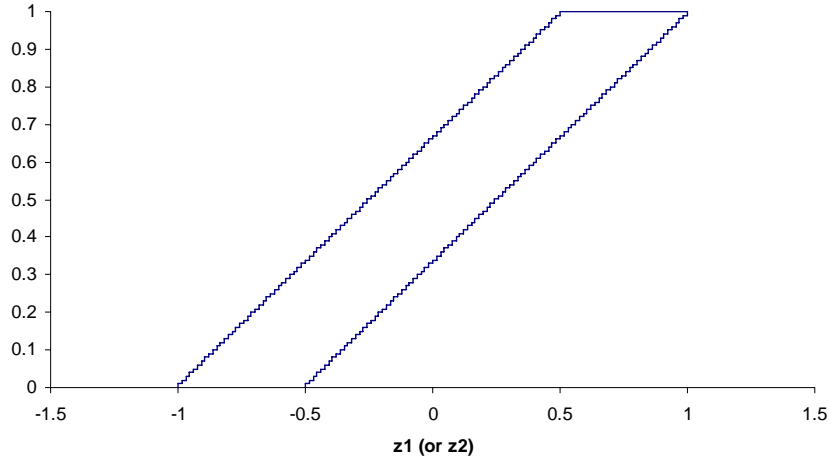


Figure 1. The p-box for  $z_1$  (and  $z_2$ ).

First, we will solve the deterministic optimization problem with both random variables replaced by 0. We get  $\mathbf{d} = (3.1139, 2.0626)$ , the same optimal solution as Zhou and Mourelatos (2008a). The objective function value is 5.1765.

Next, we let  $\mathbf{Z}^{i(0)} = (0, 0)$  for all  $i \in S_2$  and solve problem  $P$ . This yields the solution

$\mathbf{d}^{(1)} = (3.1139, 2.0626)$ .  $\bar{P}\{g_i(\mathbf{d}^{(1)}, \mathbf{Z}) \leq 0\}$  is greater than 0.02 for the first two constraints but equals zero for the third constraint.

When evaluating this probability, we have to compare 10,000 combinations, in which each combination has an interval from the p-box for  $z_1$  and an interval from the p-box for  $z_2$ . For each combination, we must find the minimal value of  $g_i$  over those values of  $z_1$  and  $z_2$ . Based on the mathematical analysis of the constraints, it is possible to develop simple rules to identify the values of  $z_1$  and  $z_2$  that give the minimum for that combination.

Let  $[z_1^{\min}, z_1^{\max}]$  and  $[z_2^{\min}, z_2^{\max}]$  be the intervals that form the combination. For the first constraint, the minimum is found at  $z_2 = z_2^{\min}$ , and  $z_1$  is either an endpoint of the interval or  $-d_1$ . For the second constraint, the minimum is found at one of the following five points:  $(z_1^{\min}, z_2^{\min})$ ,  $(6.4 - d_1 - 0.6d_2 - 0.6z_2^{\min}, z_2^{\min})$ ,  $(z_1^{\max}, z_2^{\min})$ ,  $(z_1^{\max}, 1.6 - 0.6d_1 - d_2 - 0.6z_1^{\max})$ , or  $(z_1^{\max}, z_2^{\max})$ . For third constraint, the minimum is found at  $z_2 = z_2^{\max}$ , and  $z_1$  is one of the endpoints of its interval.

From this algorithm we set  $(z_1^{1(1)}, z_2^{1(1)}) = (-0.7121, -1)$  and  $(z_1^{2(1)}, z_2^{2(1)}) = (0.8939, -0.8182)$ . We will use these two vectors in the first two constraints as we try to find a feasible solution in the next iteration. For the third constraint, which was already feasible, we let  $(z_1^{3(1)}, z_2^{3(1)}) = (0, 0)$ .

The second iteration of the problem yields the solution  $\mathbf{d}^{(2)} = (3.5836, 3.4255)$ . At this point, the objective function equals 7.0091.  $\bar{P}\{g_i(\mathbf{d}^{(2)}, \mathbf{Z}) \leq 0\}$  is again greater than 0.02 for the first two

constraints but equals zero for the third constraint. We set  $(z_1^{1(2)}, z_2^{1(2)}) = (-1, -0.5909)$  and  $(z_1^{2(2)}, z_2^{2(2)}) = (0.3939, -1)$ . For the third constraint,  $(z_1^{3(2)}, z_2^{3(2)}) = (0, 0)$ .

The third iteration of the problem yields the solution  $\mathbf{d}^{(3)} = (3.6113, 3.5240)$ . At this point, the objective function equals 7.1353.  $\bar{P}\{g_i(\mathbf{d}^{(1)}, \mathbf{Z}) \leq 0\}$  is less than 0.02 for the first two constraints and equals zero for the third constraint, so the solution is feasible.

The third example that we consider is the optimization of a thin-walled pressure vessel. Our formulation is based on the RBDO formulation of Zhou and Mourelatos (2008b). The problem was originally introduced by Lewis and Mistree (1997). The problem has three design variables: the radius  $R$ , the mid-section length  $L$ , and the wall thickness  $t$ . The objective is to maximize the volume of the pressure vessel. Five constraints ensure that the design is strong enough to resist the internal pressure (with a safety factor of 2) and meets geometric requirements.

In our formulation there are three random variables: the manufacturing error of the radius, the internal pressure  $P$ , and the material yielding strength  $Y$ .

In the terms of our general IPDO, we have  $\mathbf{d} = (R, L, t)$  and  $\mathbf{Z} = (z_1, P, Y)$  and the following relationships:

$$\begin{aligned}
f(\mathbf{d}) &= \frac{4}{3}\pi R^3 + \pi R^2 L \\
g_1(\mathbf{d}, \mathbf{Z}) &= 2tY - 2P\left(R + z_1 + \frac{1}{2}t\right) \\
g_2(\mathbf{d}, \mathbf{Z}) &= \left(2(R + z_1)t + t^2\right)Y - 2P\left(2(R + z_1)^2 + 2(R + z_1)t + t^2\right) \\
g_3(\mathbf{d}, \mathbf{Z}) &= 60 - (L + 2(R + z_1) + 2t) \\
g_4(\mathbf{d}, \mathbf{Z}) &= 12 - (R + z_1 + t) \\
g_5(\mathbf{d}, \mathbf{Z}) &= R + z_1 - 5t
\end{aligned} \tag{12}$$

We set  $p_i = 0.02$  for  $i = 1, \dots, 5$ . The bounds on the design variables are the following ranges:

$5 \leq R \leq 24$ ,  $10 \leq L \leq 48$ , and  $0.25 \leq t \leq 2$ . The last three constraints, which form set  $S_1 = \{3, 4, 5\}$ , can be rearranged as follows:

$$\begin{aligned}
g_3(\mathbf{d}, \mathbf{Z}) &= 2(h_3(z_1) - b_3(\mathbf{d})) = 2(-z_1 - (\frac{1}{2}L + R + t - 30)) \\
g_4(\mathbf{d}, \mathbf{Z}) &= h_4(z_1) - b_4(\mathbf{d}) = -z_1 - (R + t - 12) \\
g_5(\mathbf{d}, \mathbf{Z}) &= h_5(z_1) - b_5(\mathbf{d}) = z_1 - (5t - R)
\end{aligned} \tag{13}$$

Therefore,  $S_2 = \{1, 2\}$ . Each of the three random variables has an imprecise probability distribution. The imprecise probability distribution of the internal pressure  $P$  is approximately an imprecise normal distribution. The imprecise mean has a range of [975, 1025]. The standard deviation is precisely 50. The imprecise probability distribution of the material yielding strength  $Y$  is also approximately an imprecise normal distribution. The imprecise mean has a range of [253500, 266500]. The standard deviation is precisely 13000.

The actual p-boxes used for these two random variables are constructed as follows: for  $k = 1, \dots, 100$ , let  $f_k = (2k - 1)/200$ , which therefore ranges from 0.005 to 0.995. The  $k$ -th interval in the p-box for the internal pressure  $P$  is  $[975 + 50\Phi^{-1}(f_k), 1025 + 50\Phi^{-1}(f_k)]$ , and the  $k$ -th interval in the p-box for the material yielding strength  $Y$  is  $[253500 + 13000\Phi^{-1}(f_k), 266500 + 13000\Phi^{-1}(f_k)]$ .

The imprecise probability distribution of  $z_1$ , the manufacturing error of the radius, is based on data given by Zhou and Mourelatos (2008b), who assume that we have 100 sample points for the error, but the data are given only in bins as follows: 3 points are in the range [-4.5, -3], 45 points are in the range [-3, 0], 49 points are in the range [0, 3], 2 points are in the range [3, 4.5], and 1 point is in the range [4.5, 6]. Figure 2 shows the corresponding p-box for  $z_1$  and the curve we use for approximating the upper bound of this p-box.

An approximation is created for each part of the p-box. For  $x \in [z^L, z^U]$  where the lower left corner of the upper bound is  $(z^L, F^L)$  and the upper right corner of the upper bound is  $(z^U, F^U)$ , the approximation  $F_i(x) = F^U - (F^U - F^L) \sqrt{1 - \left(\frac{x - z^L}{z^U - z^L}\right)^2}$ . Figure 3 shows the corresponding p-box for

$-z_1$  and the curve we use for approximating the upper bound of this p-box. Because

$F_i(x) \geq \bar{P}\{h_i(\mathbf{Z}) \leq x\}$ , the approximation reduces the feasible region. If a solution is feasible with respect to the approximation, then it is feasible with respect to the original p-box.

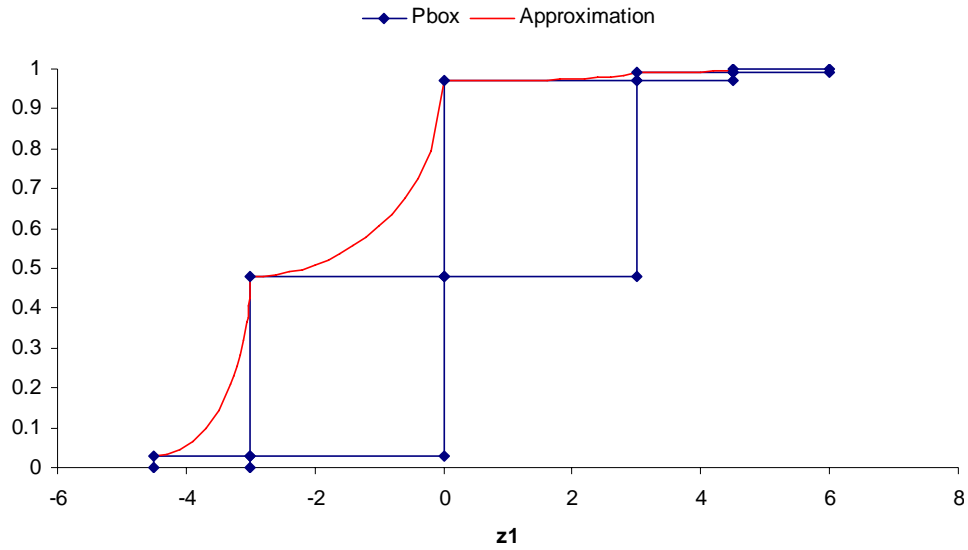


Figure 2. The p-box for  $z_1$  and the approximation for its upper bound.

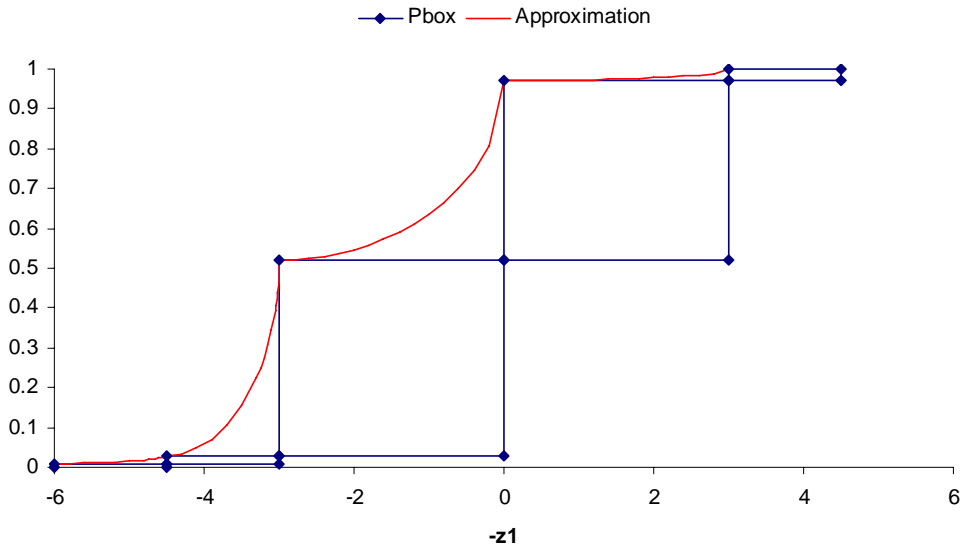


Figure 3. The p-box for  $-z_1$  and the approximation for its upper bound.

First, we will solve the deterministic optimization problem with all three random variables replaced by constants:  $\mathbf{Z} = (z_1, P, Y) = (0, 1000, 260000)$ .  $\mathbf{d} = (R, L, t) = (11.75, 36, 0.25)$  is the optimal solution that we found. The pressure vessel volume equals 22,410. Note that the probability of failure for constraints

4 and 5 equals the probability that  $-z_1$  is less than equal to zero, which is imprecise but can be quite large, so it is not a feasible solution to the IPDO problem.

Next, we let  $\mathbf{Z}^{1(0)} = \mathbf{Z}^{2(0)} = (0, 1000, 260000)$  and solve problem  $P$ . This yields the solution  $\mathbf{d}^{(1)} = (6.7606, 36.4396, 0.3186)$ . (The optimization required 348 function evaluations.) The pressure vessel volume equals 6,527. The reliability analysis technique shows that, for all  $i \in S_2$ ,  $\bar{P}\{g_i(\mathbf{d}^{(1)}, \mathbf{Z}) \leq 0\} = 0$ , so the solution is feasible, and the algorithm stops.

## 6. COMPARISON TO RBDO

The IPDO addresses situations in which probability distributions are not precise. An alternative approach is to use a traditional RBDO approach while varying the probability distributions of the random variables. The basic idea is to loop over different combinations of the distributions for the random variables. For each combination, we solve a traditional RBDO problem to get a solution. This procedure will yield a set of solutions and gives the designer some idea of where good solutions lie. But it is not clear how a designer should select a solution from this set.

Another alternative is to remove the imprecision. For instance, one can replace each imprecise probability distribution by the maximum entropy probability distribution that fits within the p-box. This yields an RBDO problem. For the first example in Section 5, we can model  $z_1$ ,  $z_2$ , and  $z_3$  with uniform distributions. The range for  $z_1$  and  $z_2$  is  $[-1, 1]$ , and the range for  $z_3$  is  $[1, 2]$ . Solving the RBDO problem yields the solution  $\mathbf{d} = (4.2227, 2.5132)$ . The objective function value is 6.7359, which is better than that of the more conservative IPDO solution, but the probability of failure of this new solution is greater than the desired target for some of the probability distributions in the p-boxes of the random variables.

## 7. SUMMARY

This paper introduced the imprecise probability design optimization (IPDO) problem, in which there is a set of deterministic design variables for which the designer chooses values and a set of imprecise random variables, which may be manufacturing errors, uncertain engineering parameters, or other sources of



uncertainty. This paper presented a sequential approach for solving this problem. To avoid unnecessary calculations, the approach partitions the constraints into two sets. By exploiting their special structure, the cumulative probability distributions for constraints in the first set are calculated only once and then replaced with an approximation. After this, the approach solves a series of deterministic optimization problems and shifts selected constraints in each iteration in order to reduce the probability of failure.

We have used examples to illustrate the usefulness of the approach. The results show that the proposed IPDO approach finds high-quality feasible solutions, though the computational effort is increased because of the computational effort of the reliability analysis technique and the iterations needed to converge to a solution.

Although this work was motivated by problems in which only imprecise probability distributions are available, the approach's use of Dempster-Shafer structures makes it compatible with other approaches within the domain of evidence theory as well (Shafer, 1976).

Future work will focus on improving the computational efficiency and stability of the approach by considering adaptive loop-methods similar to those proposed by Youn (2007).

## **ACKNOWLEDGEMENTS**

This paper reuses background material and references from previous papers written with co-author Jason Aughenbaugh, whose assistance the author greatly appreciates.

## **REFERENCES**

- Arora, Jasbir S., *Introduction to Optimum Design*, 2nd edition, Elsevier Academic Press, Amsterdam, 2004.
- Aughenbaugh, Jason M., and Jeffrey W. Herrmann, "Updating uncertainty assessments: a comparison of statistical approaches," Proceedings of IDETC/CIE 2007, ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, September 4-7, 2007, Las Vegas, Nevada, DETC2007-35158.
- Aughenbaugh, Jason M., and Jeffrey W. Herrmann, "A Comparison of Information Management using Imprecise Probabilities and Precise Bayesian Updating of Reliability Estimates," Proceedings of

- the Third International Workshop on Reliable Engineering Computing (REC), NSF Workshop on Imprecise Probability in Engineering Analysis & Design, Savannah, Georgia, February 20-22, 2008a.
- Aughenbaugh, Jason M., and Jeffrey W. Herrmann, "Reliability-Based Decision-Making: A Comparison of Statistical Approaches," submitted to *Journal of Statistical Theory and Practice*, 2008b.
- Aughenbaugh, J. M. and C. J. J. Paredis. "The Value of Using Imprecise Probabilities in Engineering Design," *Journal of Mechanical Design* 128(4):969-979, 2006.
- Berleant, D. "Automatically verified reasoning with both intervals and probability density functions," *Interval Computations*, Volume 2, pages 48-70, 1993.
- Berleant, D. "Automatically verified arithmetic on probability distributions and intervals," in *Applications of Interval Computations*, R. Baker Kearfott and Vladik Kreinovich, eds., Kluwer Academic Publishers, Dordrecht, pages 227-244, 1996.
- Coolen F.P.A., "On the use of imprecise probabilities in reliability." *Quality and Reliability in Engineering International*, 20, 193-202, 2004.
- Dempster, A.P. "Upper and lower probabilities induced by a multi-valued mapping," *Annals of Mathematical Statistics*, Volume 38, Number 2, pages 325-339, 1967.
- Du, Xiaoping, and Wei Chen, "Sequential optimization and reliability assessment method for efficient probabilistic design," *Journal of Mechanical Design*, Volume 126, pages 225-233, 2004.
- Ferson, Scott, Vladik Kreinovich, Lev Ginzburg, Davis S. Myers, and Kari Sentz, "Constructing Probability Boxes and Dempster-Shafer Structures," SAND2002-4015, Sandia National Laboratories, Albuquerque, New Mexico, January 2003.
- Ferson, Scott, Roger B. Nelsen, Janos Hajagos, Daniel J. Berleant, Jianzhong Zhang, W. Troy Tucker, Lev R. Ginzburg, and William L. Oberkampf, "Dependence in probabilistic modeling, Dempster-Shafer theory, and probability bounds analysis," SAND2004-3072, Sandia National Laboratories, Albuquerque, New Mexico, October 2004.
- de Finetti, B., *Theory of Probability Volume 1: A Critical Introductory Treatment*, Wiley, New York, 1974.
- de Finetti, B., "Foresight. Its Logical Laws, Its Subjective Sources (Translated)," *Studies in Subjective Probability*, Kyburg, H. E. and Smokler, H. E. eds., E. Krieger Publishing Company, 1980.
- Haldar, Achintya, and Sankaran Mahadevan, *Probability, Reliability, and Statistical Methods in Engineering Design*, John Wiley & Sons, Inc., New York, 2000.
- Hazelrigg, G.A., "A framework for decision-based engineering design," *Journal of Mechanical Design*, **120**, pp. 653-658, 1998.

- Liang, Jinghong, Zissimos P. Mourelatos, and Efstratios Nikolaidis, "A single-loop approach for system reliability-based design optimization," *Journal of Mechanical Design*, Volume 129, pages 1215-1224, 2007.
- Ling, Jay M., Jason Matthew Aughenbaugh, and Christiaan J. J. Paredis, "Managing the collection of information under uncertainty using information economics," *Journal of Mechanical Design*, Volume 128, pages 980-990, 2006.
- Nikolaidis, E., "Decision-Based Approach for Reliability Design," *Journal of Mechanical Design*, Vol. 129, Issue No. 5, pp. 466-475, May 2007.
- Papalambros, P.Y., and Wilde, D.J., *Principles of Optimal Design*, 2nd edition, Cambridge University Press, Cambridge, 2000.
- Ravindran, A., K.M. Ragsdell, and G.V. Reklaitis, *Engineering Optimization: Methods and Applications*, 2nd edition, John Wiley & Sons, Hoboken, New Jersey, 2006.
- Rekuc S.J., Aughenbaugh J.M., Bruns M., Paredis C.J.J., "Eliminating design alternatives based on imprecise information." In Society of Automotive Engineering World Congress. Detroit, MI, 2006.
- Renaud, J.E., and X. Gu, "Decision-based collaborative optimization of multidisciplinary systems," in *Decision Making in Engineering Design*, W. Chen, K. Lewis, and L.C. Schmidt, editors, ASME Press, New York, 2006.
- Shafer G., *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
- Tu, J., K. K. Choi, and Y. H. Park, "A new study on reliability-based design optimization," *Journal of Mechanical Design*, Volume 121, pages 557-564, 1999.
- Utkin L.V., "Interval reliability of typical systems with partially known probabilities." *European Journal of Operational Research*, 153(3 SPEC ISS), 790–802, 2004a.
- Utkin L.V., "Reliability models of m-out-of-n systems under incomplete information." *Computers and Operations Research*, 31(10), 1681–1702, 2004b.
- Walley, P., *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, New York, 1991.
- Wang, Yan, "Imprecise probabilities with a generalized interval form," Proceedings of the Third International Workshop on Reliable Engineering Computing (REC), NSF Workshop on Imprecise Probability in Engineering Analysis & Design, Savannah, Georgia, February 20-22, 2008.
- Weichselberger K. "The theory of interval probability as a unifying concept for uncertainty." *International Journal of Approximate Reasoning*, 24, 149–170, 2000.

- Williamson R.C. and T. Downs. Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* 4:89–158, 1990.
- Yager, R.R. Arithmetic and other operations on Dempster-Shafer structures. *International Journal of Man-Machine Studies* 25: 357-366, 1986.
- Youn, Byeng D., “Adaptive-loop method for non-deterministic design optimization,” *Proc. IMechE* Volume 221 Part O: *J. Risk and Reliability*, 2007.
- Zhou, Jun, and Zissimos P. Mourelatos, “A sequential algorithm for possibility-based design optimization,” *Journal of Mechanical Design*, Volume 130, January 2008a.
- Zhou, Jun, and Zissimos P. Mourelatos, “Design under uncertainty using a combination of evidence theory and a Bayesian approach,” Proceedings of the Third International Workshop on Reliable Engineering Computing (REC), NSF Workshop on Imprecise Probability in Engineering Analysis & Design, Savannah, Georgia, February 20-22, 2008b.