View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by NASA Technical Reports Server

Bounding the Failure Probability Range of Polynomial Systems Subject to P-box Uncertainties

Luis G. Crespo[a*], **Sean P. Kenny**[b], **Daniel P. Giesy**[b]
[a]**National Institute of Aerospace, Hampton, VA, 23666, USA**
[b]**NASA Langley Research Center, Hampton, VA, 23681, USA**

Abstract: This paper proposes a reliability analysis framework for systems subject to multiple design requirements that depend polynomially on the uncertainty. Uncertainty is prescribed by probability boxes, also known as p-boxes, whose distribution functions have free or fixed functional forms. An approach based on the Bernstein expansion of polynomials and optimization is proposed. In particular, we search for the elements of a multi-dimensional p-box that minimize (i.e., the best-case) and maximize (i.e., the worst-case) the probability of inner and outer bounding sets of the failure domain. This technique yields intervals that bound the range of failure probabilities. The offset between this bounding interval and the actual failure probability range can be made arbitrarily tight with additional computational effort.

Keywords: Imprecise probabilities, reliability analysis, Bernstein polynomials, extreme-case distributions.

1. INTRODUCTION

This paper studies the reliability of a system for which a parametric mathematical model is available. The acceptability of the system depends upon its ability to satisfy several design requirements. These requirements, which are represented by a set of polynomial inequality constraints on selected output metrics, depend on the uncertain parameter $\boldsymbol{p}$. The reliability analysis of the system consists of assessing its ability to satisfy all the requirements when the value of $\boldsymbol{p}$ is uncertain. In this article the uncertainty in $\boldsymbol{p}$ will be modelled as a p-box. This means that the cumulative distribution function (CDF) for $\boldsymbol{p}$ lies within the bounds set by the p-box, but that it is unknown which of the many CDFs satisfying such bounds actually describe the uncertainty in $\boldsymbol{p}$. The propagation of a p-box through the inequalities prescribing the requirements yields a range of failure probabilities. The main goal of this paper is the generation of tight outer bounds to this range.

A significant thrust of this research is the generation of inner bounding sets to the safe and failure domains using Bernstein expansions [1]. These sets are instrumental for generating intervals that bound the failure probability range resulting from propagating all the elements of a p-box from the inputs to the output. The technique proposed searches for the elements of the p-box that minimize and maximize the probability of inner and outer bounding sets of the failure domain. The corresponding probabilities are the limits of the desired intervals. Formulations applicable to p-boxes comprised of fixed-form- and arbitrary-form-distributions are proposed.

This paper is organized as follows. Basic concepts from reliability analysis, Bernstein polynomials, and imprecise probabilities are introduced in Section 2. Formulations aiming at identifying the extreme distribution functions from a p-box for a given set of polynomial requirement functions are presented and exemplified in Section 3. Finally, a few concluding remarks close the paper.

2. BACKGROUND

2.1. Basic Concepts and Notions

A probabilistic uncertainty model of $\boldsymbol{p} \in \mathbb{R}^s$ is prescribed by a random vector supported in the *support set*. This set, which is comprised of all possible uncertain parameter realizations that may occur, will be denoted as $\Delta \subseteq \mathbb{R}^s$. This model is fully prescribed by the joint cumulative distribution function $F_{\boldsymbol{p}}(\boldsymbol{p}) : \Delta \to [0, 1]$.

Consider a system that depends on the uncertain parameter $\boldsymbol{p}$. The requirements imposed upon such a system are prescribed as the vector[1] inequality $\boldsymbol{g}(\boldsymbol{p}) < \mathbf{0}$. The requirement function $\boldsymbol{g}$ is defined as $\boldsymbol{g} : \mathcal{D} \rightarrow \mathbb{R}^v$, and $\Delta \subseteq \mathcal{D} \subseteq \mathbb{R}^s$, where $\mathcal{D}$ is the *master domain*.

The *failure domain*, denoted as $\mathcal{F} \subset \mathbb{R}^s$, is comprised of the parameter realizations that fail to satisfy at least one of the requirements. Specifically, the failure domain is given by

$$\mathcal{F} = \bigcup_{i=1}^{v} \{\boldsymbol{p} : \boldsymbol{g}_i(\boldsymbol{p}) \geq 0\}, \tag{1}$$

The *safe domain*, given by $\mathcal{S} = C(\mathcal{F})$, where $C(\cdot)$ denotes the *complement* set operator $C(\mathcal{Z}) \stackrel{\Delta}{=} \mathcal{D} \setminus \mathcal{Z}$, consists of the parameter realizations satisfying all requirements. The failure probability associated with the distribution function $F_{\boldsymbol{p}}$ is given by

$$P_{F_{\boldsymbol{p}}}[\mathcal{F}] = \int_{\mathcal{F}} 1 \, dF_{\boldsymbol{p}}, \tag{2}$$

where $P_{F_{\boldsymbol{p}}}[\cdot]$ is the probability operator determined by the distribution $F_{\boldsymbol{p}}$.

Techniques for bounding $\mathcal{F}$ and $\mathcal{S}$ will be presented below. The resulting bounding sets are unions of hyper-rectangles. The *hyper-rectangle* having $\boldsymbol{m} > \mathbf{0}$ as the vector of half-lengths of the sides and $\bar{\boldsymbol{p}}$ as its geometric center, is given by

$$\mathcal{R}(\bar{\boldsymbol{p}}, \boldsymbol{m}) = \{\boldsymbol{p} : \bar{\boldsymbol{p}} - \boldsymbol{m} < \boldsymbol{p} \leq \bar{\boldsymbol{p}} + \boldsymbol{m}\} = \boldsymbol{\delta}(\bar{\boldsymbol{p}} - \boldsymbol{m}, \bar{\boldsymbol{p}} + \boldsymbol{m}), \tag{3}$$

where $\boldsymbol{\delta}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\delta}_1 \times \boldsymbol{\delta}_2 \times \cdots \times \boldsymbol{\delta}_s$, is the Cartesian product of the intervals $\boldsymbol{\delta}_i = (\boldsymbol{x}_i, \boldsymbol{y}_i]$.

A *subdivision* is the result of dividing a set in $\boldsymbol{p}$-space into pairwise disjoint subsets. Let $\rho(\cdot)$ be an operator whose input is any given set and its output are the subsets. A bisection-based subdivision of the hyper-rectangle given in (3) in the ith direction is given by

$$\rho(\mathcal{R}) = \{\mathcal{R}(\bar{\boldsymbol{p}} + \boldsymbol{w}, \boldsymbol{m} - \boldsymbol{w}), \mathcal{R}(\bar{\boldsymbol{p}} - \boldsymbol{w}, \boldsymbol{m} - \boldsymbol{w})\}, \tag{4}$$

where $\boldsymbol{w} = [0, \ldots, 0, \boldsymbol{m}_i/2, 0, \ldots, 0]$. That is, the hyper-rectangle is cut along its ith component into two sub-hyper-rectangles of equal size.

2.2. Bernstein Polynomials

The image of a hyper-rectangle when mapped by a multivariable polynomial is a bounded interval. By expanding that polynomial using a Bernstein basis over that rectangle, rigorous bounds to such an interval can be calculated using simple algebraic manipulations. Bernstein polynomials will be used for determining if a hyper-rectangle $\mathcal{R} \subset \mathbb{R}^s$ is fully contained in the failure or safe domains associated with the vector inequality $\boldsymbol{g}(\boldsymbol{p}) < \mathbf{0}$. Mathematical details of this technique can be found in [2].

The outcome of the set containment determination depends exclusively on the refinement of the partition of $\mathcal{R}$ the analyst is willing to perform. Additional computational effort in the refinement of such a partition can be used to render the tests conclusive.

2.3. Imprecise Probabilities

Imprecise probabilities is the subject of several theories involving models of uncertainty that do not assume a particular underlying probability distribution. Instead, it assumes that the underlying distribution may be any member of a family of distributions. Theories of imprecise probabilities are often expressed in terms of lower and upper probabilities for every possible event from some universal set. Interval probabilities, Dempster-Shafer structures and p-boxes can be regarded as special-cases of imprecise probabilities [3, 4].

[1]Throughout this paper, it is assumed that vector inequalities hold component-wise, super-indices denote a particular vector or set, and sub-indices refer to vector components; e.g., $\boldsymbol{p}_i^{(j)}$ is the ith component of the vector $\boldsymbol{p}^{(j)}$.

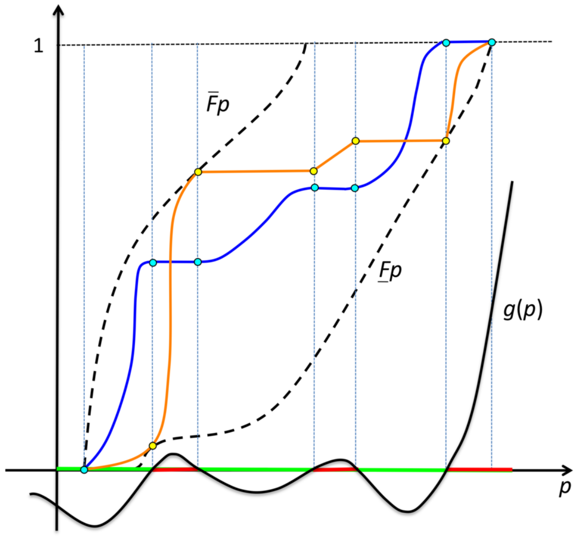

Figure 1: P-box envelopes (dashed lines), p-box members (solid lines with circles), and requirement function (thick solid line).

In this paper uncertainty will be prescribed via p-boxes. A one-dimensional p-box is a class of imprecisely known distributions functions $F_p(p)$ bounded between a pair of enveloping distribution functions $\underline{F}_p(p)$ and $\overline{F}_p(p)$ supported in $\mathcal{D}$ such that $\underline{F}_p(p) \leq F_p(p) \leq \overline{F}_p(p)$ for all p values. This means that the distribution for the random variable p, F_p, is unknown except that it satisfies these inequalities.

P-boxes result from cases (i) where the distribution is known to take a particular form but its parameters can only be specified by intervals, (ii) where envelopes of all distributions matching a given set of moments are constructed from inequalities (e.g., Markov, Chebyshev), (iii) where empirical distributions functions are built using data subject to measurement uncertainty, (iv) where the functional form of the distribution cannot be determined due to limited amount of data (e.g., Kolmogorov-Smirnov, confidence bounds), and (v) where several classes of uncertainty models are present (e.g., uncertainties are prescribed as a collection of intervals, subsets of independent random variables and subsets of dependent random variables with unknown copulas [4]).

In this paper, two types of p-boxes are considered. One type is an *arbitrary distribution p-box*. This type of p-box is defined by specifying two envelope distribution functions, $\underline{F}_p$ and $\overline{F}_p$, satisfying the restriction that $\underline{F}_p(p) \leq \overline{F}_p$ for all p. A CDF F_p belongs to this p-box if it falls between these two envelopes; i.e., if $\underline{F}_p(p) \leq F_p(p) \leq \overline{F}_p$ for all p. Another type is a *fixed form distribution p-box*. This p-box only contains members of a parametrized family of distributions. For example, to specify a fixed form distribution p-box one might require that all CDFs in it be Gaussian distributions. The size of this p-box is limited by restricting the allowable range of values for the defining parameters. In the Gaussian case, this might consist of requiring the mean and standard deviation to fall in some specified intervals. Figure 1 shows the envelope distributions of an arbitrary distribution p-box as well as two of its elements.

The set of p-boxes for a set of s uncertain parameters will be denoted as $B_{\boldsymbol{p}}$, where $B_{\boldsymbol{p}_i}$ for $i = 1, \dots s$ is an arbitrary distribution p-box or a fixed form distribution p-box as described above. Hereafter it is assumed that the uncertainty in each component of $\boldsymbol{p}$ is given by a CDF belonging to a prescribed p-box. Furthermore, any random variable associated with a possible distribution of $\boldsymbol{p}_i$ is independent of those of $\boldsymbol{p}_j$ for all $i \neq j$.

While the probability of an event for a given distribution is a scalar, the probability of an event for a p-box is a range. This range can be an interval or a disconnected set. Each point on this range corresponds to at least one element of the p-box. The reliability analysis of a system subject to the requirements $\boldsymbol{g}(\boldsymbol{p}) < \boldsymbol{0}$, where the uncertainty model of $\boldsymbol{p}$ is an unknown element of the p-box $B_{\boldsymbol{p}}$, consists of calculating or bounding the range of all possible failure probabilities

$$X = \left[\min_{F_{\boldsymbol{p}} \in B_{\boldsymbol{p}}} P_{F_{\boldsymbol{p}}}[\mathcal{F}],\ \max_{F_{\boldsymbol{p}} \in B_{\boldsymbol{p}}} P_{F_{\boldsymbol{p}}}[\mathcal{F}] \right]. \tag{5}$$

In most practical applications X cannot be calculated exactly. Instead, the inner and outer bounding intervals $\underline{X}$ and $\overline{X}$, satisfying

$$[\, \underline{l},\ \underline{u}\,] \triangleq \underline{X} \subseteq X \subseteq \overline{X} \triangleq [\, \bar{l},\ \bar{u}\,], \tag{6}$$

are calculated. The tightness of the bounds of X, measured by $(\overline{u} - \underline{u}) + (\underline{l} - \bar{l})$, will depend on a convergence parameter set by the analyst. $\underline{X}$ enables estimating the portion of $\overline{X}$ that can be reduced with additional computational effort. While the upper endpoint of X cannot be reduced more than $\overline{u} - \underline{u}$, the lower endpoint cannot be increased more than $\underline{l} - \bar{l}$.

3. RELIABILITY ANALYSIS

In this section we present an algorithm for generating outer bounds to the failure probability range by searching for the elements of the p-box $B_{\boldsymbol{p}}$ that minimize (i.e., *best-case*) and maximize (i.e., *worst-case*) the probability of inner and outer bounding sets of the failure domain. The determination of *extreme-case distributions*, by which we mean best- and worst-case distributions, has been the focus of attention recently [5]. The best- and worst-case distributions, denoted as $F_{\boldsymbol{p}}^b \in B_{\boldsymbol{p}}$ and $F_{\boldsymbol{p}}^w \in B_{\boldsymbol{p}}$ respectively, are distributions whose propagation yields the lower and upper limits of X. Specifically,

$$F_{\boldsymbol{p}}^b = \underset{F_{\boldsymbol{p}} \in B_{\boldsymbol{p}}}{\operatorname{argmin}}\ P_{F_{\boldsymbol{p}}}[\mathcal{F}], \tag{7}$$

$$F_{\boldsymbol{p}}^w = \underset{F_{\boldsymbol{p}} \in B_{\boldsymbol{p}}}{\operatorname{argmax}}\ P_{F_{\boldsymbol{p}}}[\mathcal{F}]. \tag{8}$$

These distributions, which depend on the geometry of $\mathcal{F}$, may not be unique.

The distributions shown in Figure 1 are examples of a worst-case (orange line) and best-case (blue line) distribution, respectively, for the given p-box and the requirement $g(p) < 0$. Note that the worst-case distribution rises the most over the failure domain, i.e., region over the abscissa colored in red, while the best-case distribution rises the most over the safe domain, i.e., region over the abscissa colored in green. Note that any distribution passing through the points where $F_p^w \in B_p$ (resp. $F_p^b \in B_p$) crosses the failure domain's boundary is also worst-case (resp. best-case). These points are shown as small circles over the distributions.

A strategy for determining the extreme-case distributions corresponding to bounding sets of the failure domain is presented below. An iterative procedure for generating these bounding sets is presented next.

3.1. Set Bounding

This section presents an algorithm to generate and sequentially expand subsets of the failure and safe domains. These sets are unions of disjoint hyper-rectangles chosen from a partition Q of $\mathcal{D}$. Let $\mathcal{F}^{\text{sub}}$ and $\mathcal{S}^{\text{sub}}$ denote subsets (i.e., inner approximations) of the failure and safe domains formed from selected elements of Q. Note that $\emptyset \subseteq \mathcal{F}^{\text{sub}} \subseteq \mathcal{F} \subseteq C(\mathcal{S}^{\text{sub}}) \subseteq \mathcal{D}$, where $C(\cdot)$ is the complement set operator defined earlier, and that the failure domain boundary, denoted as $\partial\mathcal{F}$, lies in the region between the interiors of $\mathcal{F}^{\text{sub}}$ and $\mathcal{S}^{\text{sub}}$.

The sequences of inner bounds $\{\mathcal{S}_1^{\text{sub}}, \mathcal{S}_2^{\text{sub}}, \ldots\}$ and $\{\mathcal{F}_1^{\text{sub}}, \mathcal{F}_2^{\text{sub}}, \ldots\}$ are generated by the algorithm below. These sequences are made to converge to the domain being bounded. In particular, the algorithm below iteratively generates indexed partitions Q^i of $\mathcal{D}$ and indexed sets $\mathcal{S}_i^{\text{sub}}$, $\mathcal{F}_i^{\text{sub}}$ and Λ_i which are unions of hyper-rectangles from Q^i, where $\mathcal{S}_i^{\text{sub}}$ is an inner approximation to the safe domain, $\mathcal{F}_i^{\text{sub}}$ is an inner approximation to the failure domain, and Λ_i is a region comprised by the rectangles of Q^i that are not in $\mathcal{S}_i^{\text{sub}}$ or $\mathcal{F}_i^{\text{sub}}$. Note that while Q^i is a list of hyper-rectangles, $\mathcal{S}_i^{\text{sub}}$, $\mathcal{F}_i^{\text{sub}}$ and Λ_i are sets comprised by the union of some of these rectangles. The algorithm proceeds by successively selecting each of the component hyper-rectangles $\mathcal{R}$ of Λ_i. The tests in Theorems 1 and 2 in [2] for $t = 1$ are applied to $\mathcal{R}$. If either of these tests determine that $\mathcal{R} \subseteq \mathcal{S}$ or $\mathcal{R} \subseteq \mathcal{F}$, then $\mathcal{R}$ is removed from Λ_i and added to $\mathcal{S}_i^{\text{sub}}$ or $\mathcal{F}_i^{\text{sub}}$, respectively. If neither of these conditions is satisfied, $\mathcal{R}$ is subdivided, and the resulting sub-rectangles replace $\mathcal{R}$ in the partition. The algorithm terminates when the volume of Λ_i is sufficiently small.

The algorithmic representation of this procedure is as follows. Let the inequality constraint $\boldsymbol{g}(\boldsymbol{p}) < \mathbf{0}$ defined over $\boldsymbol{p} \in \mathcal{D}$ prescribe the system requirements. Set $i = 1$, $Q^1 = \{\mathcal{D}\}$, $\Lambda_1 = \mathcal{D}$, $\mathcal{F}_1^{\text{sub}} = \emptyset$, and $\mathcal{S}_1^{\text{sub}} = \emptyset$. Pick a convergence criterion $\epsilon > 0$.

1. Let L contain the elements of Q^i comprising Λ_i.
2. Determine which elements of L are contained in the safe domain and which ones are contained in the failure domain. Denote by U the list of elements contained by the safe domain, by V the list of elements contained by the failure domain, and by W the list of elements that are not in U nor V. Furthermore, let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be the union of the elements in U, V and W respectively.
3. Make $\mathcal{S}_{i+1}^{\text{sub}} = \mathcal{S}_i^{\text{sub}} \cup \mathcal{U}$; $\mathcal{F}_{i+1}^{\text{sub}} = \mathcal{F}_i^{\text{sub}} \cup \mathcal{V}$; and $\Lambda_{i+1} = \Lambda_i \setminus (\mathcal{U} \cup \mathcal{V})$.
4. If $\text{Vol}(\Lambda_{i+1}) < \epsilon$ stop[2]. Otherwise, make $Q^{i+1} = (Q^i \setminus W) \cup \rho(\mathcal{W})$, increase i by one, and go to Step 1.

As the number of iterations increases, $\mathcal{S}_i^{\text{sub}}$ and $\mathcal{F}_i^{\text{sub}}$ approach the safe and failure domain. Note that the closure of Λ_i not only covers the boundary of $\mathcal{F}$ but also approaches that boundary as i increases.

3.2. Extreme-Case Distributions over Bounding Sets

In this section we present techniques which utilize the inner bounding sets of Section 3.1 to calculate the bounding intervals $\underline{X}$ and $\overline{X}$ of Equation (6). A left endpoint of $\overline{X}$ (resp. $\underline{X}$) is found by seeking distributions from the p-box which minimize the probability of an inner (resp. outer) bound to the failure set. For the right endpoint, distributions in the p-box are sought which maximize these probabilities. The tighter the bounding sets, the smaller the gap between $\underline{X}$ and $\overline{X}$.

The bounding intervals corresponding to the partition Q^i are denoted by $\underline{X}_i$ and $\overline{X}_i$. Recall that $\mathcal{F}_i^{\text{sub}}$, $\mathcal{S}_i^{\text{sub}}$, and Λ_i are comprised by elements of Q^i. While the upper limit of $\overline{X}_i$ can be used to declare a system acceptable, i.e., because it is smaller than the largest admissible failure probability, the upper limit of $\underline{X}_i$ can be used to declare it unacceptable, i.e., because it is larger than the largest admissible failure probability.

The formulations below enable calculating $\overline{X}_i$ and $\underline{X}_i$ for the partition Q^i that results from the ith iteration of the algorithm in Section 3.1. The corresponding collection of bounding intervals defines sequences. Note that the elements of the sequences $\underline{X}_i$ and $\overline{X}_i$ satisfy $\underline{X}_i \subseteq \underline{X}_{i+1} \subseteq X \subseteq \overline{X}_{i+1} \subseteq \overline{X}_i$. By design, the sequences for $\underline{X}_i$ and $\overline{X}_i$ approach X monotonically as i increases.

The developments that follow enable calculating the bounding intervals corresponding to a given partition. Formulations applicable to fixed form distribution and arbitrary distribution p-boxes are presented next.

3.2.1. Fixed Form Distribution P-Boxes

Recall that a fixed form distribution p-box assumes a parametrized functional form. Let $\boldsymbol{\theta}$ denote the corresponding vector of parameters. A fixed form p-box is fully prescribed by its functional form and the domain of $\boldsymbol{\theta}$. This domain will be denoted as Θ. The parameters of the best-case and worst-case distribution for the partition $\rho(\mathcal{D}) = \{\mathcal{F}_i^{\text{sub}}, \mathcal{S}_i^{\text{sub}}, \Lambda_i\}$ are given by

$$\boldsymbol{\theta}_i^b = \underset{\boldsymbol{\theta}\in\Theta}{\text{argmin}} \left\{ \sum_{\mathcal{R}_j \subseteq \mathcal{F}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{\theta})}}\left[\mathcal{R}_j\right] \right\}, \tag{9}$$

[2]In the context of Section 3.2, a better stopping criterion is $\text{Vol}(\overline{X}_i \setminus \overline{X}_{i+1}) < \epsilon$.

$$\boldsymbol{\theta}_i^w = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}} \left\{ \sum_{\mathcal{R}_j \subseteq \mathcal{S}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{\theta})}} \left[\mathcal{R}_j\right] \right\}. \tag{10}$$

Therefore, the best-case distribution minimizes the probability of the inner bounding set of the failure domain as if Λ_i were fully contained in the safe domain; while the worst-case distribution maximizes the probability of the inner bounding set of the safe domain as if Λ_i were fully contained in the failure domain. The solution to Equations (9) and (10) yields the outer bounding interval

$$\overline{X}_i = \left[P_{F_{\boldsymbol{p}(\boldsymbol{\theta}_i^b)}} \left[\mathcal{F}_i^{\text{sub}}\right], 1 - P_{F_{\boldsymbol{p}(\boldsymbol{\theta}_i^w)}} \left[\mathcal{S}_i^{\text{sub}}\right] \right]. \tag{11}$$

The inner bounding interval corresponding to the same partition is given by

$$\underline{X}_i = \begin{cases} [\alpha_i,\ \beta_i], & \text{if } \alpha_i \le \beta_i \\ \emptyset & \text{otherwise,} \end{cases} \tag{12}$$

where $\alpha_i = P_{F_{\boldsymbol{p}(\boldsymbol{\theta}_i^l)}} \left[\mathcal{F}_i^{\text{sub}}\right]$, $\beta_i = 1 - P_{F_{\boldsymbol{p}(\boldsymbol{\theta}_i^u)}} \left[\mathcal{S}_i^{\text{sub}}\right]$, and

$$\boldsymbol{\theta}_i^l = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmax}} \left\{ \sum_{\mathcal{R}_j \subseteq \mathcal{S}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{\theta})}} \left[\mathcal{R}_j\right] \right\}, \tag{13}$$

$$\boldsymbol{\theta}_i^u = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmax}} \left\{ \sum_{\mathcal{R}_j \subseteq \mathcal{F}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{\theta})}} \left[\mathcal{R}_j\right] \right\}. \tag{14}$$

Therefore, the lower limit of $\underline{X}_i$ corresponds to the distribution that maximizes the probability of the inner bounding set of the safe domain as if Λ_i were fully contained in the failure domain; while the upper limit corresponds to the distribution that maximizes the probability of the inner bounding set of the failure domain as if Λ_i were fully contained in the safe domain. Note that $\boldsymbol{\theta}_i^b$ and $\boldsymbol{\theta}_i^w$ closely approximate $\boldsymbol{\theta}_i^l$ and $\boldsymbol{\theta}_i^u$ when $\text{Vol}(\Lambda_i) \ll \text{Vol}(\mathcal{D})$.

When the uncertain parameters are independent, the above probabilities can be readily evaluated using

$$P_{F_{\boldsymbol{p}(\boldsymbol{\theta})}}[\mathcal{R}] = \prod_{j=1}^{s} P_{F_{\boldsymbol{p}_j(\boldsymbol{\theta})}}[\boldsymbol{\delta}_j] = \prod_{j=1}^{s} \left(F_{\boldsymbol{p}_j(\boldsymbol{\theta})}(b_j) - F_{\boldsymbol{p}_j(\boldsymbol{\theta})}(a_j) \right), \tag{15}$$

where $\boldsymbol{\delta}_j = (a_j, b_j]$ is the projection of $\mathcal{R}$ onto the $\boldsymbol{p}_j$ axis.

Example 1: In this example we calculate the sequences of $\overline{X}_i$ and $\underline{X}_i$ for the requirement functions

$$\boldsymbol{g}_1 = \boldsymbol{p}_1^2\boldsymbol{p}_2^4 + \boldsymbol{p}_1^4\boldsymbol{p}_2^2 - 3\boldsymbol{p}_1^2\boldsymbol{p}_2^2 - \boldsymbol{p}_1\boldsymbol{p}_2 + \frac{\boldsymbol{p}_1^6 + \boldsymbol{p}_2^6}{200} - \frac{7}{100}, \tag{16}$$

$$\boldsymbol{g}_2 = -\frac{\boldsymbol{p}_1^2\boldsymbol{p}_2^4}{2} - \boldsymbol{p}_1^4\boldsymbol{p}_2^2 + 3\boldsymbol{p}_1^2\boldsymbol{p}_2^2 + \frac{\boldsymbol{p}_1^5\boldsymbol{p}_2^3}{10} - \frac{9}{10}. \tag{17}$$

We assume that $B_{\boldsymbol{p}_1}$ is the family of Beta distributions supported in $\mathcal{D}$ with parameters in $20 \le a \le 30$ and $20 \le b \le 30$ while $B_{\boldsymbol{p}_2}$ is the family of Beta distributions supported in $\mathcal{D}$ with parameters in $0.5 \le c \le 0.7$ and $0.5 \le d \le 1$. In this setting $\boldsymbol{\theta} = [a, b, c, d]^\top$ where $\Theta = [20, 30] \times [20, 30] \times [0.5, 0.7] \times [0.5, 1]$. In this case $\underline{F}_{\boldsymbol{p}}$ and $\overline{F}_{\boldsymbol{p}}$ are not Beta distributions. The top of Figure 2 shows the bounding intervals $\overline{X}_i$ and $\underline{X}_i$ as a function of the total number of boxes in the elements of Q^i. The endpoints of $\overline{X}_i$ are shown as solid lines while α_i and β_i are shown as dashed lines. While the upper endpoint of X is bounded by the red lines with squares, the lower endpoint is bounded by the blue lines with circles. Note that the inner bounding interval $\underline{X}_i$ is non-empty only after the dashed lines cross, i.e., at the points where $\beta_i > \alpha_i$. Further notice that partitions with more than 4000 boxes yield very tight bounding intervals. The plot in the middle shows the convergence of the parameters corresponding to the worst-case distribution $\boldsymbol{\theta}_i^w$ while the plot at the bottom shows that for the best-case distribution $\boldsymbol{\theta}_i^b$. Note that partitions comprised of more than 1000 boxes render the very same extreme-case distributions. This suggests the possibility of stopping the search for the extreme-case distributions before stopping the expansion of the bounding sets. Figure 3 shows the p-box envelopes for each of the two uncertain parameters as well as the extreme-case distributions.

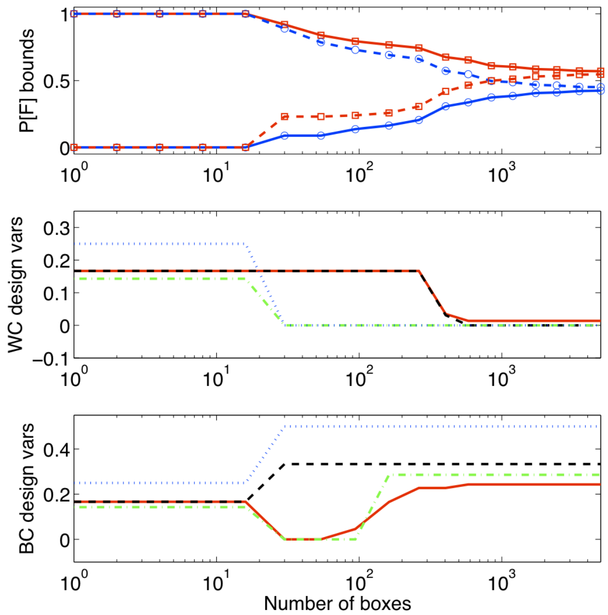

Figure 2: Convergence of the bounding intervals of X, and of design variables.

3.2.2. Arbitrary Distribution P-Boxes

The developments of this section enable calculating the bounding intervals $\underline{X}_i$ and $\overline{X}_i$ of X corresponding to an arbitrary distribution p-box $B_{\boldsymbol{p}}$ for the partition Q^i. Recall that a arbitrary distribution p-box contains all possible distributions between the envelopes $\underline{F}_{\boldsymbol{p}}$ and $\overline{F}_{\boldsymbol{p}}$. Let $Q^{i,j}$ be the ordered list of distinct components resulting from projecting the vertices of all hyper-rectangles comprising Q^i onto the $\boldsymbol{p}_j$ axis. While j can take on any integer number between 1 and s, the number of elements of the list for a fixed value of j, denoted as

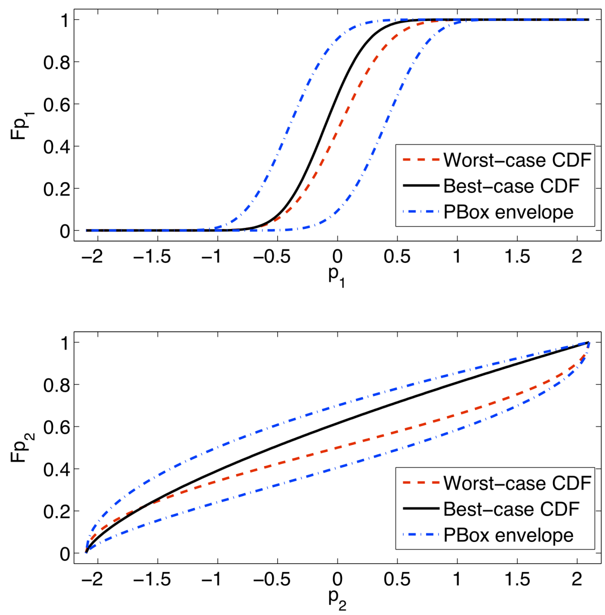

Figure 3: P-box envelopes, worst- and best-case distributions for a fixed distribution p-box.

m_j, varies. We will consider these numbers as coordinates on the $\boldsymbol{p}_j$ axis. In order to determine, for example, the left endpoint of $\overline{X}_i$, it suffices to find a distribution $F_{\boldsymbol{p}} \in B_{\boldsymbol{p}}$ for which $P_{F_{\boldsymbol{p}}}[\mathcal{F}_i^{\text{sub}}]$ is minimal.

The formulation below searches for the points

$$\boldsymbol{y}_k^{i,j} \triangleq F_{\boldsymbol{p}_j(\boldsymbol{y})}(Q_k^{i,j}), \tag{18}$$

of the extreme-case distributions leading to $\underline{X}_i$ and $\overline{X}_i$. Figure 1 displays circles at the pairs $\langle Q^k, y_k \rangle$ pairs corresponding to the best-case (blue line) and worst-case (orange line) distributions.

The search for the values $\boldsymbol{y}^{i,j}$ taken by the extreme distributions is formulated as a non-linear optimization problem subject to the constraint set

$$\begin{aligned} \mathcal{Y} \;=\; & \{\boldsymbol{y} : \boldsymbol{y}_k^{i,j} \leq \boldsymbol{y}_{k+1}^{i,j},\ k = 1, \ldots m_j - 1,\ 1 \leq j \leq s; \\ & \underline{F}_{\boldsymbol{p}_j}(Q_k^{i,j}) \leq \boldsymbol{y}_k^{i,j} \leq \overline{F}_{\boldsymbol{p}_j}(Q_k^{i,j}),\ k = 1, \ldots m_j,\ 1 \leq j \leq s\}. \end{aligned}$$

These constraints insure that the extreme-case distribution passing through these values is non-decreasing and falls within the envelopes of the p-box $B_{\boldsymbol{p}}$.

The left endpoint of $\overline{X}_i$ is given by

$$\bar{l}_i = \min_{\boldsymbol{y}} \left\{ \sum_{\mathcal{R} \in Q^i, \mathcal{R} \subseteq \mathcal{F}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{y})}}[\mathcal{R}] : \boldsymbol{y} \in \mathcal{Y} \right\}, \tag{19}$$

where $F_{\boldsymbol{p}(\boldsymbol{y})}$ is any distribution compliant with Equation (18). The value of $\boldsymbol{y}$ at which the minima occurs prescribes a finite collection of points taken by the best-case distribution function corresponding to $\mathcal{F}_i^{\text{sub}}$. In a similar fashion, the right endpoint of $\overline{X}_i$ is given by

$$\bar{u}_i = \min_{\boldsymbol{y}} \left\{ \sum_{\mathcal{R} \in Q^i, \mathcal{R} \subseteq \mathcal{S}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{y})}}[\mathcal{R}] : \boldsymbol{y} \in \mathcal{Y} \right\}. \tag{20}$$

The entries to the inner bounding interval $\underline{X}_i$ in Equation (12) are given by

$$\alpha_i = \max_{\boldsymbol{y}} \left\{ \sum_{\mathcal{R} \in Q^i, \mathcal{R} \subseteq \mathcal{S}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{y})}}[\mathcal{R}] : \boldsymbol{y} \in \mathcal{Y} \right\}, \tag{21}$$

$$\beta_i = \max_{\boldsymbol{y}} \left\{ \sum_{\mathcal{R} \in Q^i, \mathcal{R} \subseteq \mathcal{F}_i^{\text{sub}}} P_{F_{\boldsymbol{p}(\boldsymbol{y})}}[\mathcal{R}] : \boldsymbol{y} \in \mathcal{Y} \right\}. \tag{22}$$

The above expressions are analogous to those in Equations (9-14).

Some of the design variables in $\boldsymbol{y}$ may not be necessary. This can be observed in Figure 1, where the values taken by the extreme-case CDFs at points other than those marked with circles do not affect the calculation of probabilities, and hence, the resulting bounding intervals. A more efficient implementation of Equations (19-22) is obtained by only choosing the components of $\boldsymbol{y}$ corresponding to the $Q^{i,j}$ values associated with the hyper-rectangles comprising $\mathcal{F}_i^{\text{sub}}$, $\mathcal{S}_i^{\text{sub}}$, $\mathcal{S}_i^{\text{sub}}$ and $\mathcal{F}_i^{\text{sub}}$ respectively.

Note that there is an infinite number of distributions passing through the points $\langle Q_k^{i,j}, \boldsymbol{y}_k^{i,j} \rangle$. Therefore, there will be an infinite number of distributions attaining the same bounding intervals. When the uncertain parameters are independent, these probabilities are given by

$$P_{F\boldsymbol{p}(\boldsymbol{y})}[\mathcal{R}] = \prod_{j=1}^{s} P_{F_{\boldsymbol{p}(\boldsymbol{y})}}[\boldsymbol{\delta}_j] = \prod_{j=1}^{s} \left(\boldsymbol{y}_l^{i,j} - \boldsymbol{y}_k^{i,j} \right), \tag{23}$$

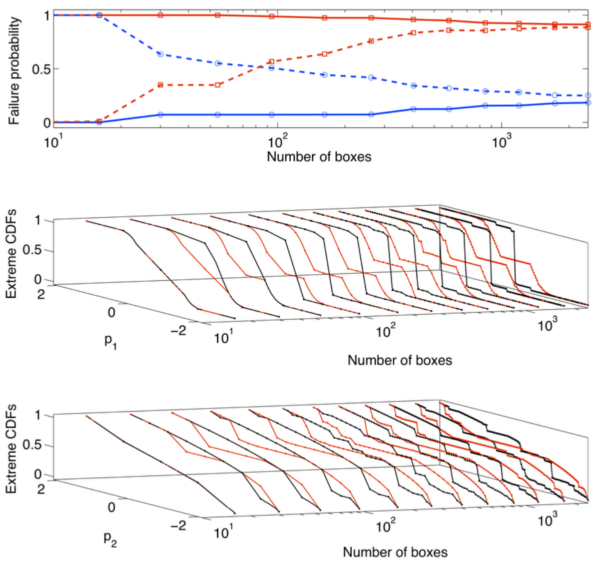

Figure 4: Convergence of the bounding intervals of X, and extreme-case distributions.

where $\delta_j = [Q_k^{i,j}, Q_l^{i,j}]$ for $l > k$ is the projection of $\mathcal{R}$ onto the $\boldsymbol{p}_j$ axis while $\boldsymbol{y}_k^{i,j}$ and $\boldsymbol{y}_l^{i,j}$ are the corresponding values of the distribution. A few remarks on the methods that can be used to solve the optimization problems above are presented next. Equations (19-22) have a multi-linear cost function of order s and linear inequality constraints. In principle, the global solution to these problems can be rigorously bounded using Bernstein expansions. Unfortunately, the computational cost of applying this method to problems of moderate size is exceedingly high. Instead, a nonlinear programming solver was used to search for the extreme distributions. Such a technique however, may fail to converge to the global optima. The best known method to find rigorous, or at least ϵ-approximate solutions to this class of optimization problems, is the spatial Branch-and-Bound algorithm [6]. This approach determines lower bounds to the cost function on each region being explored by solving convex relaxations of the cost function. Even though convex envelopes are explicitly known for low-order terms, such a description is unknown for general multi-linear terms of higher order. Note however that relaxation techniques will introduce irreducible conservatism into the bounding intervals.

Example 2: Here we consider the problem in Example 2 but use arbitrary distribution p-boxes with the same envelopes. The top of Figure 4 shows $\underline{X}_i$ and $\overline{X}_i$ as a function of the number of boxes in the elements of the partition Q^i. As before, the inner bounding set is empty for coarse partitions. As expected, the probability range for the fixed-form distribution p-boxes in Figure 2 is smaller than that for the arbitrary distribution p-boxes in Figure 4. The bottom plots show the convergence of the extreme-case distribution functions for $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. While the best-case distribution is shown as a black line, the worst-case is shown as a red line. As in the previous example, the shape of the extreme-case distributions settle for a sufficiently fine partition of the master domain. Figure 5 shows the envelopes of the p-box as well as the extreme-case distributions at the last iteration step. The coordinates of the centroid of the safe domain are near the lines $\boldsymbol{p}_1 = 0$ and $\boldsymbol{p_2} = 0$. Is at these values where the best-case distribution rises the most, i.e., a large probability is placed at the core of the safe domain, and the worst-case rises the least, i.e., a small probability is placed at the core of the safe domain. This phenomenon justifies the perpendicular intersection of the extreme-case distributions of $\boldsymbol{p}_1$ at $\boldsymbol{p}_1 = 0$.

4. CONCLUSIONS

This paper presents a reliability analysis framework applicable to systems subject to polynomial requirement functions and probability-box uncertainties. This dependency may occur naturally or artificially. Techniques

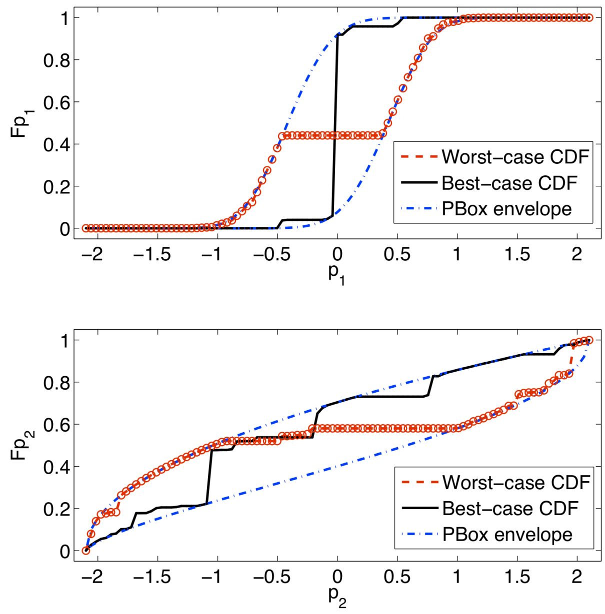

Figure 5: P-box envelopes, worst- and best-case distributions for a arbitrary distribution p-box.

for bounding the range of failure probabilities based on interval propagation and optimization are proposed. The foundation of these techniques is the Bernstein expansion of polynomials. This article presents the mathematical framework of these approaches and illustrates their application to an easily reproducible example problem. Since a single analysis, with a carefully chosen p-box, encompasses infinitely many analyses each having its own distributions for all uncertain parameters, this technique substantially mitigates the need for accurate probabilistic models of the uncertainty. When the analysis yields an inadmissibility large range of failure probabilities, the analyst can either refine the p-box, say by doing more experiments or simulations, or can redesign the system to accommodate for the existing level of knowledge.

REFERENCES

[1] M. Zettler and J. Garloff. Robustness analysis of polynomials with polynomial parameter dependency using Bernstein expansion. *IEEE Transactions in Automatic Control*, 43(3):425–431, 1998.

[2] L. G. Crespo, C. A. Munoz, A. J. Narkawicz, S. P. Kenny, and Daniel P. Giesy. Uncertainty analysis via failure domain characterization: Polynomial requirement functions. In *ESREL 2011*, number 1, Trojes, France, 18-22 September 2011.

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

[4] S. Ferson, R. Nelsen, J. Hajagos, D. Berleant, J. Zhang, W. T. Tucker, L. Ginzburg, and W. L. Oberkampf. Dependence in probabilistic modeling, Dempster-Shafer Theory, and probability bounds analysis. *Sandia National Laboratories, SAND2004-3072*, pages 1–120, 2004.

[5] A. Shapiro. Worst-case distribution analysis of stochastic programs. *Mathematical Programming Series B*, 107:91–96, 2006.

[6] C. A. Floudas and C. E. Gounaris. A review of recent advances in global optimization. *Journal of global optimization*, 45:3–38, 2009.