Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2019) 713-723

ECCOMAS Proceedia

UNCECOMP 2019
3rd ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Crete, Greece, 24-26 June 2019

BLACK-BOX PROPAGATION OF FAILURE PROBABILITIES UNDER EPISTEMIC UNCERTAINTY

Marco De Angelis[1], **Scott Ferson**[1], **Edoardo Patelli** [1], **and Vladik Kreinovich**[2]

[1]Institute for Risk and Uncertainty, School of Engineering, University of Liverpool
Liverpool, L69 7ZF
marco.de-angelis@liverpool.ac.uk

[2] Computer Science, University of Texas at El Paso
El Paso, TX 79968, USA
vladik@utep.edu

Abstract. *In engineering simulation a black-box code is often a complex, legacy or proprietary (secret) black-box software used to describe the physics of the system under study. Strategies to propagate epistemic uncertainty through such codes are desperately needed, for code verification, sensitivity, and validation on experimental data. Very often in practice, the uncertainty in the inputs is characterised by* imprecise *probability distributions or distributions with interval parameters, also known as probability boxes. In this paper we propose a strategy based on line sampling to propagate both aleatory and epistemic uncertainty through black-box codes to obtain interval probabilities of failure. The efficiency of the proposed strategy is demonstrated on the NASA LaRC UQ problem.*

Keywords: Black-box code, Cauchy-deviate method, Line sampling, Digital Twins

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2019.
doi: 10.7712/120219.6373.18699

1 INTRODUCTION

Digital twins seem to be the emerging modelling paradigm for industrial system simulation. According to ANSYS [1], a digital twin is "a complete virtual prototype of an entire system, a working system in a real world environment, a complex system integrating multiple engineering disciplines, requiring system-level simulation knowledge". ANSYS also state that digital twins " represent a new era in simulation, a new world of predictability, a new tool for engineering the future".

The rise of digital twins is justified by the progressive increase of high-fidelity methods (e.g. finite element and computational fluid dynamic) and the fast-paced growth of the computing power that has led to the solution of unprecedentedly complex models, for ever more realistic boundary conditions. The excitement about this new era of simulation must face the truth about our limitation of modeling the physics around us. Deterministic models are rarely suited to describe in detail the multifaceted reality of a system, and usually the more detailed the model the more sensitive it is to variations and uncertainty. Comparing predicted responses with measured data, however, does not generally show that the fidelity has improved as much as our ability of making more detailed models and accurate analyses.

The reason for this discrepancy is often the presence of uncertainties, for instance in the parameters of the model, which are not precisely known and must be expected to deviate from the assumed deterministic values. Another source of uncertainty is in the mathematical model, which usually involves some abstraction and simplifying assumptions to represent the actual mechanical/physical response. Given the limitations of data, quantification methods often rely on subjective judgement and assumptions and it may not always seem reasonable to characterize the uncertainties in a classical probabilistic way. To avoid the inclusion of subjective and often unjustified hypothesis, the imprecision and vagueness of the data can be treated by using generalized probabilistic methods.

The unavoidable uncertainties must be explicitly included in the computations to guarantee that the components or systems will continue to perform satisfactorily despite variability and precise models. If the effect of uncertainties in the optimized design is ignored, this design may perform unsatisfactorily in realistic conditions. Resilient/reliable systems are less sensitive to the uncertainties and hence, they reach low variability of the overall performance allowing for significant reductions in terms of e.g. the manufacturing and operating costs.

Quantifying the effect of the uncertainty is a necessary step to support decision makers. For instance, the analyst can estimate the importance of collecting additional information and identify the parameters that contribute the most to the variability of the output. One of the most important analyses is to identify the extreme response performances of the system. It is also important to determine the combination of input parameter values that causes a performance metric of interest to reach its extreme. Knowing such conditions, it might be possible to prevent those extreme performances or mitigate their consequences.

The need for efficient and robust uncertainty propagation on black-box models has been shown by the NASA Langley Uncertainty Quantification Challenge [2]. The only information that was released about the model is that it described the flight of a remotely controlled twin-jet aircraft pushed to the edge of the flight envelope, thus subject to strong uncertainties.

In this paper, a novel approach for black box failure probability propagation analysis is presented and discussed. The theoretical framework of imprecise probability is used for the representation of the uncertainty. An efficient and general computational framework based on Line Sampling simulation is proposed to perform reliability analysis [3]. The Cauchy-deviate method is used for the propagation of epistemic uncertainty [4]. We show that these two approaches can be combined to obtain interval failure probability of the aircraft performance. We show the applicability and efficiency of the methodology on the NASA Langley UQ black-box model.

2 BLACK-BOX CODES

In uncertainty quantification distinguishing between black-box models and open-source models is consequential. The research community is not united on the definition of black boxes. Researchers in machine learning often refer to black boxes as deep learning models, which are practically impenetrable due to their complexity. A more general definition of black-box code can be found in [4]. In this paper a black-box model is a quantitative model, whose source code is inaccessible.

2.1 Why such models exist

The reason why this kind of model exists can be ascribed to (i) secrecy: for example, commercial companies release to the open market only the software binaries, or the code is encrypted for verification and validation by third parties; (ii) legacy: for example, the source code is partially or totally lost and only the binaries are available, or the source code is available but does not build on the current compiler; (iii) complexity: for example, deep learning models, and large FEM and CFD models.

2.2 Need to distinguish black-box and open-source codes

In automatic uncertainty propagation the source code can be decomposed into a list of basic binary operations. This turns out to be the key for efficient rigorous propagation, as numerical strategies for interval analysis can be deployed.

2.3 Restrictions introduced by black boxes

Uncertainty quantification on black-box models is particularly challenging. This is because the model can only be queried and often the time required by a single evaluation is very long, in some cases ranging from hours to weeks.

Furthermore, when the user does not completely know the origin of the model, more evaluations are usually needed to characterise the mathematical properties of the model. For example, checking that the function is continuous or smooth may already require a significant amount of evaluations. Nevertheless, even if we know that the function is e.g. smooth, computing the partial derivatives may be computation expensive. This fact rules out the efficient use of differential algebra, local Taylor expansions, and monotonicity checks that may be alternatively possible in the case of open-source code. Another big limitation introduced by black-box codes is the difficulty of rigorous propagation of epistemic uncertainty. The rigorous propagation is only possible by means of intrusive interval analysis.

2.3.1 What can be done?

We can establish prove that the function in a black-box model is deterministic, by checking whether the outputs are different on a number of repeated computer experiments. If the black-box model is non-deterministic and single evaluations are computation expensive, then rigorous uncertainty quantification may not be practically possible. Techniques that combine surrogate modelling, and massive high performance parallel and distributed computing, may be the answer. If the model is deterministic, the aforementioned techniques may still be necessary, but it is also possible to better characterise the behaviour of the black box and deploy sampling schemes and accelerating strategies to achieve efficiency.

3 PROBLEM STATEMENT

A deterministic model can be conceptually represented in the functional form

$$\boldsymbol{y} = f(\boldsymbol{x}) \tag{1}$$

where f is a function that maps $\boldsymbol{x}$ to $\boldsymbol{y}$ [5]. The variable $\boldsymbol{x} = [x_1, ..., x_{N_X}]$ is a vector of *real* valued inputs, and $\boldsymbol{y} = [y_1, ..., y_{N_Y}]$ is a vector of real-valued model outputs. In this paper we will restrict the discussion to $N_Y = 1$, so for simplicity we assign $N = N_X$.

The uncertainty about the elements of $\boldsymbol{x}$ will be represented by a sequence of *imprecise* distributions $\mathcal{D}_1, ..., \mathcal{D}_D$, where $\mathcal{D}_j$ characterises the uncertainty associated with the element x_j of $\boldsymbol{x}$. Various correlations and other restrictions involving the elements of $\boldsymbol{x}$ may be specified. Typically, these distributions are obtained through some form of expert elicitation or expert review process.

3.1 Uncertainties

The *imprecise distribution* $\mathcal{D}_j$ fully characterises both the aleatory and epistemic uncertainty of the element x_j. The probability box or simply *p-box* [$\overline{\mathcal{D}}$, $\underline{\mathcal{D}}$] denote the set of all non-decreasing functions $\mathcal{D}$ from the real line into $[0, 1]$, such that $\underline{\mathcal{D}}(x) \leq \mathcal{D}(x) \leq \overline{\mathcal{D}}(x)$ [6]. Eq. 2 summarises in one formula the latter sentence.

$$\mathcal{D} : \mathbb{R} \rightarrow [0, 1], \quad \underline{\mathcal{D}}(x) \leq \mathcal{D}(x) \leq \overline{\mathcal{D}}(x) \tag{2}$$

So if [$\overline{\mathcal{D}}$, $\underline{\mathcal{D}}$] is a *p-box* for a random variable X whose distribution $\mathcal{D}$ is unknown except that is within the *p-box*, then $\overline{\mathcal{D}}$ and $\underline{\mathcal{D}}$ are respectively lower and upper bounds on $\mathcal{D}(x)$, which is the – imprecisely known – probability that the random variable X is smaller than x. In practical applications it is very common to construct p-boxes using known probability distributions with interval parameters. In these cases the assumption on the probability distribution type may be relaxed to include all the distributions that fall within the bounds.

3.2 Failure probability

The probability that the model output y is greater than a given, yet uncertain, threshold can be expressed as

$$P(Y > \tilde{y}) = P(f(\boldsymbol{X}) > \tilde{y}) \tag{3}$$

Often the threshold $\tilde{y}$ represents a given performance limit that the model under study should be in, with a large margin of safety. The more catastrophic are the consequences associated with exceeding the given threshold, the larger the safety margin, therefore the smaller is the failure probability.

The aim of this paper is to present a sampling strategy to compute small failure probability bounds on Y when continuous imprecise probability distributions are defined for the input variables of the black-box model.

3.3 Sampling

The probability of Eq. 3 can be approximated via sampling. The solution of the multi-dimensional integral of Eq. 4 can be obtained by averaging over the generated samples so to avoid the numerical integration.

Let $\mathcal{C}_{\boldsymbol{x}} : [0,1]^N \rightarrow [0,1]$ be the copula function of the vector of uncertain variables, and let $g(\boldsymbol{x}) = \tilde{y} - f(\boldsymbol{x})$, with $\Omega_F = g(\boldsymbol{x}) < 0$ denote the failure domain.

The probability of failure can be given as:

$$P\left(g(\boldsymbol{x}) < 0\right) = \int_{\Omega_F} \mathbf{d}\mathcal{C}_{\boldsymbol{x}} \tag{4}$$

Where, the copula expresses the aleatory dependence between the p-boxes, and can be used to represent without loss of generality, the uncertainty about the model in its entirety. Clearly, this is restricted to the case of precise aleatory dependence among the variables. The dependence among focal elements, also known as epistemic dependence, will not be treated in this paper.

In this particular formulation of the uncertainty model, every draw corresponds to a *focal element*, which is the interval counterpart of a pointwise sample. We use the *inverse transformation method* to generate focal elements from the copula function [7]. Focal elements $[\boldsymbol{x}]^{\{s\}} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}]^{\{s\}}$ are propagated through the model solving the following two problems:

$$\min_{\boldsymbol{x}\in[\boldsymbol{x}]^{\{s\}}} g([\boldsymbol{x}]^{\{s\}}), \qquad \max_{\boldsymbol{x}\in[\boldsymbol{x}]^{\{s\}}} g([\boldsymbol{x}]^{\{s\}}). \tag{5}$$

Let us define the set A of all the boxes contained in Ω_F, and the set B^c of all the boxes strictly not contained in Ω_F. Let B be the complement of B^c, then the two sets are

$$A = \{[\boldsymbol{x}] : [\boldsymbol{x}] \subseteq \Omega_F\}\,; \quad B = \{[\boldsymbol{x}] : [\boldsymbol{x}] \cap \Omega_F \neq \varnothing\} \tag{6}$$

Let us define the characteristic function χ_A for a set A as

$$\chi_A(\boldsymbol{e}_i) = \begin{cases} 1 & \text{if } \boldsymbol{e}_i \in A \\ 0 & \text{if } \boldsymbol{e}_i \notin A \end{cases} \tag{7}$$

where $\boldsymbol{e}_i$ is a box in Ω. Lower and upper bounds on the failure probability are obtained averaging the number of focal elements classified as in Eq. 6.

$$\overline{\hat{p}_F} = \sum_{s=1}^{N} \chi_A\left([\boldsymbol{x}]^{\{s\}} \subseteq \Omega_F\right), \qquad \underline{\hat{p}_F} = \sum_{s=1}^{N} \chi_B\left([\boldsymbol{x}]^{\{s\}} \cap \Omega_F \neq \varnothing\right). \tag{8}$$

4 LINE SAMPLING

Line sampling is a simulation method primarily developed to efficiently compute small failure probabilities for high dimensional problems [3]. Line sampling has been recently extended to deal with epistemic uncertainty in the form of intervals [8]. The method is generally applicable and it only requires the knowledge of the so-called "important direction", $a \in \mathrm{R}^N$, which can be any vector pointing towards the failure region. In this paper we use the line sampling scheme to produce *focal elements*. On each focal element we run the problem of Eq. 5 to solve the classification of Eq. 6.

5 PROPAGATION OF EPISTEMIC UNCERTAINTY

5.1 Why Monte Carlo is bad for epistemic propagation

Monte Carlo simulation works particularly well with aleatory uncertainty. This is because Monte Carlo is insensitive to the number of random variables. For example, the estimation of $E\left(f(\boldsymbol{x})\right)$, where E is the mean operator, can be done efficiently by Monte Carlo. However, this is quite the opposite in epistemic uncertainty, where the accuracy of the simulation method drastically decreases with the number of epistemic variables. Consider this simple example: we want to sum a number N of epistemic variables $x_{1:N}$ in the box $[a,b]^N$, and compute the bounds of the resulting *sum*. The exact bounds on the *sum* can be computed exactly and are

$$f(x) = \sum_{i=1:N} x_i = [N*a, N*b] = N*[a,b].$$

Now, let us pretend that the function $f(x)$ is a black-box model, so we do not know that behind the model is doing a simple *sum*. A Monte Carlo way to approach the problem consists in (i) generating random samples in $[a,b]^N$, (ii) summing the generated samples, (iii) getting the minimum and maximum of the *sum*. This process gets increasingly less accurate as the number of variables increases. In fact, for the central limit theorem, the obtained sum will approximate a Gaussian distribution, with mean μ and variance σ^2/N, where σ^2 is the sample variance. Given that the variance of the *sum* linearly decreases with the number of variables, it gets ever more improbable to sample close to the endpoints of the box where the exact bounds hold. Note that this applies to any probability distribution within the box $[a,b]^N$, because of the generality of the central limit theorem.

5.2 Cauchy-deviate method

The Cauchy-deviate method propagates epistemic uncertainty using sampling. [4]. The method exploits the properties of the Cauchy distribution, by which a linear combination of Cauchy random variates is also Cauchy distributed. The properties of the Cauchy distribution ensure that the samples do not get trapped around the arithmetic mean of the generated sample set. Thus the model can be treated as a black box. The method works particularly well when the intervals are small or the black-box is approximately monotonic over the interval.

6 THE NASA BLACK-BOX MODEL

The NASA Langley multidisciplinary uncertainty quantification (UQ) challenge was released in 2013 to seek responses from practitioners in the filed of UQ. Among the different challenge problems, NASA was seeking responses pertaining the *propagation of mixed aleatory and epistemic uncertainties through system models*. The challenge page quotes:

"*NASA missions often involve the development of new vehicles and systems that must be designed to operate in harsh domains with a wide array of operating conditions. These missions involve high-consequence and safety-critical systems for which quantitative data is either very sparse or prohibitively expensive to collect. Limited heritage data may exist, but is also usually sparse and may not be directly applicable to the system of interest, making uncertainty quantification extremely challenging. NASA modeling and simulation standards require estimates of uncertainty and descriptions of any processes used to obtain these estimates.*" NASA LaRC UQ Challenge 2014.

The challenge problem was based upon a model of the NASA Langley Generic Transport Model (GTM). The GTM is a 5.5% dynamically scaled, remotely piloted, twin-turbine, research aircraft used to conduct experiments for the NASA Aviation Safety Program. The multidisciplinary character of the proposed problems led the challenge scientists to release the model in the form of a *black-box*. Again quoting the challenge page:

"*Although a discipline-specific application is the focus of this challenge problem,* ***the problem was specifically structured so that specialized aircraft knowledge is not required****. We seek responses from all interested parties not only those with aircraft experience.*"

Five out-of six challenge problems involved solving a forward propagation problem with mixed aleatory and epistemic uncertainty. The epistemic uncertainty was expressed in the form of intervals, while the aleatory uncertainty in the form of beta and normal distributions.

6.1 The challenge problem

In this section we analyze a portion of the black-box model provided by the NASA Langley Research Center. The full model contains twenty-one input parameters and eight outputs, nonetheless we will limit our discussion to only the first five input parameters $\boldsymbol{p} = [p_1, p_2, p_3, p_4, p_5]$, and the first output $y = y_1$. In the original text of the challenge manifesto this output was referred to as x.

NASA provided the software binaries to evaluate h:

$$y = h(\boldsymbol{p}) = h(p_1.p_2, p_3, p_4, p_5), \tag{9}$$

Where $\boldsymbol{p}_{1:3} \in [0, 1]^3$, $\boldsymbol{p}_{4,5} \in \mathbb{R}$. Specific information about these parameters are provided in Table 1. We provide a solution to the following problem:

$$P\left(h(\boldsymbol{p}) > 0.39\right) \tag{10}$$

Or equivalently P $(g(\boldsymbol{p}) < 0)$, with $g(\boldsymbol{p}) = 0.39 - h(\boldsymbol{p})$. The problem of Eq.10 in words is: "What is the probability of $y = h(\boldsymbol{p})$ being greater than 0.39, when the uncertainty about $\boldsymbol{p}$ is given in Table 1?"

6.2 Input variables

The input variables of the problem are shown in Table 1, and are classified into Category I, II, and III.

- Category I represents random variables with a precise distributions; these variables only have aleatory component.
- Category II represents intervals, any value within the endpoints of the interval is allowed. Although interval is a pure epistemic way to characterize uncertainty, intervals do have an aleatory component. In fact, not only any value but also any distribution bounded by the endpoints of the interval is allowed.
- Category III is the more general representation of uncertainty. Dempster-Shafer structures and p-boxes fall into this category.

Table 1: Uncertain input parameters

Variable	Category	Uncertainty	Epistemic component	Marginal distribution
p_1	III	P-box Unimodal Beta	$\mu_1 = [3/5,\ 4/5]$ $\sigma_1^2 = [1/50,\ 1/25]$	$B(a_1, b_1)$ independent
p_2	II	Interval	$\Delta_2 = [0,\ 1]$	–
p_3	I	Random variable	–	$U(0, 1)$
p_4, p_5	III	P-box Normal bivariate MN($\mu_{4:5}$, $\sigma_{4:5}$, $\rho_{4:5}$)	$\mu_4 = [-5,\ 5]$ $\sigma_4^2 = [1/400,\ 4]$ $\mu_5 = [-5,\ 5]$ $\sigma_5^2 = [1/400,\ 4]$ $\rho_{4,5} = [-1,\ 1]$	$N(\mu_4, \sigma_4^2)$ dependent on p_5 $N(\mu_5, \sigma_5^2)$ dependent on p_4

7 Results

A preliminary analysis is run with 1000 samples. The model takes approximately $30s$ to produce a thousand samples on a common desktop computer.

The 1000 focal elements are propagated through the model and the classification of Figure 1 is obtained. This preliminary classification shows clear borders between the three states, namely *safe*, *plausibility*, and *belief*. In Figure 1, the set of dots that are both red and black belong to the first class of Eq.8, i.e. the focal elements that contribute to the upper bound of the failure probability $\overline{p_F}$; while the black dots only contribute to the lower bound.

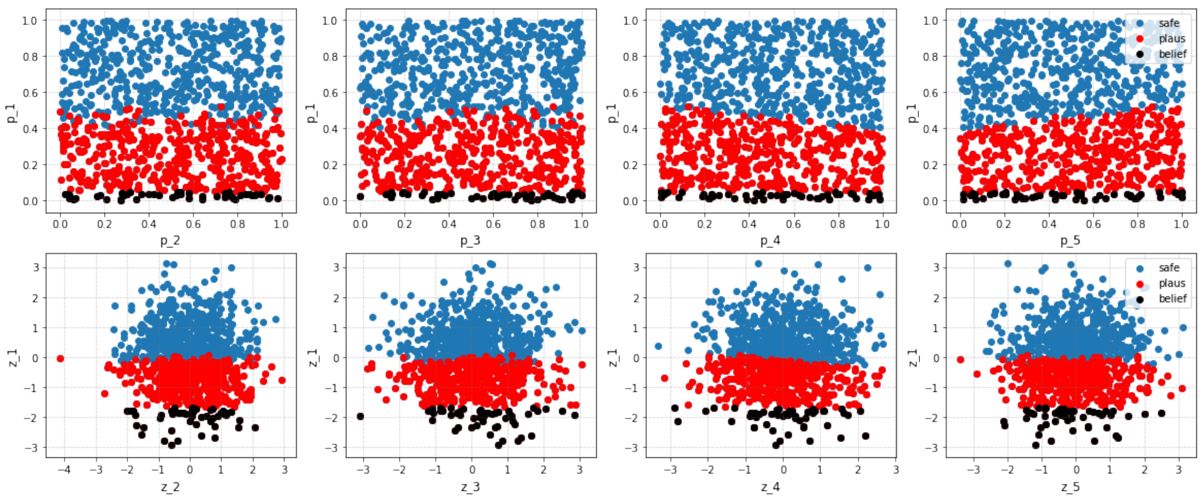

Figure 1: Focal elements p_1 v. $\boldsymbol{p}_{2:5}$ in the copula space (top row) and in the standard normal space (bottom row)

If we look at the scatter plot of classified focal elements for the remaining variables as shown in Figure 5, Figure 6, and Figure 2 we note that there is not a clear border as in the previous case. This suggests that variable p_1 is predominant with respect to the other variables. It is also to note the characteristic cross-shaped dependence between variable p_4 and p_5 depicted in Figure 2 that is typical of the case of unknown linear correlation as specified in Table 1.

With the preliminary analysis it was possible to identify the following important direction α.

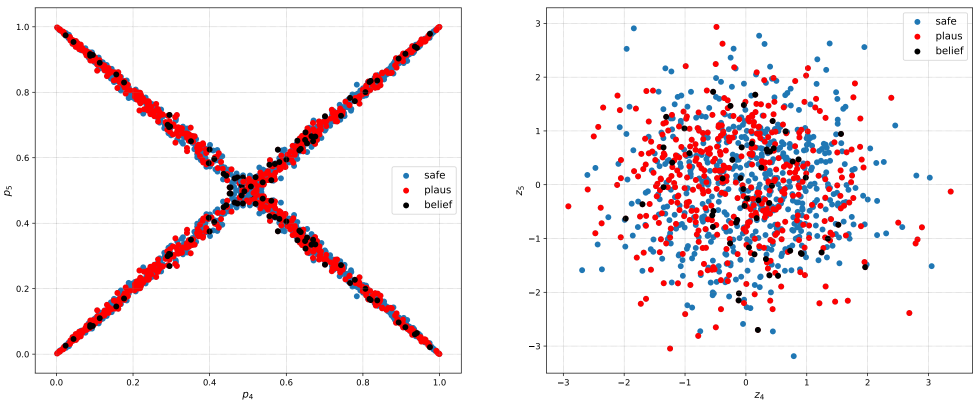

Figure 2: Classification of focal elements for variables p_4 v. p_5 in the copula (top) and the standard normal space (bottom)

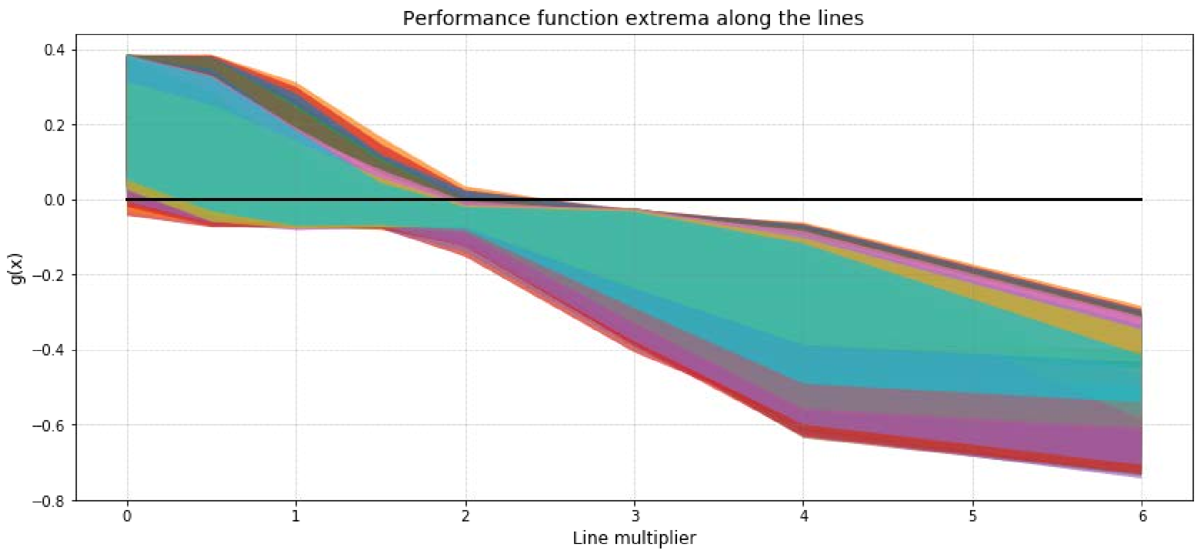

Figure 3: Performance function envelope on the line, each line displays a different color.

$$\alpha = [-0.987, -0.0087, 0.162, 0.0145, 0.00028]$$

The analysis conducted with 100 lines and a total of 500 focal elements, with the threshold of Eq. 10, led to the failure probability interval $[\hat{p}_F] = [0.0378, 0.44]$, with no update in direction. The same analysis run on the set of thresholds $t = [0.4, 0.5, 0.6, 0.7, 0.8]$ led to the upper and lower fragility curves shown in Figure 4.

t	0.4	0.5	0.6	0.7	0.8
$\overline{p_F}$	4.38E-1	1.71E-2	4.70E-3	1.06E-3	3.26E-4
$\underline{p_F}$	3.06E-2	1.17E-4	6.57E-6	3.94E-7	3.79E-8

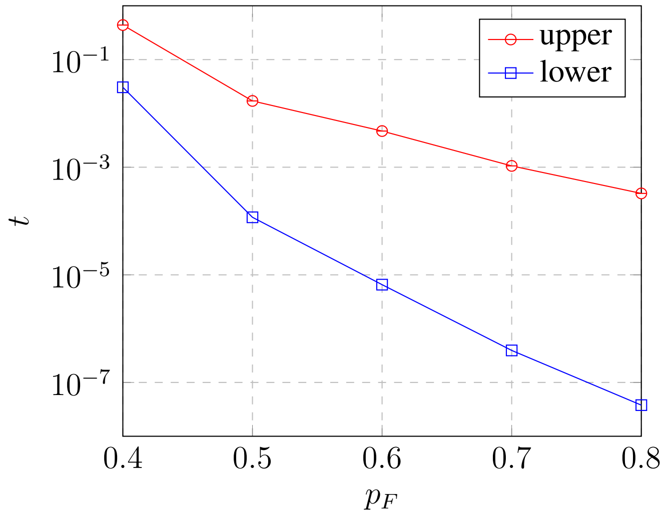

Figure 4: Upper and lower fragility curves

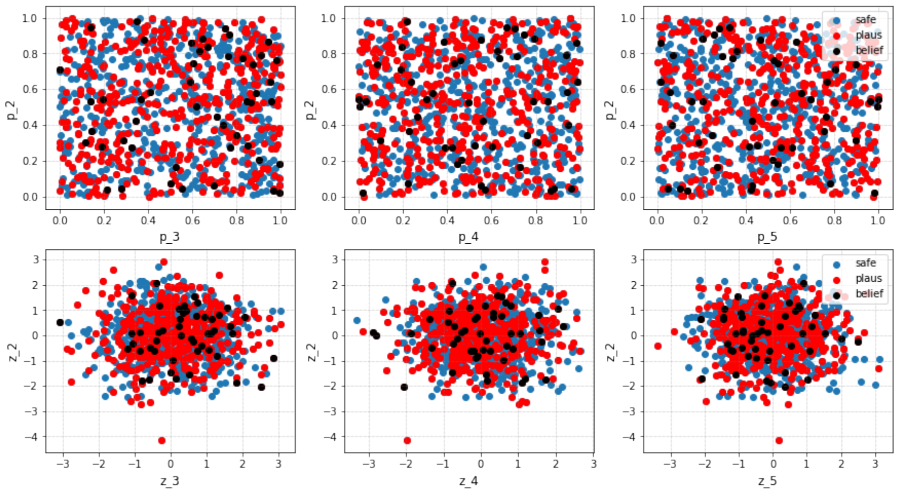

Figure 5: Classification of focal elements for variables p_2 v. $\boldsymbol{p}_{3:5}$

Acknowledgement

This research is funded by the Engineering & Physical Sciences Research Council (EPSRC) through the programme grant "Digital twins for improved dynamic design" with grant no. EP/R006768/1. EPSRC and the DigiTwin consortium are greatly acknowledged for their funding and support.

REFERENCES

[1] Ansys Inc. Delivering a digital twin. Technical report, Ansys Advantage, Issue 2, 2017.

[2] L.G. Crespo and S.P. Kenny. Special edition on uncertainty quantification of the aiaa journal of aerospace computing, information, and communication. *Journal of Aerospace Information Systems*, 12(1):9, 2015.

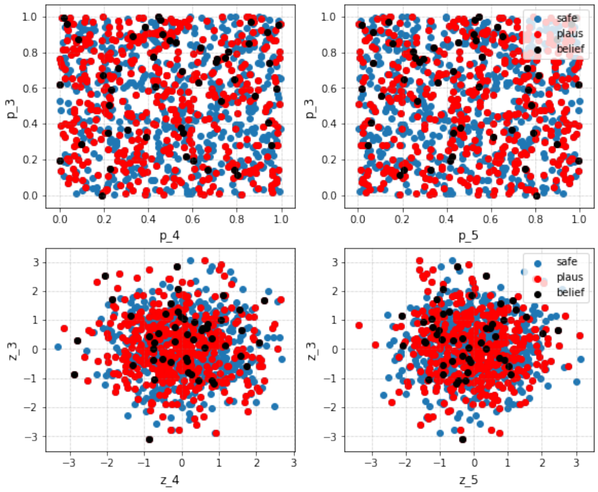

Figure 6: Classification of focal elements for variables p_3 v. $\boldsymbol{p}_{4:5}$

[3] M. de Angelis, E. Patelli, and M. Beer. Advanced line sampling for efficient robust reliability analysis. *Structural Safety*, 52(PB):170–182, 2015.

[4] V. Kreinovich and S.A. Ferson. A new cauchy-based black-box technique for uncertainty in risk analysis. *Reliability Engineering and System Safety*, 85(1-3):267–279, 2004.

[5] J.C. Helton, J.D. Johnson, W.L. Oberkampf, and C.B. Storlie. A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory. *Computer Methods in Applied Mechanics and Engineering*, 196(37):3980 – 3998, 2007. Special Issue Honoring the 80th Birthday of Professor Ivo Babuka.

[6] S. Ferson, V. Kreinovich, L. Grinzburg, D. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2015.

[7] E. Patelli, D.A. Alvarez, M. Broggi, and M. De Angelis. Uncertainty management in multidisciplinary design of critical safety systems. *Journal of Aerospace Information Systems*, 12(1):140–169, 2015. cited By 19.

[8] E. Patelli and M. De Angelis. Line sampling approach for extreme case analysis in presence of aleatory and epistemic uncertainties. pages 2585–2593, 2015.