# Augmented Reliability Analysis for Estimating Imprecise First Excursion Probabilities in Stochastic Linear Dynamics

Matthias G.R. Faes<sup>a,b,\*</sup>, Marcos A. Valdebenito<sup>c</sup>, Xiukai Yuan<sup>d</sup>, Pengfei Wei<sup>b,e</sup>, Michael Beer<sup>b,f,g</sup>

<sup>a</sup>KU Leuven, Department of Mechanical Engineering, Technology campus De Nayer, Jan De Nayerlaan 5, St.-Katelijne-Waver, Belgium

<sup>b</sup>Institute for Risk and Reliability, Leibniz Universität Hannover, Callinstr. 34, 30167 Hannover, Germany <sup>c</sup>Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Av. Padre Hurtado 750, 2562340 Viña del Mar. Chile

<sup>d</sup>School of Aerospace and Engineering, Xiamen University, Xiamen 361005, P.R. China
<sup>e</sup>School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an 710072,
China

<sup>f</sup>Institute for Risk and Uncertainty and School of Engineering, University of Liverpool, Peach Street, Liverpool L69 7ZF, UK

<sup>g</sup>International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 1239 Siping Road, Shanghai 200092, P.R. China

Abstract

Imprecise probability allows quantifying the level of safety of a system taking into account the effect of both aleatory and epistemic uncertainty. The practical estimation of an imprecise probability is usually quite demanding from a numerical viewpoint, as it is necessary to propagate separately both types of uncertainty, leading in practical cases to a nested implementation in the so-called double loop approach. In view of this issue, this contribution presents an alternative approach that avoids the double loop by replacing the imprecise probability problem by an augmented, purely aleatory reliability analysis. Then, with the help of Bayes' theorem, it is possible to recover an expression for the failure probability as an explicit function of the imprecise parameters from the augmented reliability problem, which ultimately allows calculating the imprecise probability. The implementation of the proposed framework is investigated within the context of imprecise first excursion probability estimation of uncertain linear structures subject to imprecisely defined stochastic quantities and crisp stochastic loads. The associated augmented reliability problem is solved within the context of Directional Importance Sampling, leading to an improved accuracy at reduced numerical costs. The application of the proposed approach is investigated by means of two examples. The results obtained indicate that the proposed approach can be highly efficient and accurate.

Keywords: Uncertain Linear Structure, stochastic loading, Imprecise first excursion probability, Augmented reliability problem, Directional Importance Sampling

32

34

2

8

9

10

11

12

13

14

15

<sup>\*</sup>E-mail: matthias.faes@kuleuven.be

#### 35 Highlights:

- Imprecise reliability problem is replaced by an augmented reliability problem.
- Bayes' theorem allows retrieving relevant information out of augmented reliability.
- Focus on uncertain linear structures described by imprecise random properties subject to stochastic loads.
  - Augmented reliability problem is solved with Directional Importance Sampling.

#### 1. Introduction

40

Numerical tools to approximate the solution of (sets of) differential equations have become in-42 dispensable in the design of components from the micro-scale to complete structures, all subjected 43 to loads that are representative for those they encounter during their functional life. Thanks to these tools, an engineer is now able to design, test and optimize designs long before a first proto type is built. However, despite the highly detailed numerical predictions that can be obtained, the results of these calculations often show a non-negligible discrepancy with the actual physical behaviour of the structure. At the core of this discrepancy lies epistemic (= lack of knowledge) and aleatory (= caused by inherent variation) uncertainty in the description of the model physics, as well as the governing parameters. In recent years, several highly performing methods based on stochastic analysis [1], Fuzzy set theory and Interval analysis [2] have been introduced in literature to account for respectively aleatory and epistemic uncertainties in the model parameters. Also, recent studies compared the applicability of several of these techniques in applications such as geotechnical engineering [3] or inverse uncertainty quantification for stochastic dynamics [4, 5]. Uncertainties, be it epistemic or aleatory, are commonly encountered in the context of struc-55 tural dynamics. This is especially true when considering natural phenomena such as earthquakes or wind loads, since the corresponding dynamical loads that act on the system often cannot be described in a crisp way. Similarly, when considering natural materials and/or products manufactured using highly advanced production techniques, such uncertainties may arise. Stochastic processes, see, e.g. [6, 7], provide a rigorous framework to deal with the uncertainties and space/time correlations of uncertain loads and properties by resorting to the well-documented framework of probability theory, which is highly suited to deal with aleatory uncertainties. However, in practice, the analyst is often confronted with limited, incomplete or conflicting sources of data (i.e., epistemic uncertainty). In this case, the application of a pure probabilistic framework to take this additional level of uncertainty into account is questionable since in this case, there is simply not enough information to construct an objective probabilistic uncertainty model. In this case, by introducing probability density functions to represent lack of knowledge, one automatically inserts subjective information into the analysis, which might cause further divergence from the actual physics. The framework of imprecise probabilities (see, e.g. [3]) might in this case offer a more objective framework, since, rather than assuming a specific probability measure, it incorporates credal sets of probability measures to fully represent all sources of uncertainty. In practice, this allows to explicitly consider a set of probabilistic models, each consistent with the definition of the epistemic uncertainty, and to infer bounds on the possible model behaviour [8]. For example, in the context of structural dynamics, rather than computing a crisp value of the first excursion probability given a crisp stochastic process load (e.g., following the methods proposed in [9]), an imprecise probabilistic calculation provides bounds on the first excursion probability given a set of stochastic processes that are consistent with the epistemic uncertainty. This allows both to assess the sensitivity of the model reliability to the existing epistemic uncertainty, as well as to provide an estimate of the lower bound of the reliability. Similar observations can be made when considering imprecisely defined stochastic structural properties, which is the main focus of this contribution.

66

70

71

75

76

77

80

81

85

86

91

In engineering practice, however, the application of the framework of imprecise probabilities to infer bounds on the probability of failure is hindered by the computational cost of propagating both sources of uncertainty (aleatory and epistemic) jointly towards the model responses. The high computational cost is a direct result of the fact that the propagation has to be conducted such that the effects of aleatory and epistemic uncertainty are kept separated [10]. Hence, double loop approaches are typically applied, where the outer loop takes care of epistemic uncertainty while the inner loop deals with aleatory uncertainty. Many recent works deal with improving the numerical efficiency of propagation schemes for imprecise probabilities. In general four groups of approaches can be considered: (1) series expansion methods, (2) surrogate modelling schemes, (3) decoupling approaches and (4) augmented space techniques. Series expansion methods rely on the approximation of the epistemic uncertain parameters via series expansion methods (see e.g., [11],

[12]) or orthogonal polynomial expansion schemes (see e.g., [13]), effectively enabling propagation without double loop approaches. However, in case the epistemic uncertainty is comparatively large, perturbation approaches are known to be inaccurate [2], a problem that is alleviated by resorting to Chebyshev polynomial based schemes such as presented in [14]. Considering surrogate modelling schemes, many efficient techniques for the propagation of imprecise probabilistic problems have been proposed using sparse polynomial chaos expansion representations of the model (see e.g., [15, 16]), interval predictor models [17, 18] or variants of the Sobol-Hoeffding decomposition (also known as HDMR representation) of the relation between the epistemic parameters and the 100 probability of failure [19, 20], providing an efficient and accurate approximation of the problem. 101 Decoupling approaches, such as presented in recent work by some of the authors (see [21, 22, 23]) 102 are proven to be extremely efficient, but their application is limited to linear models. Finally, 103 the idea of using augmented space methods was introduced by [24] in the context of probabilistic 104 failure analysis and further developed for sensitivity analysis and reliability-based optimization 105 in, e.g. [25, 26, 27, 28, 29]. In the context of imprecise probabilities, similar methods were intro-106 duced by [19, 20, 30] and independently by [31] in a different form. Following these approaches, 107 the main idea is to propagate the epistemic and aleatory uncertainty jointly in a purely aleatory, 108 augmented space that is optimal with respect to a certain well-defined measure, in such a way 109 that both sources of uncertainty can be decoupled again at the response side. 110

This contribution is situated in the latter category: augmented space methods, and it aims at solving a particular class of problems, that is, those of bounding the first excursion probability of a system subjected to a Gaussian excitation, where epistemic uncertainty is present in the hyper-parameters of the distributions of several uncertain model quantities (i.e., parameterized p-boxes). Imprecise stochastic loads, such as considered in [21], are not considered in this contribution. The proposed approach converts the p-box valued uncertain parameters into aleatory uncertainties in an augmented space. Then, with the help of Bayes' theorem, it is possible to recover an expression for the failure probability as an explicit function of the imprecise parameters from the augmented reliability problem, which ultimately allows calculating the imprecise probability due to the established functional relationship. The implementation of the proposed framework is investigated within the context of imprecise first excursion probability estimation of uncertain linear structures where epistemic uncertainty is present in the hyper-parameters of the distributions of several uncertain model quantities. The associated augmented reliability problem

111

112

113

114

115

116

117

118

119

120

121

122

123

is solved within the context of Directional Importance Sampling [32], leading to an improved accuracy at reduced numerical costs. Furthermore, by virtue of Directional Importance Sampling, it is possible to derive an explicit approximation of the probability of failure as a function of the interval-valued hyper-parameters of the uncertain input quantities. The paper is complementary to earlier work by the authors, presented in [21, 22, 23], in the sense that it presents a conceptually totally different framework for the propagation of imprecise probabilities towards the bounds on first excursion probabilities. Indeed, whereas the work in [21, 22, 23] focuses on the operator norm framework to decouple the double loop, the methods presented in this paper are based on solving the problem in an augmented space.

125

132

135

137

139

152

The paper is structured as follows: Section 2 introduces the general framework of augmented methods. Section 3 discusses the application of the framework to the calculation of the first excursion probability of uncertain linear dynamical systems. Section 4 provides two examples to illustrate the application, efficiency and effectivity of the proposed framework. Section 5 lists the conclusions of the work and gives some outlook for future developments.

2. General Framework

#### 2.1. Imprecise Probability

Consider a numerical model that represents the behavior of a system. It is assumed that some input parameters of this model are uncertain. These uncertain parameters are grouped into two vectors z and y, whose dimensions are  $n_z \times 1$  and  $n_y \times 1$ , respectively, and which are assumed to be independent. The uncertainty associated with these two vectors is represented in terms of probability density functions  $f_Z(z)$  and  $f_Y(y|\theta)$ , respectively, where  $\theta$  is a vector of dimension  $n_{\theta} \times 1$  that collects distribution parameters such as mean value, standard deviation, etc. Due to issues such as lack of knowledge, imprecision, conflicting sources of information, etc.,  $\theta$  cannot be identified precisely and hence, it is described as an interval valued vector, whose lower and upper bounds are  $\underline{\theta}$  and  $\overline{\theta}$ , respectively. From the above discussion, it is noted that: the uncertainty associated with z is aleatoric; the uncertainty associated with  $\theta$  is epistemic; and that the uncertainty associated with z is both aleatoric and epistemic. In fact, the uncertainty associated with z falls into the category of a parametric probability box (p-box) with specified distribution function [33].

The response of the system under analysis becomes uncertain due to the uncertainty in its in-

put parameter vectors z and y. Such response is monitored in terms of a performance function g, which assumes a value equal or smaller than zero whenever an unacceptable system behavior occurs. In other words, unacceptable behavior occurs whenever  $g(z, y) \leq 0$ . It is important to note that the performance function does not depend on  $\theta$ : recall that this is a vector collecting distribution parameters which do not affect the physical behavior of the system but just the probabilistic description associated with y. The chances that an unacceptable behavior occur are measured in terms of a probability of failure  $p_F$ , which is defined as:

$$p_F(\boldsymbol{\theta}) = \int_{\boldsymbol{y} \in \Omega_y} \int_{\boldsymbol{z} \in \Omega_z} I_F(\boldsymbol{y}, \boldsymbol{z}) f_{\boldsymbol{Z}}(\boldsymbol{z}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}) d\boldsymbol{z} d\boldsymbol{y}, \tag{1}$$

where  $\Omega_y$  and  $\Omega_z$  are the sets containing all values that  $\boldsymbol{y}$  and  $\boldsymbol{z}$  may assume, respectively; and  $I_F(\boldsymbol{y}, \boldsymbol{z})$  is the indicator function, which is equal to one in case  $g(\boldsymbol{y}, \boldsymbol{z}) \leq 0$  and zero, otherwise. It is noted from the equation above that the failure probability  $p_F$  does depend on  $\boldsymbol{\theta}$ . This is evident, as  $\boldsymbol{\theta}$  affects the probabilistic description of the problem. Furthermore, as the uncertainty associated with  $\boldsymbol{\theta}$  is described in terms of an interval, the failure probability itself becomes interval valued as well. In other words, due to the imprecision associated with the vector of distribution parameters, the failure probability is actually imprecise. Its lower and upper bounds (denoted as  $\underline{p}_F$  and  $\overline{p}_F$ , respectively) can be determined by solving the following two optimization problems.

$$\underline{p}_F = \min_{\boldsymbol{\theta} \in [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]} (p_F(\boldsymbol{\theta})), \qquad (2)$$

$$\bar{p}_F = \max_{\boldsymbol{\theta} \in [\underline{\boldsymbol{\theta}}, \bar{\boldsymbol{\theta}}]} (p_F(\boldsymbol{\theta})). \tag{3}$$

The structure of eqs. (2) and (3) indicates that for determining the bounds of the imprecise 169 failure probability, it is necessary to solve two optimization problems with respect to the epistemic 170 uncertainty associated with  $\theta$ . In turn, for a given crisp value of  $\theta$ , it is necessary to propagate 171 the aleatoric uncertainty associated with z and y in order to calculate the failure probability as shown in eq. (1). Such structure reveals the challenge associated with the calculation of an 173 imprecise probability, as it is necessary to propagate both epistemic and aleatoric uncertainty at 174 different levels. In other words, failure probabilities must be calculated for different crisp values 175 of  $\theta$  as required by the optimization. In turn, as the calculation of failure probabilities demands 176 performing repeated system analyses for different realizations of the uncertain input parameters z and y, the numerical cost associated with the estimation of an imprecise failure probability becomes insurmountable, even for simple problems. For the sake of completeness, it should be noted that surrogate modelling schemes, such as e.g., presented in [15] or [19], can alleviate the computational cost of propagating imprecise probabilities drastically. Indeed, in some cases, the application of appropriate surrogate modelling schemes might even render the application of brute-force double loop approaches tractable at acceptable levels of accuracy. Nonetheless, even when applying surrogate modelling schemes, it is advantageous to apply more efficient calculation methods to further increase the numerical efficiency. For a thorough discussion on the applicability of decoupling approaches and surrogate modelling schemes, the reader is referred to [34].

### 2.2. Augmented Reliability Problem

A possible means for avoiding the double loop problem associated with the calculation of imprecise probabilities consists of formulating an augmented reliability problem. This augmented problem was originally devised within the context of sensitivity analysis and reliability-based optimization in [24, 25, 27, 28] and was later applied to calculation of imprecise probabilities arising in problems where the performance function is affected by interval valued parameters [30]. In the following, a novel application of the augmented reliability approach is proposed, where the focus is on estimating imprecise probabilities in case imprecision affects the distribution parameters of some random variables of a reliability problem. This novel application can be seen as an extension of the approach proposed in [35] for sensitivity analysis in structural reliability.

The augmented reliability problem is defined such that an instrumental probability density function  $f_{\Theta}(\theta)$  is associated with the distribution parameter vector  $\theta$ . It is emphasized that this probability distribution is only an artifact, which has no physical meaning, as the uncertainty associated with  $\theta$  is epistemic (and not aleatoric). In principle, this instrumental probability distribution could be chosen arbitrarily. In this sense, note that the distribution does not have to be bounded between the lower and upper bounds of the associated intervals. In fact, any distribution that is capable of characterizing the effect of the epistemic uncertain parameters from a physical standpoint is admissible. In this contribution, a uniform distribution is selected, such that  $f_{\Theta}(\theta) \sim \mathcal{U}(\underline{\theta}, \overline{\theta})$ . Thus, the failure probability  $p_F^A$  associated with the augmented reliability problem is:

$$p_F^A = \int_{\boldsymbol{\theta} \in \Omega_{\theta}} \int_{\boldsymbol{y} \in \Omega_{y}} \int_{\boldsymbol{z} \in \Omega_{z}} I_F(\boldsymbol{y}, \boldsymbol{z}) f_{\boldsymbol{Z}}(\boldsymbol{z}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{z} d\boldsymbol{y} d\boldsymbol{\theta}, \tag{4}$$

where  $\Omega_{\theta} = [\underline{\theta}, \overline{\theta}]$ . This augmented failure probability can be calculated using any suitable reliability method. Let F represent the failure set, that is  $F = \{(\boldsymbol{z}, \boldsymbol{y}) : g(\boldsymbol{z}, \boldsymbol{y}) \leq 0\}$ . Applying Bayes' theorem, it is noted that [24, 25, 30]:

$$P(F|\boldsymbol{\theta}) = p_F(\boldsymbol{\theta}) = \frac{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|F)p_F^A}{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})},$$
 (5)

where  $P(\cdot)$  denotes probability of occurrence of the argument and  $f_{\Theta}(\boldsymbol{\theta}|F)$  is the (instrumental) probability distribution associated with  $\boldsymbol{\theta}$  conditioned on failure. Eq. (5) provides an expression for calculating the failure probability as a function of the imprecise distribution parameters. In this expression,  $f_{\Theta}(\boldsymbol{\theta})$  is known (as it is selected arbitrarily) and  $p_F^A$  is estimated by means of a reliability method. Hence, the only term which remains to be evaluated is  $f_{\Theta}(\boldsymbol{\theta}|F)$ . In order to calculate this term, recall the definition of a marginal distribution, that is:

$$f_{\Theta}(\boldsymbol{\theta}|F) = \int_{\boldsymbol{y}\in\Omega_y} \int_{\boldsymbol{z}\in\Omega_z} f_{\Theta,\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{z}|F) d\boldsymbol{z} d\boldsymbol{y},$$
(6)

with  $f_{\Theta,Y,Z}(\theta, y, z|F)$  the joint distribution of y, z and  $\theta$  conditional on the failure set. Applying
Bayes' theorem, it is found that:

$$f_{\Theta}(\boldsymbol{\theta}|F) = \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{z} \in \Omega_{\boldsymbol{z}}} f_{\Theta}(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{z}, F) f_{\boldsymbol{Y}, \boldsymbol{Z}}(\boldsymbol{y}, \boldsymbol{z}|F) d\boldsymbol{z} d\boldsymbol{y}. \tag{7}$$

Now, focusing on the conditional probability distribution  $f_{\Theta}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{z},F)$  and recalling that  $\boldsymbol{y}$  and  $\boldsymbol{z}$  are independent between them, it is possible to deduce the following:

$$f_{\Theta}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{z},F) = I_{F}(\boldsymbol{y},\boldsymbol{z}) \frac{f_{\Theta,\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{z})}{f_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{y},\boldsymbol{z})},$$

$$= I_{F}(\boldsymbol{y},\boldsymbol{z}) \frac{f_{\Theta,\boldsymbol{Y}}(\boldsymbol{\theta},\boldsymbol{y})}{f_{\boldsymbol{Y}}(\boldsymbol{y})},$$

$$= I_{F}(\boldsymbol{y},\boldsymbol{z}) \frac{f_{\Theta}(\boldsymbol{\theta})f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}\in[\underline{\boldsymbol{\theta}},\overline{\boldsymbol{\theta}}]} f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})f_{\Theta}(\boldsymbol{\theta})d\boldsymbol{\theta}}.$$
(8)

The last equation can be further simplified recalling that  $f_{\Theta}(\theta)$  is a constant, as it is selected as a uniform distribution. Taking into account this fact and defining  $\Delta(\boldsymbol{y}) = \int_{\boldsymbol{\theta} \in [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]} f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}) d\boldsymbol{\theta}$ ,

eq. (8) reduces to: 223

$$f_{\Theta}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{z},F) = I_{F}(\boldsymbol{y},\boldsymbol{z}) \frac{f_{\Theta}(\boldsymbol{\theta})f_{Y}(\boldsymbol{y}|\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}\in[\underline{\boldsymbol{\theta}},\overline{\boldsymbol{\theta}}]} f_{Y}(\boldsymbol{y}|\boldsymbol{\theta})f_{\Theta}(\boldsymbol{\theta})d\boldsymbol{\theta}},$$

$$= I_{F}(\boldsymbol{y},\boldsymbol{z}) \frac{f_{\Theta}(\boldsymbol{\theta})f_{Y}(\boldsymbol{y}|\boldsymbol{\theta})}{f_{\Theta}(\boldsymbol{\theta})\int_{\boldsymbol{\theta}\in[\underline{\boldsymbol{\theta}},\overline{\boldsymbol{\theta}}]} f_{Y}(\boldsymbol{y}|\boldsymbol{\theta})d\boldsymbol{\theta}},$$

$$= I_{F}(\boldsymbol{y},\boldsymbol{z}) \frac{f_{Y}(\boldsymbol{y}|\boldsymbol{\theta})}{\Delta(\boldsymbol{y})}.$$
(9)

Inserting eq. (9) into (7) leads to the following expression for  $f_{\Theta}(\theta|F)$ :

$$f_{\Theta}(\boldsymbol{\theta}|F) = \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{z} \in \Omega_{\boldsymbol{z}}} \frac{f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})}{\Delta(\boldsymbol{y})} f_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{y},\boldsymbol{z}|F) d\boldsymbol{z} d\boldsymbol{y},$$
(10)

224

227

236

238

239

240

where the indicator function  $I_F(y,z)$  has been omitted, as it assumes a value equal to one given the presence of the probability density  $f_{Y,Z}(y,z|F)$ , which ensures that integration is carried out exclusively over the failure domain.

The expression in eq. (10) is quite convenient from a practical viewpoint, as it can be evaluated by means of simulation. This is due to the following: the terms  $f_{Y}(y|\theta)$  and  $\Delta(y)$  are both known either in closed form or numerically. Moreover, samples distributed according to  $f_{Y,Z}(y,z|F)$ can be obtained as a byproduct of a reliability analysis associated with the calculation of the augmented failure probability in eq. (4), see e.g. [24]. Thus, by combining eqs. (5) and (10), it is possible to construct an explicit expression for estimating the failure probability as a function of 233 the vector of distribution parameters  $\theta$ . This expression can then be used to solve the optimization problems in eqs. (2) and (3) at reduced numerical efforts, as it is no longer required to solve a nested problem.

As a summary, it can be stated that the approach based on an augmented reliability problem can be quite convenient from a numerical viewpoint. This is due to the fact that the calculation of the bounds for the imprecise probability reduces to:

- A single reliability analysis, carried out considering the augmented problem.
- Solution of two optimization problems considering the explicit approximation of the failure 241 probability as a function of  $\theta$ 242

# 3. Application to the Calculation of First Excursion Probabilities of Uncertain Linear Dynamical Systems

#### 3.1. General Remarks

245

The framework for determining imprecise probabilities as described in Section 2.2 is quite general 246 in the sense that no assumptions have been introduced with respect to the type of system in-247 volved. In principle, the approach can cope with, e.g. linear or nonlinear systems under static or dynamic loading. Nonetheless, the choice of a specific reliability method for solving the associated 249 augmented failure probability is heavily influenced by the type of system. For the particular case 250 of this contribution, the focus is on the analysis of uncertain linear structural systems subject 251 to imprecise stochastic loading. In such case, Directional Importance Sampling can offer a con-252 venient means for calculating failure probabilities, as shown in [32, 36]. Thus, the rest of this 253 Section is organized as follows. Section 3.2 describes the type of stochastic load considered here, 254 which corresponds to a Gaussian process. Then, Section 3.3 describes the characterization of the 255 dynamic response of a linear system, where the structural matrices (that is, mass, damping and 256 stiffness) may be affected by the input parameter y. Then, Sections 3.4 and 3.5 address the solution of the augmented reliability problem in order to estimate the augmented failure probability  $(p_F^A, \text{ see eq. } (4))$  as well as the probability distribution of the vector of distribution parameters 259 conditioned on failure  $(f_{\Theta}(\theta|F), \text{ see eq. } (10)), \text{ respectively. Section 3.6 formulates the expression}$ 260 for estimating the failure probability as a function of the vector of distribution parameters (that 26 is,  $p_F(\boldsymbol{\theta})$ ) while Section 3.7 provides a summary. 262

# 263 3.2. Stochastic loading

The uncertain load acting over the structure is denoted as  $p(t, \mathbf{z})$ , where t stands for time. Its duration is T and is represented at discrete time instants  $t_k = (k-1)\Delta t$ ,  $k = 1, \ldots, n_T$ , where  $\Delta t$  is the time step and  $n_T$  is the total number of time instants considered. The load process has zero mean and it is assumed as a crisp Gaussian process (that is, it does not depend on  $\mathbf{y}$ ). Such an assumption implies that the probability density  $f_{\mathbf{z}}(\mathbf{z})$  associated with  $\mathbf{z}$  is actually Gaussian. The covariance matrix of the discretized process is denoted as  $\mathbf{\Gamma}$ , whose dimension is  $n_T \times n_T$ . Following the above assumptions, the stochastic load is represented by means of the Karhunen-Loève expansion (see, e.g. [37]):

$$p(z) = \Psi \Lambda^{1/2} z, \tag{11}$$

where p(z) is a vector containing a realization of the uncertain load of dimension  $n_T \times 1$ , where its k-th component  $p_k$  represents the load at time  $t_k$ ;  $\Lambda$  is a matrix whose diagonal contains the largest  $n_{KL}$  eigenvalues of the covariance matrix  $\Gamma$ ;  $\Psi$  is a matrix of dimension  $n_T \times n_{KL}$  containing the  $n_{KL}$  eigenvectors of the covariance matrix  $\Gamma$  associated with the aforementioned eigenvalues;  $n_{KL}$ is the number of terms retained for the Karhunen-Loève expansion  $(n_{KL} \leq n_T, \text{ see, e.g. } [37])$ ; and 276 z is a realization of a standard Gaussian random variable vector Z of dimension  $n_{KL} imes 1$ . It is  $_{277}$ easily verified that due to the above formulation,  $n_z = n_{KL}$ . 278

#### 3.3. Imprecise Structural Response

This contribution considers systems that are modeled as a linear structures with classical damping subject to dynamic loading. The associated equation of motion is [38]:

$$M(y)\ddot{x}(t,y,z) + C(y)\dot{x}(t,y,z) + K(y)x(t,y,z) = \rho(y)p(t,z),$$
 (12)

279

280

282

288

289

290

294

where M, C and K are the mass, damping and stiffness matrices of dimension  $n_D \times n_D$  each; 283  $n_D$  is the number of degrees-of-freedom of the structural model;  $\rho$  is a vector of dimension  $n_D \times 1$ coupling the stochastic load p(t, z) with the degrees-of-freedom of the structure; and where  $\ddot{x}$ ,  $\dot{x}$  and x denote the acceleration, velocity and displacement vectors, each of dimension  $n_D \times 1$ . It is noted that the system's matrices as well as the coupling vector  $\rho$  are affected by y, whose uncertainty is described by a p-box model, while the loading is affected by aleatoric uncertainty. Hence, the uncertainty associated with the response of the structure becomes a p-box as well. It is assumed that there are  $n_R$  responses of interest of the structural system that are of interest due to, e.g. practical design reasons. These responses are denoted as  $\eta_i(t, \boldsymbol{y}, \boldsymbol{z}), i = 1, \dots, n_R$ . Due to the linearity of the structure, these responses are calculated by means of a convolution integral. In fact, as shown in Appendix A, the i-th response of interest at the k-th time instant is given by the expression:

$$\eta_i(t_k, \boldsymbol{y}, \boldsymbol{z}) = \boldsymbol{a}_{i_k}(\boldsymbol{y})\boldsymbol{z},\tag{13}$$

where  $a_{i_k}(y)$  is a row vector of dimension  $1 \times n_{KL}$  which is associated with the discrete time 295 representation of the convolution integral (see Appendix A for details). Furthermore, the i-th 296 response of interest at each of the  $n_T$  time instants of analysis is equal to:

$$\eta_i(\boldsymbol{y}, \boldsymbol{z}) = \boldsymbol{A}_i(\boldsymbol{y})\boldsymbol{z}, \ i = 1, \dots, n_R,$$
 (14)

stochastic loading; and  $A_i$  is a matrix of dimension  $n_T \times n_{KL}$  whose k-th row contains the vector  $a_{i_k}(\boldsymbol{y})$ .

The structural system undergoes an unacceptable behavior whenever the absolute value of any of its responses exceeds corresponding prescribed thresholds  $b_i$ ,  $i = 1, \ldots, n_R$  at any time within the duration of the imprecise stochastic load. The occurrence of such unacceptable behavior can be expressed in mathematical terms with the help of the normalized response function  $\xi(\boldsymbol{y}, \boldsymbol{z})$  [39], which is equal to:

where  $\eta_i$  is a vector of dimension  $n_T \times 1$  containing the *i*-th response along the duration T of the

$$\xi(\boldsymbol{y}, \boldsymbol{z}) = \|\boldsymbol{A}(\boldsymbol{y})\boldsymbol{z}\|_{\infty},\tag{15}$$

where  $\|\cdot\|_{\infty}$  represents infinity norm; and  $\boldsymbol{A}$  is a matrix of dimension  $(n_R n_T) \times n_{KL}$  that allows calculating all responses of interest throughout the duration T of the load and which is calculated as shown below, assuming that  $b_i > 0$ ,  $i = 1, ..., n_R$ . The matrix  $\boldsymbol{A}$  is specifically given as:

$$\mathbf{A}(\mathbf{y}) = \begin{bmatrix} b_1^{-1} \mathbf{A}_1(\mathbf{y}) \\ \vdots \\ b_{n_R}^{-1} \mathbf{A}_{n_R}(\mathbf{y}) \end{bmatrix}. \tag{16}$$

Matrix A as shown in eq. (16) provides a means for calculating the responses of interest in a dimensionless, normalized fashion. This is a consequence of multiplying by the inverse of the corresponding threshold level. Furthermore, the infinity norm introduced in eq. (15) allows retrieving the maximum value of these dimensionless responses. Hence, whenever  $\xi \geq 1$ , one or more of the  $n_R$  responses of interest of the structure exceed their prescribed thresholds at one or more time instants within the duration of the stochastic load. Thus, the performance function associated with the failure probability in eq. (1) and its augmented counterpart in eq. (4) is equal to:

$$g(\boldsymbol{y}, \boldsymbol{z}) = 1 - \xi(\boldsymbol{y}, \boldsymbol{z}),$$
  
= 1 -  $\|\boldsymbol{A}(\boldsymbol{y})\boldsymbol{z}\|_{\infty}$ . (17)

# 3.4. Application of Directional Importance Sampling for Calculating Augmented Failure Probability

317

327

331

332

The structure of the performance function in eq. (17) reveals that the limit state function g(y, z) =0 is actually a piecewise linear function with respect to z for a fixed value of y [9, 40, 41]. In fact, linearity is due to the functional form of the i-th response with respect to z as shown in eq. (14) while the piecewise characteristic stems out of the infinity norm present in eq. (15). Figure 1 shows an schematic representation of such issue. For simplicity, it has been assumed that  $n_y = 1$ and  $n_z = n_{KL} = 2$ . The abscissa and ordinate contain coordinates  $z_1$  and  $z_2$  while coordinate yis orthogonal to the plane of the Figure. The limit state function is represented schematically for two different realizations y and  $y + \Delta y$  with solid and dashed line, respectively. It is noted that both limit state functions are different, as they correspond to two different values of y. However, they are both piecewise linear with respect to z.

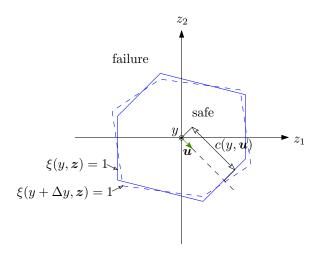


Figure 1: Schematic representation of limit state surface.

The fact that the limit state function is piecewise linear with respect to z for a fixed value of y opens the way to solve the augmented reliability problem in eq. (4) by means of Directional Importance Sampling. Such an approach has been shown to be particularly effective for addressing this type of limit state functions [32, 36]. Therefore, the application of this simulation method is investigated in the following for the problem at hand.

As a preliminary step before applying Directional Importance Sampling, it is necessary to express vector z in polar form, that is z = ru, where  $u = z/\|z\|$  is the unit vector pointing towards z,  $r = \|z\|$  is the radius and  $\|\cdot\|$  denotes Euclidean norm. Thus, the augmented failure probability in eq. (4) is expressed as:

$$p_F^A = \int_{\boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}} \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{u} \in \Omega_{\boldsymbol{U}}} \int_0^{\infty} I_F(\boldsymbol{y}, r\boldsymbol{u}) 2r f_{R^2}(r^2) f_{\boldsymbol{U}}(\boldsymbol{u}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) dr d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta},$$
(18)

where  $\Omega_U$  denotes the unit hypersphere in  $(n_{KL}-1)$  dimensions;  $f_U(u)$  is the uniform probability density over the  $(n_{KL}-1)$ -dimensional hypersphere; and  $f_{R^2}(r^2)$  is the Chi-square distribution of 338  $n_{KL}$  degrees-of-freedom. The factor 2r in eq. (18) appears due to the rule of change of variables 339 for probability distributions [42]. In this sense, note that  $r^2$  follows a Chi-square distribution of 340  $n_{KL}$  degrees-of-freedom as it is the sum of the squares of  $n_{KL}$  standard Gaussian variables. Hence, the probability density function associated with r is equal to  $f_R(r) = 2r f_{R^2}(r^2)$ . 342 The integral associated with r in eq. (18) can be solved in closed form taking advantage of the 343 linearity of the responses of interest with respect to z. In fact, let c be the scalar that solves the 344 equation g(y, cu) = 0 for fixed values of y and u. Then, c can be obtained in explicit form out of eq. (17), as shown below.

$$c(\boldsymbol{y}, \boldsymbol{u}) = \min_{i=1,\dots,n_{\eta}} \left( \min_{k=1,\dots,n_{T}} \left( c_{i,k}(\boldsymbol{y}, \boldsymbol{u}) \right) \right), \tag{19}$$

where  $c_{i,k}(\boldsymbol{y},\boldsymbol{u})$  is defined as:

$$c_{i,k}(\boldsymbol{y}, \boldsymbol{u}) = \frac{b_i}{|\boldsymbol{a}_{i_k}(\boldsymbol{y})\boldsymbol{u}|},\tag{20}$$

and where  $a_{i_k}(\boldsymbol{y})$  represents the k-th row of matrix  $A_i(\boldsymbol{y})$  (see Appendix A) and  $|\cdot|$  denotes absolute value. Note that  $c(\boldsymbol{y}, \boldsymbol{u})$  in eq. (19) represents the Euclidean distance between the origin of the standard normal space and the limit state function along direction  $\boldsymbol{u}$ . Figure 1 sketches distance  $c(\boldsymbol{y}, \boldsymbol{u})$  for a given unit vector  $\boldsymbol{u}$ . From the above discussion, it is seen that  $I_F(\boldsymbol{y}, r\boldsymbol{u}) = 1$  for  $r \in [c(\boldsymbol{y}, \boldsymbol{u}), \infty)$ . Thus, the integral associated with r in eq. (18) can be solved in closed form, leading to the following expression:

$$p_F^A = \int_{\boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}} \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{u} \in \Omega_{\boldsymbol{U}}} \int_0^{\infty} I_F(\boldsymbol{y}, r\boldsymbol{u}) 2r f_{R^2}(r^2) f_{\boldsymbol{U}}(\boldsymbol{u}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) dr d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta},$$

$$= \int_{\boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}} \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{u} \in \Omega_{\boldsymbol{U}}} \int_{c(\boldsymbol{y}, \boldsymbol{u})}^{\infty} 2r f_{R^2}(r^2) f_{\boldsymbol{U}}(\boldsymbol{u}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) dr d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta},$$

$$= \int_{\boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}} \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{u} \in \Omega_{\boldsymbol{U}}} \left( 1 - F_{R^2} \left( c(\boldsymbol{y}, \boldsymbol{u})^2 \right) \right) f_{\boldsymbol{U}}(\boldsymbol{u}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta},$$

$$(21)$$

where  $F_{R^2}(r^2)$  denotes the Chi-square cumulative density function of  $n_{KL}$  degrees-of-freedom. Eq. (21) provides a mean for estimating the augmented failure probability in a directional fashion with respect to u by means of simulation [43, 44]. However, the associated estimator may posses a large variability, which is undesirable from a numerical viewpoint. This issue can be addressed by introducing an Importance Sampling density (ISD) function associated with u that takes advantage of the geometry of the limit state function, that is, piecewise linear for a fixed value  $\boldsymbol{y}$ , as shown in Figure 1. For this purpose, consider the so-called elementary failure event  $F_{i,k}$ [9], where the i-th response of interest exceeds its prescribed threshold level  $b_i$  at time  $t_k$  for a given value y. Recalling the linearity of the response with respect to z (see eq. (13)), the polar representation z = ru and the definition of quantity  $c_{i,k}(y, u)$  in eq. (20), the elementary failure event is defined in mathematical terms as:

$$F_{i,k}(\boldsymbol{y}) = \{ \boldsymbol{z} \in \Omega_z : |\boldsymbol{a}_{i_k}(\boldsymbol{y})^T \boldsymbol{z}| \ge b_i \},$$

$$= \{ \boldsymbol{u} \in \Omega_u \ \land r \in [0, \infty) : |\boldsymbol{a}_{i_k}(\boldsymbol{y})^T (r \boldsymbol{u})| \ge b_i \},$$

$$= \{ \boldsymbol{u} \in \Omega_u \ \land r \in [0, \infty) : r \ge c_{i_k}(\boldsymbol{y}, \boldsymbol{u}) \}.$$
(22)

364

371

According to Bayes' theorem, the probability density associated with u conditioned on the occurrence of the elementary failure event  $F_{i,k}(\mathbf{y})$  is the following [9, 32]. 366

$$f_{U}(\boldsymbol{u}|F_{i,k}(\boldsymbol{y})) = \frac{f_{U}(\boldsymbol{u})P(F_{i,k}(\boldsymbol{y})|\boldsymbol{u})}{P(F_{i,k}(\boldsymbol{y}))}.$$
(23)

From the above equation, it is noted that  $f_{U}(u)$  is available in closed form (see, e.g. [45]). The 367 term  $P(F_{i,k}(y)|u)$  measures the probability of occurrence of the elementary failure event for a 368 fixed value of u. Given the definition of the elementary failure event in eq. (22) and recalling that  $r^2$  follows a Chi-square distribution of  $n_{KL}$ -degrees-of-freedom, this probability can be determined  $^{370}$ in closed form, that is:

$$P(F_{i,k}(\boldsymbol{y})|\boldsymbol{u}) = P(r \ge c_{i,k}(\boldsymbol{y},\boldsymbol{u})) = 1 - F_{R^2}(c_{i,k}(\boldsymbol{y},\boldsymbol{u})^2).$$
(24)

Moreover, the denominator in eq. (23) can also be determined in closed form recalling that 372  $a_{i,k}(y)^T z$  follows a Gaussian distribution with zero mean and variance  $a_{i,k}(y)^T a_{i,k}(y)$  for a fixed 373 value y. Thus, it is noted that:

$$P(F_{i,k}(\boldsymbol{y})) = P(|\boldsymbol{a}_{i,k}(\boldsymbol{y})^T \boldsymbol{z}| \ge b_i) = 2F_Z(-\beta_{i,k}(\boldsymbol{y})),$$
(25)

where  $F_Z(\cdot)$  is the standard Gaussian cumulative distribution and  $\beta_{i,k}(\boldsymbol{y})$  is the so-called reliability index, which is equal to  $\beta_{i,k}(\boldsymbol{y}) = b_i/\sqrt{a_{i,k}(\boldsymbol{y})^Ta_{i,k}(\boldsymbol{y})}$  [9, 40]. As the terms  $P(F_{i,k}(\boldsymbol{y})|\boldsymbol{u})$  and  $P(F_{i,k}(\boldsymbol{y}))$  can be calculated in closed form as shown in eqs. (24) and (25), the conditional probability density function  $f_U(\boldsymbol{u}|F_{i,k}(\boldsymbol{y}))$  in eq. (23) is available as well. A salient characteristic of this density function is that it assigns no weight to direction vectors  $\boldsymbol{u}$  which do not lead to the occurrence of the elementary failure event. Thus, the ISD function is chosen as a weighted summation of the density functions  $f_U(\boldsymbol{u}|F_{i,k}(\boldsymbol{y}))$  associated with each elementary failure event (that is,  $F_{i,k}$ ,  $i=1,\ldots,n_\eta$ ,  $k=1,\ldots,n_T$ ) [9, 36, 32]. The weight associated with each density function is denoted as  $w_{i,k}$ ,  $i=1,\ldots,n_\eta$ ,  $j=1,\ldots,n_T$ , and is selected proportional to  $P(F_{i,k}(\boldsymbol{y}))$ , as done customarily in Importance Sampling [9, 46]. Hence:

$$w_{i,k}(\mathbf{y}) = \frac{2F_Z(-\beta_{i,k}(\mathbf{y}))}{\hat{p}_F(\mathbf{y})},$$
(26)

where  $\hat{p}_F(\boldsymbol{y}) = \sum_{i_1=1}^{n_\eta} \sum_{k_1=1}^{n_T} 2F_Z(-\beta_{i_1,k_1}(\boldsymbol{y}))$ . Inserting eqs. (24) and (25) into eq. (23) and taking the summation of all conditional density functions weighted by the expression in eq. (26), the ISD function becomes the following [32, 36].

$$f_{\boldsymbol{U}}^{\mathrm{IS}}(\boldsymbol{u}|\boldsymbol{y}) = \sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} w_{i,k}(\boldsymbol{y}) f_{\boldsymbol{U}}(\boldsymbol{u}|F_{i,k}(\boldsymbol{y})),$$

$$= \frac{f_{\boldsymbol{U}}(\boldsymbol{u})}{\hat{p}_{F}(\boldsymbol{y})} \sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} \left(1 - F_{R^{2}}\left(c_{i,k}(\boldsymbol{y}, \boldsymbol{u})^{2}\right)\right). \tag{27}$$

It is emphasized that the ISD function as shown in the above equation is valid for a given value of y. Now, this ISD function can be considered for estimating the augmented failure probability

defined in eq. (21), leading to:

395

397

$$p_F^A = \int_{\boldsymbol{\theta} \in \Omega_{\theta}} \int_{\boldsymbol{y} \in \Omega_{y}} \int_{\boldsymbol{u} \in \Omega_{U}} \left( 1 - F_{R^2} \left( c(\boldsymbol{y}, \boldsymbol{u})^2 \right) \right) \frac{f_{\boldsymbol{U}}(\boldsymbol{u})}{f_{\boldsymbol{U}}^{\mathrm{IS}}(\boldsymbol{u}|\boldsymbol{y})} f_{\boldsymbol{U}}^{\mathrm{IS}}(\boldsymbol{u}|\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta}$$

$$= \int_{\boldsymbol{\theta} \in \Omega_{\theta}} \int_{\boldsymbol{y} \in \Omega_{y}} \int_{\boldsymbol{u} \in \Omega_{U}} \frac{\hat{p}_{F}(\boldsymbol{y}) \left( 1 - F_{R^2} \left( c(\boldsymbol{y}, \boldsymbol{u})^2 \right) \right)}{\sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} \left( 1 - F_{R^2} \left( c(\boldsymbol{y}, \boldsymbol{u})^2 \right) \right)} f_{\boldsymbol{U}}^{\mathrm{IS}}(\boldsymbol{u}|\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta},$$

$$\approx \tilde{p}_{F}^{A} = \frac{1}{N} \sum_{j=1}^{N} \frac{\hat{p}_{F} \left( \boldsymbol{y}^{(j)} \right) \left( 1 - F_{R^2} \left( c(\boldsymbol{y}^{(j)}, \boldsymbol{u}^{(j)})^2 \right) \right)}{\sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} \left( 1 - F_{R^2} \left( c_{i,k} \left( \boldsymbol{y}^{(j)}, \boldsymbol{u}^{(j)} \right)^2 \right) \right)},$$

$$\boldsymbol{\theta}^{(j)} \sim f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}), \ \boldsymbol{y}^{(j)} \sim f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}^{(j)}), \ \boldsymbol{u}^{(j)} \sim f_{\boldsymbol{U}}^{\mathrm{IS}}(\boldsymbol{u}|\boldsymbol{y}^{(j)}), \ j = 1, \dots, N,$$
(28)

where  $\tilde{(\cdot)}$  denotes an estimator of a quantity. From the above equation, it is important to note that samples of  $\theta$  are generated such that they follow the instrumental uniform probability density  $f_{\Theta}(\theta)$ ; then, samples of y are generated conditional on those samples of  $\theta$ ; and finally, samples of u are generated conditional on the samples of y. For details on how samples of u are generated, it is referred to Appendix B.

3.5. Application of Directional Importance Sampling for Calculating Conditional Probability Density Function  $f_{\Theta}(\boldsymbol{\theta}|F)$ 

The next challenge for applying the formulation presented in Section 2 is the calculation of the conditional probability density function  $f_{\Theta}(\theta|F)$  in eq. (10). The objective is estimating this density function by means of simulation with samples of  $\boldsymbol{y}$  and  $\boldsymbol{z}$  distributed according to  $f_{Y,Z}(y,z|F)$ . In this sense, the samples used for estimating the augmented failure probability in eq. (28) could be used for that purpose. Nonetheless, these samples cannot be used directly, as they are distributed according to the ISD function, which is not necessarily equal to  $f_{Y,Z}(y,z|F)$ . 403 Therefore, special care must be taken for using those samples, as described in the following. 404 According to the definition of a marginal distribution and Bayes' theorem,  $f_{Y,Z}(y,z|F)$  is equal 405 to: 406

$$f_{Y,Z}(y,z|F) = \int_{\boldsymbol{\theta} \in [\underline{\boldsymbol{\theta}},\overline{\boldsymbol{\theta}}]} f_{\Theta,Y,Z}(\boldsymbol{\theta}, y, z|F) d\boldsymbol{\theta}$$

$$= \frac{1}{p_F^A} \int_{\boldsymbol{\theta} \in [\underline{\boldsymbol{\theta}},\overline{\boldsymbol{\theta}}]} I_F(y,z) f_Z(z) f_Y(y|\boldsymbol{\theta}) f_{\Theta}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{29}$$

Recalling the definition of polar coordinates z = ru, the latter expression is further simplified to:

$$f_{Y,Z}(y,z|F) = f_{Y,R,U}(y,r,u|F) = \frac{1}{p_F^A} \int_{\boldsymbol{\theta} \in [\underline{\boldsymbol{\theta}},\overline{\boldsymbol{\theta}}]} I_F(y,ru) 2r f_{R^2}(r^2) f_U(u) f_Y(y|\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (30)$$

Note that eq. (30) provides an expression for the conditional probability distribution of  $f_{Y,Z}(y,z|F)$  as a function of the polar coordinates, which has been considered for the implementation of Directional Importance Sampling in Section 3.4. Now, this expression is inserted into eq. (10) for the calculation of  $f_{\Theta}(\theta|F)$ . However, special consideration should be paid to one detail: in eq. (10),  $\theta$  refers to an independent variable, that is,  $\theta$  may assume any value within  $[\underline{\theta}, \overline{\theta}]$ . On the contrary, in eq. (30),  $\theta$  corresponds to an integration variable over the domain  $[\underline{\theta}, \overline{\theta}]$ . Hence, and in order to avoid confusion, the variable  $\theta$  appearing in eq. (30) is written in the following as  $\theta'$ . Having taken this special consideration, the expression for  $f_{\Theta}(\theta|F)$  becomes:

$$f_{\Theta}(\boldsymbol{\theta}|F) = \frac{1}{p_F^A} \int_{\boldsymbol{\theta}' \in [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]} \int_{\boldsymbol{y} \in \Omega_y} \int_{\boldsymbol{u} \in \Omega_u} \int_0^{\infty} \frac{f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})}{\Delta(\boldsymbol{y})} I_F(\boldsymbol{y}, r\boldsymbol{u}) 2r f_{R^2}(r^2) f_{\boldsymbol{U}}(\boldsymbol{u}) f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}') f_{\Theta}(\boldsymbol{\theta}') d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta}'.$$
(31)

This last equation can be simplified by recalling that  $I_F(\boldsymbol{y}, r\boldsymbol{u}) = 1$  for  $r \in [c(\boldsymbol{y}, \boldsymbol{u}), \infty)$  and that for its evaluation, it is possible to introduce the ISD function defined in eq. (27). Thus, the expression for  $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|F)$  is as follows.

$$f_{\Theta}(\boldsymbol{\theta}|F) = \frac{1}{p_F^A} \int_{\boldsymbol{\theta}' \in [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]} \int_{\boldsymbol{y} \in \Omega_y} \int_{\boldsymbol{u} \in \Omega_u} \frac{f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})}{\Delta(\boldsymbol{y})} \frac{\hat{p}_F(\boldsymbol{y}) \left(1 - F_{R^2} \left(c(\boldsymbol{y}, \boldsymbol{u})^2\right)\right)}{\sum_{i=1}^{n_\eta} \sum_{k=1}^{n_T} \left(1 - F_{R^2} \left(c_{i,k}(\boldsymbol{y}, \boldsymbol{u})^2\right)\right)} f_{IJ}^{\text{IS}}(\boldsymbol{u}|\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}') f_{\Theta}(\boldsymbol{\theta}') d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta}'.$$
(32)

3.6. Explicit Approximation of the Failure Probability as a Function of the Distribution Parameter

Vector

Considering all previous deductions, it is possible to determine the following expression for the failure probability as a function of the distribution parameter vector by inserting eq. (32) into eq. (5).

$$p_{F}(\boldsymbol{\theta}) = \frac{1}{f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})} \int_{\boldsymbol{\theta}' \in [\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}]} \int_{\boldsymbol{y} \in \Omega_{\boldsymbol{y}}} \int_{\boldsymbol{u} \in \Omega_{\boldsymbol{u}}} \frac{f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta})}{\Delta(\boldsymbol{y})} \frac{\hat{p}_{F}(\boldsymbol{y}) \left(1 - F_{R^{2}} \left(c(\boldsymbol{y}, \boldsymbol{u})^{2}\right)\right)}{\sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} \left(1 - F_{R^{2}} \left(c_{i,k}(\boldsymbol{y}, \boldsymbol{u})^{2}\right)\right)} f_{\boldsymbol{U}}^{IS}(\boldsymbol{u} | \boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y} | \boldsymbol{\theta}') f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}') d\boldsymbol{u} d\boldsymbol{y} d\boldsymbol{\theta}'.$$
(33)

It is noted that the above expression can be estimated by using samples of u, y and  $\theta'$  generated when conducting augmented reliability analysis with Directional Importance Sampling, as performed in Section 3.4. This estimator is equal to: 426

$$p_{F}(\boldsymbol{\theta}) \approx \tilde{p}_{F}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{j=1}^{N} \frac{f_{Y}\left(\boldsymbol{y}^{(j)}|\boldsymbol{\theta}\right)}{\Delta\left(\boldsymbol{y}^{(j)}\right)} \frac{\hat{p}_{F}\left(\boldsymbol{y}^{(j)}\right)\left(1 - F_{R^{2}}\left(c(\boldsymbol{y}^{(j)}, \boldsymbol{u}^{(j)})^{2}\right)\right)}{\sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} \left(1 - F_{R^{2}}\left(c_{i,k}\left(\boldsymbol{y}^{(j)}, \boldsymbol{u}^{(j)}\right)^{2}\right)\right)},$$

$$\boldsymbol{\theta}^{\prime(j)} \sim f_{\boldsymbol{\Theta}^{\prime}}(\boldsymbol{\theta}^{\prime}), \ \boldsymbol{y}^{(j)} \sim f_{Y}(\boldsymbol{y}|\boldsymbol{\theta}^{\prime(j)}), \ \boldsymbol{u}^{(j)} \sim f_{U}^{IS}(\boldsymbol{u}|\boldsymbol{y}^{(j)}), \ j = 1, \dots, N.$$
(34)

As  $p_F(\theta)$  is estimated with independent, identically distributed samples, it is straightforward to estimate the standard deviation of this estimator, which is equal to: 428

$$\sigma_{\tilde{p}_{F}(\boldsymbol{\theta})} = \sqrt{\frac{1}{N(N-1)} \sum_{j=1}^{N} \left( \frac{f_{\boldsymbol{Y}}(\boldsymbol{y}^{(j)}|\boldsymbol{\theta})}{\Delta(\boldsymbol{y}^{(j)})} \frac{\hat{p}_{F}(\boldsymbol{y}^{(j)}) \left(1 - F_{R^{2}}\left(c(\boldsymbol{y}^{(j)}, \boldsymbol{u}^{(j)})^{2}\right)\right)}{\sum_{i=1}^{n_{\eta}} \sum_{k=1}^{n_{T}} \left(1 - F_{R^{2}}\left(c_{i,k}(\boldsymbol{y}^{(j)}, \boldsymbol{u}^{(j)})^{2}\right)\right)} - \tilde{p}_{F}(\boldsymbol{\theta}) \right)^{2}}.$$
(35)

3.7. Summary 429

As a summary of the material presented in this Section, it is seen that it is possible to formulate an explicit approximation of the failure probability as a function of the vector of distribution parameters  $\theta$ , as shown in eq. (34). This explicit approximation is constructed by carrying out 432 a single run of Directional Importance Sampling for solving the associated augmented reliability problem. Then, this explicit approximation can be coupled with any appropriate optimization algorithm in order to estimate the bounds of the imprecise probability. In addition, it is possible to estimate the standard deviation of the probability estimator by means of eq. (35). This is most useful from a practical viewpoint, as it is possible to assert whether or not the estimator of the failure probability  $p_F(\boldsymbol{\theta})$  is sufficiently accurate and take proper action if necessary: for example, increasing the number of samples in order to improve the accuracy.

437

439

As a side remark, it should be noted that the formulation of Directional Importance Sampling as presented here considers an ISD function related with the direction vector  $\boldsymbol{u}$  only. Nonetheless, references in the literature suggest that for the class of problems considered here, that is, first excursion probability of uncertain linear systems subject to imprecise dynamic load, it is also feasible to establish an ISD function associated with the uncertain input parameters y, as discussed in [27, 47, 48, 41]. Such an approach has the potential to improve even further accuracy when 445 solving the augmented reliability problem. In spite of this issue, the possibility of introducing this additional ISD function is not further discussed here, in order to focus more on the application of the augmented reliability problem and its connection with the estimation of imprecise first excursion probabilities.

#### 450 4. Examples

452

454

455

456

457

459

460

461

#### 4.1. General Remarks

This Section illustrates the application of the augmented reliability approach for estimating the bounds of first excursion probabilities. A test example involving a single-degree-of-freedom oscillator and an application example comprising the finite element model of a composite wing are considered. For both cases, aleatoric and epistemic uncertainties affect the structural parameters while the loading is modeled as purely aleatoric. In this sense, it should be noted that the scheme proposed in Section 3 can consider epistemic uncertainty on the loading as well. Studies on the effect of imprecise stochastic loads on the bounds on the first excursion probability of linear structures can be found in previous work of the authors [21, 22, 23]. These studies acknowledge the importance of including imprecision in the definition of stochastic loads modelled by advanced auto-correlation methods. However, this possibility is not explored further in order to simplify the presentation of the performance of the method and focus on the application of the augmented reliability framework for calculating imprecise probabilities.

464 4.2. Test Example 1: Single-degree-of-freedom shear beam model subject to stochastic ground ac-465 celeration

The first example involves the single-story shear beam model depicted in Figure 2. This model is subjected to a stochastic ground acceleration and its stiffness is characterized as a p-box. The objective is estimating the bounds for the first excursion probability.

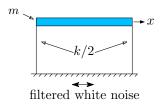


Figure 2: Example 1 – Single story shear beam model subject to filtered white noise excitation

The mass of the shear beam model is equal to  $m = 10^4$  [kg] while its (classical) damping ratio is d = 5%. The lateral stiffness k of the model is characterized by means of a lognormal p-box,

such that the expected value is represented by the interval  $\mu_k \in [0.7 \times 10^6, 1.3 \times 10^6]$  [MN/m] and the standard deviation is  $\sigma_k = 10^5$  [MN/m]. The stochastic ground acceleration is modeled 472 as a white noise process of duration T=20 [s] and spectral density  $S=5\times 10^{-3}$  [m<sup>2</sup>/s<sup>3</sup>], which is represented at discrete time steps of  $\Delta t = 0.01$  [s]. This discrete white noise is modulated considering the Shinozuka-Sato envelope, with parameters  $c_1 = 0.14$  and  $c_2 = 0.16$  [7]. In addition, 475 the modulated white noise is filtered considering the Clough-Penzien model (see, e.g. [49]), with circular frequencies  $\omega_{g,1}=6\pi$  [rad/s] and  $\omega_{g,2}=0.6\pi$  [rad/s] and damping ratios  $d_{g,1}=d_{g,2}=0.6\pi$ 60% for the primary and secondary filters, respectively. This stochastic ground acceleration is represented by means of the Karhunen-Loève expansion, considering all  $n_{KL}=n_z=2001$  terms. The responses of interest of the shear beam model are its relative displacement with respect to the ground as well as its absolute acceleration. Failure occurs whenever any of these responses exceeds the prescribed thresholds  $b_1 = 0.07$  [m] and  $b_2 = 7.5$  [m/s<sup>2</sup>] within the duration of the acceleration.

479

483

485

486

490

In order to estimate the bounds of the first excursion probability, the augmented reliability problem is solved first with Directional Importance Sampling. For this purpose, a total of N=2000 samples are considered. The results obtained for the estimate  $\tilde{p}_F^A$  as well as its coefficient of variation  $\delta_{\tilde{p}_F^A}$ are shown in Figure 3. It is noted from this figure that the estimate of the augmented failure probability stabilizes after about 500 samples and that a coefficient of variation of less than 10% is attained at about 2000 samples. These results suggest that the level of precision of the estimator is adequate.

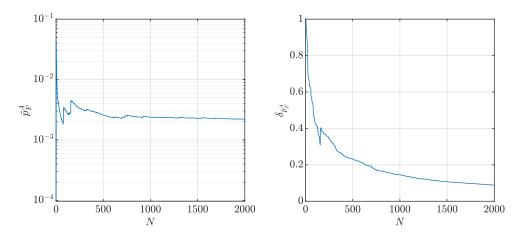


Figure 3: Example 1 – Evolution of the estimator of the augmented failure probability  $\tilde{p}_F^A$  with respect to the number of samples.

In a next step, the first excursion probability is estimated as a function of the interval variable 491

 $\mu_k$  applying eq. (34), using the samples generated when estimating the augmented failure probability. This probability  $\tilde{p}_F(\mu_k)$  is plotted in Figure 4 with blue, solid line. In addition, the blue 493 shaded area in that figure shows the standard deviation associated with the estimate of the failure probability, that is  $\tilde{p}_F(\mu_k) \pm \sigma_{\tilde{p}_F(\mu_k)}$ . For comparison and validation purposes, the first excursion 495 probability is also calculated for ten different crisp values of  $\mu_k$  using Directional Importance Sam-496 pling. In other words, ten different simulations of Directional Importance Sampling are carried 497 out, each of them considering N = 1000 samples. Hence, the total number of samples involved in such an approach is equal to  $N_T = 10N = 10^4$ . The probability estimates along the grid of crisp 499 values of  $\mu_k$  are shown with red x marks; in addition, the standard deviation of these probability 500 estimates is marked with a bar. It is seen that there is an excellent agreement between the two 501 approaches, suggesting that the framework provided by augmented reliability is appropriate for 502 estimating the first excursion probability as a function of the imprecise distribution parameter. 503 Furthermore, it should be noted that the relation between the failure probability and the interval-504 valued  $\mu_k$  is non-monotonic. In this sense, it is quite remarkable that the proposed approach can 505 produce this type of approximations, as no particular assumptions have to be introduced to arrive 506 to such result. The non-monotonic behavior can be explained as follows. For small crisp values 507 of  $\mu_k$ , the shear beam model is quite flexible and failure due to displacements is likely. As the crisp value of  $\mu_k$  increases, the shear beam possesses more stiffness and it is capable of controlling 509 displacements better, which decreases the failure probability. However, as the crisp value of  $\mu_k$ 510 continues increasing, the shear beam becomes too stiff, thus increasing the absolute acceleration 511 and in turn, the failure probability. 512

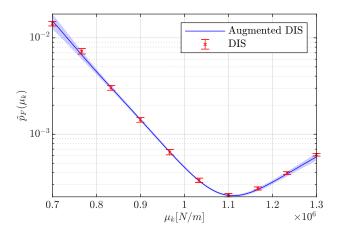


Figure 4: Example 1 – Failure probability as a function of the interval parameter  $\mu_k$ . DIS: Directional Importance Sampling

Finally, the bounds on the imprecise failure probability are determined by minimizing/maximizing<sub>13</sub>  $\tilde{p}_F(\mu_k)$ . The numerical cost associated with this step is negligible, as the failure probability as function of  $\mu_k$  is available in closed form. It is found that  $p_F(\mu_k) \in [2.3 \times 10^{-4}, 1.5 \times 10^{-2}]$ , for revealing that the uncertainty on  $\mu_k$  has a major impact in the first excursion probability.

517

518

521

522

523

524

530

#### 4.3. Example 2: Composite wing

#### 4.3.1. Case introduction and physical description

The second example comprises of a Finite Element of a fictitious aircraft wing that is produced in a laminated composite. The structure has a total length of 30 [m]. This wing is produced using a multi-layer laminar composite material, with deterministic orthotropic ply material properties  $E_1 = 231$  [GPa],  $E_2 = 77$  [GPa],  $\nu_{12} = 0.31$  and  $G_{12} = G_{23} = G_{13} = 42.7$  [GPa]. The wing consists of a composite outer shell (top, leading edge, bottom, trailing edge), as well as two vertical stiffening ribs in the centre for structural stiffness. A total of 15 different composite lay-ups are present throughout the structure, which are summarized in Table 1. The dynamic behavior of the structure is modeled using a Finite Element model containing 621 nodes, 606 bi-linear shell elements, 573 rigid connections, 10 concentrated masses and 132 rod elements. The finite element model of this structure is shown in Figure 5. A single evaluation of this model takes approx. 30 [s] on a server equipped with two AMD EPYC 7601 CPUs running at 2.65 [GHz] and 512 [GB] of RAM.

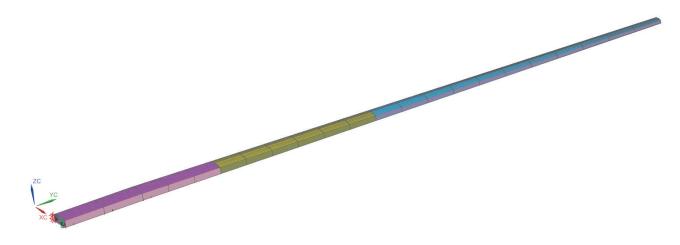


Figure 5: Finite element model of the composite blade (taken from [50])

In this case, the equation of motion of the wing can be represented as in Eq. (12), where  $n_D$  531 of the unconstrained model is  $n_D = 3726$ .

Table 1: Composite lay-up structure of the wing. Left means at y = 0 in Figure 5 and the leading edge is depicted at the back side of figure 5. The + and - signs in front of  $45^0$  denote an alternating layering sequence of laminae oriented at  $+45^{\circ}$  and  $-45^{\circ}$ .

Nr	Location	Lay-up (symmetrical)	thickness
			per layer [mm]
1	top and bottom left	$+ - 45^{\circ}$	7.5 [mm]
2	leading edge left	$+-45^{\circ}$	$7.5 [\mathrm{mm}]$
3	front-middle edge left	$+ - + - + - +45^{\circ}$	$7.5 [\mathrm{mm}]$
4	back-middle vertical left	$+ - + - + - +45^{\circ}$	$7.5 [\mathrm{mm}]$
5	trailing edge left	$+ - + - + - +45^{\circ}$	$7.5 [\mathrm{mm}]$
6	top and bottom middle	$+-45^{\circ}$	$7.5 [\mathrm{mm}]$
7	leading edge middle	$+-45^{\circ}$	$7.5 [\mathrm{mm}]$
8	front-middle edge middle	$+ - + - + - +45^{\circ}$	$7.5 [\mathrm{mm}]$
9	back-middle vertical middle	$+ - + - + - +45^{\circ}$	7.5  [mm]
10	trailing edge middle	$+ - + - + - +45^{\circ}$	$7.5 [\mathrm{mm}]$
11	top and bottom right	$+-45^{\circ}$	$7.5 [\mathrm{mm}]$
12	leading edge right	$+ - 45^{\circ}$	$7.5 [\mathrm{mm}]$
13	front-middle edge right	$+ - + - + - +45^{\circ}$	$7.5 [\mathrm{mm}]$
14	back-middle vertical right	$+ - + - + - +45^{\circ}$	7.5 [mm]
_15	trailing edge right	$+ - + - + - +45^{\circ}$	7.5 [mm]

In this case study, we consider the case where the wing is subjected to a turbulence wind load, which is described as:

$$p(t, \boldsymbol{\xi}) = p_0(t) + p'(t, \boldsymbol{\xi}), \tag{36}$$

where  $p_0(t)$  is the mean load of the turbulence load and  $p'(t, \boldsymbol{\xi})$  is a zero-mean Gaussian stochastic process governed by the transverse Dryden spectrum [51]:

$$S_{p'p'}(\omega) = \frac{c\sigma_{p'}^2}{2\pi} \frac{1 + 3c^2\omega^2}{(1 + c^2\omega^2)^2},$$
(37)

where c = L/v is a scale for the turbulence and  $\sigma^2$  is the variance, with L the turbulence length scale and v the true air speed. Applying the well-known Wiener-Kinchin transformation, following auto-correlation model is obtained:

$$\Gamma_{p,p'}(\tau) = (1 - 0.5 \frac{\tau}{L}) \exp(-\tau/L),$$
(38)

with  $\tau$  a positive time-lag factor, which can be used directly in combination with the Karhunen-Loève expansion, as explained in Section 3.2. In this case study, failure is conceived as the first passage of the wing displacement over a threshold value of b = 7 [mm] under a unit variance load (i.e.,  $\sigma_{p'}^2 = 1$ ). It is important to note that in this case study,  $p_0(t)$  is considered to be zero, which is permissible due to linearity in the FE model. In all proceeding computations, the effect of mode-crossover and -veering, which may occur during the calculation of the impulse response functions, is accounted for by tracking the numerically computed mode shapes via the modal assurance criterion.

# 4.4. Uncertainty propagation

The presented approach is applied to this model considering the thickness t of the  $1^{st}$ ,  $5^{th}$  and  $6^{th}$  laminae as being modelled by a p-box, whereas the other 12 thickness values are considered to have a thickness of 7.5 [mm]. This selection is made based on the observation that the thickness of these three laminae has the largest influence on the wing tip displacement (see also [22]). Specifically, these thickness values are modelled as a lognormal distribution with interval-valued mean  $\mu_{t1} = \mu_{t5} = \mu_{t6} = [0.006, 0.009]$  [m] and standard deviation  $\sigma_t = 0.00075$  [m]. This imprecision in the mean of the distributions can come for instance from limited experimental data set sizes, or a lack of precision in the employed measurement equipment such that only the bounds of one deterministic measurement are known [34]. The aleatory (random) part of the uncertainty is assumed to come from variations during the manufacturing of the laminae. The bounds on the probability of failure are estimated by solving the augmented reliability problem using Directional Importance Sampling with a sample size of N = 2000 samples.

Then, the first excursion probability is estimated as a function of the interval variables  $\mu_t = [\mu_{t1}, \mu_{t5}, \mu_{t6}]$  applying eq. (34), using the samples generated when estimating the augmented failure probability. This probability  $\tilde{p}_F(\mu_t)$  is plotted in Figure 6 with blue, solid line, assuming that  $\mu_t = [\mu_{t1}, \mu_{t5}, \mu_{t6}]$ . Please note that during the analysis to determine the bounds on  $P_f$ , the intervals are considered to be fully independent, and all samples in the augmented space are drawn from the corresponding uniform distributions independently. It is only for the sake of visualisation in Figure 6 that fully dependent samples are generated within the interval bounds. In addition, the blue shaded area in that figure shows the standard deviation associated with the estimate of the failure probability, that is  $\tilde{p}_F(\mu_t) \pm \sigma_{\tilde{p}_F(\mu_t)}$ . It is important to notice that this blue line is obtained in a single run of Directional Importance Sampling. In addition, for the purpose of validation, the first excursion probability is also determined using 10 runs of Directional Importance Sampling for several crisp values of  $\mu_{t1} = \mu_{t5} = \mu_{t6}$ , evenly distributed throughout  $\mu_t^I$ . Each of these 10 runs requires a full Directional Importance Sampling propagation involving

1000 deterministic model analyses. In Figure 6, the corresponding first excursion probabilities are indicated as red crosses, and their standard deviation is indicated using error bars.

As can be noted, also in the case of a realistic Finite Element model, the method gives an excellent agreement with a more conventional approach, with a reduced cost of a factor 5. Finally the bounds on the probability of failure can be determined by minimizing/maximizing over  $\tilde{p}_F(\boldsymbol{\mu}_t)$ , which are found to be  $\tilde{p}_F(\boldsymbol{\mu}_t) \in [5.62 \times 10^{-4}, 7.67 \times 10^{-2}]$ . From Figure 6, it can be noted that concerning the upper bound of  $\tilde{p}_F(\boldsymbol{\mu}_t)$ , a slight deviation exists with respect to the result obtained for that particular value of  $\boldsymbol{\mu}_t$  using conventional DIS, which is caused by the relatively big size of the interval in combination with the more complicated dynamical behavior of the wing.

Finally, as a word of caution, numerical experience indicates that the performance of the method degrades rapidly with the number of uncertain epistemic parameters. For instance, if one considers all 15 thickness values of this model to be uncertain, with the same uncertainty model as prescribed above, and using the same augmented DIS estimator, the bounds on the probability of failure are found to be  $[1.04 \times 10^{-4}, 6.00 \times 10^{-3}]$ , as compared to the correct interval  $[2.3 \times 10^{-3}, 1.43 \times 10^{-2}]$  which is obtained via a double-loop implementation. This performance drop is explained by the comparatively high number of samples that is required to fully explore the functional behaviour of a response with respect to a high-dimensional input space, and hence, to correctly establish the functional relationship between  $\mu_t$  and  $\tilde{p}_F(\mu_t)$ . This is a well-documented shortcoming of this class of methods (see e.g., [29]), and remains an open research question.

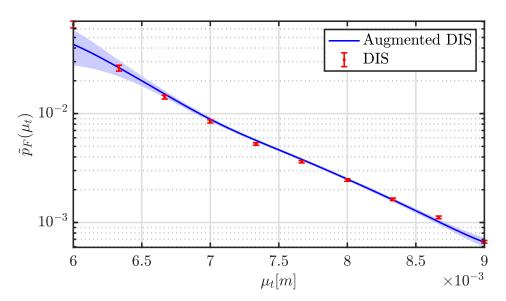


Figure 6: Estimated  $\tilde{p}_F(\mu_t)$  as a function of the mean value  $\mu_t$  of the thickness values t, as computed by Augmented DIS (blue) and individual runs of DIS for selected values of  $\mu_t$  (red).

#### 5. Conclusions and Outlook

This paper presents an efficient augmented space method to propagate parametric p-boxes towards the bounds on the first excursion probability of a structure subjected to Gaussian excitation. The method establishes a functional relationship between the interval-valued hyper-parameters of the p-box valued uncertain input quantities and the first excursion probability by representing the problem in an augmented space. Then, by virtue of Bayes' theorem, an expression for the failure probability as an explicit expression of the imprecise parameters from the augmented reliability problem can be recovered, which ultimately allows calculating the imprecise probability by means of Directional Importance Sampling. Following conclusions can be made concerning the proposed approach:

- The method allows for a highly efficient and accurate calculation of the bounds on the probability of failure for both a small-scale academic case, as for a realistic finite element model
- In case the dimension of the input epistemic space is small, the approach is highly accurate, however, numerical experience suggests that the accuracy of the calculated bounds degrades quickly when the dimension of the input epistemic space increases. This is caused by the comparatively high number of samples that is required to fully establish the required functional relationship between the probability of failure and the interval-valued hyperparameters of the uncertain input quantities. This is consistent with conclusions drawn by other researchers in the area of, e.g. optimal design under uncertain conditions (see, e.g. [29, 27]). Further research however has to be performed in order to define a precise (preferably a priori) criterion in which the augmented reliability problem does not bring substantial advantages compared to double loop approaches.

As a last, more practical, remark, it can be noted that in both case studies the bounds of the probability of failure span several orders of magnitude. As an analyst, such observation shows that additional data should be gathered to reduce the epistemic uncertainty on the definition of the p-boxes that represent the uncertainty on the input parameters to shrink the bounds on the estimation of the probability of failure. Alternatively, in case no further data can be gathered, one should consider the upper bound of the analysis and use this to base decisions on, as this is the most conservative estimate given the epistemic uncertainty.

A future challenge concerning the application of the framework proposed here involves includ-623 ing epistemic uncertainty in the characterization of stochastic load. Such epistemic uncertainty 624 may have a large impact on the failure probability, as discussed in [21]. In addition, the application of the framework to problems involving structural non linearities should be explored as well. For that purpose, it would be necessary to consider simulation techniques which are more general 627 than Directional Importance Sampling, such as Subset Simulation. Other path for future development consists of considering instrumental probability distributions (associated with the epistemic parameters) different from uniform distributions, which may eventually bring advantages for deducing the functional relationship between epistemic parameters and the failure probability. As 631 a final comment, it can be noted that the principles laid out in this paper also can be applied to 632 combinations of parametric p-boxes and 'regular' random variables. Such extension will also be 633 pursued in future work.

# 35 6. Acknowledgments

Matthias Faes gratefully acknowledges the financial support of the Research Foundation Flanders (FWO) under grant number 12P3519N. Matthias Faes and Pengfei Wei acknowledge the support of the Alexander von Humboldt foundation. Marcos Valdebenito acknowledges the support of ANID (National Agency for Research and Development, Chile) under its program FONDECYT, grant number 1180271.

#### 641 Appendix A. Calculation of Structural Response

In view of the linearity of the structural system, the responses of interest are calculated by means of a convolution, that is:

$$\eta_i(t, \boldsymbol{y}, \boldsymbol{z}) = \int_0^t h_i(t - \tau, \boldsymbol{y}) p(\tau, \boldsymbol{z}) dt, \ i = 1, \dots, n_R$$
(A.1)

where  $h_i(t, \boldsymbol{y})$  is the unit impulse response function associated with the *i*-th response. For the case where the *i*-th response of interest is expressed as a linear combination of the displacement vector, the corresponding unit impulse response function is equal to:

$$h_i(t, \boldsymbol{y}) = \sum_{v=1}^{n_D} \frac{\boldsymbol{q}_i^T \boldsymbol{\phi}_v(\boldsymbol{y}) \boldsymbol{\phi}_v(\boldsymbol{y})^T \boldsymbol{\rho}(\boldsymbol{y})}{\boldsymbol{\phi}_v(\boldsymbol{y})^T \boldsymbol{M}(\boldsymbol{y}) \boldsymbol{\phi}_v(\boldsymbol{y})} \frac{e^{-\zeta_v(\boldsymbol{y})\omega_v(\boldsymbol{y})t}}{\omega_{d,v}(\boldsymbol{y})} \sin(\omega_{d,v}(\boldsymbol{y})t), \ i = 1, \dots, n_R$$
(A.2)

where  $\omega_v$ ,  $v = 1, ..., n_D$  are the natural frequencies;  $\zeta_v$ ,  $v = 1, ..., n_D$  are the damping ratios;  $\omega_{d,v} = \omega_v \sqrt{(1 - \zeta_v^2)}$ ,  $v = 1, ..., n_D$  are the damped frequencies;  $\phi_v$ ,  $v = 1, ..., n_D$  are the eigenvectors associated with the eigenproblem of the undamped equation of motion; and  $q_i$  is a vector such that  $\eta_i = q_i^T x$ , where  $(\cdot)^T$  denotes transpose. Note that while the above equation applies to a unit impulse response function associated with displacement, similar expressions can be deduced for other quantities such as, e.g. accelerations.

In view of the discrete time representation of the imprecise stochastic loading as shown in eq. (11), 653 the convolution integral in eq. (A.1) reduces to the following expression.

$$\eta_{i}(t_{k}, \boldsymbol{y}, \boldsymbol{z}) = \sum_{l_{1}=1}^{k} \epsilon_{l_{1}} h_{i}(t_{k} - t_{l_{1}}, \boldsymbol{y}) p(t_{l_{1}}, \boldsymbol{z}) \Delta t$$

$$= \sum_{l_{1}=1}^{k} \epsilon_{l_{1}} h_{i}(t_{k} - t_{l_{1}}, \boldsymbol{y}) \left( \sum_{l_{2}=1}^{n_{KL}} \psi_{l_{1}, l_{2}} \sqrt{\lambda_{l_{2}}} z_{l_{2}} \right) \Delta t$$

$$= \boldsymbol{a}_{i_{k}}(\boldsymbol{y}) \boldsymbol{z}, \ i = 1, \dots, n_{\eta}, \ k = 1, \dots, n_{T} \tag{A.3}$$

where  $\psi_{l_1,l_2}$  is the  $(l_1,l_2)$ -th element of matrix  $\Psi$ ;  $\epsilon_{l_1}$  is a coefficient depending on the quadrature scheme used to approximate the convolution integral; and  $\boldsymbol{a}_{i_k}$  is a row vector of dimension  $1 \times n_{KL}$  defined as:

$$\mathbf{a}_{i_k}(\mathbf{y}) = \left[ \sum_{l_1=1}^k \Delta t \epsilon_{l_1} h_i(\mathbf{y}, t_k - t_{l_1}) \psi_{l_1, 1} \sqrt{\lambda_1}, \dots, \sum_{l_1=1}^k \Delta t \epsilon_{l_1} h_i(\mathbf{y}, t_k - t_{l_1}) \psi_{l_1, n_{KL}} \sqrt{\lambda_{n_{KL}}} \right]$$

$$i_k = (i-1)n_T + k, \ i = 1, \dots, n_R, \ k = 1, \dots, n_T \quad (A.4)$$

Each of the rows of matrix  $A_i$  introduced in eq. (14) contains the row vectors  $a_{i_k}$ , that is:

$$\mathbf{A}_{i}(\mathbf{y}) = \begin{bmatrix} \mathbf{a}_{i_{1}}(\mathbf{y}) \\ \vdots \\ \mathbf{a}_{i_{n_{T}}}(\mathbf{y}) \end{bmatrix}, i = 1, \dots, n_{R}$$
(A.5)

658

The coefficients  $\epsilon_{l_1}$  are chosen following the trapezoidal rule for integration (see, e.g. [52]), yielding  $\epsilon_{l_1} = 1/2$  in case  $l_1 = 1$  or  $l_1 = k$ ; otherwise,  $\epsilon_{l_1} = 1$ .

## 661 Appendix B. Sample generation of Direction Vector

- Samples  $\mathbf{u}^{(j)}$ , j = 1, ..., N required for evaluating the augmented failure probability in eq. (28) are generated by means of the following procedure [9, 32, 45].
- 1. Set j = 1.
- 2. Sample  $\boldsymbol{\theta}^{(j)}$  from  $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ .
- 3. Sample  $\boldsymbol{y}^{(j)}$  from  $f_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}^{(j)})$ .
- 4. Draw a pair of indices (I, K) from the set  $\Omega = \{(i, k) : i \in \{1, \dots, n_{\eta}\}, k \in \{1, \dots, n_T\}\}$ with probability proportional to the weights  $w_{i,k}(\boldsymbol{y}^{(j)}), i = 1, \dots, n_{\eta}, k = 1, \dots, n_T$ .
- 5. Generate a sample z of the random variable vector Z and two realizations  $\kappa_1$  and  $\kappa_2$  following a uniform distribution between 0 and 1.
- 6. Calculate  $\alpha = -F_Z^{-1}\left((1-\kappa_1)F_Z\left(-\beta_{I,K}\left(\boldsymbol{y}^{(j)}\right)\right)\right)$ , where  $F_Z^{-1}(\cdot)$  is the inverse cumulative standard Gaussian distribution.
- 7. Compute  $\boldsymbol{a}_{I_{K}}^{*} = \boldsymbol{a}_{I_{K}}\left(\boldsymbol{y}^{(j)}\right) / \left\|\boldsymbol{a}_{I_{K}}\left(\boldsymbol{y}^{(j)}\right)\right\|$ . See eq. (A.4) for the calculation of  $\boldsymbol{a}_{I_{K}}\left(\boldsymbol{y}^{(j)}\right)$ .
- 8. Define  $z^*$  as:

$$\boldsymbol{z}^* = \begin{cases} \boldsymbol{z} + (\alpha - \boldsymbol{z}^T \boldsymbol{a}_{I,K}^*) \boldsymbol{a}_{I_K}^* & \text{if } \kappa_2 \le 1/2 \\ -\boldsymbol{z} - (\alpha - \boldsymbol{z}^T \boldsymbol{a}_{I_K}^*) \boldsymbol{a}_{I,K}^* & \text{otherwise} \end{cases}$$
(B.1)

- 9. Calculate the sought sample as  $\boldsymbol{u}^{(j)} = \boldsymbol{z}^* / \|\boldsymbol{z}^*\|$ .
- 10. In case j = N, stop the procedure. Otherwise, return to step 2 with j = j + 1.

#### 677 References

- [1] G. Schuëller, Computational Methods in Stochastic Dynamics, Springer Netherlands, 2011, Ch. Model Reduction and Uncertainties in Structural Dynamics, pp. 1–24.
- [2] M. Faes, D. Moens, Recent Trends in the Modeling and Quantification of Non-probabilistic
   Uncertainty, Arch. Comput. Methods Eng. (feb 2019).
- [3] M. Beer, Y. Zhang, S. T. Quek, K. K. Phoon, Reliability analysis with scarce information:

  Comparing alternative approaches in a geotechnical engineering context, Struct. Saf. 41 (2013)

  1–10.
- [4] M. Broggi, M. Faes, E. Patelli, Y. Govers, D. Moens, M. Beer, Comparison of Bayesian and interval uncertainty quantification: Application to the AIRMOD test structure, in: 2017

	IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings, Vol. 2018-Janua, 2018, pp. 1–8.	687 688
[5]	M. Faes, M. Broggi, E. Patelli, Y. Govers, J. Mottershead, M. Beer, D. Moens, A multivariate interval approach for inverse uncertainty quantification with limited experimental data, Mech. Syst. Sig. Process. 118 (2019) 534–548.	689 690 691
[6]	G. Deodatis, Non-stationary stochastic vector processes: seismic ground motion applications, Probab. Eng. Mech. 11 (3) (1996) 149 – 167.	692 693
[7]	M. Shinozuka, Y. Sato, Simulation of nonstationary random process, Journal of the Engineering Mechanics Division 93 (1) (1967) 11–40.	694 695
[8]	S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, K. Sentz, Constructing Probability Boxes and Dempster-Shafer Structures, Tech. Rep. January, Technical report, Sandia National Laboratories (2003).	696 697 698
[9]	S. Au, J. Beck, First excursion probabilities for linear systems by very efficient importance sampling, Probab. Eng. Mech. 16 (3) (2001) 193–207.	699 700
[10]	D. Moens, D. Vandepitte, An interval finite element approach for the calculation of envelope frequency response functions, Int. J. Numer. Methods Eng. 61 (14) (2004) 2480–2507.	701 702
[11]	W. Gao, D. Wu, C. Song, F. Tin-Loi, X. Li, Hybrid probabilistic interval analysis of bar structures with uncertainty using a mixed perturbation monte-carlo method, Finite Elem. Anal. Des. 47 (7) (2011) 643–652.	703 704 705
[12]	B. Xia, D. Yu, J. Liu, Hybrid uncertain analysis for structural-acoustic problem with random and interval parameters, J. Sound Vib. 332 (11) (2013) 2701–2720.	706 707
[13]	S. Yin, D. Yu, Z. Luo, B. Xia, Unified polynomial expansion for interval and random response analysis of uncertain structure-acoustic system with arbitrary probability distribution, Comput. Methods Appl. Mech. Eng. 336 (2018) $260-285$ .	708 709 710
[14]	J. Wu, Z. Luo, Y. Zhang, N. Zhang, L. Chen, Interval uncertain method for multibody mechanical systems using Chebyshev inclusion functions, Int. J. Numer. Methods Eng. 95 (7) (2013) 608–630.	711 712 713

- [15] R. Schöbi, B. Sudret, Structural reliability analysis for p-boxes using multi-level meta-models,
   Probab. Eng. Mech. 48 (2017) 27–38.
- 716 [16] R. Schöbi, B. Sudret, Global sensitivity analysis in the context of imprecise probabilities (p717 boxes) using sparse polynomial chaos expansions, Rel. Eng. Syst. Saf. 187 (2019) 129 141,
  718 sensitivity Analysis of Model Output.
- [17] M. Faes, J. Sadeghi, M. Broggi, M. de Angelis, E. Patelli, M. Beer, D. Moens, On the Robust
   Estimation of Small Failure Probabilities for Strong Nonlinear Models, ASCE-ASME J Risk
   and Uncert in Engrg Sys Part B Mech Engrg 5 (4) (dec 2019).
- <sup>722</sup> [18] J. Sadeghi, M. de Angelis, E. Patelli, Robust propagation of probability boxes by interval <sup>723</sup> predictor models, Struct. Saf. 82 (2020) 101889.
- [19] P. Wei, J. Song, S. Bi, M. Broggi, M. Beer, Z. Lu, Z. Yue, Non-intrusive stochastic analysis
   with parameterized imprecise probability models: I. performance estimation, Mech. Syst. Sig.
   Process. 124 (2019) 349 368.
- [20] P. Wei, J. Song, S. Bi, M. Broggi, M. Beer, Z. Lu, Z. Yue, Non-intrusive stochastic analysis
   with parameterized imprecise probability models: II. reliability and rare events analysis,
   Mech. Syst. Sig. Process. 126 (2019) 227 247.
- [21] M. G. Faes, M. A. Valdebenito, D. Moens, M. Beer, Bounding the first excursion probability
   of linear structures subjected to imprecise stochastic loading, Computers & Structures 239
   (2020) 106320.
- [22] M. G. R. Faes, M. A. Valdebenito, Fully Decoupled Reliability-Based Design Optimization
   of Structural Systems Subject to Uncertain Loads, Computer Methods in Applied Mechanics
   and Engineering (2020).
- [23] M. G. R. Faes, M. A. Valdebenito, D. Moens, M. Beer, Operator norm theory as an efficient tool to propagate hybrid uncertainties and calculate imprecise probabilities, Mechanical
   Systems and Signal Processing (2021).
- [24] S. Au, Probabilistic failure analysis by importance sampling Markov chain simulation, J. Eng.
   Mech. 130 (3) (2004) 303-311.

[25]	J. Ching, Y. Hsieh, Local estimation of failure probability function and its confidence interval
	with maximum entropy principle, Probab. Eng. Mech. 22 (1) (2007) 39–49.

- [26] J. Ching, Y. Hsieh, Approximate reliability-based optimization using a three-step approach based on subset simulation, J. Eng. Mech. 133 (4) (2007) 481–493.
- [27] H. Jensen, M. Valdebenito, G. Schuëller, An efficient reliability-based optimization scheme for uncertain linear systems subject to general Gaussian excitation, Comput. Methods Appl. Mech. Eng. 198 (1) (2008) 72–87.
- [28] A. Taflanidis, J. Beck, An efficient framework for optimal robust stochastic system design using stochastic simulation, Comput. Methods Appl. Mech. Eng. 198 (1) (2008) 88–101.
- [29] P. Koutsourelakis, Design of complex systems in the presence of large uncertainties: A statistical approach, Comput. Methods Appl. Mech. Eng. 197 (49-50) (2008) 4092–4103.
- [30] J. Song, P. Wei, M. Valdebenito, S. Bi, M. Broggi, M. Beer, Z. Lei, Generalization of non-intrusive imprecise stochastic simulation for mixed uncertain variables, Mech. Syst. Sig. Process. 134 (2019) 106316.
- [31] J. Zhang, M. D. Shields, On the quantification and efficient propagation of imprecise probabilities resulting from small datasets, Mech. Syst. Sig. Process. 98 (June) (2018) 465–483.
- [32] M. A. Misraji, M. A. Valdebenito, H. A. Jensen, C. F. Mayorga, Application of directional importance sampling for estimation of first excursion probabilities of linear structural systems subject to stochastic Gaussian loading, Mechanical Systems and Signal Processing 139 (2020) 106621. doi:10.1016/j.ymssp.2020.106621.
  URL https://linkinghub.elsevier.com/retrieve/pii/S0888327020300078
- [33] M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses, Mech. Syst. Sig. Process. 37 (1-2) (2013) 4–29.
- [34] M. G. R. Faes, M. Daub, S. Marelli, E. Patelli, M. Beer, Engineering analysis with probability boxes: a review on computational methods (2021).
- [35] X. Yuan, S. Liu, M. Valdebenito, J. Gu, M. Beer, Efficient framework for failure probability function estimation in augmented space, Structural Safety (2021, submitted).

- [36] O. Ditlevsen, P. Bjerager, R. Olesen, A. Hasofer, Directional simulation in Gaussian processes,
   Probab. Eng. Mech. 3 (4) (1988) 207 217.
- 770 [37] G. Stefanou, The stochastic finite element method: Past, present and future, Comput. Methods Appl. Mech. Eng. 198 (9-12) (2009) 1031–1051.
- <sup>772</sup> [38] A. Chopra, Dynamics of structures: theory and applications to earthquake engineering, Pren-<sup>773</sup> tice Hall, 1995.
- [39] S. Au, J. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, Probab. Eng. Mech. 16 (4) (2001) 263–277.
- <sup>776</sup> [40] A. Der Kiureghian, The geometry of random vibrations and solutions by FORM and SORM, <sup>777</sup> Probab. Eng. Mech. 15 (1) (2000) 81–90.
- [41] M. Valdebenito, H. Jensen, A. Labarca, Estimation of first excursion probabilities for uncertain stochastic linear systems subject to gaussian load, Computers & Structures 138 (2014)
   36–48.
- <sup>781</sup> [42] A. Ang, W. Tang, Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering, Wiley, 2007.
- [43] I. Deák, Three digit accurate multiple normal probabilities, Numerische Mathematik 35 (4)
   (1980) 369–380.
- [44] P. Bjerager, Probability integration by directional simulation, J. Eng. Mech. 114 (8) (1988)
   1285–1297.
- <sup>787</sup> [45] L. Katafygiotis, S. Cheung, Domain decomposition method for calculating the failure probability of linear dynamic systems subjected to Gaussian stochastic loads, J. Eng. Mech. 132 (5) (2006) 475–486.
- [46] G. Schuëller, R. Stix, A critical appraisal of methods to determine failure probabilities, Struct.
   Saf. 4 (4) (1987) 293–309.
- [47] H. Jensen, M. Valdebenito, Reliability analysis of linear dynamical systems using approximate
   representations of performance functions, Struct. Saf. 29 (3) (2007) 222–237.

[48]	H. Pradlwarter, G. Schuëller, Uncertain linear structural systems in dynamics: Efficient	794
	stochastic reliability assessment, Computers & Structures 88 (1-2) (2010) 74–86.	795
[49]	A. Zerva, Spatial Variation of Seismic Ground Motions – Modeling and Engineering Appli-	796
	cations, CRC Press, 2009.	797
[50]	M. Faes, D. Moens, Multivariate dependent interval finite element analysis via convex hull	798
	pair constructions and the extended transformation method, Comput. Methods Appl. Mech.	799
	Eng. 347 (2019) 85 – 102.	800
[51]	L. J. Howell, Y. Lin, Response of flight vehicles to nonstationary atmospheric turbulence,	801
	AIAA Journal 9 (11) (1971) 2201–2207. doi:10.2514/3.50026.	802
[52]	W. Gautschi, Numerical Analysis, 2nd Edition, Birkhäuser Boston, 2012.	803