

# Uncertainty management in multidisciplinary design of critical safety systems

Edoardo Patelli\*    Diego A. Alvarez<sup>†</sup>    Matteo Broggi<sup>‡</sup>    Marco de Angelis\*

Managing the uncertainty in multidisciplinary design of safety critical systems requires not only the availability of a single approach or methodology to deal with uncertainty but a set of different strategies and scalable computational tools (i.e. by making use of the computational power of a cluster and grid computing). The availability of multiple tools and approaches for dealing with uncertainties allows to cross-validate the results and increase the confidence in the performed analysis.

This paper presents a unified theory and an integrated and open general purpose computational framework to deal with scarce data, aleatory and epistemic uncertainties. It allows to solve different tasks necessary to manage the uncertainty, such as: uncertainty characterization, sensitivity analysis, uncertainty quantification and robust design. The proposed computational framework is generally applicable to solve different problems in different fields and numerically efficient and scalable allowing for a significant reduction of the computational time required for uncertainty management and robust design.

The applicability of the proposed approach is demonstrated by solving a multidisciplinary design of a critical system proposed by NASA Langley in the multidisciplinary uncertainty quantification challenge problem.

---

\*Institute for Risk and Uncertainty, University of Liverpool, UK, email: edoardo.patelli@liverpool.ac.uk

<sup>†</sup>Department of Civil Engineering, Universidad Nacional de Colombia at Manizales, Colombia

<sup>‡</sup>Virtual Engineering Centre, University of Liverpool, UK

## I. Introduction

In order to design safe components and systems, the explicit inclusion of uncertainties from different sources is an indispensable step. In fact, under realistic conditions, these products are affected by uncertainties, caused by the lack of sufficient knowledge and/or by natural unpredictable external events. Uncertainty analysis is essential for modellers to obtain a robust representation of model predictions consistent with the state-of-knowledge. If the effects of the uncertainties in the “optimized” products are ignored, these products may perform unsatisfactorily in realistic conditions; for instance, they can show a very low reliability, high reparation and maintenance costs. On the other hand, in a robust design a product or system is less sensitive to the uncertainties and hence, it reaches low variability of the overall performances that can allow significant reductions in terms of e.g. the manufacturing and operating costs).

The design of safety critical systems faces very complex problems due to the presence of varying levels of aleatory and epistemic uncertainty. Aleatory uncertainty is inherent in many natural systems, and therefore cannot be reduced, but can be described and its effect quantified. Epistemic uncertainty is not completely avoidable, since it is not possible to perfectly model or predict real world situations, although epistemic uncertainty can be reduced, better characterized and quantified by using available knowledge. Despite the different levels of uncertainty, decision makers still need to make clear choices based on the available information. They need to trust the methodology adopted to propagate uncertainties through multi-disciplinary analysis, in order to quantify the risk with the current level of information and avoid wrong decisions due to artificial restrictions introduced by the modelling.

Multiple mathematical concepts can be used to characterize variability and uncertainty. Probability distributions can be used to represent the relative frequency of a given state of the system, or they can represent the degree of belief or confidence that a given state of the system exists.<sup>1</sup> Often, very limited information is available, and collecting more data or samples might not be possible or too expensive. Given the limitations of amount of data, quantification methods often rely on subjective judgment 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 concepts of imprecise probabilities. Imprecise probability combines probabilistic and set theoretical components in a unified theory allowing the identification of bounds on probabilities for the

48 events of interest.<sup>2</sup>

49 Random set theory is specially suited to model under the same framework uncertainty represented as  
50 cumulative distribution functions (CDFs), intervals,<sup>3</sup> probability boxes,<sup>4</sup> possibility distributions<sup>5</sup> (they are  
51 closely related to normalized fuzzy sets) and Dempster-Shafer<sup>6,7</sup> structures without making any implicit or  
52 explicit assumptions. In other words, random set theory is a technique that permits to model naturally the  
53 aforementioned representations of uncertainty.

54 In this work, novel and efficient strategies are proposed to deal with aleatory and epistemic uncertainty.  
55 Random set theory is used as a unifying theoretical framework, to model different representations of the  
56 uncertainty. Additionally, the developed procedures have been implemented in an integrated computational  
57 framework allowing to solve realistic problems using a number of different approaches and methodologies.  
58 This provides an excellent tool for cross-validating the results obtained at each stage of the analysis and  
59 hence to increase the confidence in the adopted methodology and in the results. The applicability of approach  
60 is demonstrated by solving the NASA Langley multidisciplinary uncertainty quantification (UQ) challenge  
61 problem.<sup>8</sup>

## 62 **Motivation of the study**

63 The development of safety-critical systems that must be designed to operate in harsh environments with a  
64 wide array of operating conditions (e.g. new vehicles, aircraft, nuclear power plants etc.) is a challenging  
65 problem. Furthermore, the failure of such systems might have high consequences for which quantitative  
66 data is either very sparse or prohibitively expensive to collect. Hence, uncertainty management is necessary  
67 to provide support to the decision makers through a series of different and interconnected analyses. For  
68 instance, estimating the importance of collecting additional information allows to characterize and reduce  
69 uncertainty; by performing sensitivity analysis, it is possible to identify the parameters that contribute the  
70 most to the variability of the output; uncertainty propagation allows to study the effects of uncertainty on  
71 the performance of the system and to identify extreme-case scenarios; finally, optimizing the design explicitly  
72 taking into account the effect of uncertainties allows to design a robust system.

73 Recent reports have clearly shown that the risk assumed by the decision maker is often wrongly estimated  
74 due to inadequate assessment of uncertainty.<sup>8</sup> Modelling and simulation standards require estimates of

75 uncertainty (and descriptions of any processes used to obtain these estimates) in order to increase confidence  
76 and consistency in safety predictions and encourage the development of improved methods for quantifying  
77 and managing uncertainty. In this context, the NASA Langley multidisciplinary uncertainty quantification  
78 (UQ) challenge problem has been addressed in order to determine limitations and ranges of applicability of  
79 existing UQ methodology and to advance the state of the practice in UQ problem.<sup>9</sup>

80 The NASA challenge problem has represented a unique opportunity to test, validate and advance the  
81 capability of the computational framework, namely OPENCROSSAN.<sup>10</sup> This computational framework is able  
82 to deal with different representations of uncertainty and has been adopted to solve all the tasks proposed by  
83 the challengers.

84 Since many of the employed methods are rooted on random set theory, a brief introduction of the theory  
85 for the representation of the joint existence of aleatory and epistemic uncertainty is presented in Section  
86 **II**. The developed approaches for uncertainty quantification and management are presented in Section **III**.  
87 Section **IV** presents the computational framework and some details on computational complexity. Section  
88 **V** summarizes the main aspects, goals and difficulties of the NASA Langley multidisciplinary UQ challenge  
89 problem as well as the results of the various tasks of the challenge problem. Finally, the potentiality and  
90 applicability of the developed computational framework and the proposed approaches are discussed.

## 91 **II. Theoretical background**

92 Random set theory is specially suited to model, under the same framework, uncertainty represented as  
93 cumulative distribution functions (CDFs), intervals, distribution-free probability boxes, possibility distribu-  
94 tions and Dempster-Shafer structures<sup>6,7</sup> without making any implicit or explicit assumption at all.<sup>11,12</sup> In  
95 other words, random set theory is a technique that permits to model the aforementioned representations of  
96 uncertainty. Random sets can be understood as random variables that sample, instead of points, sets (called  
97 focal elements) as realizations.

98 In this context, many of the proposed solutions to the challenge problem make strong use of this kind of  
99 representation. In consequence, in the following, a brief review of the main concepts of random set theory  
100 that will be required in the subsequent discussion is provided. Also some new concepts developed in order  
101 to solve the challenge problem will be introduced.

102 **A. A succinct review of random set theory**

103 Let us consider a universal set  $\mathcal{X} \neq \emptyset$  and its power set  $\mathcal{P}(\mathcal{X})$ . Let  $(\Omega', \sigma_{\Omega'}, P_{\Omega'})$  be a probability space and  
 104  $(\mathcal{F}, \sigma_{\mathcal{F}})$  be a measurable space where  $\mathcal{F} \subseteq \mathcal{P}(\mathcal{X})$ . A *random set*  $\Gamma$  is a  $(\sigma_{\Omega'} - \sigma_{\mathcal{F}})$ -measurable mapping  
 105  $\Gamma : \Omega' \rightarrow \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$ . We will say that every  $\gamma := \Gamma(\alpha) \in \mathcal{F}$  is a *focal element* while  $\mathcal{F}$  is a *focal set*.

106 Analogously to the definition of a random variable, this mapping can be used to define a probability  
 107 measure on  $(\mathcal{F}, \sigma_{\mathcal{F}})$  given by  $P_{\Gamma} := P_{\Omega'} \circ \Gamma^{-1}$ . That is, an event  $\mathcal{R} \in \sigma_{\mathcal{F}}$  has the probability

$$P_{\Gamma}(\mathcal{R}) = P_{\Omega'}\{\alpha \in \Omega' : \Gamma(\alpha) \in \mathcal{R}\}. \quad (1)$$

108 The random set  $\Gamma$  will be also referred to as  $(\mathcal{F}, P_{\Gamma})$ . When all the focal elements of  $\mathcal{F}$  are singletons, then  
 109  $\Gamma$  becomes a random variable  $X$ ; hence,  $\Gamma(\alpha) = X(\alpha)$  and the probability of occurrence of the event  $\mathcal{F}$ , is  
 110  $P_X(\mathcal{F}) := (P_{\Omega'} \circ X^{-1})(\mathcal{F}) = P_{\Omega'}\{\alpha : X(\alpha) \in \mathcal{F}\}$  for every  $\mathcal{F} \in \sigma_X$ . In the case of random sets, it is not  
 111 possible to compute exactly  $P_X(\mathcal{F})$  but its upper and lower probability bounds. Dempster<sup>6</sup> defined those  
 112 upper and lower probabilities by,

$$\text{LP}_{(\mathcal{F}, P_{\Gamma})}(\mathcal{F}) := P_{\Omega'}\{\alpha : \Gamma(\alpha) \subseteq \mathcal{F}, \Gamma(\alpha) \neq \emptyset\} \quad (2a)$$

$$\text{UP}_{(\mathcal{F}, P_{\Gamma})}(\mathcal{F}) := P_{\Omega'}\{\alpha : \Gamma(\alpha) \cap \mathcal{F} \neq \emptyset\} \quad (2b)$$

113 where  $\text{LP}_{(\mathcal{F}, P_{\Gamma})}(\mathcal{F}) \leq P_X(\mathcal{F}) \leq \text{UP}_{(\mathcal{F}, P_{\Gamma})}(\mathcal{F})$ .

114 **COPULAS** A *copula* is a function  $C : [0, 1]^d \rightarrow [0, 1]$  that relates a joint cumulative density functions  
 115 (CDFs) with its marginals, carrying in this way the dependence information in the joint CDF such that each  
 116 of its marginal CDFs is uniform on the interval  $[0, 1]$ . According to Sklar's theorem (see Refs. 13, 14), a  
 117 multivariate CDF  $F_{X_1, X_2, \dots, X_d}(x_1, \dots, x_d) = P[X_1 \leq x_1, \dots, X_d \leq x_d]$  of a random vector  $(X_1, X_2, \dots, X_d)$   
 118 with marginals  $F_{X_i}(x_i) = P[X_i \leq x_i]$  can be written as  $F_{X_1, X_2, \dots, X_d}(x_1, \dots, x_d) = C(F_{X_1}(x_1), \dots, F_{X_d}(x_d))$ ,  
 119 where  $C$  is a copula. The copula  $C$  is itself a CDF and it contains all information on the dependence structure  
 120 between the components of  $(X_1, X_2, \dots, X_d)$  whereas the marginal cumulative distribution functions  $F_{X_i}$   
 121 contain all information on the marginal distributions.

122 The reader is referred to Ref. 15 for an excellent introduction to copulas.

## 123 B. Random sets, CDFs, distribution-free probability boxes and intervals

124 The original definition of random sets is very general; Alvarez<sup>11,12</sup> showed that making the particularizations  
125  $\Omega' := (0, 1]^d$ ,  $\sigma_{\Omega'} := (0, 1]^d \cap \mathcal{B}^d$ , where  $\mathcal{B}$  stands for the Borel  $\sigma$ -algebra on  $\mathbb{R}$ , and  $P_{\Gamma} \equiv \mu_C$  for some  
126 copula  $C$  that contains the dependence information within the joint random set, and using intervals and  
127  $d$ -dimensional boxes as elements of  $\mathcal{F}$ , it is enough to model possibility distributions, distribution-free prob-  
128 ability boxes, intervals, CDFs and Dempster-Shafer structures or their joint combinations (for a definition of  
129 joint Dempster-Shafer structure and joint random set the reader is referred to Ref. 12). Here,  $P_{\Gamma} \equiv \mu_C$  de-  
130 notes the fact that  $P_{\Gamma}$  is the probability measure generated by  $P_{\Omega'}$  which is defined by the Lebesgue-Stieltjes  
131 measure corresponding to the copula  $C$ , i.e.  $\mu_C$ . In other words,  $P_{\Gamma}(\Gamma(G)) = \mu_C(G)$  for  $G \in \sigma_{\Omega'}$ .

132 In the rest of this subsection,  $(\Omega', \sigma_{\Omega'}, P_{\Omega'})$  will stand for a probability space with  $\Omega' := (0, 1]$ ,  $\sigma_{\Omega'} :=$   
133  $(0, 1] \cap \mathcal{B} := \cup_{\theta \in \mathcal{B}} \{(0, 1] \cap \theta\}$  and  $P_{\Omega'}$  will be a probability measure corresponding to the CDF of a random  
134 variable  $\tilde{\alpha}$  uniformly distributed on  $(0, 1]$ , i.e.  $F_{\tilde{\alpha}}(\alpha) := P_{\Omega'}[\tilde{\alpha} \leq \alpha] = \alpha$  for  $\alpha \in (0, 1]$ ; that is,  $P_{\Omega'}$  is a  
135 Lebesgue measure on  $(0, 1]$ .

### 136 1. Cumulative distribution functions

137 When a variable is expressed as a random variable on  $X \subseteq \mathbb{R}$ , the probability law of the random variable  
138 can be expressed using a CDF  $F_X$  (recall  $F_X(x) = P_{\Gamma}(X \leq x)$  for  $x \in X$ ). That CDF can be represented as  
139 the random set  $\Gamma : \Omega' \rightarrow \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$  where  $\mathcal{F}$  is the system of focal elements  $\Gamma(\alpha) := F_X^{-1}(\alpha)$  for  $\alpha \in \Omega'$   
140 (the *inverse* of the CDF  $F_X$  is defined by  $F_X^{-1}(\alpha) := \inf\{x : F_X(x) \geq \alpha, \alpha \in (0, 1]\}$ ; take into account that  
141 this definition uses the infimum since CDFs are weakly monotonic and right-continuous). Note that the  
142 representation of the CDF as a random set only contains an aleatory component, which is given either by  
143  $\alpha$ , or by its corresponding sample  $x = F_X^{-1}(\alpha)$ ; there is not an epistemic component in this representation.

### 144 2. Intervals

145 An *interval*  $I = [l, u]$  can be represented as the random set  $\Gamma : \Omega' \rightarrow \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$  (i.e.  $(\mathcal{F}, P_{\Gamma})$ ) defined  
146 on  $\mathbb{R}$  where the focal set contains the unique focal element  $[l, u]$ , that is,  $\mathcal{F} = I$  and  $\alpha \in (0, 1] \equiv \Omega'$ ; in this  
147 case,  $P_{\Gamma}$  is specified by Eq. (1). In other words, all the samplings of  $\alpha \in \Omega'$  draw the interval  $[l, u]$ . Note  
148 that the representation of intervals as a random set does not contain an aleatory component, inasmuch as

149 it does not matter which value  $\alpha$  takes, because all  $\alpha$ -s map to the same focal element  $I$ . In this case, the  
 150 epistemic component is given by the interval itself,  $I$ .

### 151 3. Probability boxes

152 A *probability box* or *p-box* (term coined by Ferson et al.<sup>4</sup>)  $\langle \underline{F}, \overline{F} \rangle$  is a set of CDFs  $\{F : \underline{F} \leq F \leq$   
 153  $\overline{F}, F \text{ is a CDF}\}$ , delimited by upper and lower CDF bounds  $\underline{F}$  and  $\overline{F} : \mathbb{R} \rightarrow [0, 1]$ , which collectively  
 154 represent the epistemic uncertainty about the CDF of a random variable. This class of functions may not  
 155 have additional restrictions or may belong, as well, to a reduced class of CDFs; using that discrimination,  
 156 probability boxes can be naturally grouped into two disjoint groups: free and distributional.

157 DISTRIBUTION-FREE P-BOXES Distribution-free p-boxes (also known as *non-parametric* p-boxes) appear  
 158 when the CDF of a random variable cannot be specified precisely, given that the CDF family is unknown; in  
 159 this case only the upper and lower CDF bounds  $\underline{F}$  and  $\overline{F}$  bounds of the probability box are specified. These  
 160 bounds can either be defined in advance or can be estimated using for example the methods listed in Zhang  
 161 et. al.<sup>16</sup> and references therein. Note that distribution-free p-boxes do not make any assumption about the  
 162 family or shape of the uncertain CDFs that belong to the p-box.

163 There are two alternatives but equivalent methods to represent distribution-free p-boxes using random  
 164 set theory.

165 The first method was proposed in Refs. 11, 12. Using this method, a *distribution-free probability box*  
 166 delimited by lower and upper CDF bounds  $\underline{F}$  and  $\overline{F}$  can be represented as the random set  $\Gamma : \Omega' \rightarrow$   
 167  $\mathcal{F}, \alpha \mapsto \Gamma(\alpha)$  (i.e.  $(\mathcal{F}, P_\Gamma)$ ) defined on  $\mathbb{R}$  where  $\mathcal{F}$  is the class of focal elements  $\Gamma(\alpha) := \langle \underline{F}, \overline{F} \rangle^{-1}(\alpha) :=$   
 168  $[\overline{F}^{-1}(\alpha), \underline{F}^{-1}(\alpha)]$  for  $\alpha \in (0, 1] \equiv \Omega'$  with  $\underline{F}^{-1}(\alpha)$  and  $\overline{F}^{-1}(\alpha)$  denoting the inverses of  $\underline{F}$  and  $\overline{F}$  and  $P_\Gamma$  is  
 169 specified by Eq. (1).

170 The second alternative method, proposed here, considers a random variable which follows a CDF  $F$  with  
 171 parameters  $\theta_i$  that belong to the interval  $I_i$  for  $i = 1, 2, \dots, m$  (i.e.  $F(\cdot; \theta_1, \theta_2, \dots, \theta_m)$ ). This representation  
 172 is in comparison to the first method, which models a p-box using only its lower and upper CDF bounds  $\underline{F}$  and  
 173  $\overline{F}$ . Using the random set representation, a focal element of the probability box  $\langle \underline{F}, \overline{F} \rangle$  can be represented  
 174 as the image through the function  $F^{-1}$  of the input intervals  $\{I_i : i = 1, 2, \dots, m\}$  together with the sample  
 175 of  $\alpha$  which is a uniform random variable on  $(0, 1] \equiv \Omega'$ . In consequence, it can be represented as the

176 random set  $\Gamma : \Omega' \rightarrow \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$  (i.e.  $(\mathcal{F}, P_\Gamma)$ ) defined on  $\mathbb{R}$  where  $\mathcal{F}$  is the system of focal elements  
 177  $\{F^{-1}(\alpha; I_1, I_2, \dots, I_m) : \alpha \in \Omega'\}$  and  $P_\Gamma$  is specified by Eq. (1). Observe that each focal element has an  
 178 aleatory component  $\alpha$  and an epistemic component in the Cartesian product  $\times_{i=1}^m I_i \triangleq I_1 \times I_2 \times \dots \times I_m$ .

179 This representation of distribution-free p-boxes shows that for a single realization of the aleatory com-  
 180 ponent  $\alpha$ , a focal element contains the image through  $F^{-1}$  of all the possible combinations of values of  
 181 the intervals for the parameters of the parental CDF  $F$ . It derives from the fact that a focal element  
 182 is defined as:  $\langle \underline{F}, \overline{F} \rangle^{-1}(\alpha) = \{x : F(x) = \alpha, F \in \langle \underline{F}, \overline{F} \rangle\}$ ,  $\overline{F}^{-1}(\alpha) = \inf_{\theta \in \times_{i=1}^m I_i} F^{-1}(\alpha; \theta_1, \dots, \theta_m)$  and  
 183  $\underline{F}^{-1}(\alpha) = \sup_{\theta \in \times_{i=1}^m I_i} F^{-1}(\alpha; \theta_1, \dots, \theta_m)$ .

184 Note that only distribution-free probability boxes can be represented using random set theory. However,  
 185 in the analysis of the challenge problem a different approach has been used to represent distributional  
 186 probability boxes as will be explained in the following lines.

187 DISTRIBUTIONAL P-BOXES AND THE DOUBLE LOOP MONTE CARLO STRATEGY Distributional p-boxes  
 188 (also known as *parametric* p-boxes) appear when there is uncertainty in the representation of the parameters  
 189 of a given CDF (hereafter called the *parental CDF*.) These parameters are imprecisely specified as intervals.  
 190 For instance, consider a quantity that is known to be Gaussian with mean within the interval  $[1, 2]$  and  
 191 standard deviation somewhere in  $[3, 4]$ ; Ferson et. al.<sup>4</sup> describes how to obtain such probability boxes. All  
 192 CDFs that are normal and have means and standard deviations inside these respective intervals will belong  
 193 to this probability box. The upper and lower CDF bounds  $\underline{F}$  and  $\overline{F}$  of the p-box enclose many non-normal  
 194 distributions, but these would be excluded from the p-box by specifying the normal CDF as the *parental*  
 195 *distribution family*.

196 According to the second representation of distribution-free p-boxes, the focal element corresponding to  
 197 a realization  $\alpha$  of the aleatory component contains the image through  $F^{-1}$  of all possible  $\theta \in \times_{i=1}^m I_i$ . As  
 198 consequence, a set of focal elements of the probability box would be a family of intervals each of them being  
 199 a mapping of  $\times_{i=1}^m I_i$  through  $F^{-1}$ . Hence, for a fixed value of  $\theta$  it is not possible to identify in that set of  
 200 intervals the points that would belong to some CDF. For this reason, random set theory can not be used to  
 201 model distributional p-boxes.

202 Distributional p-boxes can be dealt with using a double loop Monte Carlo strategy, in which the outer  
 203 loop draws  $\theta$ -s from  $\times_{i=1}^m I_i$  and the inner loop samples  $\alpha$ -s from a uniform distribution in  $(0, 1]$ . In this case,

204 using the principle of maximum entropy, we will assume a uniform distribution in  $\times_{i=1}^m I_i$ . This approach  
 205 has been used to solve some of the tasks of the challenge problem. Please note that the outer loop can be  
 206 used to drive an optimization/search process in  $\times_{i=1}^m I_i$  to identify the lower and upper bounds. In this case,  
 207 it is not necessary to assume a uniform distribution in  $\times_{i=1}^m I_i$ .

### 208 C. Sampling from a random set

209 A sample from a random set is simply obtained by generating an  $\alpha$  from a uniform distribution on  $(0, 1]$  and  
 210 then, retrieving the corresponding focal element  $\Gamma(\alpha)$ ; for example, for sampling from a distributional-  
 211 free probability box an  $\alpha$  uniformly distributed in  $(0, 1]$  is drawn and then its corresponding “ $\alpha$ -cut”  
 212  $[\overline{F}^{-1}(\alpha), \underline{F}^{-1}(\alpha)]$  is obtained. In the case of multivariate random sets, a sample  $\alpha \in \Omega'$  is drawn from  
 213 the copula  $C$  that models the dependence between the input variables. Then, the corresponding marginal  
 214 focal elements are obtained and combined as explained in the next subsection. Take into account that  $n$   
 215 samples of a random set form the Dempster-Shafer structure  $(\mathcal{F}_n, m)$ ; here  $\mathcal{F}_n$  denotes the set of all sampled  
 216 focal elements; the basic mass assignment  $m$  associated to each focal element is equal to  $1/n$ ; note that a  
 217 Dempster-Shafer structure is itself a finite random set.<sup>11,12</sup>

218 Samples from distributional p-boxes can be obtained resorting to a double Monte Carlo loop as explained  
 219 in the previous Section.

### 220 D. Combination of focal elements

221 After sampling each input variable, a combination of the sampled focal elements is carried out. Usually,  
 222 the joint focal elements are given by the Cartesian product  $\times_{i=1}^d \gamma_i \subseteq \mathcal{X}$  where  $d$  is the number of input  
 223 variables,  $\gamma_i := \Gamma^i(\alpha_i)$  are the sampled focal elements from every input variable (that is,  $\gamma_i$  represents a  
 224 sampled marginal focal element). Some of these  $\gamma_i$  are intervals, some other, points. Inasmuch as every  
 225 sample of a input variable can be represented by  $\gamma_i$  or by the corresponding  $\alpha_i$ , the joint focal element can  
 226 be represented either by the  $d$ -dimensional box  $\gamma := \times_{i=1}^d \gamma_i \subseteq \mathcal{X}$  or by the point  $\alpha := [\alpha_1, \alpha_2, \dots, \alpha_d] \in \Omega'$   
 227 (see Figure 1).

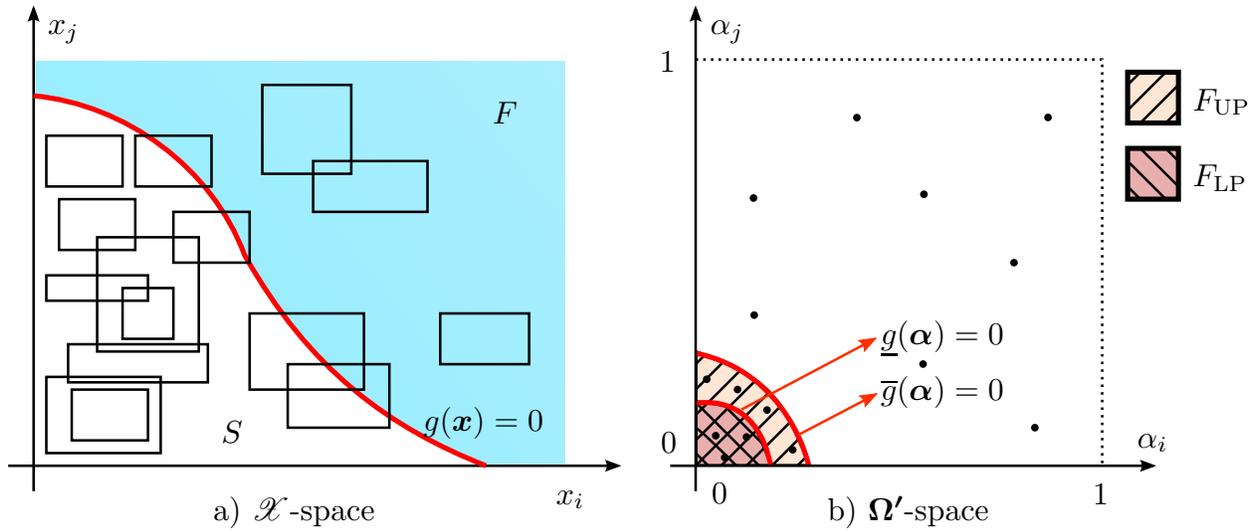


Figure 1: Focal elements in the  $\mathcal{X}$  (Panel a) and in the  $\Omega'$ -space (Panel b), respectively. The focal elements are the realizations of input variables which are depicted either as the points  $\alpha$  in the  $\Omega'$ -space or as (multi-dimensional) boxes, corresponding to the focal element  $\Gamma(\alpha)$ , in the  $\mathcal{X}$  space. The figure (a) shows also the failure surface,  $g(\mathbf{x}) = 0$ , that defines the safe  $\mathcal{S}$  and failure  $\mathcal{F}$  domains. In the  $\Omega'$ -space (b) are defined the regions  $F_{LP}$  and  $F_{UP}$  together with the failure surfaces  $\underline{g}(\alpha) = 0$  and  $\bar{g}(\alpha) = 0$ , where  $\underline{g}(\alpha) := \min_{\mathbf{x} \in \Gamma(\alpha)} g(\mathbf{x})$  and  $\bar{g}(\alpha) := \max_{\mathbf{x} \in \Gamma(\alpha)} g(\mathbf{x})$ . Those boxes in  $\mathcal{X}$  which contain at least one point of the failure region  $\mathcal{F}$  have a corresponding  $\alpha$  point in the region  $F_{UP}$ ; while those boxes in  $\mathcal{X}$  which are completely contained in the region  $\mathcal{F}$  have a corresponding  $\alpha$  point in the region  $F_{LP}$ .

## 228 E. The epistemic and the aleatory spaces

229 Along this paper, two spaces are defined for modelling the aleatory and the epistemic uncertainties and  
 230 which are called the *aleatory space*  $\Omega$  and the *epistemic space*  $\Theta$ , respectively (see Figure 2).

231 **THE ALEATORY SPACE** A sample from a random set is obtained by drawing an  $\alpha \in \Omega'$  from the copula  $C$ .  
 232 Since a sample from an interval does not contain an aleatory component, if we strip from space  $\Omega'$  all those  
 233 components which belong to intervals, then a subspace  $\Omega$  of  $\Omega'$  is obtained. The subspace  $\Omega$  contains only  
 234 probabilistic information without spurious random variables. This set is called from now on the *aleatory*  
 235 *space*  $\Omega$ . Without loss of generality all copulas in our discussion will be defined on  $\Omega$ , and all subsequent  
 236 discussion will be performed with respect to the set  $\Omega$ .

237 THE EPISTEMIC SPACE The epistemic space  $\Theta$  is formed by the Cartesian product of all intervals  $\{I_i, i =$   
238  $1, 2, \dots, q\}$  that contain epistemic uncertainty, that is  $\Theta = \times_{i=1}^q I_i$ . Since the epistemic uncertainty can be  
239 reduced when additional information is available, we will assume that a point  $\theta^* \in \Theta$  in the epistemic space  
240 will represent the “true uncertainty model”, which will result once all epistemic uncertainty is removed from  
241  $\Theta$ . When new information is available, the epistemic space will shrink to a subset of it called the *reduced*  
242 *epistemic space*.

243 For example, let’s consider a problem with four input variables: two correlated random variables  $X$  and  $Y$   
244 modelled as a bivariate normal distribution and two independent variables  $W$  and  $Z$  which are modelled by  
245 the intervals  $I_W$  and  $I_Z$ , correspondingly. The joint CDF of  $X$  and  $Y$  is defined by the mean vector  $[\mu_X, \mu_Y]^T$ ,  
246 variances  $[\sigma_X^2, \sigma_Y^2]^T$  and a Pearson correlation coefficient  $\rho_{XY}$ . If we assume that all those five parameters  
247 are also unknown and represented by intervals, namely,  $I_{\mu_X}$ ,  $I_{\mu_Y}$ ,  $I_{\sigma_X^2}$ ,  $I_{\sigma_Y^2}$  and  $I_{\rho_{XY}}$ , respectively, then,  
248 the aleatory space  $\Omega$  is  $(0, 1]^2$  while  $\Omega'$  is  $(0, 1]^4$ ; in addition, a Gaussian copula is defined on the aleatory  
249 space  $\Omega$ . Finally, the epistemic space  $\Theta$  is a seven-dimensional space formed by the Cartesian product  
250  $I_{\mu_X} \times I_{\mu_Y} \times I_{\sigma_X^2} \times I_{\sigma_Y^2} \times I_{\rho_{XY}} \times I_W \times I_Z$ . Notice that the point  $\theta^*$  belongs to that space.

## 251 F. The system representation as a function of the aleatory and the epistemic uncertainty

252 Let us denote by  $\mathcal{G} : \mathcal{X} \rightarrow \mathbb{R}$  a function that represents the system; this function maps from the input  
253 space  $\mathcal{X}$  of input variables to the real line and let  $\mathcal{W} : \Omega \times \Theta \rightarrow \mathcal{X}$  be a function which returns the point  
254 in  $\Gamma(\alpha)$  after reducing the epistemic uncertainty in  $\Theta$  to  $\theta$ . The function  $\mathcal{W}$  exists only if the random set  
255  $\Gamma$  models intervals, CDFs, p-boxes or their joint combination. This function does not exist if  $\Gamma$  models  
256 Dempster-Shafer structures or possibility distributions, but this is not the case in this paper. Note on the  
257 one hand, that the image of  $\Theta$  through  $\mathcal{W}(\alpha; \cdot)$  is the focal element  $\Gamma(\alpha)$ ; on the other hand, the image of  $\Omega$   
258 through  $\mathcal{W}(\cdot; \theta)$  can be modelled as a CDF with parameter vector  $\theta$ , that is,  $F(\cdot; \theta)$ . Take into account that  
259 the definition of function  $\mathcal{W}$  uses, in the case of CDFs and p-boxes, the inverse CDF of the input variable  
260 in consideration.

261 We will define the function  $\mathcal{H} : \Omega \times \Theta \rightarrow \mathbb{R}$  as  $\mathcal{H} = \mathcal{G} \circ \mathcal{W}$ , that is,  $\mathcal{H}$  represents the system as well, but  
262 its domain is the Cartesian product of the aleatory and epistemic spaces (see Figure 2).

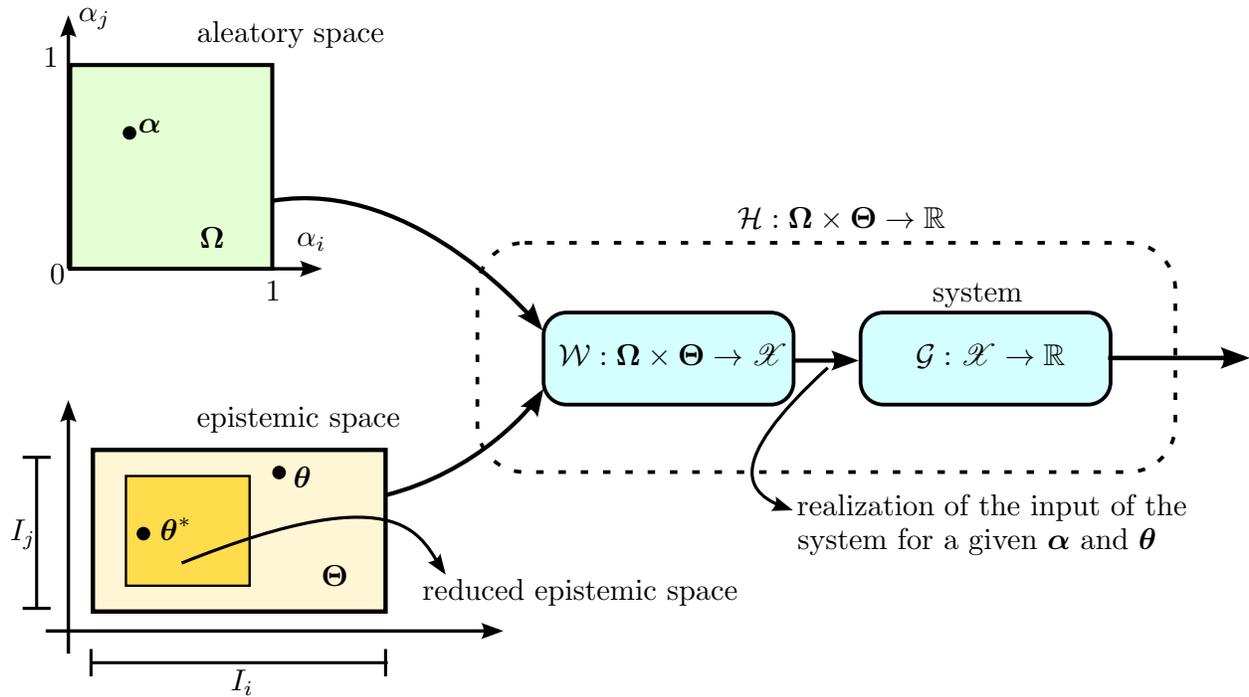


Figure 2: Representation of the aleatory and epistemic spaces and their propagation through the model. Here  $\theta^*$  represents the *true uncertainty model* after all epistemic uncertainty has been removed. The subset of the epistemic space that appears after new information is available is the *reduced epistemic space*. The function  $\mathcal{W}$  produces a realization for a given  $\alpha$  and  $\theta$ ; this output becomes the input for the system  $\mathcal{G}$ . The composition of both functions forms the function  $\mathcal{H}$ .

## 263 G. Mapping of a focal element through a system: the extension principle of random sets

264 The capability to propagate intervals, CDFs, p-boxes, and their combination through a system represents  
 265 the core of the developed computational framework. In order to find the image of a focal element,  $\gamma_i \subseteq \mathcal{X}$ ,  
 266 through a function  $\mathcal{G} : \mathcal{X} \rightarrow \mathbb{R}$ , the extension principle of random sets is used (this principle states how  
 267 to propagate a random set through a function – see Ref. 17). This can be done by means of optimization  
 268 methods,<sup>18</sup> sampling methods,<sup>19</sup> a vertex method,<sup>20</sup> or the interval arithmetic method.<sup>21,3</sup> In the following,  
 269 the optimization and the sampling methods will be explained in detail, since both methods have been  
 270 employed to solve the NASA UQ challenge problem.

271 THE OPTIMIZATION METHOD If a focal element  $\gamma_i := \Gamma(\alpha_i)$  is connected and compact and  $\mathcal{G}$  is continuous,  
 272 then the image of the set  $\gamma_i$  through  $\mathcal{G}$ , can be calculated as

$$\mathcal{G}(\Gamma(\alpha_i)) = [\underline{\mathcal{G}}(\alpha_i), \overline{\mathcal{G}}(\alpha_i)] \quad (3)$$

273 where,

$$\underline{\mathcal{G}}(\alpha_i) := \min_{\mathbf{x}_i \in \Gamma(\alpha_i)} \mathcal{G}(\mathbf{x}_i) \quad \overline{\mathcal{G}}(\alpha_i) := \max_{\mathbf{x}_i \in \Gamma(\alpha_i)} \mathcal{G}(\mathbf{x}_i); \quad (4)$$

274 are limit state functions defined in  $\Omega$ . Using the function  $\mathcal{H}$  defined in Section F, Eqs. (4) can be written  
 275 as an optimization over the epistemic space:

$$\underline{\mathcal{G}}(\alpha_i) := \min_{\theta \in \Theta} \mathcal{H}(\alpha_i, \theta) \quad \overline{\mathcal{G}}(\alpha_i) := \max_{\theta \in \Theta} \mathcal{H}(\alpha_i, \theta). \quad (5)$$

276 This approach is usually employed when  $\mathcal{G}$  is a nonlinear function of the system parameters. The main  
 277 drawback of this method is that it requires a high computational effort in a complex and large scale system.

278 THE SAMPLING METHOD (RANDOM SEARCH) The image of the focal element  $\gamma_i$  through  $\mathcal{G}$  can also be  
 279 estimated using a sampling technique (this should not be confused with the double loop simulation used  
 280 to deal with distributional p-boxes). Remembering that the focal element  $\gamma_i$  is a multi-dimensional box;  
 281 random samples can be generated inside that box and then they are mapped through  $\mathcal{G}$ ; then Eqs. (4)  
 282 are approximated by the smallest and largest values of the images of those samples. This method is easy  
 283 to implement but it requires huge number of samples (due to the curse of dimensionality). The sampling  
 284 method gives does not guarantee that the true minimum and maximum are identified even using a very large  
 285 number of samples.

### 286 III. Proposed approach for uncertainty management and quantification

287 The robust design of safety-critical systems requires not only the explicit treatment of different forms and  
 288 representations of uncertainty but also, performing a number of different tasks. Generally, the design of such  
 289 systems requires inputs and criteria of different disciplines and one of the main challenges in uncertainty  
 290 management is how to propagate the uncertainty and understand how the uncertainty in one field affects  
 291 other disciplines. More specifically,

- 292 • the first task required is to refine the current uncertainty model using new available information. This  
293 task is often called *model updating* (see e.g. Refs. 22,23);
- 294 • usually *sensitivity analysis* is performed for the identification of those parameters whose uncertainty is  
295 the most/least consequential. This allows to drive the collection of new data and information focusing  
296 on those parameters that affect mostly the variability of the outputs;
- 297 • the *propagation of mixed aleatory and epistemic uncertainties* of the refined/improved model and the  
298 extreme-case system performance assessment are performed in order to identify the combinations of  
299 parameters that lead to the worst performance;
- 300 • finally the *design in the presence of uncertainty* is achieved. This task is computationally demanding  
301 since it requires the propagation of the uncertainty through the system for each candidate solution.

302 Different tools and approaches exist for uncertainty quantification and characterization that can be  
303 potentially used in the design of safety critical systems. Each method is based on some assumptions that  
304 often cannot be verified *a priori*. Moreover, the simulation strategies are able to produce accurate results  
305 only if the right set of parameters is selected and this often cannot be verified. Finally, the numerical  
306 implementation might contain errors.

307 For these reasons, it is necessary to perform the analysis using different strategies and hypotheses in  
308 order to be able to *cross-validate* the results. Hence, different strategies implemented in a flexible and open  
309 computational framework are briefly summarized in the next sections.

## 310 A. Model updating

311 The aim of model updating is to reduce the epistemic uncertainty on the output of the model  $x = \mathcal{H}(\alpha; \theta)$   
312 based on the availability of a limited set of data (observations)  $\mathcal{D}_e := \{x_k^e : k = 1, 2, \dots, n_e\}$ . These  
313 observations of the “true uncertainty model”  $\theta^* \in \Theta$  can be used to improve the uncertainty model, i.e. to  
314 reduce the original intervals of the epistemic uncertainties by excluding those combinations of parameters  
315 that fail to describe the observations as shown in Figure 2. Two different approaches will be used for model  
316 updating: a non-parametric model based on some statistical tests and a Bayesian method.

317 Along the rest of this paper hats ( $\hat{F}$ ) and tildes ( $\tilde{F}$ ) will be used for referring to empirical CDFs and a  
 318 kernel density estimations of CDFs, respectively.

319 *1. Non-parametric statistic method based on the Kolmogorov-Smirnov test*

320 A simple and fast approach to improve the uncertainty model is based on the comparison of the CDFs of  
 321 the observations of the true uncertainty model and those obtained by means of random combinations of the  
 322 input parameters in order to identify tighter intervals which form a reduced epistemic space and which are  
 323 in agreement with the observations.

324 Let us consider the epistemic space  $\Theta$  of the involved variables. Random realizations  $\theta_i$  in the epistemic  
 325 space  $\Theta$  are generated assuming, for example, a uniform PDF on  $\Theta$  (in agreement with the Laplace's principle  
 326 of indifference). Thereafter the points  $\{\alpha_j, j = 1, 2, \dots, n\}$  are sampled from the aleatory space  $\Omega$  according  
 327 to the copula  $C$  (Nelsen<sup>15</sup> provides methods to do it), in order to simulate  $n$  observations from the system  
 328  $\mathcal{H}$  as  $x_j^i = \mathcal{H}(\alpha_j, \theta_i)$ . For a single realization  $\theta_i$ , the Kolmogorov-Smirnov statistic, which is defined as

$$D_i = \sup_x |\hat{F}(x|\theta_i) - \hat{F}_e(x)|, \quad (6)$$

329 is used to measure the similarity between the CDFs obtained with the sampled set  $\{x_j^i, j = 1, 2, \dots, n\}$  and  
 330 the set of observations  $\mathcal{D}_e$ . Here  $\hat{F}(\cdot|\theta_i)$  and  $\hat{F}_e$  are the empirical CDFs obtained using the random samples  
 331 drawn according to the epistemic parameters  $\theta_i$  and the provided experimental data, respectively.

332 The Kolmogorov-Smirnov test is used to obtain confidence limits on  $\hat{F}(\cdot|\theta_i)$  by choosing different critical  
 333 values of the test statistic  $D$ . This implies that a band of width  $\pm D$  around  $\hat{F}_e(x)$  will entirely contain  
 334  $\hat{F}(\cdot|\theta_i)$  with probability  $1 - c$ . This allows to identify those combinations of epistemic parameters such that  
 335  $P(D_i > D) = c$ .  $c = 0$  means that all the CDFs  $\hat{F}(\cdot|\theta_i)$  are accepted and the refinement of the input  
 336 intervals is not possible, whereas  $c = 1$  implies that  $\hat{F}(\cdot|\theta_i)$  comes exactly from the same model that has  
 337 generated the target distribution  $\tilde{F}_e(x)$ , i.e. no epistemic uncertainty is present.

338 The selection of  $D$  is a critical task and generally depends of the amount of available information (i.e.  
 339 number of observations). A practical approach is to use two different data sets that come from the same  
 340 process to estimate the critical level of the measure of similarity  $D_{\hat{v}}$  (using Eq. (6)). The computed validation  
 341 distance  $D_{\hat{v}}$  can be used to set the required confidence level, accepting all the combinations of epistemic  
 342 parameters with  $D_i < D_{\hat{v}}$ . When an independent validation data set is not available, a cross validation data

343 set can be constructed to test the model in order to limit problems such as overfitting. This cross validation  
 344 data set can be obtained by means of re-sampling techniques.<sup>24</sup> Cross-validation is important to protect  
 345 against hypotheses suggested by the data<sup>25</sup> specially where further samples are costly or simply impossible  
 346 to collect.

347 The non-parametric approach based on the Kolmogorov-Smirnov test is a simple and fast method for  
 348 performing uncertainty characterization (and model updating). However, it is important to keep in mind  
 349 the limitations of the approach. In fact, the method assumes that the measure of similarity  $D_i$  is distributed  
 350 according to the Kolmogorov distribution,<sup>26</sup> which is strictly true only for large sample sets. It is possible  
 351 to use some smoother techniques such as the Gaussian kernel density estimation to overcome this limitation.  
 352 Gaussian kernel density estimates for  $\mathcal{D}_e$  are given by

$$\tilde{F}_e(x) = \frac{1}{n\sigma\sqrt{2\pi}} \int_{-\infty}^x \sum_{j=1}^{n_e} \exp\left(-\frac{1}{2}\left(\frac{x' - x_j^e}{\sigma}\right)^2\right) dx'; \quad (7)$$

353 here  $\sigma$  stands for the standard deviation of the Gaussian kernels that represents the smoothing parameter,  
 354 proportional to the so-called bandwidth. Assuming  $x$  is a continuous random variable, for  $n_e \rightarrow \infty$  the  
 355 Gaussian kernel density estimate converges to the true underlying density. The support of the associated  
 356 PDFs  $\tilde{f}_e(x)$  (i.e.  $\{x : \tilde{f}_e(x) > 0\}$ ) and the bandwidth of the kernel have strong influence on the resulting  
 357 estimate. We suggest to use the approach in Ref. 27 to estimate the support of the PDF and Silverman's  
 358 rule of thumb<sup>28</sup> to estimate the bandwidth of the kernels. Using realizations from Eq. (7) the measure of  
 359 similarity can be calculated via Eq. (6) where  $\hat{F}_e(x)$  is replaced by  $\tilde{F}_e(x)$ . Please note that the Gaussian  
 360 kernels can be used to define a new critical measure level indicated with  $D_{\tilde{v}}$ .

361 To summarize, the following pseudo-algorithm is used:

- 362 1. Estimate the parameters  $\sigma$  and the Gaussian kernel CDF  $\tilde{F}_e$  using Eq. (7);
- 363 2. Estimate  $D_{\tilde{v}}$  and  $D_{\tilde{v}}$ ;
- 364 3. Generate realizations on the epistemic space,  $\boldsymbol{\theta}_i$ ;
- 365 4. Draw  $n$  points from the aleatory space  $\boldsymbol{\Omega}$ , using copula  $C$ ; we will call these samples  $\{\boldsymbol{\alpha}_j : j = 1, \dots, n\}$ ;
- 366 5. Evaluate the model  $x_j^i := \mathcal{H}(\boldsymbol{\alpha}_j; \boldsymbol{\theta}_i)$  for  $j = 1, \dots, n$ ;

- 367 6. Estimate the empirical CDF  $\hat{F}(\cdot|\theta_i)$  of the set of samples  $\{x_j^i, j = 1, 2, \dots, n\}$ ;
- 368 7. Using Eq. (6), compute the measure of similarity  $D_i$ ;
- 369 8. If  $D_i < D_{\bar{v}}$  (or  $D_i < D_{\bar{v}}$ ) collect  $\theta_i$ . The set of collected points identify a reduced space in the original
- 370 epistemic space.

371 2. *Bayesian updating on the epistemic space*

372 Bayesian inference is a statistical method in which the Bayes' rule is used to update the probability estimate

373 for a hypothesis as additional information is available.

374 Suppose we are given a set of observed data points  $\mathcal{D}_e := \{x_k^e : k = 1, 2, \dots, n_e\}$  called the *evidence*, and

375 which are sampled from a PDF  $p(\cdot; \theta^*)$  which belongs to a certain family of PDFs  $\{p(\cdot; \theta) : \theta \in \Theta\}$  called

376 the *parametric model*. The idea of Bayesian inference is to update our belief about the vector of parameters

377  $\theta$  provided that  $\theta^*$ , the true set of parameters of the PDF, is unknown. Bayes' theorem updates that belief

378 using two antecedents:

- 379 • a *prior* PDF  $p(\theta)$ , which indicates all available knowledge about  $\theta^*$  before the evidence  $\mathcal{D}_e$  is observed;
- 380 • and the *likelihood function*  $P(\mathcal{D}_e|\theta)$ , which is a function related to the probability of observing the
- 381 samples  $\mathcal{D}_e$  assuming that the true parameter underlying the model PDF  $p(x; \theta)$  is  $\theta$ ; it is defined as

$$P(\mathcal{D}_e|\theta) = \prod_{k=1}^{n_e} p(x_k^e; \theta), \quad (8)$$

382 when a set of independent and identically distributed observations  $\mathcal{D}_e$  is available. Please note that

383 in practice (i.e. for the numerical implementation) the log-likelihood is used instead of the likelihood.

384 The updated belief about the vector of parameters  $\theta$  after observing the evidence  $\mathcal{D}_e$ , is modelled by the

385 so-called *posterior* PDF  $p(\theta|\mathcal{D}_e)$  which is calculated by:

$$p(\theta|\mathcal{D}_e) = \frac{P(\mathcal{D}_e|\theta)p(\theta)}{P(\mathcal{D}_e)}; \quad (9)$$

386 where the probability of the evidence,

$$P(\mathcal{D}_e) = \int_{\Theta} P(\mathcal{D}_e|\theta)p(\theta) d\theta \quad (10)$$

387 can be understood as a normalizing constant. Bayesian updating hopes that after using the evidence  $\mathcal{D}_e$   
 388 the posterior PDF  $p(\boldsymbol{\theta}|\mathcal{D}_e)$  is sharply peaked about the true value of  $\boldsymbol{\theta}^*$ . We will update our belief about the  
 389 true set of parameters  $\boldsymbol{\theta}^* \in \Theta$  propagating the evidence through the Bayes' equation numerically. Samples  
 390 of the posterior PDF can be generated without the necessity to evaluate  $p(\boldsymbol{\theta}|\mathcal{D}_e)$ , using an algorithm called  
 391 Transitional Markov Chain Monte Carlo (TMCMC).<sup>29</sup>

392 As the prior PDF, we will use a uniform distribution on the epistemic space  $\Theta$ , that is  $\boldsymbol{\theta} \sim \text{Unif}(\Theta)$ , in  
 393 accordance to the Laplace's principle of indifference (or more generally, the principle of maximum entropy).

394 Different likelihood functions can be used, based on different mathematical assumptions; in the following  
 395 two methods will be proposed: a method that used a kernel density estimator to approximate  $p(\cdot|\boldsymbol{\theta}_i)$  and a  
 396 approximate Bayesian computational method.

397 **BAYESIAN COMPUTATIONAL METHOD** In this case, the likelihood is estimated through kernel density. As-  
 398 suming that the samples  $\mathcal{D}_e$  were drawn from  $p(x; \boldsymbol{\theta}_i)$ , the likelihood  $P(\mathcal{D}_e|\boldsymbol{\theta}_i)$  is defined in the following  
 399 way:

- 400 1. Draw  $n$  points,  $\{\boldsymbol{\alpha}_j : j = 1, \dots, n\}$ , from the aleatory space  $\Omega$ , using copula  $C$ ;
- 401 2. Calculate  $x_j^i := \mathcal{H}(\boldsymbol{\alpha}_j; \boldsymbol{\theta}_i)$  for  $j = 1, \dots, n$ ;
- 402 3. Using kernel density estimation and the samples  $\{x_j^i : j = 1, \dots, n\}$ , estimate the CDF  $\tilde{F}(\cdot|\boldsymbol{\theta}_i)$  and  
 403 its associated PDF  $\tilde{p}(x|\boldsymbol{\theta}_i) \equiv p(x; \boldsymbol{\theta}_i)$ . This step is required because  $\tilde{p}(x|\boldsymbol{\theta}_i)$  cannot be obtained  
 404 analytically;
- 405 4. Calculate the likelihood function  $P(\mathcal{D}_e|\boldsymbol{\theta}_i)$  as in Eq. (8).

406 **APPROXIMATE BAYESIAN COMPUTATIONAL METHOD** The likelihood calculated by means of the "Bayesian  
 407 computational method" applies Bayes' theorem directly and without strong assumptions. However it requires  
 408 a large number of model evaluations and a relative larger data set to converge.<sup>30</sup> Recently, approximate  
 409 Bayesian computational methods have been proposed to reduce the computational costs of the expensive  
 410 or intractable likelihood function.<sup>31,32</sup> The likelihood can be for instance approximated with the following  
 411 expression:

$$P(\mathcal{D}_e|\boldsymbol{\theta}_i) = \prod_{k=1}^{n_e} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{\delta_k}{\sigma}\right)^2\right) \quad (11)$$

412 where  $\delta_k$  is the absolute value of the difference between the empirical CDF  $\hat{F}(\cdot|\boldsymbol{\theta}_i)$  obtained for an individual  
 413 realization  $\boldsymbol{\theta}_i$  of the epistemic space  $\Theta$ , evaluated at each point  $\{x_k^e, k = 1, 2, \dots, n_e\}$  and the empirical CDFs  
 414 of the experimental dataset  $\mathcal{D}_e$ , that is:

$$\delta_k = \left| \hat{F}(x_k|\boldsymbol{\theta}_i) - \hat{F}_e(x_k^e) \right| \quad (12)$$

415 for  $k = 1, 2, \dots, n_e$ . Please note that the Bayesian updating approach is generally applied to identify a fixed  
 416 estimate of  $\boldsymbol{\theta}$  as close as possible to  $\boldsymbol{\theta}^*$ . Here, the approach has been used to identify a reduced epistemic  
 417 space containing the true values of the unknown parameters. If a constant  $\sigma$  is used, the Bayesian updating  
 418 formulation, here introduced, is equivalent to a minimization in the least square sense of the distance between  
 419 the CDFs  $\hat{F}(\cdot|\boldsymbol{\theta}_i)$  and  $\hat{F}_e$ . However the value of  $\sigma$  is unknown and hence it represents an additional parameter  
 420 that needs to be estimated.<sup>33</sup>

421 This last approach is indeed based more on practical considerations than on a sound mathematical basis,  
 422 and is open to criticisms since the differences  $\delta_k$  are assumed to be independent and normally distributed  
 423 with zero mean and unit variance, and that even though  $\delta_k$  is normally distributed, it will only take values  
 424 in the interval  $[0, 1]$  since the CDF ranges between 0 and 1.

425 Using the above defined prior PDF and likelihood functions, the TMCMC algorithm<sup>29</sup> is employed in  
 426 order to find samples of the posterior  $p(\boldsymbol{\theta}|\mathcal{D}_e)$ . The likelihood  $P(\mathcal{D}_e|\boldsymbol{\theta}_i)$  is calculated, using the approximate  
 427 Bayesian computational method, by the following procedure:

- 428 1. Draw  $n$  points ( $\{\boldsymbol{\alpha}_j : j = 1, \dots, n\}$ ) from the aleatory space  $\Omega$ , using copula  $C$ ;
- 429 2. Calculate  $x_j^i := \mathcal{H}(\boldsymbol{\alpha}_j; \boldsymbol{\theta}_i)$  for  $j = 1, \dots, n$ ;
- 430 3. Using the samples  $\{x_j^i : j = 1, \dots, n\}$ , estimate the empirical CDF  $\hat{F}(\cdot|\boldsymbol{\theta}_i)$ ;
- 431 4. Compute  $\delta_k = \left| \hat{F}(x_k^e|\boldsymbol{\theta}_i) - \hat{F}_e(x_k^e) \right|$  at each point  $x_k^e \in \mathcal{D}_e$ ;
- 432 5. Calculate the likelihood function  $P(\mathcal{D}_e|\boldsymbol{\theta}_i)$  as in Eq. (11).

## 433 B. Sensitivity analysis

434 The aim of sensitivity analysis is to identify and rank the parameters that contribute mostly to the variability  
 435 of the output of a system  $\mathcal{H}$ . Two approaches can be used: the Hartley-like measure of nonspecificity and

436 the global sensitivity analysis based on Sobol' and total sensitivity measures. Both approaches can be used  
 437 to perform global sensitivity analysis in presence of epistemic uncertainty.

### 438 1. Nonspecificity technique

439 Before delving into this method of sensitivity analysis, a small introduction to the nonspecificity measure is  
 440 presented.

441 The nonspecificity, proposed by Klir and coworkers,<sup>34,35</sup> is a measure of the amount of information  
 442 required to remove the epistemic uncertainty; it is used in cases when we have to select a unique element  
 443 from a set, but we are totally indifferent about which element of the provided ones to choose.

444 The nonspecificity is based on the so called *Hartley-like measure*, which for a  $d$ -dimensional box (or focal  
 445 element)  $A = \times_{i=1}^d [l_i, u_i]$ , like the ones that we are considering in this paper is given by:

$$\text{HL}(A) = \log_2 \left( \prod_{i=1}^d (1 + u_i - l_i) \right). \quad (13)$$

446 The nonspecificity of a random set with an infinite number of focal elements is given by (see Ref. 36):

$$\text{HL}((\mathcal{F}, P_{\Gamma})) = \int_{\Omega} \text{HL}(\Gamma(\boldsymbol{\alpha})) dC(\boldsymbol{\alpha}); \quad (14)$$

447 two special cases of Eq. (14) are:

- 448 • the nonspecificity of a Dempster-Shafer structure  $(\mathcal{F}_n, m)$  with focal set  $\mathcal{F}_n = \{A_1, \dots, A_n\}$  and basic  
 449 mass assignment  $m$ :

$$\text{HL}((\mathcal{F}_n, m)) = \sum_{i=1}^n \text{HL}(A_i) m(A_i); \quad (15)$$

- 450 • the nonspecificity of a distribution-free probability box  $\langle \underline{E}, \overline{F} \rangle$ :

$$\text{HL}(\langle \underline{E}, \overline{F} \rangle) = \int_0^1 \log_2 \left( 1 + \underline{E}^{-1}(\alpha) - \overline{F}^{-1}(\alpha) \right) d\alpha. \quad (16)$$

451 The nonspecificity is a measure of epistemic uncertainty, and in consequence, it is useful for assessing the  
 452 variability of the output due to the epistemic uncertainty in the input of the model.

453 The method, which is detailed in Ref. 37, calculates a Dempster-Shafer structure that is the result of  
 454 propagating the epistemic uncertainty through the system  $\mathcal{H}$  (using the extension principle for random sets).  
 455 Then, the Hartley-like measure of nonspecificity of that output Dempster-Shafer structure is evaluated. More  
 456 specifically:

457 1.  $n$  samples  $\alpha_i \in \Omega$  are drawn from copula  $C$ . Thereafter, the image of the focal element  $\alpha_i$  through  
 458  $\mathcal{H}$  is calculated by means of Eq. (5) as  $[\min_{\theta \in \Theta} \mathcal{H}(\alpha_i, \theta), \max_{\theta \in \Theta} \mathcal{H}(\alpha_i, \theta)]$ . This set of  $n$  focal  
 459 elements is used to construct a Dempster-Shafer structure  $(\mathcal{F}_n, m)$  (as explained in Section II-C)  
 460 which represents the propagation of the aleatory and epistemic uncertainty through the system. The  
 461 basic mass assignment  $m$  of each focal element is  $1/n$ .

462 2. Compute  $\text{HL}_{\text{unpinched}} = \text{HL}((\mathcal{F}_n, m))$  according to Eq. (15).

463 3. Consider a set of points  $\{p_r : 0 \leq p_1 < p_2 < \dots < p_r < \dots < p_z \leq 1\}$  which are evenly distributed in  
 464 the interval  $[0, 1]$ .

465 4. For each point  $p_r$ , do the following:

466 (a) Each interval,  $I_j = [l_j, u_j]$  that forms the epistemic space  $\Theta$ , is pinched (or reduced) to the value  
 467 given by  $l_j + p_r \cdot (u_j - l_j)$ , while leaving all other intervals unchanged (take into account that  
 468 pinching of groups of input variables can be performed as well). After pinching the  $j$ -th input  
 469 variable, a subset of  $\Theta$ , namely  $\Theta_{rj}$ , which includes the pinched inputs is formed.

470 (b)  $n$  samples  $\alpha_i \in \Omega$  are drawn from copula  $C$ ; thereafter, the image of the focal element  $\alpha_i$  through  
 471  $\mathcal{H}$  is calculated by means of Eq. (5) as  $[\min_{\theta \in \Theta_{rj}} \mathcal{H}(\alpha_i, \theta), \max_{\theta \in \Theta_{rj}} \mathcal{H}(\alpha_i, \theta)]$ . This set of  $n$   
 472 focal elements is used to construct a Dempster-Shafer structure  $(\mathcal{F}_n^{rj}, m)$  for each pinching (as  
 473 explained in Section II). The basic mass assignment  $m$  of each focal element is  $1/n$ .

474 (c) The nonspecificity  $\text{HL}_{rj} := \text{HL}((\mathcal{F}_n^{rj}, m))$  of each of those output Dempster-Shafer structures is  
 475 computed, as in Eq. (15).

476 (d) The nonspecificity measure of the output Dempster-Shafer structure is normalized against the  
 477 nonspecificity measure computed before pinching. In this way, the index  $I_{rj} = \text{HL}_{rj} / \text{HL}_{\text{unpinched}}$   
 478 is calculated.

479 5. The index  $I_j$  is calculated as the mean square of all indexes  $I_{rj}$  (remember that  $\text{E}[I_j^2] = \text{E}[I_j]^2 + \text{Var}[I_j]$ );  
 480 those indexes are ranked according to their mean square.

481  $I_j$  is used as a measure of the propagation of the epistemic uncertainty to the output of the system  $\mathcal{H}$ .  
 482 The smaller  $I_j$  is, the larger is the sensitivity of the system to the epistemic uncertainty in the input variable

483 *j*. The mean squared has been chosen as a ranking criterion in order to account for not only the bias but also  
 484 the variance of the estimator. Note that a precise estimation of  $I_j$  is not necessary, since only the ranking  
 485 of the variables  $I_j$  is required; therefore,  $n$  is usually a small number.

## 486 2. Global Sensitivity analysis

487 The second approach is based on global sensitivity analysis to estimate the Sobol' and the total indices.<sup>38</sup>  
 488 The global sensitivity approach cannot be applied directly to solve the problems where the uncertainty is  
 489 described as a distributional/free p-boxes and intervals. In fact, this method requires the exact knowledge  
 490 of the PDF of the input variables and the variance of a measurable model output. In consequence, an  
 491 alternative mathematical model to  $\mathcal{H}$  (as defined in Section II-F) has to be defined in the next.

492 Consider a model  $\mathcal{H}^* : \Theta \rightarrow \mathbb{R}$ , that is,  $Y = \mathcal{H}^*(\theta)$ , where  $\theta = [\theta_1, \dots, \theta_q]$  is a vector of random variables  
 493 and  $Y$  is a chosen univariate model output.

494 Let us associate  $\Theta$  with the epistemic space; for a given value of  $\theta_j \in \Theta$ , the function  $\mathcal{H}^*$  returns the  
 495 area between a CDF  $F(\cdot|\theta_j)$  and a reference CDF  $F(\cdot|\bar{\theta})$ :

$$y_i := \mathcal{H}^*(\theta_i) := \int_{-\infty}^{+\infty} |F(x|\theta_i) - F(x|\bar{\theta})| dx; \quad (17)$$

496 here  $\bar{\theta}$  denotes the center of gravity of  $\Theta$  (in other words,  $\bar{\theta}$  is a vector formed by the mean value of each  
 497 input epistemic parameter), and  $F(\cdot|\theta)$  represents the CDF obtained after mapping all aleatory uncertainty  
 498 through the system  $\mathcal{H}$ , for a given set of epistemic parameters  $\theta$ . Since the global sensitivity analysis is  
 499 based on the variance decomposition, any reference CDF can be used in the model  $\mathcal{H}^*$ .

500 The procedure to estimate the empirical CDF  $\hat{F}(\cdot|\theta)$  as an approximation to  $F(\cdot|\theta)$  is as follows:

- 501 1. Draw  $n$  points ( $\{\alpha_j : j = 1, \dots, n\}$ ) from the aleatory space  $\Omega$ , using copula  $C$ ;
- 502 2. Evaluate the model  $x_j^i := \mathcal{H}(\alpha_j; \theta)$  for  $j = 1, \dots, n$ ;
- 503 3. Estimate the empirical CDF  $\hat{F}(\cdot|\theta)$  of the set of samples  $\{x_j^i, j = 1, 2, \dots, n\}$ .

504 Using the above procedure, a sample from the random variable  $Y$ , namely  $y_i$  can be estimated by means  
 505 of Eq. (17) for each realization of input  $\theta_i$  by using the empirical CDFs  $\hat{F}(\cdot|\theta)$  and  $\hat{F}(\cdot|\bar{\theta})$ . Please note that  
 506 when the model produces a scalar value for each realization of the input  $\Theta_i$ , e.g. it returns the expected

507 value ( $y_i = E(x|\theta_i)$ ) or a quantile of a distribution, it is not necessary to evaluate Eq. (17) but the model  
508 output can be used directly (see Section B-2).

509 Finally, the first order *Sobol' indices* are calculated as follows<sup>39</sup>

$$S_i = \frac{\text{Var}_{\theta_i}[\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i)]}{\text{Var}[Y]} \quad (18)$$

510 where  $\text{Var}[Y]$  represents the unconditional variance of the quantity of interest and  $\text{Var}_{\theta_i}[\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i)]$  the  
511 variance of conditional expectation. The *total sensitivity index*,  $T_i$ , measures the contribution to the output  
512 variance of  $\theta_i$  of the input factors including all interactions with any other input variables,

$$T_i = 1 - \frac{\text{Var}_{\theta_{\sim i}}(\mathbb{E}_{\theta_i}(Y | \theta_{\sim i}))}{\text{Var}(Y)}. \quad (19)$$

513 Note that unlike the first order indices, the sum of total indices can exceed one.

514 The proposed approach allows to decompose the variance of the output  $Y$  into parts attributable to  
515 the variance of the input variables  $\theta$ ; in other words, it allows to identify and rank the contribution of the  
516 epistemic uncertainty, i.e. interval of the parameters, on the p-boxes of quantity of interest. The magnitude  
517 of the sensitivity indices are proportional to the contribution to the output variance, i.e. input factor  
518 associated with a large sensitivity index contributes most to the variance of the output. Hence, adopting the  
519 approach proposed here, the global sensitivity analysis allows us to identify the contribution of the epistemic  
520 uncertainty of input factors on the variance of the model.

521 Different techniques exist to compute the sensitivity indices such as the extended-“Fourier Amplitude  
522 Sensitivity Test” (FAST)<sup>40,41</sup> and the Saltelli’s method.<sup>38</sup> The FAST method allows to estimate first order  
523 Sobol’ indices, whereas Saltelli’s method computes also the total indices.

### 524 C. Uncertainty Propagation

525 The focus of the uncertainty propagation analysis is to quantify the effect of the uncertain model parameters  
526 on quantities of interest such as the mean, variance and quantiles of the system’s response or its failure prob-  
527 ability. The generalized probabilistic model makes the UQ rather challenging task in terms of computational  
528 cost. The challenge is to compute the lower and upper bounds of the quantities of interest. Monte Carlo  
529 method remains the most versatile and simple tool to propagate epistemic and aleatory uncertainty.

530 1. *Optimization in the epistemic space (standard approach)*

531 In this approach, the quantity of interest (e.g. mean or failure probability estimation) defines the objective  
 532 function; and the bounds on that objective function are calculated by means of a global search in the  
 533 epistemic space  $\Theta$ . On one hand, the lower and upper bounds of the mean are obtained as:

$$\underline{\mu} = \min_{\theta \in \Theta} \mu(\theta) \qquad \bar{\mu} = \max_{\theta \in \Theta} \mu(\theta) \qquad (20)$$

534 where the mean of the response model is given by:

$$\mu(\theta) = \int_{\Omega} \mathcal{H}(\alpha; \theta) dC(\alpha). \qquad (21)$$

535 On the other hand, the lower and upper bound of the failure probability, defined as the exceedance of a  
 536 critical threshold level  $\mathcal{H}^{crit}$  of the model response, are obtained as

$$\underline{P}_f = \min_{\theta \in \Theta} P_f(\theta) \qquad \bar{P}_f = \max_{\theta \in \Theta} P_f(\theta); \qquad (22)$$

537 here  $P_f(\theta)$  stands for the failure probability, that is,

$$P_f(\theta) := \int_{\Omega} \mathcal{I}[\mathcal{H}(\alpha; \theta) > \mathcal{H}^{crit}] dC(\alpha). \qquad (23)$$

538

539 Monte Carlo method is used to calculate the bounds Eq. (20) and Eq. (22), by means of a double loop  
 540 simulation:

- 541 • The outer loop drives an optimization/search process in the epistemic space  $\Theta$  to identify the lower  
 542 and upper bounds Eq. (20) and Eq. (22). This search is performed by Monte Carlo sampling taking  
 543 into account that this optimization method is very inefficient in high dimensional spaces since the  
 544 search space grows exponentially with the number of variables. Better optimization strategies such as  
 545 Genetic Algorithms can also be adopted as shown in Section V-C.
- 546 • The inner loop propagates the aleatory uncertainty and estimates the statistical quantities of interest  
 547 (e.g. expected value, failure of probability, CDF, etc). In this way, several  $\alpha_j$  are sampled from copula  
 548  $C$  in order to estimate integrals of Eq. (21) and Eq. (23). Take into account that this Monte Carlo  
 549 integration in the aleatory space  $\Omega$  is insensitive to the dimensionality of the problem although it can

550 be inefficient in case of the calculation of integral of Eq. (23), when the probability of failure is very  
 551 small. The estimation of the integrals can be speed up by adopting the so called Advanced Monte  
 552 Carlo methods such as Importance Sampling, Subset Simulation and Line Sampling.<sup>19</sup>

553 2. Propagation of focal sets (counter approach)

554 The second approach for uncertainty propagation, which is proposed in Refs. 11, 12, 42, is based on the  
 555 propagation of focal sets through a function. Using random set theory, as explained in Section II, it can  
 556 be seen that the aleatory space  $\Omega$  contains the regions  $F_{LP} := \{\alpha \in \Omega : \Gamma(\alpha) \subseteq F, \Gamma(\alpha) \neq \emptyset\}$  and  
 557  $F_{UP} := \{\alpha \in \Omega : \Gamma(\alpha) \cap F \neq \emptyset\}$  which are correspondingly formed by all those points whose respective  
 558 focal elements are completely contained in the failure set  $F = \{\mathbf{x} \in \mathcal{X} : g(\mathbf{x}) > \mathcal{H}^{crit}\}$  or have in common  
 559 at least one point with  $F$  correspondingly (see Figure 1b). Notice that the set  $F$  is defined in the space of  
 560 input variables  $\mathcal{X}$ ; in this case, the lower Eq. (2a) and upper Eq. (2b) probability measures of  $F$  can be  
 561 calculated by:

$$\underline{\underline{P}}_f = LP_{(\mathcal{F}, P_T)}(F) = \int_{\Omega} \mathcal{I}[\alpha \in F_{LP}] dC(\alpha) \quad \overline{\overline{P}}_f = UP_{(\mathcal{F}, P_T)}(F) = \int_{\Omega} \mathcal{I}[\alpha \in F_{UP}] dC(\alpha) \quad (24)$$

562 provided that  $F_{LP}$  and  $F_{UP}$  are  $\mu_C$ -measurable sets; here  $\mathcal{I}$  stands for the indicator function.

563 Eq. (24) can be evaluated by means of simple Monte Carlo method sampling  $n$  points from the copula  $C$ ,  
 564 namely  $\alpha_1, \alpha_2, \dots, \alpha_n \in \Omega$ , and then retrieving the corresponding focal elements  $\gamma_j := \Gamma(\alpha_j), j = 1, \dots, n$   
 565 from  $\mathcal{F}$ . Afterwards, integrals Eq. (24) are computed by the unbiased estimators  $\underline{\underline{\hat{P}}}_f$  and  $\overline{\overline{\hat{P}}}_f$ , which are  
 566 given by:

$$\underline{\underline{\hat{P}}}_f = \frac{1}{n} \sum_{j=1}^n \mathcal{I}[\alpha_j \in F_{LP}] \quad \overline{\overline{\hat{P}}}_f = \frac{1}{n} \sum_{j=1}^n \mathcal{I}[\alpha_j \in F_{UP}]. \quad (25)$$

567

568 The image of  $\Gamma(\alpha_i)$  through the function  $\mathcal{G}$  can be computed using the optimization method, as described  
 569 by equations (3) and (4). Since,  $\mathcal{I}[\mathcal{G}(\Gamma(\alpha_i)) \subseteq F] = \mathcal{I}[\overline{\mathcal{G}}(\alpha_i) > \mathcal{H}^{crit}] = \mathcal{I}[\alpha_i \in F_{LP}]$  and  $\mathcal{I}[\mathcal{G}(\Gamma(\alpha_i)) \cap F \neq$   
 570  $\emptyset] = \mathcal{I}[\underline{\mathcal{G}}(\alpha_i) > \mathcal{H}^{crit}] = \mathcal{I}[\alpha_i \in F_{UP}]$  it follows that Eqs. (25) can be written as:

$$\underline{\underline{\hat{P}}}_f = \frac{1}{n} \sum_{i=1}^n \mathcal{I}[\overline{\mathcal{G}}(\alpha_i) > \mathcal{H}^{crit}] \quad \overline{\overline{\hat{P}}}_f = \frac{1}{n} \sum_{i=1}^n \mathcal{I}[\underline{\mathcal{G}}(\alpha_i) > \mathcal{H}^{crit}]. \quad (26)$$

571

572 Observe that this approach operates by inverting the order of execution of the loops in the double loop  
573 described above:

- 574 • the outer loop propagates the aleatory uncertainty by sampling the points  $\alpha_1, \alpha_2, \dots, \alpha_n \in \Omega$  using  
575 copula  $C$ .
- 576 • the inner loop drives an optimization/search process in  $\Gamma(\alpha_i)$  in order to find the image of the input  
577 focal element through the system  $\mathcal{G}$ ; this step is performed when evaluating Eqs. (3) and (4).

578 One of the main advantages of the random set theory is that, for a problem where inputs are defined using  
579 any possible imprecise probability framework (CDFs, intervals, distribution-free probability boxes, possibility  
580 distributions, Dempster-Shafer structures, etc.), it allows to employ the methods developed by the community  
581 of stochastic mechanics for estimating the failure probabilities of the two limit state functions  $\underline{\mathcal{G}}$  and  $\overline{\mathcal{G}}$ , i.e.  
582 calculating of bounds on probability  $[\underline{\hat{P}}_f, \overline{\hat{P}}_f]$ . In case that the calculation of very small probability bounds is  
583 requested, the plain Monte Carlo simulation described here is not efficient. Advanced Monte Carlo methods  
584 can be used to estimate small probabilities of failure as described in e.g. Ref. 43,19.

585 It is worth noting that although the random set theory is in general not applicable in the case of dis-  
586 tributional p-boxes, the method presented in Section C-1 can still be used as far as the bounding CDFs of  
587 the input p-boxes can be identified. However, applying this approach to distributional p-boxes treats those  
588 p-boxes as distribution-free ones. This inevitably leads to loss of information which results in the underes-  
589 timation and overestimation of the lower and upper bounds respectively, when compared to the method of  
590 optimization in the epistemic space (standard approach).

### 591 3. Numerical considerations

592 Two degrees of error can be identified using both approaches for UQ. The first error concerns the estimation  
593 of the statistics and failure probability, which can be reduced by increasing the number of samples or by  
594 implementing an efficient sampling technique. 43,19 In reliability analysis, the limited set of samples may  
595 lead to both an underestimation and to an overestimation. The confidence of the estimator can always be  
596 improved adopting a larger set of samples but at the cost of increasing the computational demand.

597 The second error concerns the global search. In general it is not possible to guarantee the identification  
598 of global optima. Only when the feasible (search) domain of the input variables is small ( $\approx 5$  variables), a

599 thorough search can lead to a good approximation of the global optima. The search error can only affect the  
600 results in one direction. For example, if a global minimum is searched, the identified minimum can only be  
601 greater than the global one; in the same way, the identified maximum can only be smaller than the global  
602 maximum.

603 Under the assumption that the sampling error for estimating the failure probability is very small, the  
604 “optimization in the epistemic space” approach (Section C-1) always results in an overestimation of the lower  
605 bound and an underestimation of the upper bound, which may lead to an optimistic decision.

#### 606 D. Extreme case analysis

607 The extreme case analysis consists in identifying the the combinations of epistemic realizations  $\theta$  that leads  
608 to the worst/best behaviour of the system. This analysis can be seen as an inverse problem of the uncertainty  
609 propagation, the forward problem, described in Section C.

610 This problem is a by-product of the uncertainty propagation but the ability to solve it depends on the  
611 approach used to perform the forward problem. The extreme case analysis can not be performed using  
612 the approach “propagation of focal sets” presented in Section C. This is because distributional p-boxes are  
613 treated as distribution-free p-boxes. Hence, extreme cases might result associated with distributions that  
614 lay inside the p-boxes but that do not comply with the associated parental distributions. Only the approach  
615 “optimization of the epistemic space” can be used because the approach holds a bijective mapping between  
616 the inputs in the epistemic domain and the quantity of interest.

617 Solving Eqs. (20) and (23)) it is possible to identify directly realizations of the epistemic space  $\theta$  that  
618 produce the bounds of quantity of interest, as

$$\theta_{\underline{\mu}} = \arg \min_{\theta \in \Theta} \mu(\theta) \qquad \theta_{\overline{\mu}} = \arg \max_{\theta \in \Theta} \mu(\theta) \qquad (27)$$

$$\theta_{\underline{P}_f} = \arg \min_{\theta \in \Theta} P_f(\theta) \qquad \theta_{\overline{P}_f} = \arg \max_{\theta \in \Theta} P_f(\theta). \qquad (28)$$

619 Unfortunately, the uncertainty makes the inverse problem an ill-posed and difficult to solve problem. For  
620 instance, the objective of the optimization can involve the calculation of some statistics. These are generally  
621 estimated by means of samples and those statistics are not exact but approximate. Stochastic optimization  
622 methods<sup>44</sup> are specially suited to make optimization with random objective functions.

623 In addition, the necessity to separate epistemic and aleatory uncertainty makes the extreme case anal-  
 624 ysis even more difficult. This is because an extreme case can derive from different combinations of epis-  
 625 temic/aleatory uncertainty  $(\boldsymbol{\alpha}, \boldsymbol{\theta})$ .

## 626 E. Robust Design

627 The final task in the design of a safety critical system is to perform a robust design optimization. The main  
 628 aim of the robust design is to consider explicitly the effects of the uncertainties in the optimization problem.  
 629 A solution of this problem can be obtained by performing an optimization analysis able to identify the design  
 630 point with improved robustness and reliability characteristics.<sup>9</sup>

631 This requires to repeatedly evaluate the performance of the system that can be defined as e.g. expected  
 632 values, probability of failure. The approach described in Section C can be adopted for the estimation of  
 633 these quantities (inner loop) and it generally requires considerable numerical efforts. In addition, it has to  
 634 be performed for each candidate solutions of the optimization procedure (the outer loop).

635 Generally in robust design only one bound is of interest. For instance we would like to reduce the  
 636 probability of failure. In this sense, the optimal design point  $\mathbf{d}^{opt}$ , would be given for example by:

$$\mathbf{d}^{opt} = \arg \min_{\mathbf{d} \in \mathbf{D}} \overline{P}_f(\mathbf{d}) = \arg \min_{\mathbf{d} \in \mathbf{D}} \max_{\boldsymbol{\theta} \in \Theta} \int_{\Omega} \mathcal{I}[\mathcal{H}(\boldsymbol{\alpha}; \boldsymbol{\theta}; \mathbf{d}) > \mathcal{H}^{crit}] dC(\boldsymbol{\alpha}) \quad (29)$$

637 or by

$$\mathbf{d}^{opt} = \arg \min_{\mathbf{d} \in \mathbf{D}} \overline{\mu}(\mathbf{d}) = \arg \min_{\mathbf{d} \in \mathbf{D}} \max_{\boldsymbol{\theta} \in \Theta} \int_{\Omega} \mathcal{H}(\boldsymbol{\alpha}; \boldsymbol{\theta}; \mathbf{d}) dC(\boldsymbol{\alpha}), \quad (30)$$

638 where  $\mathbf{D}$  is the design space and  $\mathcal{H}$  becomes a function  $\Omega \times \Theta \times \mathbf{D} \rightarrow \mathbb{R}$ .

639 Nevertheless, the estimation of bounds of the system performance for this subproblem, remains a compu-  
 640 tational challenge. Thus, due to the tremendous numerical cost involved, caused by the repeated assessment  
 641 of the system response for different candidate solutions, the direct solution of this subproblem may render  
 642 the computational task unfeasible, even for academic problems. Then, it is necessary to resort to specific  
 643 techniques such as the use of surrogate models in order to decrease the computational costs.<sup>45,46</sup> Surrogate-  
 644 models mimic the behaviour of the original model, by means of an analytical expression with negligible  
 645 computational cost. The approximation is constructed by selecting some predefined interpolation points in  
 646 the design space, at which the maximal failure probability is estimated; then, a surrogate model is adjusted

647 to the data collected in a least squares sense. As the construction of this approximation over the entire  
648 domain can be demanding, it may be easier to generate an approximation of the failure probabilities over  
649 a sub-domain,<sup>47</sup> i.e. to generate a local surrogate model. Local surrogate model might require generally  
650 less evaluation points to be constructed although they have to be continuously updated in order to follow  
651 the current values of the design variables. Artificial neural networks are very versatile surrogate models;  
652 other methods such as kriging can be used as well.<sup>45,46</sup> Surrogate models should not introduce unnecessary  
653 approximations and errors. Hence, only the most computational expensive part of the model should be  
654 replaced keeping the original models for the less demanding parts.

## 655 IV. Numerical Implementation

656 The uncertainty quantification and management require the availability of flexible numerical tools able  
657 to deal with the different representations of uncertainty. Furthermore, since the non-deterministic analysis is  
658 computationally quite demanding, such numerical tools need to be very efficient and scalable. In fact, since  
659 such analyses need to be repeated a large number of times, the computational cost could be excessive even  
660 when the solver is reasonably fast (e.g. the computation of  $g$  in the challenge problem requires 2 seconds on  
661 a common desktop computer). For these reasons, the proposed approach has been developed and integrated  
662 into the OPENCOSSAN framework.<sup>10</sup>

### 663 A. OpenCossan

664 OPENCOSSAN is a collection of open source algorithms, methods and tools released under the LGPL licence,<sup>48</sup>  
665 and under continuous development at the Institute for Risk and Uncertainty at the University of Liverpool,  
666 UK. The source code is available upon request at the web address <http://www.cossan.co.uk>.

667 OPENCOSSAN is also the computational core of a general purpose software, namely COSSAN-X, originally  
668 developed by the research group of Prof. G.I. Schuëller at the University of Innsbruck, Austria.<sup>49,50</sup> As a  
669 general purpose software, it means that a reasonably wide range of engineering and scientific problems can  
670 be treated by the software.

671 This computational core, developed in MATLAB<sup>®</sup> using an object-oriented programming paradigm,  
672 includes several predefined solution sequences to solve a number of different problems. The framework is

673 organized in classes, i.e. data structures consisting of data fields and methods together with their interac-  
674 tions and interfaces. Thanks to the modular nature of `OPENCROSSAN`, it is possible to define specialized  
675 solution sequences including reliability methods, optimization strategies and surrogate modelling or parallel  
676 computing strategies to reduce the overall cost of the computation.

677 `OPENCROSSAN` provides intuitive, clear, well documented and human readable interfaces to the classes.  
678 Furthermore, the developed numerical methods are highly scalable and parallelizable, thanks to its integra-  
679 tion with distributed resource management, such as `openlava` and `GridEngine`. These job management tools  
680 allow to take advantages of high performance computing, as shown in the next sub-session.

## 681 **B. High performance computing**

682 The proposed strategies for solving the challenge problem are generally very demanding in terms of compu-  
683 tational resources. For instance, sensitivity analysis and uncertainty quantification might require more than  
684  $10^6$  up to  $10^9$  model evaluations (see e.g. Section [V-B](#))).

685 Even though the computational cost to evaluate the model might be low, the huge number of model  
686 evaluations required by the analyses represents a computational challenge. A possible way to reduce the  
687 execution time of the analyses is to employ surrogate models to approximate the input/output relations  
688 with faster analytical approximations. This, however, introduces loss of accuracy in the analysis, and such  
689 surrogate models have to be accurately calibrated before being employed in the analysis.

690 Alternatively, multiple independent instances of the solver can be executed simultaneously for different  
691 values of the input to the system, allowing for a reduction of the analyses time without any loss of accuracy.

692 Hence, in order to reduce the computational wall-clock time required by the analyses two types of  
693 parallelization can be used. The first type of parallelization is used to speed-up the analysis of most internal  
694 loop required by the simulations. In this case, a special job on a pool of MATLAB workers is created on  
695 each multi-core machine, connecting the MATLAB client to the parallel pool (e.g. using the command  
696 `parpool`). Features from the MATLAB parallel toolbox e.g., `parfor`, can be used to distribute the tasks  
697 on the MATLAB clients. This type of parallelism can be implemented on each single computational node.  
698 Clearly such kind of parallelization can only be used if the model is evaluated in MATLAB. In case the analysis  
699 of the inner loop requires the call of an external solver (such as a FE/CFD analysis) the multi-thread, shared

700 memory parallelism capabilities of the external software need to be adopted in order to enable *the first level*  
701 of parallelization. The second level of parallelization exploits *cluster and grid computing*, i.e. the availability  
702 of machines connected in an heterogeneous network. In this case, the total number of simulations is slitted  
703 in a multiple number of independent batch jobs. The jobs are then submitted to the job scheduler/manager  
704 and distributed efficiently on the available machines of the grid/cluster.

705 As a final consideration, these two types of parallelization can be combined together. As an example,  
706 the model evaluation required by global sensitivity analysis can be spread using batch jobs along multiple  
707 computational nodes. Then, for each batch processed on each node of the cluster, a subset of analyses is  
708 performed in parallel on the cores of the node in order to compute the quantity of interest, e.g., Monte  
709 Carlo simulation can be performed to evaluate a stochastic model and to compute the empirical CDFs of  
710 the quantity of interest.

711 Although, using `OPENCROSSAN` framework, the parallelization of the analysis is straightforward, the  
712 parallelization of a generic model might be quite challenging. In fact, independent multiple stream and sub-  
713 stream should be generated by the master node and distributed to the workers. In `MATLAB`, a combined  
714 multiple recursive generator (`mrg32k3a`) can be used to generate such independent sub-streams. When user-  
715 supplied code is involved, the standard approach, taken by `OPENCROSSAN` to parallelize `MATLAB` functions  
716 with independent jobs, is to compile such functions using `mcc` and then distribute the compiled code to  
717 the node of the cluster (workers). Hence, it is possible to execute in parallel `MATLAB` code without the  
718 necessity to install `MATLAB` on each computational node of the cluster, but only accessing the `MATLAB`  
719 runtime libraries. When this approach is not possible, for instance due to license limitations to deploy code,  
720 multiple headless instances of `MATLAB` are executed (available `MATLAB` licenses on each cluster node are  
721 necessary).

## 722 V. Numerical application

### 723 NASA Langley multidisciplinary uncertainty quantification challenge

724 The necessity to determine limitation and range of applicability of existing uncertainty quantification (UQ)  
725 methodologies and to advance the state of the practice in UQ problem of direct interest of NASA has lead to

726 the development of a challenge problem. The reader is referred to Ref. 9 for a full description of the NASA  
 727 UQ challenge problem.

728 A mathematical model that describes the dynamic of a remotely operated twin-jet aircraft developed by  
 729 NASA Langley Research Center is analyzed (see Figure 3). The model, provided as a “Black Box”, contains  
 730 21 parameters,  $\mathbf{p}$ , 16 design variables,  $\mathbf{d}$  and 8 outputs,  $\mathbf{g}$ . Furthermore, a set of intermediate variables,  $\mathbf{x}$ ,  
 731 that can be interpreted as outputs of the so-called fixed discipline analysis,  $\mathbf{x} = h(\mathbf{p})$ , are the inputs of the  
 732 cross discipline analysis  $\mathbf{g} = f(\mathbf{x}, \mathbf{d})$ . One of the main objectives of the proposed problem is to identify the  
 733 design parameters,  $\mathbf{d}$ , that provide optimal worst case probabilistic performance in presence of the model  
 734 parameters uncertainty,  $\mathbf{p}$  i.e. perform a robust optimization. This requires to solve a series of subproblems,  
 735 such as uncertainty characterization, sensitivity analysis, among others, in order to improve the model.

736 In the following, the term “original model” is used to describe the uncertainty model as provided in the  
 737 challenge problem; “reduced model” refers to the model with reduced uncertainty after the solution of the  
 738 subproblem A and “improved model” refers to the reduced model with four parameters with the smallest  
 739 ranges of uncertainty obtained from NASA. Only the main findings are reported and the reader is referred  
 740 to Ref. 30 for detailed results of the challenge problem.

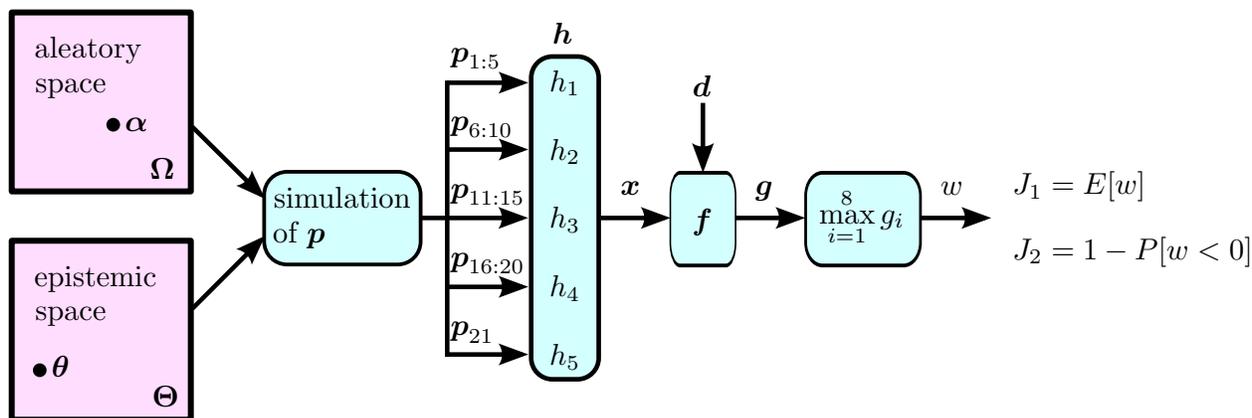


Figure 3: Relationship between the variables and functions of the NASA Langley multidisciplinary uncertainty quantification challenge problem.<sup>8</sup>

741 DECOMPOSITION OF VARIABLES  $\mathbf{p}$  INTO ITS ALEATORY AND EPISTEMIC COMPONENTS Table 1 lists all  
 742 variables of vector  $\mathbf{p}$  decomposed into an aleatory component and an epistemic component. Note that on  
 743 the one hand, the aleatory component of a random variable or distributional p-box can be represented as a

744 uniform random variable in  $(0, 1]$ ; on the other hand, the epistemic component of a distributional p-box is  
 745 given by the intervals that describe the parameters of the parental CDF; in this way, the aleatory  $\Omega$  and the  
 746 epistemic  $\Theta$  spaces have respectively 17 and 31 dimensions.

747 REPRESENTATION OF VARIABLE  $p_1$  In Section II it has been shown how to represent p-boxes. However,  
 748 variable  $p_1$  requires special considerations in its representation. These are discussed in detail in the following:  
 749 The input variable  $p_1$  is represented as a unimodal beta distribution whose mean  $\mu$  and variance  $\sigma^2$  are  
 750 uncertain, but are known to lie in the intervals  $[3/5, 4/5]$  and  $[1/50, 1/25]$  respectively. Instead beta  
 751 distributions are characterized by shape parameters  $a$  and  $b$  which are related to  $\mu$  and  $\sigma^2$  by:

$$\mu = \frac{a}{a+b} \qquad \sigma^2 = \frac{ab}{(a+b+1)(a+b)^2} \qquad (31)$$

752 that is,

$$a = -\frac{\mu(\sigma^2 + \mu^2 - \mu)}{\sigma^2} \qquad b = \frac{(\mu - 1)(\sigma^2 + \mu^2 - \mu)}{\sigma^2} \qquad (32)$$

753 The required unimodality implies that  $a$  and  $b$  are greater than 1. For shape parameters lower than 1 the  
 754 beta distribution assume the U-shaped bimodal distributions.

755 REPRESENTATION OF VARIABLES  $p_4$  AND  $p_5$  One drawback of the proposed approach is that the copula  
 756 must be perfectly modelled, without any epistemic uncertainty in its parameters. The copula that relates  
 757 variables  $p_4$  and  $p_5$  has an interval parameter, namely  $I_8$ , which models the correlation  $\rho(p_4, p_5)$ . Variables  
 758  $p_4$  and  $p_5$  are modelled using the following formulation, which permits to split uncertainty into the aleatory  
 759 and the epistemic spaces while representing the dependence with an independent copula, which does not  
 760 have any epistemic component at all:

- 761 • The aleatory part of the joint probability box is given by  $\alpha_3$  and  $\alpha_4$  which are independent and uniform  
 762 random variables on  $(0, 1]$ . Note that  $z_3 = \Phi^{-1}(\alpha_3)$  and  $z_4 = \Phi^{-1}(\alpha_4)$  where  $\Phi$  represents the standard  
 763 normal CDF.
- 764 • The epistemic part of the joint distribution is given by the 5-dimensional box  $\times_{i=4}^8 I_i$ .

765 A simulation from variables  $p_4$  and  $p_5$  can be performed by using the vector  $\mathbf{z} = [z_3, z_4]^T$  and a parameter  
 766 vector  $\boldsymbol{\theta} \in \times_{i=4}^8 I_i$ ; the simulation uses the standard procedure for sampling from a multivariate normal  
 767 PDF. This method employs the Cholesky decomposition of the covariance matrix.

Table 1: Aleatory and epistemic components of the input variables  $p_i$ , The first column provides the parameter's symbol, the second one its category (see above for a description of the categories), the third and fourth one describe its aleatory and epistemic uncertainty model. Here  $\rho(\cdot, \cdot)$ ,  $E[\cdot]$  and  $\text{Var}[\cdot]$ , denote the correlation, expected value, and variance operators respectively.

Variable	Category	Aleatory component	Epistemic component	Description
$p_1$	III	$\alpha_1 \sim \text{Unif}(0, 1]$	$I_1 = [3/5, 4/5]$	Interval of $E[p_1]$
		(distribution type: unimodal Beta)	$I_2 = [1/50, 1/25]$	Interval of $\text{Var}[p_1]$
$p_2$	II		$I_3 = [0, 1]$	Interval
$p_3$	I	$\alpha_2 \sim \text{Unif}(0, 1]$		Random variable
$p_4, p_5$	III	$\alpha_3 \sim \text{Unif}(0, 1]$	$I_4 = [-5, 5]$	Interval of $E[p_4]$
		$\alpha_4 \sim \text{Unif}(0, 1]$	$I_5 = [1/400, 4]$	Interval of $\text{Var}[p_4]$
		(distribution type: multivariate gaussian)	$I_6 = [-5, 5]$	Interval of $E[p_5]$
			$I_7 = [1/400, 4]$	Interval of $\text{Var}[p_4]$
		$I_8 = [-1, 1]$	Interval of $\rho(p_4, p_5)$	
$p_6$	II		$I_9 = [0, 1]$	Interval
$p_7$	III	$\alpha_5 \sim \text{Unif}(0, 1]$	$I_{10} = [0.982, 3.537]$	Interval of $a$
		(distribution type: Beta)	$I_{11} = [0.619, 1.080]$	Interval of $b$
$p_8$	III	$\alpha_6 \sim \text{Unif}(0, 1]$	$I_{12} = [7.450, 14.093]$	Interval of $a$
		(distribution type: Beta)	$I_{13} = [4.285, 7.864]$	Interval of $b$
$p_9$	I	$\alpha_7 \sim \text{Unif}(0, 1]$		Random variable
$p_{10}$	III	$\alpha_8 \sim \text{Unif}(0, 1]$	$I_{14} = [1.520, 4.513]$	Interval of $a$
		(distribution type: Beta)	$I_{15} = [1.536, 4.750]$	Interval of $b$
$p_{11}$	I	$\alpha_9 \sim \text{Unif}(0, 1]$		Random variable
$p_{12}$	II		$I_{16} = [0, 1]$	Interval
$p_{13}$	III	$\alpha_{10} \sim \text{Unif}(0, 1]$	$I_{17} = [0.412, 0.737]$	Interval of $a$
		(distribution type: Beta)	$I_{18} = [1.000, 2.068]$	Interval of $b$

768 Consequently, the joint distribution-free probability box formed by variables  $p_4$  and  $p_5$  can be represented  
769 as the random set  $\Gamma : (0, 1]^2 \rightarrow \mathcal{F}, \alpha \mapsto \Gamma(\alpha)$  where  $\alpha = (\alpha_3, \alpha_4)$ ,  $\mathcal{F}$  is the system of focal elements given  
770 by the preimages of  $\{\alpha_3 \times \alpha_4 \times I_4 \times I_5 \times \dots \times I_8 : (\alpha_3, \alpha_4) \in (0, 1]^2\}$  through  $F_{p_4 p_5}$ . Since  $\alpha_3$  and  $\alpha_4$  are  
771 independent uniform random variables in  $(0, 1]$ , they can be considered as the realization of a bidimensional  
772 product copula, defined on  $(0, 1]^2$ . For the interpretation of  $\alpha_3, \alpha_4, I_4, \dots, I_8$  the reader is referred to  
773 Table 1.

## 774 A. Subproblem A

775 The aim of the uncertainty characterization or subproblem A is to reduce the epistemic uncertainty compo-  
776 nents of the category II ( $p_2$ ) and III parameters ( $p_1, p_4, p_5$ ) that are inputs of a subsystem  $h_1$ . The subsystem  
777 provides a scalar output  $x_1$  as a function of those five uncertain parameters, that is,

$$x_1 = h_1(p_1, p_2, p_3, p_4, p_5). \quad (33)$$

778

779 In this subproblem, the vector  $[p_1, \dots, p_5]$  is the output of the system  $\mathcal{W}$ , the system  $h_1$  is equivalent to  
780 the function  $\mathcal{G}$  defined on Section II-F, the epistemic space is the Cartesian product  $\Theta := \times_{i=1}^8 I_j$  and the  
781 aleatory space, which models variables  $\alpha_1$  to  $\alpha_4$ , is defined by  $\Omega := (0, 1]^4$  (see Table 1).

782 Two sets of 25 observations of the “true uncertainty model”  $\theta^* \in \Theta$  are available to reduce the uncertainty  
783 in  $\Omega$ . The approaches described in Section III-A are here adopted.

784 One of the main challenges of this subproblem is provided by the limited available information (25  
785 observation points for each dataset) and the relatively large dissimilarity of the empirical CDFs associated  
786 with those datasets as shown in Figure 4.

### 787 1. Non-parametric statistic method based on the Kolmogorov-Smirnov test

788 The procedure presented in Section III-A has been used to solve the subproblem A. First, the validation  
789 similarity level has been calculated after using a Gaussian KDE to compute the CDF  $\tilde{F}_e$  for the observation  
790 sets. A validation similarity level  $D_{\tilde{v}} = 0.18$  has been obtained calculating the maximum distance between  
791 the two KDEs adjusted to the two datasets respectively  $\tilde{F}_e$  (i.e. using Eq. (6)). The measure of similarity

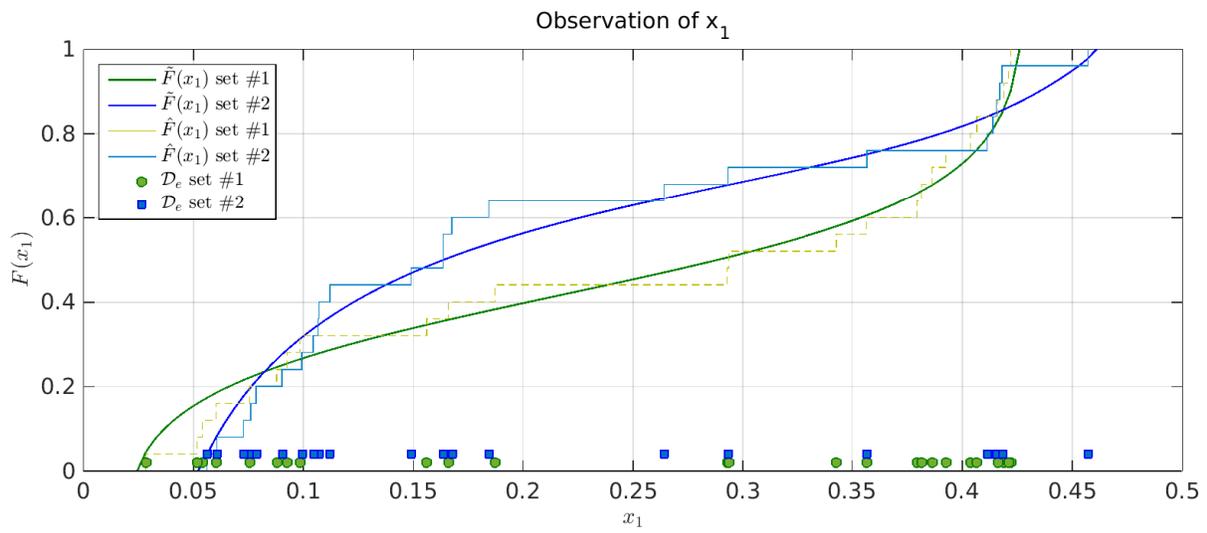


Figure 4: Empirical CDF,  $\hat{F}$ , of the two set of observation points and CDF obtained adopting the Gaussian kernel density of Eq. (7),  $\tilde{F}$ . The dots and squares show the two datasets  $\mathcal{D}_e$ , respectively.

792 obtained comparing the two empirical CDFs,  $\hat{F}_e$ , of the datasets is  $D_{\hat{v}} = 0.24$  as shown in Figure 5. This  
 793 allows to identify those points  $\theta_i \in \Theta$  that conform with the observations such that  $D_i < D_{\hat{v}}$ .

794 Assuming a uniform distribution on  $\Theta$ , 10000 samples  $\theta_i$  are drawn and for each  $\theta_i$ ,  $n = 5000$  samples  
 795 from the aleatory space  $\Omega$  are used to propagate the aleatory uncertainty through the model (using the  
 796 function `p_to_x1`). Finally, using the empirical CDF of  $x_1$  ( $\hat{F}(x_1|\theta_i)$ ), the measure of similarity  $D_i$  is  
 797 calculated against  $\tilde{F}_e$  according to Eq. (6) (i.e.  $D_i = \sup_x |\hat{F}(x|\theta_i) - \tilde{F}_e(x)|$ ). Please note that due to the  
 798 large number of samples used  $\hat{F}_i(x_1|\theta_i) \approx \tilde{F}_i(x_1|\theta_i)$ . The histograms of the measure of similarity  $D_i$  are  
 799 shown in Figure 5 computed for the dataset of 25 and 50 observations, respectively. It is possible to observe  
 800 that  $D_i$  is smaller when the KDE  $\tilde{F}_i(x_1|\theta_i)$  and all 50 observations are used.

801 The measure of similarity  $D_{\hat{v}} = 0.18$  identifies model outputs,  $x_1$  obtained from the realizations in the  
 802 epistemic space,  $\theta_i$ , that are in agreement with the observations (represented in Figure 5 by the bars on the  
 803 left of  $D_{\hat{v}}$ ). Calculating  $P(D_i > D_{\hat{v}}) = c$ , two confidence levels have been obtained:  $c_{\hat{v}(25)} = 0.8031$  and  
 804  $c_{\hat{v}(50)} = 0.547$  when  $D_i$  is calculated against the  $\tilde{F}_e$  obtained using 25 and 50 observations, respectively.

805 Figure 6 shows the parallel coordinate plot of the epistemic realizations. Please note that for readability  
 806 purposes, only 1000 realizations are shown. In a parallel plot a multi-dimensional quantity is shown graphi-  
 807 cally and represented as a polyline with vertices on the parallel axes. The vertex on the  $m$ -axis corresponds  
 808 to the  $i$ -th realization of the  $m$ -coordinate (i.e.  $\theta_m^{(i)}$ ). The axes of the plot have been normalized, between 0  
 809 and 1. The top panel of Figure 6 shows combination of epistemic realizations for different level of similarity  
 810 measure computed against  $\tilde{F}_e$  constructed from 25 observations. The Figure shows all the combinations of  
 811 all epistemic realizations ( $c = 0$ ), those with a similarity measure  $D_i < D_{\hat{v}}$  (i.e.  $c = 0.547$ ) and  $D_i < D_{\hat{v}}$   
 812 (i.e.  $c = 8031$ ), respectively. The top panel of Figure 6 shows the parallel plot with measures of similarity  
 813 calculated using all the 50 observations.  $c = 0.0547$  correspond to a similarity measure  $D_i < D_{\hat{v}}$  while  
 814  $c = 0.0547$  correspond an arbitrary level  $D_i < 0.1$ .

815 The parallel coordinate plot allows to identify the epistemic uncertainty that can be reduced. For instance,  
 816 all the realizations of  $E[p_5]$  with similarity level lower  $D_{\hat{v}}$  are in the normalized interval  $[0, 0.6]$  while  $E[p_1]$   
 817 is in the normalized interval  $[0, 0.7]$ . On the contrary, the intervals of  $\text{Var}[p_1]$ ,  $p_2, E[p_4]$ ,  $\text{Var}[p_4]$ ,  $\text{Var}[p_5]$   
 818 and  $\rho(p_4, p_5)$  cannot be improved based on the current available data. Although the resulting model for  
 819  $\theta$  obtained are collection of points, the identified realizations cover connected ranges (remember that only

1000 over 10000 realizations are shown in Figure 6). The results are summarized in Table 2.

## 2. Bayesian updating on the epistemic space

The Bayesian inference is the second approach used to reduce the epistemic uncertainty as explained in Section III-A-2. In this method, Transitional Monte Carlo Markov Chains have been used to sample 1000 realizations from the posterior PDF  $p(\boldsymbol{\theta}|\mathcal{D}_n)$ . Two strategies have been employed to estimate the likelihood  $P(\mathcal{D}_e|\boldsymbol{\theta}_i)$ : the standard Bayesian and an approximate Bayesian computational method.

**BAYESIAN COMPUTATIONAL METHOD (BC)** In this case, the likelihood is computed using Eq. (8) and  $p(x|\boldsymbol{\theta}_i)$  is estimated by means of a KDE, computed with  $n = 1000$  points from the aleatory space. Figures 7 shows the posterior distributions sampled using TMCMC with 25 and 50 observation points as evidence, respectively. Histograms of the posterior samples are normalized, assigning a value of 1 to the number of counts in the bin containing the majority of samples. After normalizing the histograms, it is possible to set a general limit of normalized counts used to exclude outliers of the TMCMC algorithm and indicated by the horizontal red lines in Figures 7.

**APPROXIMATE BAYESIAN COMPUTATIONAL METHOD (ABC)** In this case, 200 samples are used to evaluate  $\hat{F}(\cdot|\boldsymbol{\theta}_i)$  and the quantities  $\delta_k$  by means of Eq. (12). Thereafter, the likelihood (11) is computed. Figures 8 shows the normalized posterior distributions sampled using TMCMC with 25 and 50 observation points as evidence, respectively.

**THE RESULTS** The proposed method has been able to identify a reduced epistemic space associated to  $E[p_1]$  and  $E[p_5]$  but no conclusions can be drawn for the other input parameters. The updated ranges of the epistemic uncertainties are summarized in Table 2.

The Bayesian updating procedure successfully managed to reduce the uncertainty associated to the output  $x_1$  as shown for example in Figure 9, for the approximate Bayesian computational method. Figure 9 shows different p-boxes of  $x_1$  obtained with the updated epistemic uncertainty parameters, using the first set of 25 observations and the full set of 50 observations, respectively. The approximated p-boxes have been obtained using the following procedure. First, 10000 samples  $\boldsymbol{\theta}_i$  of the epistemic variable are drawn from uniform distributions defined by the full range of the updated bounds (light gray) and by the updated

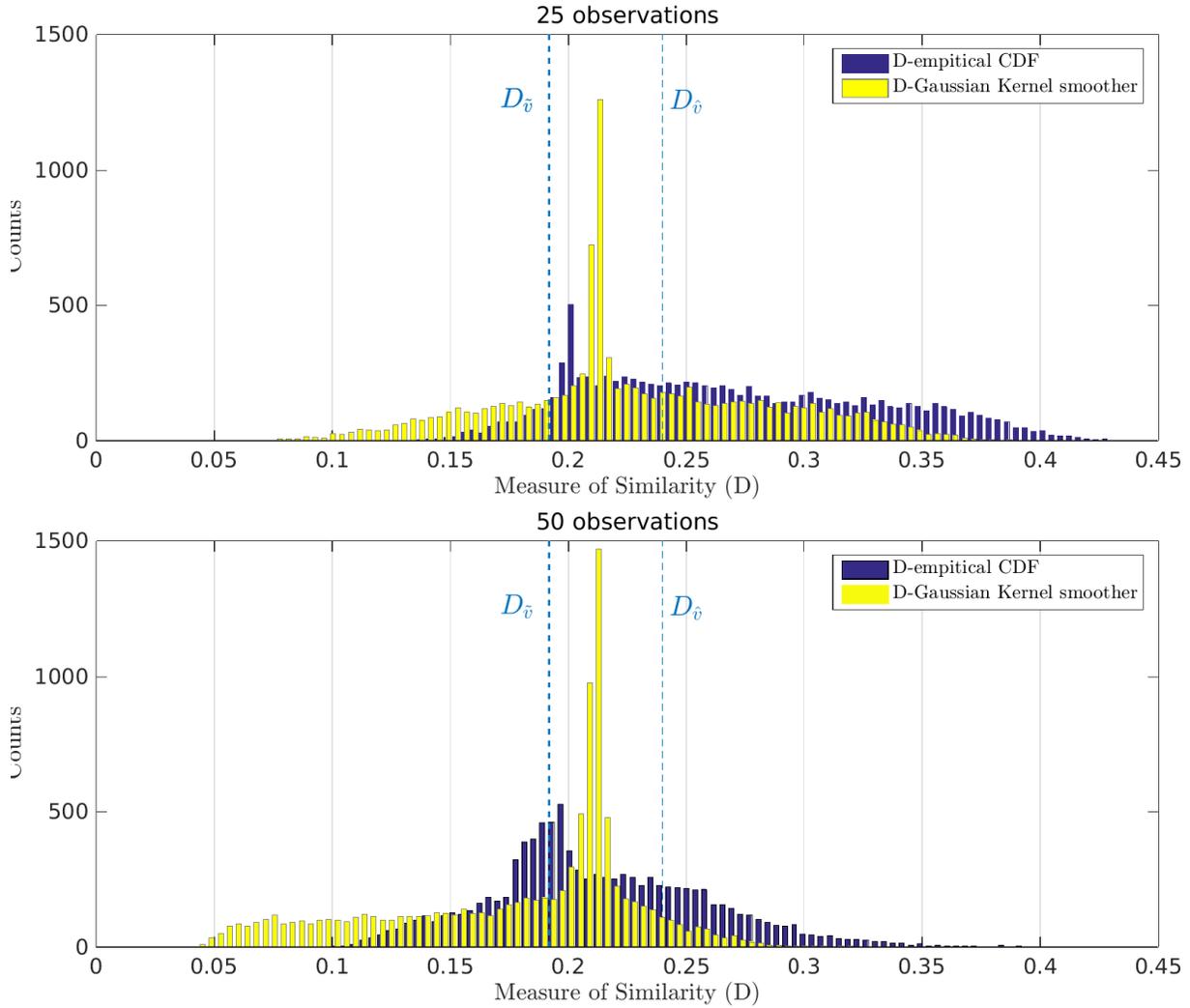


Figure 5: Histogram of the measure of similarity,  $D_i$ , between the CDF calculated sampling randomly in the epistemic space and the observations, for 25 (top panel) and 50 (bottom panel) observations ( $\mathcal{D}_e$ ).  $D_i$  has been computed using the empirical CDF of the experimental data (blue bars) and the CDF obtained using Gaussian kernel smoother functions (yellow bars). The figure also shows the values of the measure of similarity between the two set of observation data computed using Gaussian kernel smoother techniques,  $D_{\bar{v}}$ , and empirical CDF,  $D_{\hat{v}}$ , respectively.

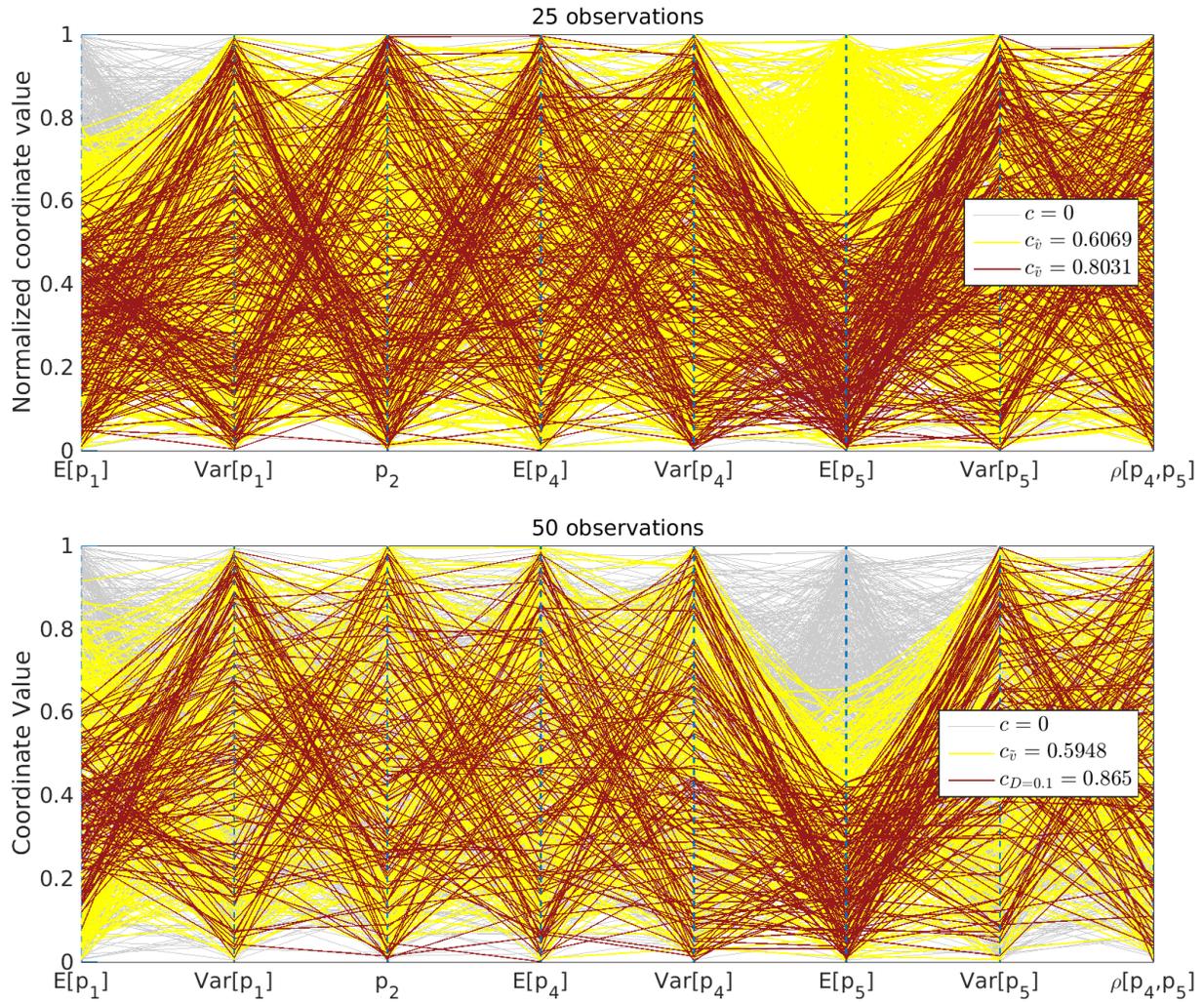
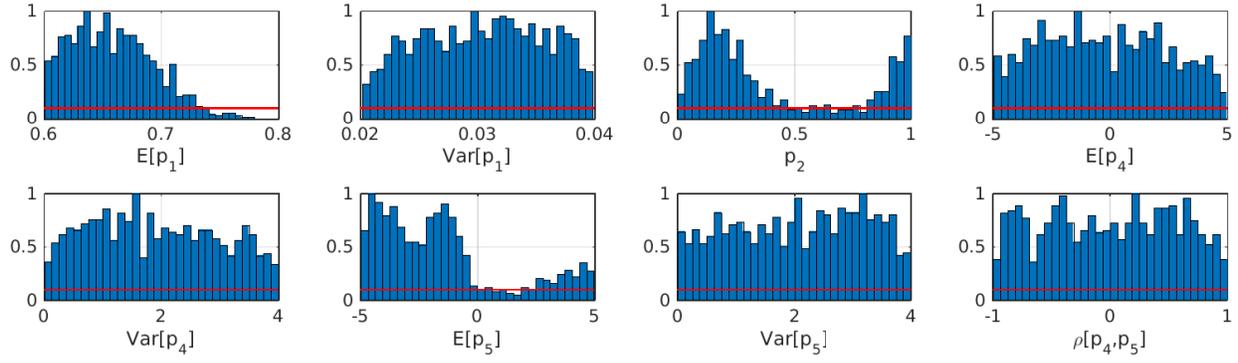
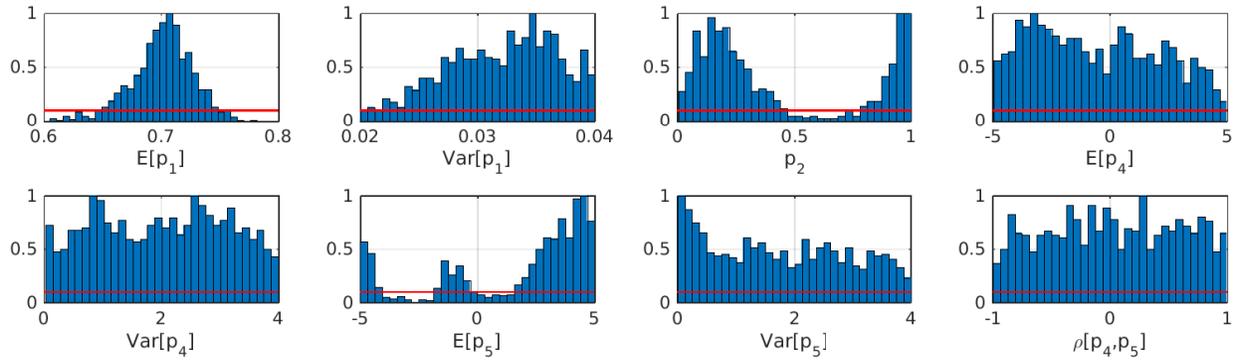


Figure 6: Parallel coordinates plot of the 8 category II and III parameters of the input factors of  $h_1$  (i.e.  $p_i, i = 1, \dots, 5$ ) for 25 (top panel) and 50 (bottom panel) observations ( $\mathcal{D}_e$ ). The figure shows only 1000 realizations (over a total sample of 10000) of the epistemic space for different significant levels  $c$  of the Kolmogorov-Smirnov test.  $c = 0$  represents of all the realizations.  $c_{\bar{v}}$  represents realizations of  $\theta$  with a measure of similarity  $D_i < D_{\bar{v}}$ .  $c_{\bar{v}}$  represents realizations of  $\theta$  with a measure of similarity  $D_i < D_{\bar{v}}$  and  $c_{D=0.1}$  realizations with a measure of similarity  $D_i < 0.1$ .

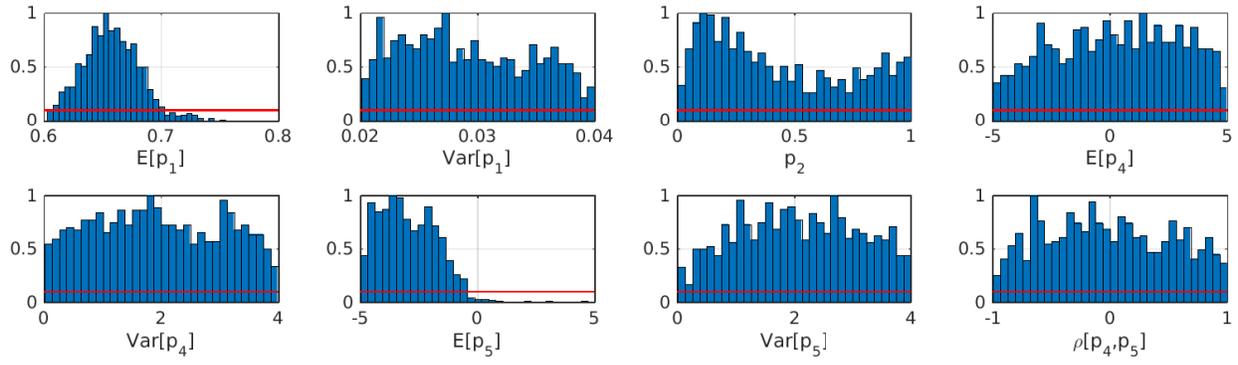


(a) 25 observations.

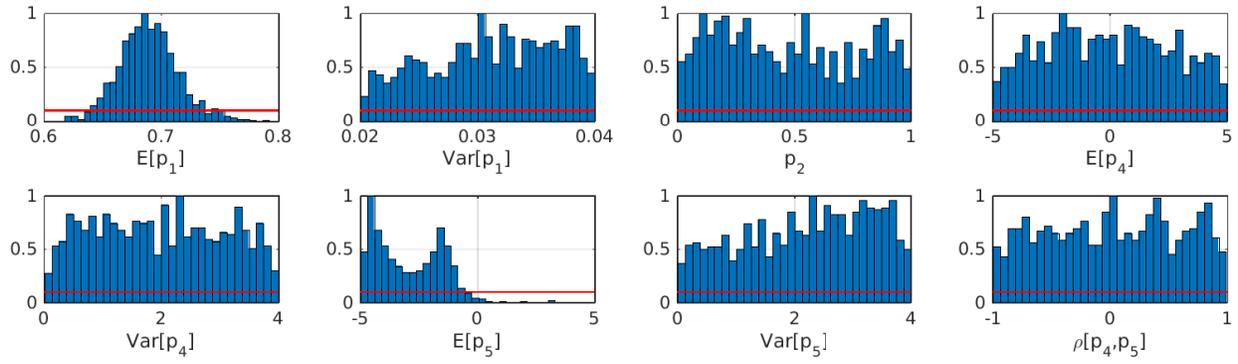


(b) 50 observations.

Figure 7: Normalized histogram of  $p(\theta|\mathcal{D}_e)$  obtained using Bayesian Computational method with (a) 25 experimental observations and (b) 50 experimental observations (b) of  $x_1$ , respectively. The normalization assigns a value of 1 to the bin with the highest number of counts. The red line represent the cut-off value to determine the updated range.

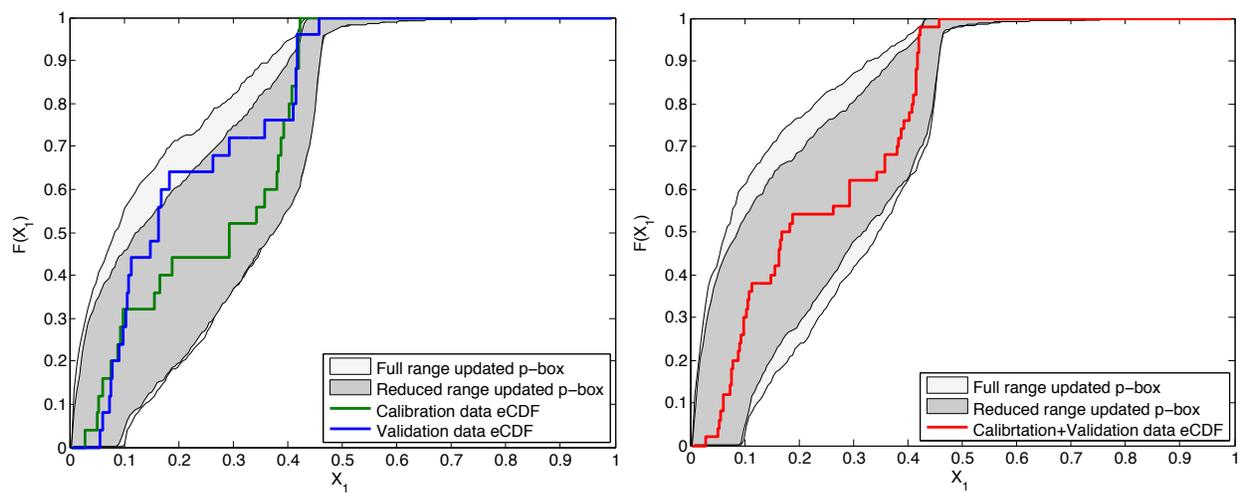


(a) 25 observations.



(b) 50 observations.

Figure 8: Normalized histogram of  $p(\theta|\mathcal{D}_e)$  obtained using Approximate Bayesian Computational method with (a) 25 experimental observations and (b) 50 experimental observations (b) of  $x_1$ , respectively. The normalization assigns a value of 1 to the bin with the highest number of counts. The red line represent the cut-off value to determine the updated range.



(a) 25 observations.

(b) 50 observations.

Figure 9: P-boxes of  $x_1$  and the empirical CDFs of the experimental data. The p-boxes have been obtained using the full range of the posterior parameters and using the range that excludes the outliers, respectively.

846 bounds obtained excluding the outliers (dark gray). Then, the CDF  $\hat{F}(\cdot|\boldsymbol{\theta}_i)$  is computed for each epistemic  
 847 realization. Finally the curves enveloping all the CDFs are obtained and shown in Figure 9. It is possible  
 848 to notice that the updated p-box of  $x_1$  is tighter when all the 50 experimental observations are used.  
 849 Additionally, the experimental CDFs of the calibration data set are fully contained in the light gray area  
 850 (i.e. the p-boxes obtained excluding the outliers). However, the validation data lay inside the updated p-box  
 851 only when the full intervals of updated parameters are considered.

852 The reduced uncertainty model identified by the non-parametric approach and by the Bayesian inference  
 853 approach are summarized in Table 2, respectively. Although only the uncertainty of two parameters can be  
 854 significantly reduced, the results provided by the proposed approaches are in agreement providing a cross  
 855 validation of the developed procedures used to solve the subproblem A.

## 856 B. Subproblem B

857 The aim of this subproblem is to identify and rank the input parameters of category II and III (i.e. intervals  
 858 and distributional p-boxes) according to degree of refinement in the output p-boxes which one could hope  
 859 to obtain by refining their uncertainty models. More specifically, in problem B1 the focus is to rank the 4  
 860 input factors that affect the variability the output  $x_i$  of each model  $h_i(\cdot), i = 1, \dots, 4$ , respectively. In tasks  
 861 B2-B3, 17 parameters need to be ranked according to the reduction in the range of  $J_1 = E[w(\mathbf{p}, \mathbf{d}_{\text{baseline}})]$   
 862 (task B2) and  $J_2 = 1 - P[w(\mathbf{p}, \mathbf{d}_{\text{baseline}}) < 0]$  (task B3), respectively. In those expressions, the worst-case  
 863 requirement metric  $w$  is defined by  $w(\mathbf{p}, \mathbf{d}) = \max_{1 \leq i \leq 8} g_i(\mathbf{p}, \mathbf{d})$ . The strategy presented in Section III-B  
 864 will be used.

### 865 1. Problem B1

866 NONSPECIFICITY TECHNIQUE By means of the nonspecificity measure, each interval  $[I_i, \overline{I}_i]$  is reduced  
 867 to the value given by  $I_i + p_r \cdot (\overline{I}_i - I_i)$ , where  $p_r \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ . For instance, interval  
 868  $I_1 = [\underline{E}[p_1], \overline{E}[p_1]] = [3/5, 4/5]$  is reduced to the constants 0.62, 0.66, 0.70, 0.74 and 0.78 and  $n = 50$   
 869 samples from the product copula that links aleatory variables,  $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ , are employed to construct  
 870 the output Dempster-Shafer structure for each reduction. Note that  $\alpha_3$  and  $\alpha_4$  are used to model the  
 871 variables  $p_4$  and  $p_5$ , according to the transformation explained at the beginning of Section V. Then, the

Table 2: Reduced uncertainty model using the non-parametric approach ( $c = 0.547$ ) or 25 observations and  $c = 8031$  for 50 observations) and the Bayesian inference, respectively. A – means that the method could not reduce the epistemic uncertainty for the referred variable.

Variable	Original	Nonparametric	Bayesian methods	
	interval	method	BC	ABC
25 observations				
$E[p_1]$	[ 0.6000, 0.80]	[ 0.6000, 0.72]	[ 0.6000, 0.73]	[0.6030, 0.755 ]
$\text{Var}[p_1]$	[ 0.0200, 0.04]	–	–	–
$p_2$	[ 0.0000, 1.00]	–	–	–
$E[p_4]$	[-5.0000, 5.00]	–	–	–
$\text{Var}[p_4]$	[ 0.0025, 4.00]	–	–	–
$E[p_5]$	[-5.0000, 5.00]	[-5.0000, 0.78]	–	[-5.0000, 4.50 ]
$\text{Var}[p_5]$	[ 0.0025, 4.00]	–	–	–
$\rho(p_4, p_5)$	[-1.0000, 1.00]	–	–	–
50 observations				
$E[p_1]$	[ 0.6000, 0.80]	[0.63, 0.76]	[0.60, 0.75]	[0.618, 0.791]
$\text{Var}[p_1]$	[ 0.0200, 0.04]	[0.0260, 0.04]	–	–
$p_2$	[ 0.0000, 1.00]	–	–	–
$E[p_4]$	[-5.0000, 5.00]	[-4.50, 4.80]	–	–
$\text{Var}[p_4]$	[ 0.0025, 4.00]	–	–	[0.097, 3.943]
$E[p_5]$	[-5.0000, 5.00]	[-4.90, 0.30]	–	[-5.00, 4.45 ]
$\text{Var}[p_5]$	[ 0.0025, 4.00]	–	–	–
$\rho(p_4, p_5)$	[-1.0000, 1.00]	–	–	–

872 nonspecificity which is a measure of epistemic uncertainty, of each of those Dempster-Shafer structures is  
 873 calculated. Following a similar procedure, the rankings of input variables have been calculated (see Table 3)  
 874 according to the output nonspecificity for the systems  $h_2$ ,  $h_3$  and  $h_4$ , respectively. In all cases, the evaluation  
 875 of equation (5) was performed for each focal element using a genetic algorithm with a population of 30000  
 876 individuals and 10 generations.

877 GLOBAL SENSITIVITY ANALYSIS TECHNIQUE The global sensitivity analysis has been performed on a re-  
 878 defined mathematical model  $h^*$  of the original  $h$  as detailed in Section III-B-2.  $h^*$  takes as inputs only  
 879 uniform distributions (that represents the epistemic space  $\Theta$ ) and returns a scalar output  $y_i$  (the area of  
 880 distribution-free p-boxes) as shown in Figure 10.

881 For each combination  $\theta_i \in \Theta$  of the input parameters, the model  $h^*$  performs an internal Monte Carlo  
 882 simulation using  $n = 500$  samples  $\alpha_j$  to calculate an empirical CDF of  $x_j^i$ ,  $\hat{F}(\cdot|\theta_i)$ . A sample size of 500 is  
 883 sufficient to rank unequivocally the most important parameters with respect to the outputs  $x_k, k = 1, \dots, 5$ ,  
 884 as shown in Figure 11. Then,  $\hat{F}(\cdot|\theta_i)$  is compared with a “reference CDF”,  $F(\cdot|\bar{\theta})$  and the final output  $y_i$  is  
 885 returned (see Eq. (17)).

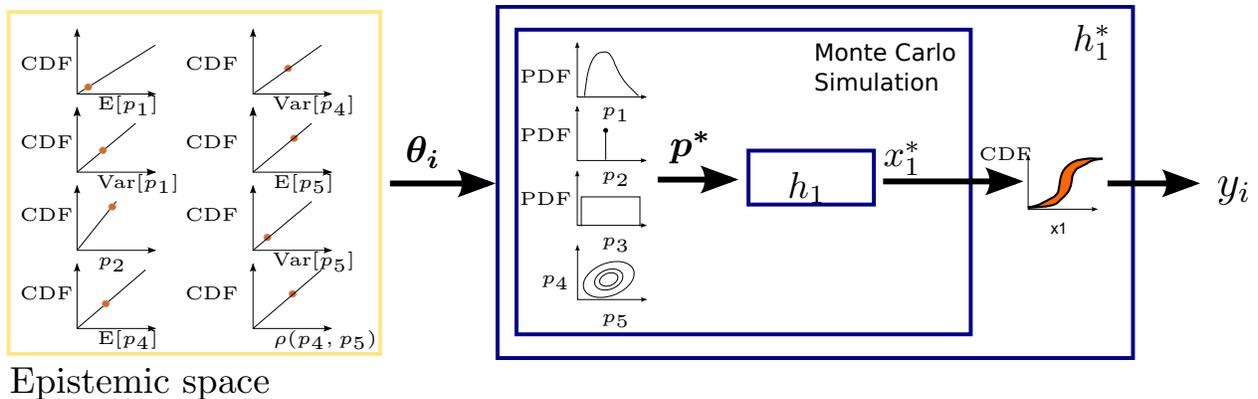


Figure 10: Redefined model  $h_1^*$  used for performing the Global sensitivity analysis with aleatory and epistemic uncertainty.

886 In order to reduce the computational noise of  $h^*$  (i.e. the variance of the output), the common random  
 887 number technique<sup>51</sup> has been used to propagate the aleatory uncertainty (i.e. performing the internal Monte  
 888 Carlo simulation for the model  $h^*$ ). The extended-FAST method has been used with 2048 samples of  $\theta_i$  for  
 889 each of the 8 input factors of the refined model, and in consequence, 16384 simulations are required for each

890 measure of  $x_k$  while the Saltelli's method has been run with 16384 samples for a total cost of 540672 model  
 891 evaluations for each  $x_k$ .

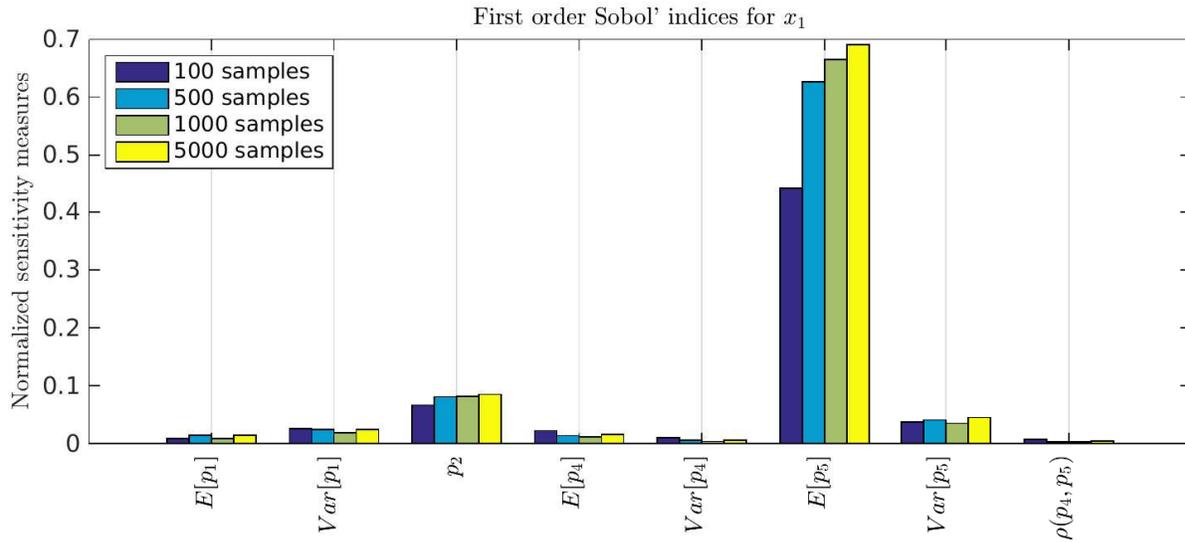


Figure 11: Effect of different samples size for the internal Monte Carlo simulation on the estimation of the Sobol' indices with respect to  $x_1$  by means of extended-FAST method method.

892 Since the global sensitivity procedure computes the sensitivity measure of the individual components for  
 893 the category III parameters (e.g.  $E[p_1]$ ,  $Var[p_1]$ , the numerical values for the input parameters have been  
 894 calculated as:  $S(p_1) = S(E[p_1]) + S(Var[p_1])$ ,  $S(p_4) = S(E[p_4]) + S(Var[p_4]) + S(\rho(p_4, p_4))$ , and  $S(p_5) =$   
 895  $S(E[p_5]) + S(Var[p_5]) + S(\rho(p_4, p_5))$ .

896 From the results summarized in Table 3, it is possible to see that the results obtained applying the two  
 897 approaches are in agreement.

## 898 2. Problems B2 and B3

899 Similar strategies applied in the solution of of task B1 have been here applied. In this case, for the nonspeci-  
 900 ficity technique, the variables are mapped through the system  $w(\mathbf{p}, \mathbf{d}_{\text{baseline}})$  with only 10 focal elements and  
 901 the range of the interval was measured instead of the nonspecificity of each focal element.

902 The redefined model  $h^*$  has also been adopted for performing global sensitivity analyses. Here,  $h^*$  takes  
 903 as input uniform distributions representing the epistemic uncertainties and returns the output  $J_1$  and  $J_2$ .  
 904 For each realization of the epistemic uncertainty, a Monte Carlo simulation with 500 samples is performed

Table 3: Ranking of the 4 category II-III parameters according to the nonspecificity technique (NST) and global sensitivity analysis (GSA) for  $x_i, i = 1, \dots, 4$ , respectively. Note that using the global sensitivity analysis, the larger the value of the “first Sobol’ index” is, the more important the input factor is. On the other hand for nonspecificity technique parameters with the lower values are more important than parameter with larger values.

Output	Rank #1	Rank #2	Rank #3	Rank #4	Strategy
$x_1$	$p_1(\mu)(0.235)$	$p_5(\mu)(0.757)$	$p_4(\mu)(0.808)$	$p_2(0.850)$	NST
	$p_1(0.684)$	$p_5(0.145)$	$p_4(0.056)$	$p_2(0.02)$	GSA
$x_2$	$p_6(0.063)$	$p_7(a)(0.596)$	$p_8(a)(0.922)$	$p_{10}(b)(0.993)$	NST
	$p_6(0.701)$	$p_7(0.153)$	$p_8(0.021)$	$p_{10}(< 0.001)$	GSA
$x_3$	$p_{12}(0.026)$	–	–	–	NST
	$p_{12}(0.835)$	$p_{15}(0.017)$	$p_{14}(< 0.001)$	$p_{13}(< 0.001)$	GSA
$x_4$	$p_{16}(0.121)$	$p_{17}(a)(0.779)$	$p_{18}(a)(0.786)$	$p_{20}(a)(0.938)$	NST
	$p_{16}(0.761)$	$p_{18}(0.073)$	$p_{17}(0.025)$	$p_{20}(0.001)$	GSA

905 to propagate the aleatory uncertainty Although the distribution of  $J_1$  is very sensitive to the number of  
 906 aleatory samples  $\alpha_j$ , a sample size of 500 has been demonstrated to be sufficient for ranking unequivocally  
 907 the most important parameters as shown in Figure 12..

908 The values of  $\mathbf{g}$  are computed via  $\mathbf{f}$  function and the CDF of  $w$  computed. Finally,  $J_1$  and  $J_2$  are  
 909 calculated from  $F(w)$ . Finally, the sensitivity indices of  $J_1$  and  $J_2$  are calculated. The extended-FAST  
 910 method and the Saltelli’s method has been used to estimate the sensitivity measure. The extended-FAST  
 911 method has been computed using 1000 samples  $\theta_i$  for each input factor of the model  $h^*$  (i.e. the intervals of  
 912 the epistemic space) for a total cost of simulations 31000 whereas the Saltelli’s method has been performed  
 913 with 8192 samples for a total cost of 270336 model evaluations. Figures 13-14 show the sensitivity measures  
 914 of the input factors  $\mathbf{p}$  with respect to  $J_1$  and  $J_2$ , respectively. The most important factor that contributes  
 915 to the variance of  $J_1$  is  $p_{21}$  and in particular its variance. The total indices for  $p_4$  and  $p_5$  show that their

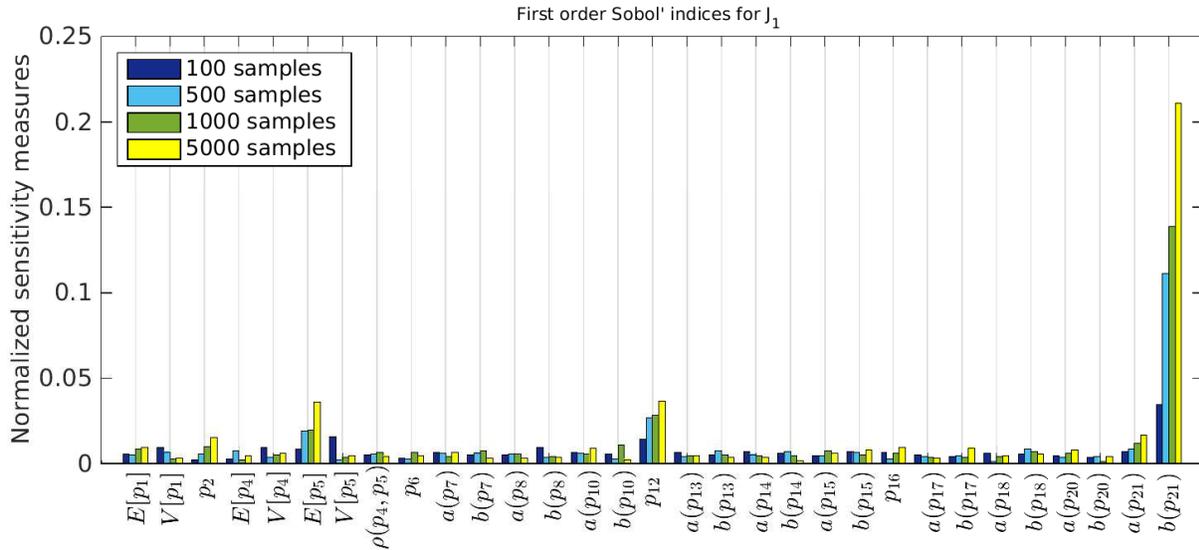


Figure 12: Effect of different samples size for the internal Monte Carlo simulation on the estimation of the Sobol' indices with respect to  $J_1$  by means of extended-FAST method method.

916 interaction also contributes to the variance of  $J_1$  but it is not possible to discriminate the single contribution  
 917 of the parameters  $p_4$  and  $p_5$ . All the other components provide similar (small) contributions to the variance  
 918 of  $J_1$ . Regarding the variance of  $J_2$ , the first order and total indices indicate that the parameters  $p_{12}$  and  
 919  $p_1$  are the most important parameters. The first order index indicates also a contribution from  $p_4$ . All the  
 920 other components provide similar (small) contribution to the variance of  $J_2$ .

921 The results of the sensitivity analysis are summarized in Table 4. The most important variables in the  
 922 reduction of uncertainty on  $J_1$  are  $p_{21}$  and the one that reduces the uncertainty on  $J_2$  is  $p_{12}$ . Again, the  
 923 different approaches have provided consistent results.

924 Four variables have been selected ( $p_1$ ,  $p_4$ ,  $p_{12}$  and  $p_{21}$ ). An improved uncertainty model for these variables  
 925 has been obtained from NASA.  $p_{21}$  is the most important parameter for  $J_1$ ,  $p_{12}$  and  $p_1$  are the most important  
 926 parameters for  $J_2$ . The sensitivity analyses indicated that the parameters  $p_4$  and  $p_5$  are also important  
 927 however without been able to discriminate between the two parameters. Since the parameter  $p_5$  has been  
 928 already reduced during the Uncertainty Characterisation analysis (see Section A), it has been decided to  
 929 ask for an improvement of the parameter  $p_4$ , in case  $p_4$  and  $p_5$  where strongly correlated. The improved  
 930 uncertainty intervals cannot be disclosed, as requested by the challengers.

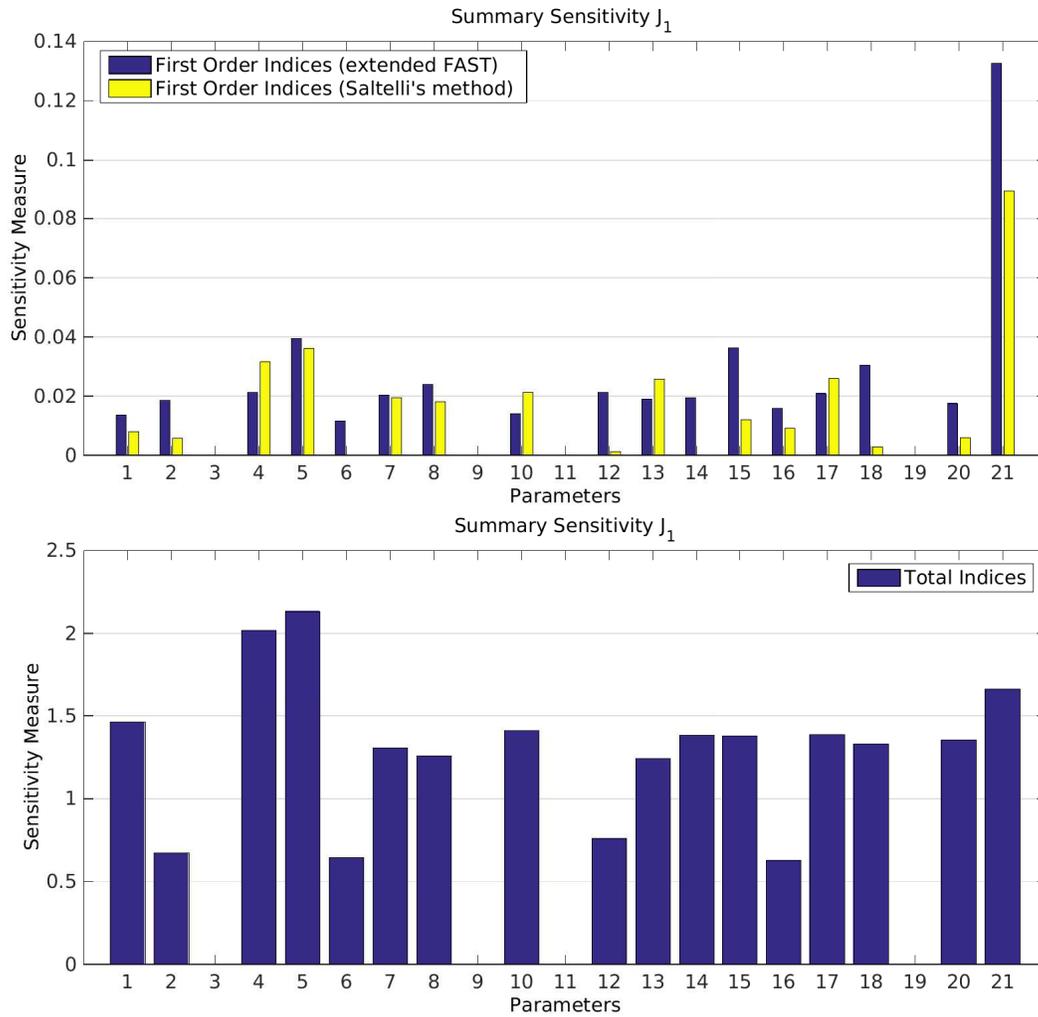


Figure 13: First and total sensitivity measure of the  $p$  parameters respect to  $J_1$ .

Table 4: Ranking of the category II-III parameters for  $J_1$  and  $J_2$  computed by means of the nonspecificity technique and global sensitivity analysis (GSA).

Output	Rank #1	Rank #2	Rank #3	Rank #4	Strategy
$J_1$	$p_{21}(b)(0.726)$	$p_6(0.751)$	$p_1(\mu)(0.763)$	$p_7(a)(1.007)$	NST
$J_1$	$p_{21}(0.089)$	$p_5(0.036)$	$p_4(0.031)$	–	GSA
$J_2$	$p_{12}(0.189)$	$p_1(\mu)(0.571)$	$p_5(\mu)(0.945)$	–	NST
$J_2$	$p_{12}(0.666)$	$p_1(0.393)$	$p_4(0.201)$	$p_5(0.179)$	GSA

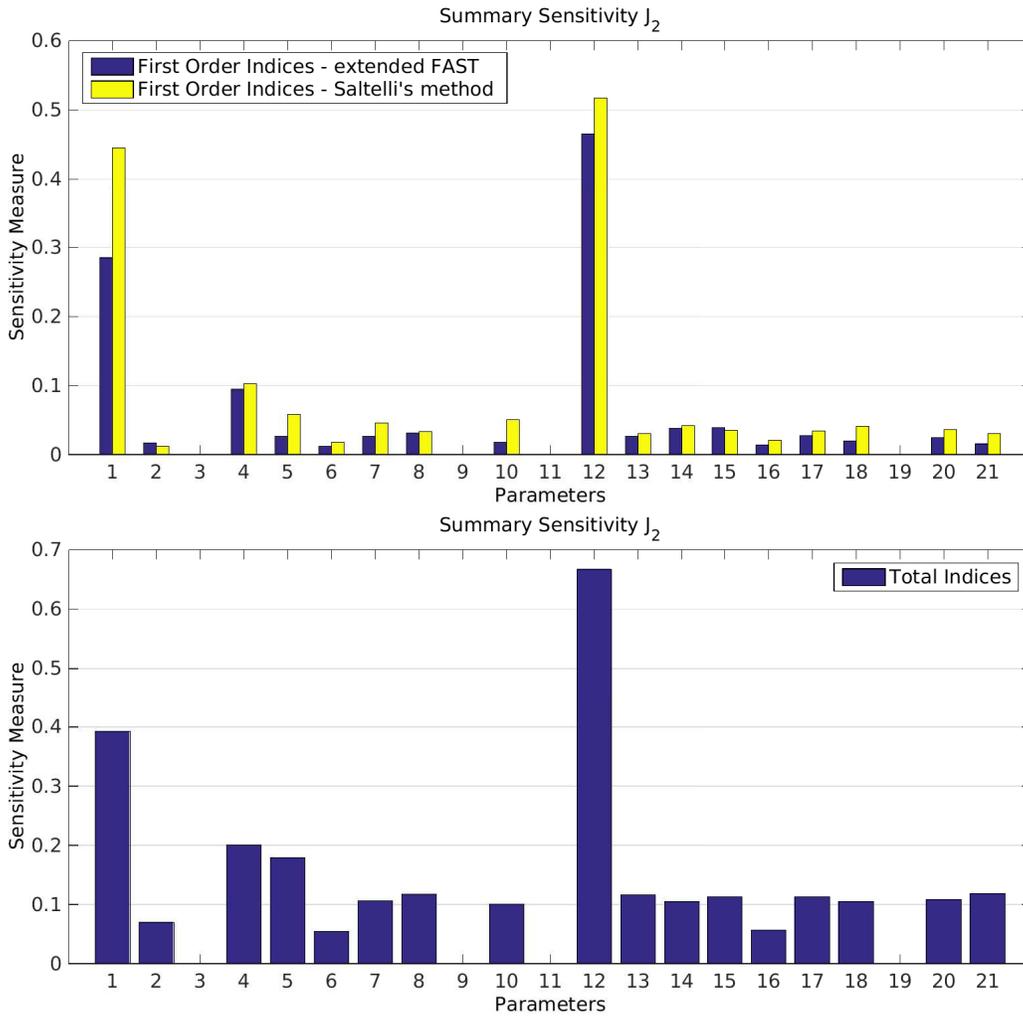


Figure 14: First and total sensitivity measure of the  $p$  parameters respect to  $J_2$ .

### 931 C. Subproblem C

932 For this subproblem, we were asked to find the range of the metrics  $J_1 = E[w(\mathbf{p}, \mathbf{d}_{\text{baseline}})]$  and  $J_2 =$   
933  $1 - P[w(\mathbf{p}, \mathbf{d}_{\text{baseline}}) < 0]$ , both with the reduced and with the improved uncertainty models. The metric  
934  $J_1$  is the expected value of the worst-case requirement metric  $w$ , while the metric  $J_2$  represents the failure  
935 probability of the system. For solving this problem two different strategies, introduced in in Section III-C,  
936 have been employed.

937 OPTIMIZATION IN THE EPISTEMIC SPACE (STANDARD APPROACH) A global optimization is performed in  
938 the epistemic space  $\Theta \equiv \times_{i=1}^{31} I_i$ , in order to find those points in  $\Theta$  that produce the upper and lower bounds  
939 on  $J_1$  and  $J_2$ . For any candidate solution provided by the optimization algorithm, i.e.  $\theta_i \in \Theta$ , a set of  
940  $n = 1000$  random points  $\{\alpha_j, j = 1, 2, \dots, n\}$  is drawn from the aleatory space  $\Omega \equiv (0, 1]^{17}$  to estimate the  
941 metrics. The number of samples from the aleatory space has been selected after performing a convergence  
942 test. More specifically, in this test, both  $J_1$  and  $J_2$  are estimated with increasing values of  $n$  (i.e. 100, 500,  
943 1000, 5000 and 10000) for 5 representative realizations of the epistemic space, as shown in Figure 15. From  
944 the figure, it can be seen that  $n = 1000$  points are sufficient for estimating  $J_1$  and  $J_2$ , with a C.o.V. of 0.1  
945 and 0.05 respectively. The confidence of these estimates can be improved by using a larger sample size at  
946 the expense of increasing even more computational cost of the analysis.

947 The search for lower and upper bounds is performed by means of Monte Carlo optimization using Latin  
948 Hypercube sampling, with approximately 50000 samples. A total of  $5 \times 10^7$  evaluations of the function  $\mathbf{x\_to\_g}$   
949 (model  $\mathbf{f}$ ) are thus, required to complete the analysis. Here, Monte Carlo is a convenient method to solve the  
950 optimization, as the objective functions  $J_1$  and  $J_2$  can be quite noisy, varying approximately between  $\mp 10\%$   
951 of the true value. In order to reduce the effect of the estimation error introduced by using finite sample sets,  
952 the objective functions *maximum* and *minimum* of  $J_{i=1,2}$ , are redefined as lower  $J_i (1 - t_{\alpha/2} \text{ C. o. V.})$  and  
953 upper  $J_i (1 + t_{\alpha/2} \text{ C. o. V.})$  estimations, respectively, where  $\alpha = 0.14$  and  $t_{\alpha/2} = 1.48$  is the 86th t-Student  
954 percentile (see also<sup>52</sup>).

955 Note that, in order to run the analysis within a reasonable time, parallelization lies at the foundations  
956 of this approach. On a common dual-core personal computer, a single estimation of  $J_i$  takes approximately  
957 3.4 minutes, thus a total of  $\sim 120$  days for a complete analysis. By means of a double parallelization, as

described in Section IV, it has been possible to reduce the running time by two orders of magnitude, making it possible to complete the analysis in just  $\sim 80$  hours.

PROPAGATION OF FOCAL SETS (COUNTER APPROACH) Using the propagation of focal sets method,  $n = 1000$  random vectors  $\{\alpha_j, j = 1, 2, \dots, n\}$  are drawn from the aleatory space  $\Omega \equiv (0, 1]^{17}$ . Thereafter, the procedure described in Section III-C-2 was applied. In order to evaluate equations (3) and (4), genetic algorithms with a population of 125 individuals and 50 generations are adopted requiring a total computational cost of  $5 \times 10^6$  evaluations of  $w$ . Figure 16 shows the convergence of the genetic algorithms for two representative focal elements. The convergence is achieved using 30 generations for the identification of the minimum/maximum of the Eq. (4).

For this approach, parallelization is also essential. In fact, approximately  $5 \times 10^6$  evaluations of the function `x_to_g` are required to complete a full analysis. Although, in this case, the use of GA makes the parallelization a little more articulated (jobs need to be sent at any iteration of the algorithm), it is still possible to significantly reduce the running time up to two orders of magnitude (as in the standard approach). It is worth noting that the overall number of function evaluations makes this approach about 10 times more efficient than the standard approach.

RESULTS The results of the reduced uncertainty model and the improved model are summarized in Table 5. Using the proposed methods, it has been possible to bound the actual solution for the targeted metrics. As expected, the improved uncertainty model is far more informative than the reduced model, which is shown by a sensible reduction in the upper bound of  $J_1$ . An even more significant difference is documented for the range of  $J_2$  (see Table 5), where the model of uncertainty from being totally uninformative,  $J_2 \in [0, 1]$ , is reduced to  $J_2 \in [0.20, 0.41]$ . Note also that the optimization in the epistemic space (standard approach) provided tighter bounds than the propagation of focal sets (counter approach). This result was expected inasmuch as, the random set methodology cannot cope with distributional probability boxes and has to treat them as distribution-free p-boxes, as discussed in Section III-C-2.

The computational costs using the optimization approach in the epistemic space is less intensive than the propagation of focal sets inasmuch as only four optimization tasks are required to find the lower and upper bounds of  $J_1$  and  $J_2$  while the counter approach requires a pair of optimization tasks for each focal

Table 5: Bounds of the variable  $J_1$  and  $J_2$  for the reduced and improved uncertainty model obtained by means of the two proposed approaches (Optimization in the epistemic space and Propagation of focal sets), respectively.

Reduced Uncertainty model	Improved Uncertainty model	Strategy
$J_1 = [1.37 \times 10^{-2}, 4.97]$	$J_1 = [2.88 \times 10^{-2}, 1.11]$	Optimization in the epistemic space
$J_2 = [6.4 \times 10^{-2}, 0.82]$	$J_2 = [0.24, 0.38]$	
$J_1 = [-1.57 \times 10^{-4}, 54.05]$	$J_1 = [-1.10 \times 10^{-4}, 3.05]$	Propagation of focal sets
$J_2 = [0, 1]$	$J_2 = [0.20, 0.41]$	

985 element and for each quantity of interest (i.e.  $J_1$  and  $J_2$ ). Both approaches are based on global optimization  
986 strategies and hence, they both suffer from the curse of dimensionality. The approaches proposed require  
987 an increasingly larger sample size (number of individuals and generations) in order to explore properly the  
988 optimization domain. In consequence, it is no longer guaranteed that the calculated optima are actually  
989 the global ones. In uncertainty propagation problems, missing the global optima means computing ranges  
990 of the targeted variables that are narrower than the sought ones. In this case, the methods result in an  
991 under(inner)-estimation of the actual solution, which may lead to an under-prediction of e.g. the failure  
992 probability of the system.

#### 993 **D. Subproblem D**

994 Subproblem D aims at identifying the epistemic realizations that lead to the smallest and largest values of  $J_1$   
995 (task D1) and  $J_2$  (task D2). The extreme case analysis has been performed both for the reduced uncertainty  
996 model and the improved uncertainty model, as requested. However, for conciseness, only results from the  
997 improved model will be herein presented.

998 The extreme case analysis in presence of uncertainty is an ill posed inverse problem. The direct identi-  
999 fication of the epistemic realizations,  $\theta$ , leading to the maximum/minimum of  $J_1$  and  $J_2$  from the forward  
1000 simulation has not been possible. Further, due to the complexity of the problem (in terms of nonlinearity  
1001 and computational costs), a specific strategies has been developed as explained in the following section.

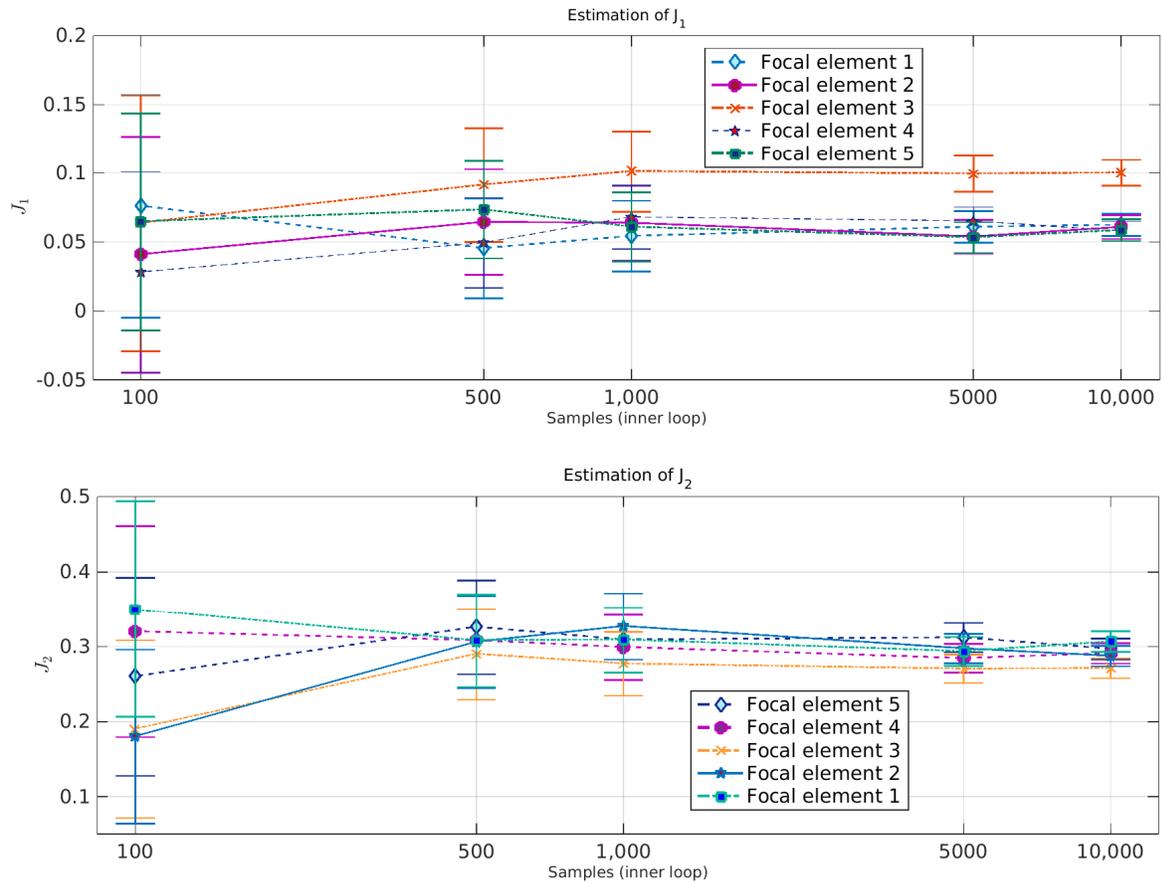


Figure 15: Effect of the number of samples sampled from the aleatory space in the inner loop on the estimation of  $J_1$  and  $J_2$ , respectively, in the optimization in the epistemic space approach applied for the solution of subproblem C.

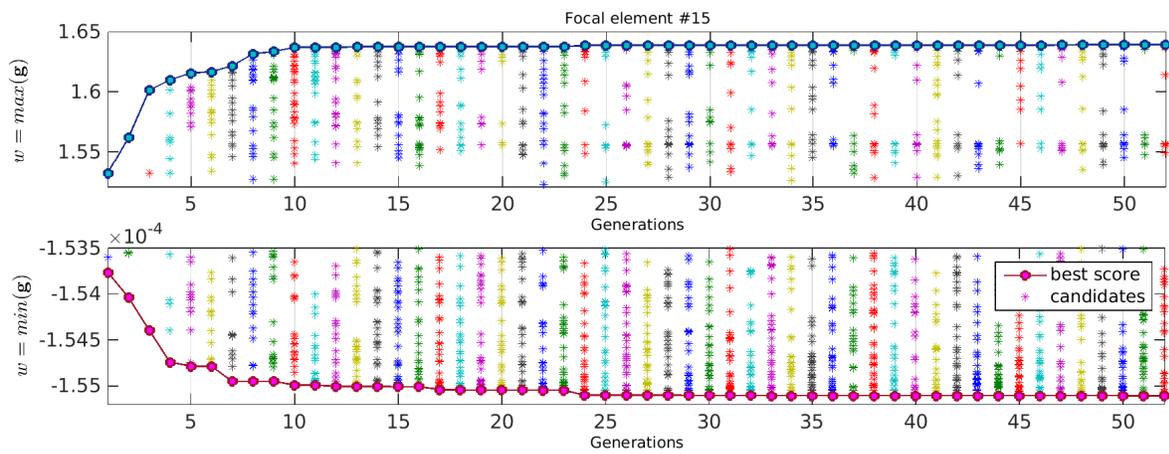


Figure 16: Convergence of the objective function  $w$  to the *minimum* and *maximum* for a representative focal element. Genetic Algorithms have been used with a population of 1000 individuals to identify the realizations in the epistemic space that minimize and maximize the objective function  $w$ .

1002 1. Extreme values of  $J_1$  (task D1)

1003 In this task we are focusing on  $J_1 = E[w]$  that is the expectation (mean) of the worst-case requirement  
 1004 metric:  $w = \max_{i=1:8}(g_i)$ . In order to be able to identify the realizations of the inputs  $\mathbf{p}$  that produce the  
 1005 extreme values of  $J_1$ , the relationships among intermediate variables,  $\mathbf{g}$ ,  $\mathbf{x}$  and  $\mathbf{p}$  are analyzed.

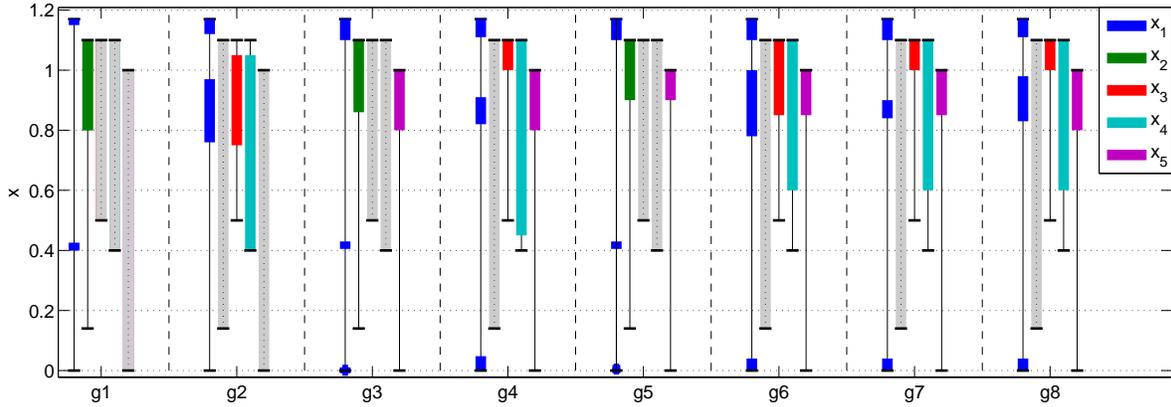


Figure 17: Analysis of the performance function  $\mathbf{g}$  with respect to the output of the subdisciplines,  $\mathbf{x}$ . In the plot the ranges of  $x_{i=1:5}$  leading to large positive values of  $g_{i=1:8}$  are shown using coloured bars. Grey bars (and dashed lines) indicate variables that are not important for the maximum of the corresponding performance  $g_i$ .

1006 DEPENDENCE OF  $J_1$  ON  $w$  The extreme values of  $J_1$  depend on the presence of very large (but rare) values  
 1007 of  $w$  (hereafter indicated as *outliers* of  $w$ ). The outliers of  $w$  can assume values  $w > 1000$ , while the most  
 1008 probable values of  $w$  are limited to values around 0. Two well distinct classes for  $w$  have been identified. A  
 1009 first class identifies values where  $w < 3$ , and a second class identifies the outliers, where  $w > 100$  and have  
 1010 values as high as 1000. Hence,  $J_1$  may assume its smallest value only if no outliers are present. On the other  
 1011 hand, the more outliers are present, the larger the value of  $J_1$ .

1012 DEPENDENCE BETWEEN  $\mathbf{g}$  AND  $\mathbf{x}$  Next, the dependence between the performance functions of the system  
 1013  $\mathbf{g}$  and the output of subdisciplines  $\mathbf{x}$  is analysed. The interest is to identify values (and ranges) of  $\mathbf{x}$  that  
 1014 produce the maxima of the performance functions  $\mathbf{g}$ .

1015 This study is performed by means of an optimization procedure where  $g_{i=1:8}$  are the objective functions

1016 to be maximized and  $\mathbf{x}$  are the search variables. Genetic Algorithm with 243 individuals and 50 generations  
 1017 is used for analysing each performance function  $g_i$ . The results are shown in Figure 17. The analysis of  
 1018 the function  $x\_to\_g$  (i.e. the model  $f$ ) has revealed that only the performance functions  $g_{i=3:8}$  yield values  
 1019  $w > 100$ , while  $g_1$  and  $g_2$  are always lower than 1 and 2.8, respectively.

1020 Then, the individuals that produce  $g_1 > 0.1$ ,  $g_2 > 0.1$  and  $g_{i=3:8} > 100$  are collected and shown in  
 1021 Figure 17 using coloured bars.

1022 Some variables, shown in the Figure using grey color and dashed line, do not influence the maximum of  
 1023 the performance functions (i.e. they can assume any value within their bounds).

1024 From Figure 17 critical sets (or regions) for each variable  $x_i$  can be identified. For instance, there are  
 1025 three sets of  $x_1$  able to produce values of  $g_4 > 100$ , namely  $x_1 \in [0, 0.05] \cup [0.82, 0.91] \cup [1.11, 1.17]$ .  
 1026 However, these sets have been found without taking into account the probability distributions associated to  
 1027 the inputs  $p_{i=1:5}$ . The most probable regions of  $\mathbf{x}$  has been identified by means of the double loop Monte  
 1028 Carlo simulation used in Section C.

1029 Interestingly, the most probable realizations of  $\mathbf{x}$  that produce outliers of  $w$  belong to a very clear pattern  
 1030 of coordinates, as shown in Figure 18.

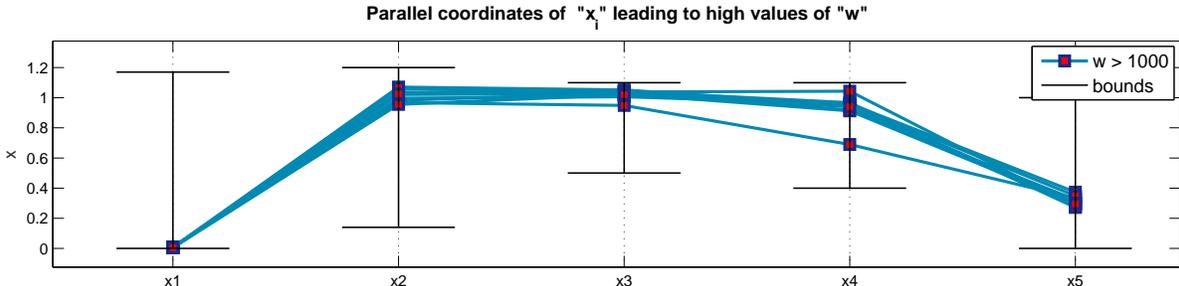


Figure 18: Parallel coordinates of  $x_{i=1:5}$  leading to the outliers of  $w$ . The plot shows also the bounds of the variables  $x_{i=1:5}$  identified for the improved uncertainty model.

1031 **DEPENDENCE BETWEEN  $\mathbf{x}$  AND  $\mathbf{p}$**  Once the regions of  $\mathbf{x}$  that produce the outliers of  $w$  have been identified,  
 1032 it is necessary to establish if such critical sets can be produced by any feasible realizations of inputs  $\mathbf{p}$ . This  
 1033 analysis has been performed by studying the functions  $p\_to\_x$  (i.e. the model  $h$ ) by using a double loop  
 1034 Monte Carlo approach, with an outer loop of 10000 Latin Hypercube samples (for the epistemic uncertainty,

1035  $\theta$ ) and an internal loop (for the aleatory uncertainty,  $\alpha$ ) of 1000 samples.

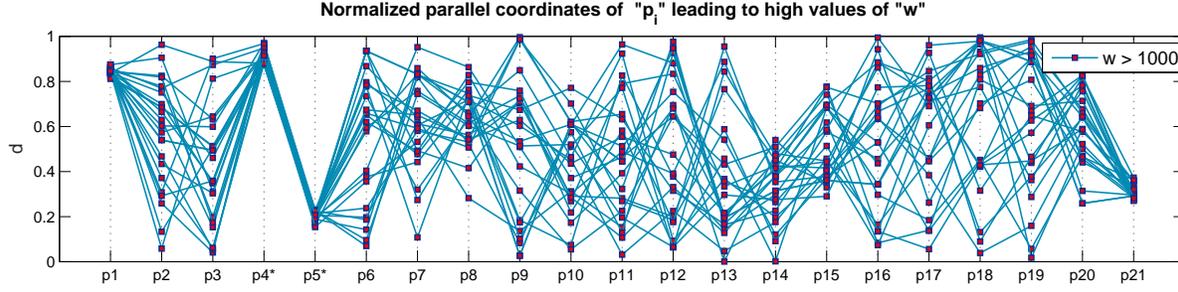


Figure 19: Parallel coordinates of the inputs  $p_{i=1:21}$  leading to values of  $w > 1000$  and  $J_1 > 1.0$ . The y-axis has been normalized between the lower and upper bound of the inputs  $p_{i=1:3,6:21}$ .  $p_4$  and  $p_5$  have been normalized between  $*[-5, 5]$ .

1036 EPISTEMIC REALIZATION THAT PRODUCE MAXIMUM OF  $J_1$  Figure 19 shows the identified realizations of  
 1037  $\mathbf{p}$  that produce critical values of  $\mathbf{x}$  (as shown in Figure 17). Only some inputs can lead unequivocally to  
 1038 the critical values of  $\mathbf{x}$ , namely  $p_1, p_4, p_5, p_{14}, p_{15}$  and  $p_{21}$ . In the matter of  $p_4$  and  $p_5$ , only values in the  
 1039 region where  $3.72 < p_4 < 4.70$  and  $-3.46 < p_5 < -2.70$  can produce  $x_1$  in the critical set and hence leading  
 1040 to large values of  $w$ . Since  $p_4$  and  $p_5$  are normally distributed, it is possible to select distributions peaked  
 1041 around the identified region as shown in Table 6. The epistemic realizations of Table 6 are calculated by  
 1042 maximizing the joint probability  $\pi_{p_4 p_5} = P[3.72 < p_4 < 4.70, -3.46 < p_5 < -2.70]$ . Using the distribution  
 1043 parameters reported in Table 6, such target maximum probability is  $\max_{\Theta}(\pi_{p_4 p_5}) = 0.9912$ .

Table 6: Epistemic realizations of  $p_4$  and  $p_5$  leading to the maximum of  $J_1$ . The parameters of the multivariate distribution are calculated maximizing the probability  $\pi_{p_i}$  of being inside the specified ranges (i.e. Critical range  $R_c$ ).

	$\pi_{p_i}^{\max}$	Critical range ( $R_c$ )	Epistemic real.		
$p_4$	0.9912	$3.72 < p_4 < 4.70$	$E(p_4) = 4.21$	$V(p_5) = \underline{V}(p_4)$	$\rho = 0$
$p_5$	0.9912	$-3.46 < p_5 < -2.70$	$E(p_5) = -3.04$	$V(p_5) = \underline{V}(p_5)$	$\rho = 0$

1044 Epistemic realizations corresponding to parameters  $p_1, p_{14}, p_{15}$  and  $p_{21}$  are also calculated in a similar  
 1045 way. Table 7 show the epistemic realizations of these inputs corresponding to the critical values, and the

1046 second column shows the corresponding values of the maximum probabilities  $\pi_{p_i}$ .

Table 7: Epistemic realizations of  $p_1$ ,  $p_{14}$ ,  $p_{15}$  and  $p_{21}$  leading to the maximum of  $J_1$ . These realizations maximize the probability of the input parameter  $\pi_{p_i}$  of being inside the specified ranges ( $R_c$ ).

	$\pi_{p_i}^{\max}$	Critical range ( $R_c$ )	Epistemic real.	
$p_1$	0.141	$0.81 < p_1 < 0.9$	$E(p_1) = \bar{E}(p_1)$	$V(p_1) = \underline{V}(p_1)$
$p_{14}$	0.854	$0.00 < p_{14} < 0.54$	$a(p_{14}) = \underline{a}(p_{14})$	$b(p_{14}) = \bar{b}(p_{14})$
$p_{15}$	0.940	$0.29 < p_{15} < 0.78$	$a(p_{15}) = \bar{a}(p_{15})$	$b(p_{15}) = 6.498$
$p_{21}$	0.077	$0.27 < p_{21} < 0.45$	$a(p_{21}) = \bar{a}(p_{21})$	$b(p_{21}) = \underline{b}(p_{21})$

1047  $p_1$  and  $p_{21}$  are somehow problematic inputs in the determination of the epistemic realization. By analysing  
 1048 the realizations from the input parameters  $p_1$  and  $p_{21}$ , it can be seen that critical values of  $\mathbf{x}$  are obtained  
 1049 when  $0.805 < p_1 < 0.902$  and  $0.27 < p_{21} < 0.45$ , respectively. However, from the p-boxes associated to these  
 1050 inputs (see Figure 20), it is not possible to select any CDF within in the p-box of  $p_1$  and  $p_{21}$  that permits  
 1051 to exclude (or include) completely the critical realizations (shown as round dots in Figure 20).

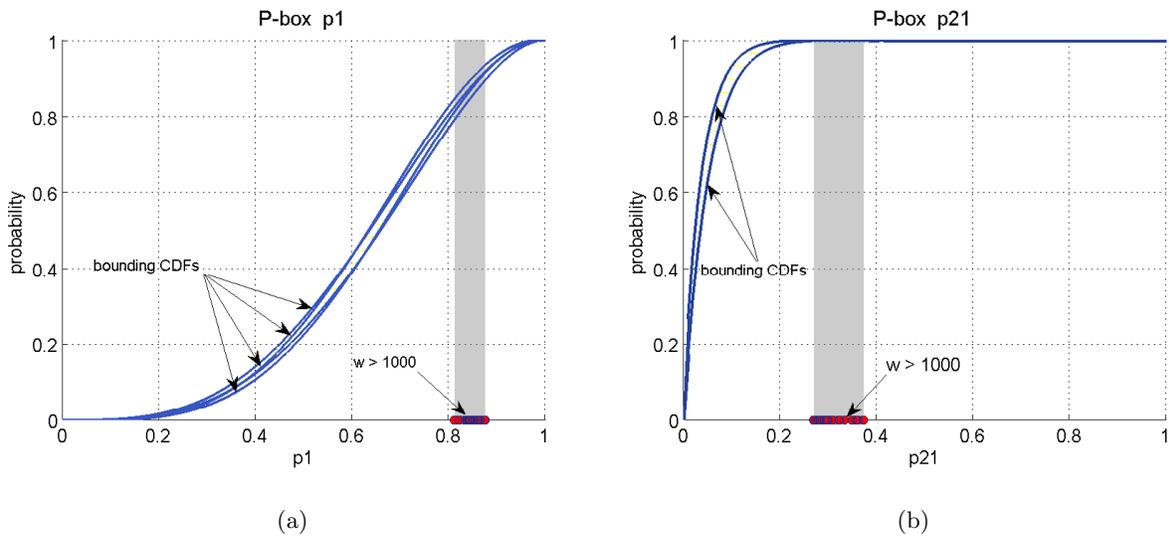


Figure 20: P-box representation of parameter  $p_1$  and  $p_{21}$ , respectively. The figures show the ranges of values that produce critical values of  $\mathbf{x}$  (and in turn large values of  $w$ ).

1052 Epistemic realization of the remaining parameters  $\mathbf{p}$ , which do not appear to have influence in the

1053 generation of the critical values of  $w$  (see Figure 19), have been obtained by maximizing the probability  
 1054  $\nu_c = P[w > 1000 \mid p_i \in R_c(p_i)]$  for  $i = 1, 4, 5, 14, 15, 21$ . A random search for the maximum values of the  
 1055 mean of  $p_i$  has been performed. 1000 aleatory samples have been used to calculate the above conditional  
 1056 probability. The results are reported in Table 8. The maximum identified frequencies is  $\nu_c^{\max} = 0.572$  and  
 1057 minimum  $\nu_c^{\min} = 0.261$ . These values are quite close meaning that the epistemic uncertainty may play a  
 1058 secondary role for the extreme value of  $J_1$ .

1059 The parameters of the p-boxes have been calculated using the identified values of  $E[p_i]$  and the maximum  
 1060 admissible value for  $V[p_i]$ .

Table 8: Epistemic realization that are very likely to produce the maximum of  $J_1$ . The realization has been identified maximizing the probability  $\nu_c$ .

Parameter	Epistemic real.	Parameter	Epistemic real.
$p_2$	0.719	$p_{12}$	$\underline{p_{12}}$
$p_6$	0.760	$p_{13}$	$a = 0.45, b = \bar{b}$
$p_7$	$a = \underline{a}, b = 0.73$	$p_{16}$	0.590
$p_8$	$a = \bar{a}, b = \underline{b}$	$p_{17}$	$a = \underline{a}, b = 1.32$
$p_{10}$	$a = 3.55, b = \bar{b}$	$p_{18}$	$a = 3.26, b = \underline{b}$
$p_{12}$	$\underline{p_{12}}$	$p_{20}$	$a = 10.68, b = \bar{b}$

1061 The realization leading to the minimum of  $J_1$  can be directly identified from results of task C1 (see  
 1062 Section C). The results are summarized in Figure 21.

## 1063 2. Extreme values of $J_2$

1064 The task D2 asks to identify the extreme case for metric  $J_2$ , where  $J_2 = P[w \geq 0]$  is the failure probability  
 1065 of the worst-case requirement metric  $w = \max_{i=1:8}(g_i)$ . Differently from  $J_1$ , this metric is not sensitive to the  
 1066 largest values of  $w$ . A double loop Monte Carlo approach has been adopted to solve this problem. 1000  
 1067 aleatory samples have been used to compute the failure probability  $J_2$ . It is known from Section C that  
 1068 both lower and upper bounds of  $J_2$  are greater than  $10^{-1}$ , hence 1000 samples are enough for a sufficiently

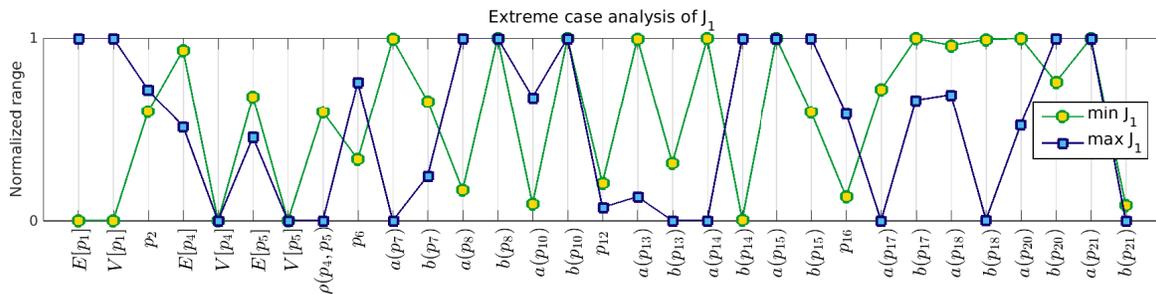


Figure 21: Extreme case analysis of  $J_1$ : parallel plot of the epistemic parameters. The y-axis represents normalized values of the epistemic variables.

robust estimation of  $J_2$  in the analysis.

The realizations of the input parameters  $\mathbf{p}$  that produce the extreme values of  $J_2$  are shown in Figure 22. Results from this analysis show, as expected, that realizations leading to the maximum (minimum) of  $J_2$  are generally different from those leading to the maximum (minimum) of  $J_1$ . It is also noted that many realizations are very close to the bounds of the epistemic domain.

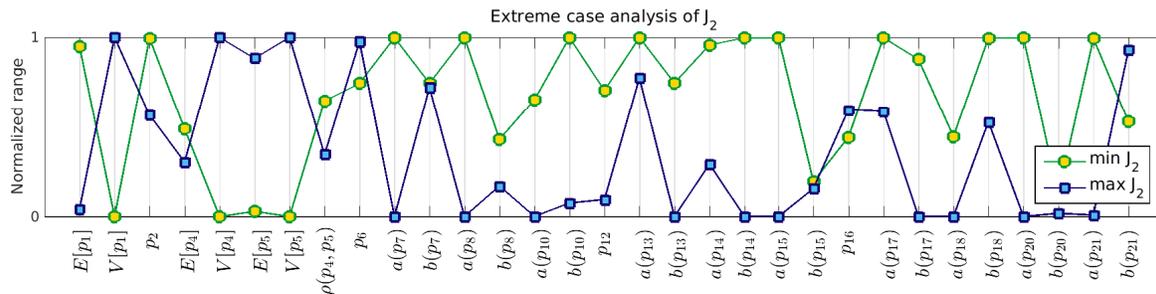


Figure 22: Extreme case analysis of  $J_2$ : parallel plot of the epistemic parameters. The y-axis represents normalized values of the epistemic variables.

### 3. Solution of task D3

In task D3, it is asked to identify some representative realizations of  $\mathbf{x}$  that typify different failure scenarios. The results of this task have already been discussed in Section D-1 and visualised in Figure 17. Overall, the following failure scenarios have been identified:

- Values of  $x_{i=1:5}$  close to their upper bounds lead to large values of  $g_{i=1:8}$ ;
- Small values of  $x_1$  combined with large values of  $x_{i=2,3,4,5}$ , lead to values of  $g_{i=3,4,5,6,7,8} > 1000$ ;

- Values of  $x_1 \in [0.84, 0.9]$  combined with large values of  $x_{i=2,3,4,5}$ , lead to values of  $g_{i=4,6,7,8} > 1000$
- Values of  $x_1 \in [0.4, 0.425]$ , combined with large values of  $x_{2,3,4,5}$ , lead to values of  $g_1 > 0.1$  and  $g_{i=3,5} > 1000$ .

Analysing the results of the simulations used in Section C ( Genetic Algorithm with 125 individuals and 45 generations), it is also possible to study the relationship between  $\mathbf{x}$  and  $\mathbf{g}$ . For example, large positive values of  $g_5$ , whose maximum is  $g_5^{\max} = 1021$ , are insensitive to  $x_3$  and  $x_4$ . This can be appreciated in Figure 23, where the evolution of objective function  $g_5$  and search variables  $x_i$  are represented. During the optimization the values of variables  $x_3$  and  $x_4$  change frequently, despite that the value of the objective remains the same. Analogously, for the other performances, it is found that large positive values of  $g_4$ ,  $g_6$ ,  $g_7$  and  $g_8$  are totally insensitive to  $x_2$  and slightly insensitive to  $x_4$ .

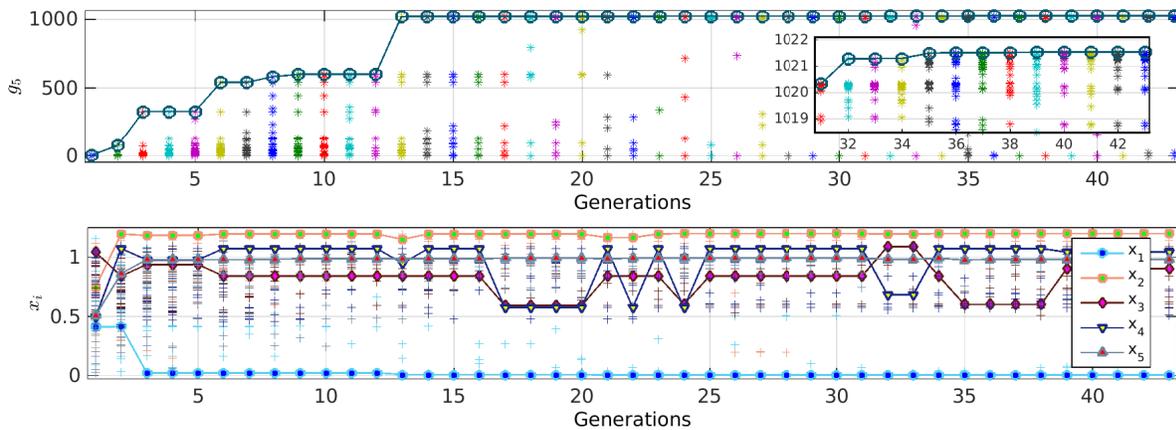


Figure 23: Evolution of the objective function  $g_5$  and search variables  $x_i$ . Note that the values of variables  $x_3$  and  $x_4$  change frequently during the optimization, despite the values of the objective remain the same.

## 1090 E. Subproblem E

1091 The last task of the challenge problem is to perform a robust design of the multidisciplinary system: per-  
 1092 forming an optimization able to identify the design point  $\mathbf{d}$  with improved robustness and reliability charac-  
 1093 teristics. This requires to perform UQ for each candidate solutions leading to unmanageable computational  
 1094 costs. Hence, it is necessary to adopt surrogate models. Here, it has been decided to replace with surro-  
 1095 gate models only the computational costly part of the model and keeping the original functions for the less

1096 demanding parts. In fact, training a surrogate model to approximate the non-linear, an noisy, functions  $J_1$   
1097 and  $J_2 : \mathbb{R}^{21} \times \mathbb{R}^{14} \rightarrow \mathbb{R}$  would have required a huge number of training samples with no warranties on the  
1098 quality of the approximation.

1099 The subproblems E1 and E2 require the calculation of  $w(\mathbf{p}, \mathbf{d}) = \max_{i=1, \dots, 8} g_i(\mathbf{x})$ , where  $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}, \mathbf{d})$   
1100 and  $\mathbf{x} = \mathbf{f}(\mathbf{p})$  where the most computationally expensive part is the evaluation of  $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x}, \mathbf{d})$ . Multilayer  
1101 perceptron artificial neural networks<sup>53</sup> are suggested to speed up the calculation of function  $\mathbf{g} = \hat{\mathbf{f}}(\mathbf{x}, \mathbf{d}) :$   
1102  $\mathbb{R}^5 \times \mathbb{R}^{14} \rightarrow \mathbb{R}^8$ ; in other words, the artificial neural networks act here as nonlinear response surfaces.

1103 To train the artificial neural networks, training examples of  $\{\mathbf{x}, \mathbf{d}, \mathbf{g}\}$  are passed to an error backprop-  
1104 agation algorithm. A set  $\mathbf{d}^{(i)} : i = 1, 2, \dots, 2000$  of Latin Hypercube quasi-random points were generated  
1105 in the 14-dimensional space of the design variables. The design variables  $\mathbf{d}$  can theoretically assume any real  
1106 valued quantity, but they have been actually generated in a bounded space to generate a local surrogate  
1107 model. For the first local meta-model, the following bounds has been assigned to each design variable,  
1108  $d_i : [\min(0.5 * d_{i,baseline}, 1.2 * d_{i,baseline}), \max(0.5 * d_{i,baseline}, 1.2 * d_{i,baseline})]$ . Please note that the base-  
1109 line  $d_i$  can also be negative, and this definitions guarantees that the baseline is included in the bounds. In  
1110 case the optimization procedure would have found a optimum design laying on one of the bounds of the  
1111 training region, a new local surrogate model would have needed to be trained, around the identify optimum.  
1112 Then, the optimization procedure is restarted.

1113 The generation of samples of  $\mathbf{x}$  is more involved. One possible approach is to determine the bounding  
1114 box of  $\mathbf{x}$  using an optimization procedure over the function  $\mathbf{x} = \mathbf{h}(\mathbf{p})$  and then draw samples from this box.  
1115 However, with this approach samples will be drawn from regions where it is less probable to obtain values  
1116 of  $\mathbf{x}$ , and where the neural network does not need to give an accurate prediction.

1117 To concentrate the generation of training samples only in the region of space of higher probability of  $\mathbf{x}$ ,  
1118 a set  $\boldsymbol{\theta}^{(i)} : i = 1, 2, \dots, 2000$  of Latin Hypercube quasi-random points was generated in the 31-dimensional  
1119 box of the epistemic space. For each set  $\{\boldsymbol{\theta}, \mathbf{d}\}^{(i)}$ , 200 Monte Carlo samples are generated in the aleatory  
1120 uncertainty space, obtaining  $\mathbf{x}^{(j)} : j = 1, 2, \dots, 200$  realizations of the function  $h(\mathbf{p})$ . The main draw back of  
1121 this procedure is that few samples will be generated in the tails of the distributions, thus the neural network  
1122 will perform badly in the prediction of the extreme values.

1123 In the end, 400000 points  $\mathbf{x}, \mathbf{d}$  are available to compute the model outputs  $\mathbf{f}(\mathbf{x}, \mathbf{d})$  and then train the

1124 surrogate model  $\hat{\mathbf{f}}(\mathbf{x}, \mathbf{d})$ . Since the minimum values  $g_j$ , for variables  $j = 3$  to 8 are very close to zero and  
 1125 very small compared to the respective maximum value of variable  $g_j$ , the following nonlinear transformation  
 1126 of variables  $g_j$  for  $j = 3, 4, \dots, 8$  is employed:

$$z(g_j) := \frac{1}{200|\min(g_j)|} - \frac{1}{100(g_j + 2|\min(g_j)|)} \quad (34)$$

1127 This nonlinear transformation of variables  $g_j$  stretches the interval of  $g_j$  for those values that are close to  
 1128 zero but preserving the sign of  $g_j$ . This is a very important characteristic since zero represents the limit  
 1129 value between the failure and the safe region.

1130 Using nonlinear transformation Eq. (34), we set map  $g_j$  to:

$$y_j = \begin{cases} g_j & \text{for } j = 1, 2 \\ 100z(g_j) & \text{for } j = 3, \dots, 8 \end{cases} \quad (35)$$

1131 in this way,  $\mathcal{T} = \{(\mathbf{x}^{(s)}, \mathbf{y}^{(s)}) : s = 1, 2, \dots, 400000\}$  served as the set of samples that were used for training,  
 1132 validating and testing the artificial neural networks.

1133 Given different levels of non-linearities in the relations between the inputs and each  $g_j$ , one multi layer  
 1134 perceptron has been trained for each  $g_j$ , and the optimal network architecture, i.e. characterized by the  
 1135 smallest regression error, has been identified for each output quantity. The first 300000 samples of  $\mathcal{T}$  have  
 1136 been used to train each multilayer perceptron using the Levenberg-Marquardt learning algorithm, a least  
 1137 squares curve fitting algorithm. The rest of the samples were used for validating and testing the artificial  
 1138 neural networks. The LGPL library FANN (Fast Artificial Neural Network),<sup>53</sup> integrated in OPENCOSSAN,  
 1139 has been used. Finally, the surrogate model will approximate  $\mathbf{g}(\mathbf{x})$  by applying the invers of the non-linear  
 1140 transformation of Equation 34

$$g_j = \begin{cases} y_j & \text{for } j = 1, 2 \\ \left( \frac{1}{2|\min(g_j)|} - y_j \right)^{-1} - 2|\min(g_j)| & \text{for } j = 3, \dots, 8 \end{cases} \quad (36)$$

1141 where  $y_j$  is the output of the artificial neural network. Tanks to the non-linear transformation, the neural  
 1142 network will provide a very accurate response for very small values of  $g$ , e.g., centered around 0, at the  
 1143 expenses of a less accurate prediction for values of bigger magnitude.

1144 The robust design requires to minimize the upper bound of  $J_1$  and  $J_2$  and those values need to be  
1145 estimated for each candidate design  $\mathbf{d}$ .

Genetic algorithms have been used to identify the optimal  $\mathbf{d}$  that minimize the largest value of  $J_1$  using a population size of 50 individuals. 100  $\alpha$ -cuts are randomly generated for the input parameter  $\mathbf{p}$  and 2500-sample internal Monte Carlo simulations are used to identify the upper and lower bounds of  $w_\alpha$ . This allows to estimate  $\overline{J_1}$  for each candidate design in approximately 25s, and hence leading to approximately 10 minutes of computational time for each generation using a local parallelization strategy (as explained in Section IV-B) on a Intel Xeon Processor E5-2450-v2 (8 cores at 2.5GHz). The identified optimum is:

$$\mathbf{d}_{E1} = [0.0140, -0.2568, -0.0944, -0.4405, -0.1508, -0.1029, -0.0713, \dots \\ 0.2002, -0.4431, 0.2579, 0.0044, -0.2086, 0.6330, -0.0166] \quad (37)$$

1146 corresponding to an optimum value of  $\overline{J_{1,opt}} = 0.0044$ . Subsequent run of the optimization algorithm  
1147 demonstrated that the optimum found is robust.

1148 The optimal design identified is better than the baseline in respect to the range of  $J_1$ . In fact, the range  
1149 of  $J_1$  at the optimum design is  $[\underline{J_1}, \overline{J_1}] = [1.798 \cdot 10^{-4}, 0.0044]$  and it is narrower than the range identified in  
1150 Table 5. A optimum design point  $\mathbf{d}$  that minimize the largest value of  $J_2$  has also been identified. Monte-  
1151 Carlo simulation has been used to compute the upper value of the probability of failure,  $\max(J_2)$ . In order  
1152 to reduce the coefficient of variation of the probability of failure estimator, 1500  $\alpha$ -cuts have been used. In  
1153 order to asses the robustness of the identified optimum, the optimization have been performed 3 times using  
1154 different initial populations and the results are shown in Table 9.

1155 Although, the maximum of  $J_2$  is very close, the identified design variables shows a large variability, in  
1156 particular variables  $d_1$ ,  $d_5$  and  $d_{14}$ . In order to asses the importance of these design variables with respect to  
1157 the computation of the upper bound of  $J_2$ , a global sensitivity analysis of the model with respect the design  
1158 variables has been carried out. The first order indices has been computed by means of the extended-FAST  
1159 method with 1000 samples. The most important (design) variables are  $d_9$ ,  $d_4$ ,  $d_{12}$  and  $d_{14}$ , and the remaining  
1160 variables has a lower, similar importance. It can be noticed that the range of  $J_2$  at the optimum is larger  
1161 than the range of  $J_2$  using the base design (Table 5). However, it is important to keep in mind that the aim  
1162 of the robust design is to reduce the upper bound of  $J_2$  and not its range.

run	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$	$d_6$	$d_7$	$d_8$	$d_9$
1	-0.0013	-0.2322	-0.0993	-2.0426	-0.2417	-0.1681	-0.0979	-0.4362	-0.5958
2	0.0052	-0.2658	-0.0874	-1.0996	0.2852	-0.1798	-0.0981	-0.4362	-0.5958
3	-0.0001	-0.2722	-0.1003	-1.6712	0.3191	-0.1640	-0.0981	-0.4362	-0.5958

run	$d_{10}$	$d_{11}$	$d_{12}$	$d_{13}$	$d_{14}$	$J_2$
1	0.0730	0.0053	-0.2012	0.5144	-0.0083	[0.0053 0.2973]
2	0.3337	0.0053	-0.2014	0.5875	0.0054	[0.0047 0.2993]
3	0.3230	0.0053	-0.2127	0.5641	0.0187	[0.0033 -0.3073 ]

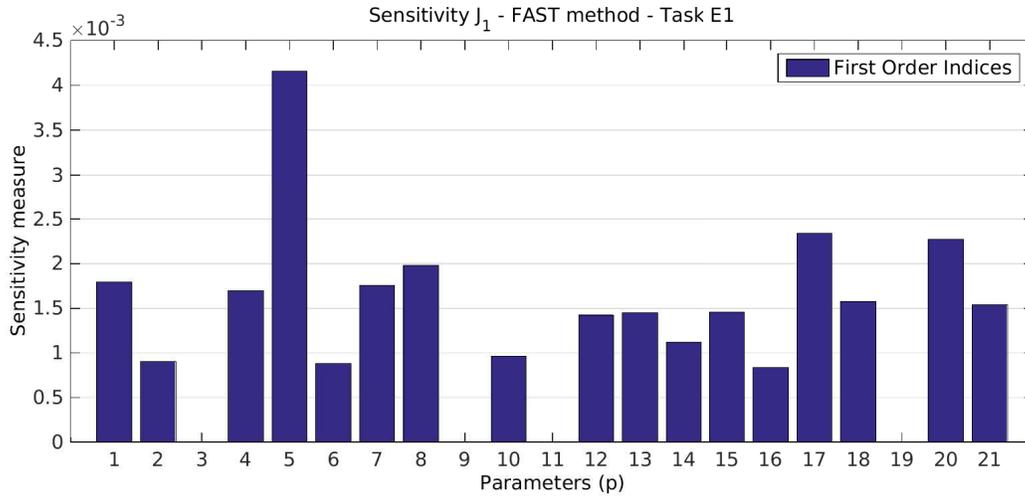
Table 9: Robust desing  $\mathbf{d}_{E2}$  with respect the upper bound of  $J_2$

1163 Finally, the sensitivity analysis for the identified design points that minimize the  $\overline{J_1}$  and  $\overline{J_2}$ , have been  
1164 rerun. The extended-FAST method have been computed using 16384 samples for each input factor for a  
1165 total cost of 507904 simulations. Each model evaluation requires the propagation of the aleatory uncertainty  
1166 and 500 Monte Carlo samples have been used. The sensitivity analysis has been performed using the real  
1167 model  $\mathbf{f}(\mathbf{x}, \mathbf{d})$  and not the surrogate model.

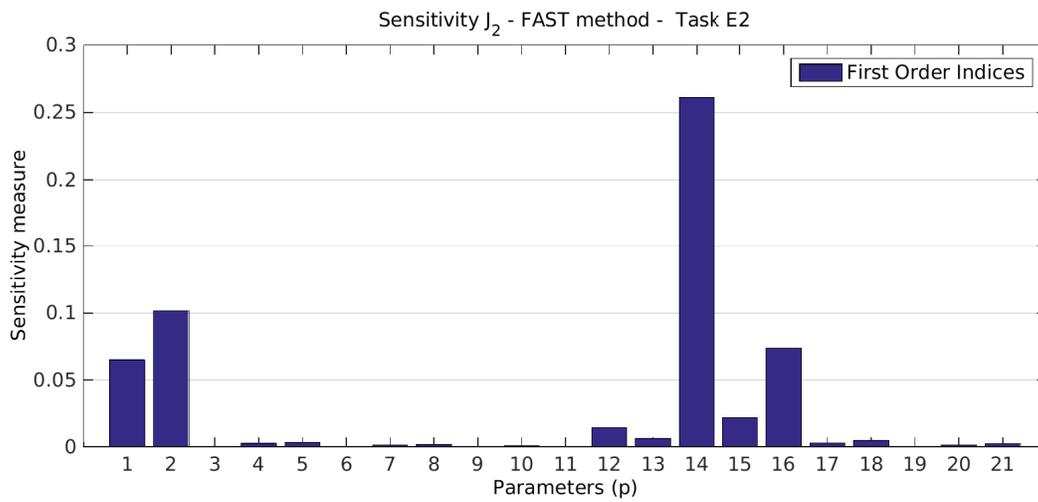
1168 Figure 24 shows the sensitivity analysis of the epistemic input factors respect the performances  $J_1$  and  $J_2$   
1169 evaluated at the design points  $\mathbf{d}_{E1}$  and  $\mathbf{d}_{E2}$ , respectively. The design point  $\mathbf{d}_{E1}$  seems to be very robust since  
1170 all the sensitivity measures are all very small (1 order of magnitude smaller compared to the sensitivity of  
1171 Section B. The most important factor that contribute the the variance of  $J_1$  is the expected value of  $p_5$ . The  
1172 most important factor that contribute the the variance of  $J_2$  is the parameter  $p_{14}$  followed by parameters  
1173  $p_2, p_1$  and  $p_{16}$ . It is interesting to notice that the sensitivity analysis using the baseline for  $\mathbf{d}$  does not show  
1174 the importance of  $p_{14}$  and  $p_2$ .

## 1175 VI. Conclusions

1176 The development and design of robust safety-critical systems is a challenging problem since in general  
1177 quantitative data is either very sparse or prohibitively expensive to collect. Moreover, the failure of such



(a) Respect  $J_1$  at  $d_{E1}$



(b) Respect  $J_2$  at  $d_{E2}$

Figure 24: First order sensitivity measure of the category II and III input factors evaluated. The first order sensitivity computed using extended-FAST method.

1178 systems might have severe consequences. In order to increase confidence and consistency in safety predictions,  
1179 modelling and simulation standards require estimates of uncertainty and descriptions of any processes used  
1180 to obtain these estimates.

1181 In this paper, a unified theory and an integrated and open general purpose computational framework to  
1182 deal with scarce data, aleatory and epistemic uncertainties has been presented. The proposed computational  
1183 tools are generally applicable to solve a reasonable large number of different problems and numerically  
1184 efficient and scalable. The applicability of the proposed strategy has been shown solving addressing the  
1185 NASA Langley UQ challenge problem.

1186 The presented results of this challenge problem clearly show that there are many ways of performing  
1187 analysis when different types of uncertainties, namely epistemic and aleatory, are present. All of these  
1188 methods have at some point made some weak or strong assumptions in order to find an answer. This  
1189 forms a sound basis for future improvements and developments. In fact, it is envisaged that this quantitative  
1190 comparison of the approaches will be most instrumental and useful for the engineering community, since it will  
1191 highlight the advantages and disadvantages of existing methods for the handling joint existence of epistemic  
1192 and aleatory uncertainty. As a general remark, the suggested procedures reveal the capability of random set  
1193 theory to represent without any assumption epistemic and aleatory uncertainty. A major drawback of the  
1194 proposed techniques was that many of them were based, up to some point in global optimization algorithms,  
1195 which is known to be difficult for noisy and high dimensional objective functions, and will lead to spurious  
1196 results when convergence to non-global optima occurs. It is left as an open problem how to circumvent that  
1197 optimization step when mapping focal elements through a function. Furthermore, the proposed techniques  
1198 are still very computational demanding requiring up to millions of model evaluations. Clearly this can only  
1199 be archived resorting to some sort of parallelizations strategies and to the computational power of cloud and  
1200 cluster computing.

1201 Considering different approaches to solve the same engineering problem might be seen a waste of resources  
1202 and time. However, all the existing approaches for dealing with epistemic and aleatory uncertainty require  
1203 fine tuning of their parameters in order to be efficient and accurate. Hence, it is of paramount importance to  
1204 be able to verify and cross-validate the results against different procedures. In this respect, the availability  
1205 of an open, flexible and modular computational framework implementing a number of different numerical



1235 <sup>9</sup>Crespo, L. G., Kenny, S. P., and Giesy, D. P., “The NASA Langley Multidisciplinary Uncertainty Quantification Chal-  
1236 lenge,” 2013.

1237 <sup>10</sup>Patelli, E., Broggi, M., and de Angelis, M., “OpenCossan: An efficient open tool for dealing with epistemic and aleatory  
1238 uncertainties,” *Vulnerability, Uncertainty, and Risk: Analysis, Modeling, and Management*, edited by I. S. A. . J. W. H.  
1239 Michael Beer, American Society of Civil Engineers, 2014.

1240 <sup>11</sup>Alvarez, D. A., “On the calculation of the bounds of probability of events using infinite random sets,” *International*  
1241 *Journal of Approximate Reasoning*, Vol. 43, 2006, pp. 241–267.

1242 <sup>12</sup>Alvarez, D. A., *Infinite random sets and applications in uncertainty analysis*, Ph.D. thesis, Arbeitsbereich für Technische  
1243 Mathematik am Institut für Grundlagen der Bauingenieurwissenschaften. Leopold-Franzens-Universität Innsbruck, Innsbruck,  
1244 Austria, 2007, Available at <https://sites.google.com/site/diegoandresalvarezmarin/RStthesis.pdf>.

1245 <sup>13</sup>Sklar, A., “Fonctions de répartition à  $N$  dimensions et leurs marges,” *Publ. Inst. Statist. Univ. Paris*, Vol. 8, 1959,  
1246 pp. 229–231.

1247 <sup>14</sup>Sklar, A., “Random variables, distribution functions, and copulas - a personal look backward and forward,” *Distributions*  
1248 *with fixed marginals and related topics*, edited by L. Rüschendorf, B. Schweizer, and M. Taylor, Institute of Mathematical  
1249 Statistics, Hayward, CA, 1996, pp. 1–14.

1250 <sup>15</sup>Nelsen, R. B., *An introduction to copulas*, Springer Series in Statistics, Springer, New York, 2nd ed., 2010.

1251 <sup>16</sup>Zhang, H., Dai, H., Beer, M., and Wang, W., “Structural reliability analysis on the basis of small samples: An interval  
1252 quasi-Monte Carlo method,” *Mechanical Systems and Signal Processing*, Vol. 37, No. 1 –2, 2013, pp. 137 – 151.

1253 <sup>17</sup>Dubois, D. and Prade, H., “Random sets and fuzzy interval analysis,” *Fuzzy Sets and Systems*, Vol. 42, No. 1, 1991,  
1254 pp. 87–101.

1255 <sup>18</sup>Arora (Ed.), J., *Optimization of Structural and Mechanical Systems*, World Scientific, 2007.

1256 <sup>19</sup>de Angelis, M., Patelli, E., and Beer, M., “Advanced line sampling for efficient robust reliability analysis,” *Structural*  
1257 *safety*, 2014 (accepted).

1258 <sup>20</sup>Dong, W. and Shah, H. C., “Vertex method for computing functions of fuzzy variables,” *Fuzzy Sets and Systems*, Vol. 24,  
1259 No. 1, 1987, pp. 65 – 78.

1260 <sup>21</sup>Zhang, H., Mullen, R. L., and Muhanna, R. L., “Interval Monte Carlo methods for structural reliability,” *Structural*  
1261 *Safety*, Vol. 32, No. 3, 2010, pp. 183–190.

1262 <sup>22</sup>Mares, C., Mottershead, J., and Friswell, M., “Stochastic model updating: Part 1–theory and simulated example,”  
1263 *Mechanical Systems and Signal Processing*, Vol. 20, No. 7, Oct. 2006, pp. 1674–1695.

1264 <sup>23</sup>Soize, C., “A comprehensive overview of a non-parametric probabilistic approach of model uncertainties for predictive  
1265 models in structural dynamics,” *Journal of Sound and Vibration*, Vol. 288, No. 3, Dec. 2005, pp. 623–652.

1266 <sup>24</sup>Efron, B. and Tibshirani, R., *An Introduction to the Bootstrap*, Chapman & Hall, New York, 1993.

1267 <sup>25</sup>Mosteller, F., “A k-sample slippage test for an extreme population,” *Annals of Mathematical Statistics*, Vol. 19, No. 1,  
1268 1948, pp. 58–65.

1269 <sup>26</sup>Kolmogorov, A., “Sulla determinazione empirica di una legge di distribuzione (In Italian),” *Giornale dell’Istituto Italiano*  
1270 *degli Attuari*, Vol. 4, 1933, pp. 83–91.

1271 <sup>27</sup>Pradlwarter, H. and Schuëller, G., “The Use of Kernel Densities and Confidence Intervals to cope with Insufficient Data in  
1272 Validation Experiments,” *Computer Methods in Applied Mechanics and Engineering*, Vol. 197, No. 29-32, 2008, pp. 2550–2560.

1273 <sup>28</sup>Silverman, B., *Density Estimation for Statistics and Data Analysis*, Chapman & Hall/CRC, 1998.

1274 <sup>29</sup>Ching, J. and Chen, Y., “Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class  
1275 Selection, and Model Averaging,” *Journal of Engineering Mechanics*, Vol. 133, No. 7, 2007, pp. 816–832.

1276 <sup>30</sup>Patelli, E., Alvarez, D. A., Broggi, M., and de Angelis, M., “An integrated and efficient numerical framework for uncer-  
1277 tainty quantification: application to the NASA Langley multidisciplinary Uncertainty Quantification Challenge,” *16th AIAA*  
1278 *Non-Deterministic Approaches Conference (SciTech 2014)*, AIAA SciTech, American Institute of Aeronautics and Astronautics,  
1279 2014.

1280 <sup>31</sup>Barber, S., Voss, J., and Webster, M., “The Rate of Convergence for Approximate Bayesian Computation,” *arXiv preprint*  
1281 *arXiv:1311.2038*, 2013.

1282 <sup>32</sup>Chiachio, M., Beck, J. L., Chiachio, J., and Rus, G., “Approximate Bayesian Computation by Subset Simulation,” *arXiv*  
1283 *preprint arXiv:1404.6225*, 2014.

1284 <sup>33</sup>Beck, J. L. and Katafygiotis, L. S., “Updating Models and their Uncertainties. I: Bayesian Statistical Framework,”  
1285 *Journal of Engineering Mechanics, ASCE*, Vol. 124, No. 4, 1998, pp. 455–461.

1286 <sup>34</sup>Klir, G. J. and Wierman, M. J., *Uncertainty-Based Information: Elements of Generalized Information Theory*, Vol. 15  
1287 of *Studies in Fuzziness and Soft Computing*, Physica-Verlag, Heidelberg, Germany, 1998.

1288 <sup>35</sup>Klir, G. J., *Uncertainty and Information : Foundations of Generalized Information Theory*, John Wiley and Sons, New  
1289 Jersey, 2006.

1290 <sup>36</sup>Alvarez, D. A., “Nonspecificity for infinite random sets of indexable type,” *Fuzzy Sets and Systems*, Vol. 159, No. 3,  
1291 2008, pp. 289 – 306.

1292 <sup>37</sup>Alvarez, D. A., “Reduction of uncertainty using sensitivity analysis methods for infinite random sets of indexable type,”  
1293 *International Journal of Approximate Reasoning*, Vol. 50, No. 5, 2009, pp. 750 – 762.

1294 <sup>38</sup>Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Salsana, M., and Tarantola, S., *Global*  
1295 *sensitivity analysis - The primer*, Wiley, 2008.

1296 <sup>39</sup>Sobol’, I., “Sensitivity estimates for nonlinear mathematical models.” *Math. Modeling & Computational Experiments*,  
1297 Vol. 1, 1993, pp. 407–414.

1298 <sup>40</sup>Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., and Schaibly, J. H., “Study of the sensitivity of coupled  
1299 reaction systems to uncertainties in rate coefficients. I Theory,” *The Journal of Chemical Physics*, Vol. 59, No. 8, 1973,  
1300 pp. 3873–3878.

1301 <sup>41</sup>Saltelli, A. and Bolado, R., “An alternative way to compute Fourier amplitude sensitivity test (FAST),” *Computational*  
1302 *Statistics & Data Analysis*, Vol. 26, No. 4, 1998, pp. 445 – 460.

1303 <sup>42</sup>Alvarez, D. A., “A Monte Carlo-based method for the estimation of lower and upper probabilities of events using infinite  
1304 random sets of indexable type,” *Fuzzy Sets and Systems*, Vol. 160, 2009, pp. 384–401.

1305 <sup>43</sup>Alvarez, D. A. and Hurtado, J. E., “An efficient method for the estimation of structural reliability intervals with random  
1306 sets, dependence modelling and uncertain inputs,” *Computers & Structures*, Vol. 142, 2014, pp. 54 – 63.

1307 <sup>44</sup>Zhigljavsky, A. and Zilinskas, A., *Stochastic global optimization*, Springer Optimization and Its Applications, Springer  
1308 Science & Business Media, 2007.

1309 <sup>45</sup>Patelli, E., Valdebenito, M. A., and Schuëller, G. I., “General Purpose Stochastic Analysis Software for Optimal Maintenance  
1310 Scheduling: application to a fatigue-prone structural component,” *International Journal Reliability and Safety, Special  
1311 Issue on: "Robust Design - Coping with Hazards Risk and Uncertainty"*, 2011.

1312 <sup>46</sup>Jin, R., Du, X., and Chen, W., “The use of metamodeling techniques for optimization under uncertainty,” *Structural and  
1313 Multidisciplinary Optimization*, Vol. 25, No. 2, July 2003, pp. 99–116.

1314 <sup>47</sup>Jensen, H. A., “Structural optimization of linear dynamical systems under stochastic excitation: a moving reliability  
1315 database approach,” *Computer Methods in Applied Mechanics and Engineering*, Vol. 194, No. 12-16, April 2005, pp. 1757–  
1316 1778.

1317 <sup>48</sup>Free Software Foundation, “GNU Lesser General Public License, <http://www.gnu.org/licenses/lgpl.html>,” .

1318 <sup>49</sup>Patelli, E., “The legacy of Prof. G.I. Schuëller and recent advances on uncertainty quantification and risk assessment,”  
1319 *Eurodyn 2014*, 2014 (in press).

1320 <sup>50</sup>Patelli, E., H.M.Panayirci, Broggi, M., Goller, B., Beaurepaire, P., Pradlwarter, H. J., and Schuëller, G. I., “General  
1321 purpose software for efficient uncertainty management of large finite element models,” *Finite Elements in Analysis and Design*,  
1322 Vol. 51, 2012, pp. 31–48.

1323 <sup>51</sup>Hammersley, J. and Handscomb, D., *Monte Carlo Methods*, Methuen & Co Ltd, London, New York, 1964.

1324 <sup>52</sup>Schneider, J. J. and Kirkpatrick, S., *Stochastic optimization*, Springer, 2006.

1325 <sup>53</sup>Nissen, S., “Implementation of a Fast Artificial Neural Network Library (fann),” Tech. rep., Department of Computer  
1326 Science University of Copenhagen (DIKU), 2003, <http://fann.sf.net>.