

# 1 First-passage probability estimation of high-dimensional nonlinear stochastic 2 dynamic systems by a fractional moments-based mixture distribution approach

3 Chen Ding<sup>a</sup>, Chao Dang<sup>a,\*</sup>, Marcos A. Valdebenito<sup>b</sup>, Matthias G.R. Faes<sup>c</sup>, Matteo Broggi<sup>a</sup>, Michael Beer<sup>a,d,e</sup>

4 <sup>a</sup> *Institute for Risk and Reliability, Leibniz University Hannover, Callinstr. 34, Hannover 30167, Germany*

5 <sup>b</sup> *Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Av. Padre Hurtado 750, 2562340 Viña del Mar, Chile*

6 <sup>c</sup> *Chair for Reliability Engineering, TU Dortmund University, Leonhard-Euler-Str. 5, Dortmund 44227, Germany*

7 <sup>d</sup> *Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Liverpool L69 7ZF, United Kingdom*

8 <sup>e</sup> *International Joint Research Center for Resilient Infrastructure & International Joint Research Center for Engineering Reliability and Stochastic  
9 Mechanics, Tongji University, Shanghai 200092, PR China*

---

## 10 Abstract

11 First-passage probability estimation of high-dimensional nonlinear stochastic dynamic systems is a significant task to  
12 be solved in many science and engineering fields, but remains still an open challenge. The present paper develops a  
13 novel approach, termed ‘fractional moments-based mixture distribution’, to address such challenge. This approach is  
14 implemented by capturing the extreme value distribution (EVD) of the system response with the concepts of fractional  
15 moment and mixture distribution. In our context, the fractional moment itself is by definition a high-dimensional  
16 integral with a complicated integrand. To efficiently compute the fractional moments, a parallel adaptive sampling  
17 scheme that allows for sample size extension is developed using the refined Latinized stratified sampling (RLSS).  
18 In this manner, both variance reduction and parallel computing are possible for evaluating the fractional moments.  
19 From the knowledge of low-order fractional moments, the EVD of interest is then expected to be reconstructed. Based  
20 on introducing an extended inverse Gaussian distribution and a log extended skew-normal distribution, one flexible  
21 mixture distribution model is proposed, where its fractional moments are derived in analytic form. By fitting a set  
22 of fractional moments, the EVD can be recovered via the proposed mixture model. Accordingly, the first-passage  
23 probabilities under different thresholds can be obtained from the recovered EVD straightforwardly. The performance  
24 of the proposed method is verified by three examples consisting of two test examples and one engineering problem.

### 25 *Keywords:*

26 First-passage probability, Stochastic dynamic system, Extreme value distribution, Fractional moment, Mixture  
27 distribution

---

## 28 1. Introduction

29 Stochastic dynamic systems which involve the randomness in internal system properties and/or external dynamic  
30 loads are widespread in various science and engineering fields, such as meteorology, quantum optics, circuit theory and  
31 structural engineering [1]. To assess the effects of input randomness on the system performance, dynamic reliability  
32 analysis has drawn increasing attention. Generally, dynamic reliability analysis for stochastic dynamic systems can be

33 classified as the first-passage probability evaluation and the fatigue failure probability estimation [2]. In the literature,  
34 the first-passage probability evaluation has been extensively studied over the past several decades. However, finding  
35 efficient and accurate solutions to the first-passage problem still remains challenging. The reason is twofold: (1) the  
36 high-dimensional input randomness and strongly nonlinear behavior of stochastic dynamic systems may be encountered  
37 simultaneously; (2) the first-passage probabilities of such systems under certain thresholds may be relatively small.

38 The existing approaches for first-passage probability estimation can be broadly divided into four kinds: the out-  
39 crossing rate approaches, the diffusion process approaches, the stochastic simulation approaches and the extreme  
40 value distribution (EVD) estimation approaches. For the out-crossing rate approaches, the first-passage probability  
41 is evaluated considering the time of out-crossing within a time duration on the basis of Rice's formula [3–6]. Such  
42 approaches are based on the Poisson assumption that level-crossing events are mutually independent and each happens  
43 at most once, or the Markovian assumption that the next crossing event only relates to the present event [7]. Although  
44 these solutions can be accurate in some special cases, they may be not applicable for general cases. Besides, it is  
45 hard to derive the joint probability density function (PDF) and its derivatives of the system response of interest when  
46 complicated nonlinear stochastic dynamic systems are encountered. The diffusion process approaches evaluate the  
47 first-passage probability by solving a partial differential equation, such as the Kolmogorov backward equation [8] or  
48 the Fokker Planck equation [9]. Solutions to such equations could be derived via the path integration method [10–12],  
49 stochastic average technique [13, 14], ensemble-evolving-based generalized density evolution equation [2, 15], etc.  
50 Nevertheless, this kind of approach is mostly applicable for nonlinear stochastic dynamic systems enforced by white  
51 noise. For the stochastic simulation approach, the extensively used Monte Carlo simulation (MCS) [16] is able to  
52 address problems regardless of their dimensions and nonlinearities. However, MCS is inefficient and even infeasible  
53 to assess a small probability for an expensive-to-evaluate model since a considerably large number of simulations  
54 are required. Although some variants of MCS have been developed, such as important sampling [17–20] and subset  
55 simulation [21–23], they still suffer their respective limitations concerning efficiency, accuracy and applicability, etc.

56 Recently, the EVD estimation approaches have attracted lots of attention. This is because once the EVD of system  
57 response of interest is obtained, the first-passage probability can be straightforwardly and conveniently evaluated [24].  
58 Nevertheless, the analytical solution to the EVD is difficult and even impossible to be obtained for a general nonlinear  
59 stochastic dynamic system. Therefore, various approximation methods have been developed to estimate the EVD, which  
60 can be roughly classified as probability conservation-based methods and moments-based methods. According to the  
61 principle of probability conservation, the probability density evolution method (PDEM) [7, 24] and direct probability  
62 integral method (DPIM) [25] are derived, which can be used for the purpose of EVD estimation. However, since such  
63 methods are typically dependent on the partition of random variable space, their application for high-dimensional  
64 problems may be challenging. Moment-based methods, on the other hand, estimate the first-passage probability by  
65 fitting an appropriate parametric distribution model to the EVD, and the free parameters of the distribution model are  
66 obtained from the estimated moments of the EVD. The integer moments-based methods can be adopted to recover the  
67 EVD [26, 27], where high-order integer moments, i.e., skewness and kurtosis, need to be considered. Yet it is difficult  
68 to evaluate such high-order integer moments using a small sample size, due to their large variability [28]. To alleviate  
69 such difficulty, a series of methods based on non-integer moments, such as fractional moments and linear moments,  
70 have been developed. The fractional moments-based maximum entropy methods [29–32] can estimate the first-passage  
71 probabilities of nonlinear stochastic dynamic systems from low to high dimensions. However, it is difficult to solve the  
72 non-convex optimization problem that is typically encountered, and the obtained results can be easily trapped into local

73 optimum. Besides, due to the polynomials involved in the maximum entropy density, the recovered EVD can have  
74 unexpected oscillating distribution tail, which then leads to an inaccurate evaluation of the first-passage probability. Two  
75 mixture parametric distribution methods in conjunction with fractional moments [33] or moment-generating function  
76 [34] are developed. These methods enable to evaluate first-passage probabilities of high-dimensional and strongly  
77 nonlinear stochastic dynamic systems from a small number of simulations. Furthermore, a fractional moments-based  
78 shifted generalized lognormal distribution method [35] is utilized to assess seismic reliability of a practical bridge  
79 subjected to spatial variate ground motions. Besides, the linear moments-based polynomial normal transformation  
80 distribution method [36] is developed to analyze high-dimensional dynamic systems with deterministic structural  
81 parameters subjected to stochastic excitations.

82 Overall, the fractional moments-based methods offer the possibility to deal with both high-dimensional and strongly  
83 nonlinear stochastic dynamic systems from a reduced number of simulations, even with small first-passage probabilities.  
84 In view of this, the present paper mainly focuses on such methods. Despite those attractive features, the fractional  
85 moments-based methods still have two main problems to be solved. On one hand, the sample size for evaluating  
86 fractional moments is usually empirically fixed. This is primarily because the sampling-based schemes adopted  
87 by the existing methods do not allow for the sample size extension. However, the optimal sample size should be  
88 problem-dependent. With a predetermined sample size, the adopted sampling methods may encounter over-sampling or  
89 under-sampling, leading to a waste of over-all computational efforts or unsatisfactory accuracy of estimated fractional  
90 moments. On the other hand, the success of fractional moments-based methods for first-passage probability evaluation  
91 also depends on the selection of an appropriate distribution model. Although the existing distribution models are  
92 capable of representing EVDs for some problems, their flexibility and applicability are limited. Hence, for a wide  
93 range of problems, they may still lack the ability to accurately recover the EVDs over the entire distribution domain,  
94 especially for the tails.

95 In this paper, we propose a fractional moments-based mixture distribution approach to estimate the first-passage  
96 probabilities of high-dimensional and strongly nonlinear stochastic dynamic systems. It is worth mentioning that the  
97 randomness from both internal system properties and external excitations is taken into account. The main contributions  
98 of this study are summarized as follows. First, a parallel adaptive sampling scheme is proposed for estimating the  
99 fractional moments, as opposed to the traditional fixed sample size scheme. Such a new scheme enables to extend  
100 the sample size sequentially, i.e., one at a time or several at a time. The optimal sample size for fractional moment  
101 estimation is determined by introducing a convergence criterion. In fact, a sequential sampling method with the ability  
102 to effectively reduce variance in high-dimensional problems, named Refined Latinized stratified sampling (RLSS) [37],  
103 is suitable for achieving our purposes and is employed within the proposed scheme. Second, one novel and versatile  
104 mixture distribution model is proposed to reconstruct the EVD with the knowledge of its estimated fractional moments.  
105 This model is based on the extension of the conventional inverse Gaussian distribution and the log transformation of the  
106 extended skew-normal distribution. The analytical expression of the fractional moments for such mixture distribution is  
107 derived, and a fractional moments-based parameter estimation technique is developed.

108 The remainder of this paper is organized as follows. Section 2 outlines the first-passage probability estimation of a  
109 stochastic dynamic system from the perspective of EVD. In section 3, the proposed fractional moments-based mixture  
110 distribution approach is described in detail, including a parallel adaptive scheme for fractional moments evaluation and  
111 a flexible mixture distribution model for EVD reconstruction. Three examples are given in section 4 to demonstrate the  
112 performance of the proposed method. The paper is closed with some concluding remarks in section 5.

## 113 2. First-passage probability estimation of stochastic dynamic systems

### 114 2.1. Stochastic dynamic systems

115 Consider a stochastic dynamic system that is governed by the following state-space equation:

$$\dot{\mathbf{Y}}(t) = \mathbf{Q}(\mathbf{Y}(t), \mathbf{U}, t), \quad (1)$$

116 with an initial condition

$$\mathbf{Y}(0) = \mathbf{y}_0, \quad (2)$$

117 where  $\mathbf{Y} = (Y_1, Y_2, \dots, Y_{n_d})$  is a  $n_d$ -dimensional state vector;  $\mathbf{Q} = (Q_1, Q_2, \dots, Q_{n_d})$  is a dynamics operator vector;  
 118  $\mathbf{U} = (U_1, U_2, \dots, U_{n_s})$  is a  $n_s$ -dimensional random parameter vector with a known joint probability density function  
 119 (PDF)  $p_{\mathbf{U}}(\mathbf{u})$ ;  $\mathbf{u} = (u_1, u_2, \dots, u_{n_s})$  denotes a realization of  $\mathbf{U}$ ;  $t$  denotes the time. Note that Eq. (1) can be strongly  
 120 nonlinear, which may be caused by material, geometrical, or contact nonlinearities inherent in the stochastic dynamic  
 121 system. In addition, hundreds or thousands of random variables can be included in the vector  $\mathbf{U}$  due to the randomness  
 122 from system properties and external excitations.

123 For a well-posed stochastic dynamic system, the solution to Eq. (1) is unique and depends on the vector  $\mathbf{U}$ , which  
 124 can be assumed to be:

$$\left[ \mathbf{Y}(t), \dot{\mathbf{Y}}(t) \right] = \left[ \mathbf{H}_{\mathbf{Y}}(\mathbf{U}, t), \frac{\partial \mathbf{H}_{\mathbf{Y}}(\mathbf{U}, t)}{\partial t} \right], \quad (3)$$

125 where  $\mathbf{H}_{\mathbf{Y}}$  and  $\frac{\partial \mathbf{H}_{\mathbf{Y}}}{\partial t}$  are the deterministic operators.

126 If we consider the system responses of interest for reliability analysis, say  $\mathcal{W}(t) = \{\mathcal{W}_1(t), \mathcal{W}_2(t), \dots, \mathcal{W}_{n_d}(t)\}$ ,  
 127 they can be evaluated from their relations to the state vectors:

$$\mathcal{W}(t) = \Psi \left[ \mathbf{Y}(t), \dot{\mathbf{Y}}(t) \right] = \mathcal{H}(\mathbf{U}, t), \quad (4)$$

128 where  $\Psi$  is the transfer operator; and  $\mathcal{H}$  denotes the mapping relation from  $\mathbf{U}$  and  $t$  to  $\mathcal{W}(t)$ . Accordingly, the  $q$ -th  
 129 component of  $\mathcal{W}(t)$  is denoted by  $\mathcal{W}_q(t) = \mathcal{H}_q(\mathbf{U}, t)$ ,  $q = 1, \dots, n_d$ . For notational simplicity, the subscript  $q$  is  
 130 omitted hereafter, and only a component  $\mathcal{W}(t)$  is considered in the following.

### 131 2.2. First-passage probability estimation by EVD

132 For a stochastic dynamic system, the first-passage probability is the probability that the system response of interest  
 133 exceeds a certain safe domain for the first time within a given time range. Accordingly, assuming  $T$  is the time duration,  
 134 we have

$$P_f = \Pr \{ \mathcal{W}(t) \notin \Omega_{\text{safe}}, \exists t \in [0, T] \}, \quad (5)$$

135 where  $P_f$  is first-passage probability;  $\Pr$  is probability operator;  $\Omega_{\text{safe}}$  denotes the safe domain. According to different  
 136 application backgrounds, the boundary of  $\Omega_{\text{safe}}$  can be different, such as one boundary, double boundary, and circle  
 137 boundary [7]. In the case of symmetric double boundary problem, the first-passage probability can be further written  
 138 as:

$$P_f = \Pr \{ |\mathcal{W}(t)| > b_{\text{lim}}, \exists t \in [0, T] \}, \quad (6)$$

139 where  $b_{\text{lim}}$  is the given threshold that limits the symmetric bounds of  $\Omega_{\text{safe}}$ , and  $|\cdot|$  is the absolute value operator. In the  
 140 present study, the first-passage probability defined by Eq. (6) is of concern.

141 Note that if the system response in the time period  $[0, T]$  remains below the boundary of  $\Omega_{\text{safe}}$ , the first-passage  
 142 probability will be equal to zero. From this perspective, once the extreme value of system response exceeds the  
 143 boundary, the system fails. Accordingly, Eq. (6) can be rewritten as

$$P_f = \Pr \{ \max \{ |\mathcal{W}(t)| \} > b_{\text{lim}}, \forall t \in [0, T] \} = \Pr \{ \mathcal{Z} > b_{\text{lim}} \}, \quad (7)$$

144 where  $\mathcal{Z} = \max_{t \in [0, T]} \{ |\mathcal{W}(t)| \}$ . Note that  $\mathcal{Z}$  is always positive, and depends on the random parameter vector  $\mathbf{U}$ . If we de-  
 145 note the functional relationship between  $\mathcal{Z}$  and  $\mathbf{U}$  as  $G$ , then we have  $\mathcal{Z} = G(\mathbf{U})$  and  $P_f = \Pr \{ \mathcal{Z} = G(\mathbf{U}) > b_{\text{lim}} \}$ .

146 According to classical probability theory, once the probability distribution of  $\mathcal{Z}$ , which is also referred to as extreme  
 147 value distribution (EVD), is obtained, Eq. (7) can be straightforwardly calculated from the EVD. Let  $f_{\mathcal{Z}}(z)$  and  $F_{\mathcal{Z}}(z)$   
 148 be the PDF and cumulative distribution function (CDF) of  $\mathcal{Z}$ . Then the first-passage probability reads

$$P_f = \int_{b_{\text{lim}}}^{+\infty} f_{\mathcal{Z}}(z) dz = 1 - F_{\mathcal{Z}}(b_{\text{lim}}). \quad (8)$$

149 It should be pointed out that the first-passage probability is easy to be obtained from Eq. (8) once the PDF or CDF  
 150 of  $\mathcal{Z}$  is known. However, how to estimate the EVD of  $\mathcal{Z}$  is quite challenging. This is because deriving an analytical  
 151 expression for the EVD is intractable even for some simple stochastic responses, not to mention the stochastic responses  
 152 of high-dimensional and strong-nonlinear stochastic dynamic systems. Therefore, to tackle such challenge, an EVD  
 153 estimation method is proposed in the following section.

154 **Remark 1.** The above-mentioned first-passage probability estimation method can also be applied to evaluate  
 155 the system failure probability for the first-passage problem considering multiple responses. According to the  
 156 theory of equivalent extreme-value event [38], the system failure probability for a first-passage problem can be  
 157 equivalent to the probability of an extreme-value event. Such extreme-value event is defined in terms of the  
 158 logical relationships between multiple inequalities corresponding to multiple responses. Besides, the correlation  
 159 information between each components is inherent in the equivalent extreme-value event. To illustrate, suppose  
 160  $\mathcal{Z}_1 = \max_{t \in [0, T]} \{ |\mathcal{W}_1(t)| \}$  and  $\mathcal{Z}_2 = \max_{t \in [0, T]} \{ |\mathcal{W}_2(t)| \}$ . Then, we can derive  $\Pr \{ (\mathcal{Z}_1 > b_1) \cup (\mathcal{Z}_2 > b_2) \} =$   
 161  $\Pr \left\{ \left( \mathcal{Z}_1 - b_1 > \hat{b} \right) \cup \left( \mathcal{Z}_2 - b_2 > \hat{b} \right) \right\} = \Pr \left\{ \max_{1 \leq q \leq 2} \{ \mathcal{Z}_q \} > \hat{b} \right\}$ , where  $b_1$  and  $b_2$  are the thresholds corre-  
 162 sponding to  $\mathcal{Z}_1$  and  $\mathcal{Z}_2$ , and  $\hat{b}$  is the common threshold obtained by a linear transformation. Accordingly, similar  
 163 to Eq. (8), the first-passage system probability can be computed as  $P_f = \Pr \left\{ \hat{\mathcal{Z}} > \hat{b} \right\} = \int_{\hat{b}}^{+\infty} f_{\hat{\mathcal{Z}}}(\hat{z}) d\hat{z}$ , where  
 164  $\hat{\mathcal{Z}} = \max_{1 \leq q \leq 2} \{ \mathcal{Z}_q \}$ .

### 165 3. A fractional moments-based mixture distribution approach

166 In this section, we propose a novel fractional moments-based mixture distribution approach to approximate the  
 167 EVD in an efficient and accurate way. The proposed method consists of two main parts. First, a parallel adaptive  
 168 scheme is proposed for fractional moments estimation, which allows sequential sample size extension until a prescribed  
 169 convergence criterion is satisfied. Second, from the knowledge of estimated fractional moments, an eight-parameter  
 170 mixture distribution model with increased flexibility is developed to capture the main body and distribution tail of the  
 171 EVD.

172 *3.1. Characterizing EVD by fractional moments*

173 The analytical expression of EVD can not be directly obtained for a general high-dimensional and nonlinear  
 174 stochastic dynamic system, as discussed earlier. To this end, we have to resort to some indirect methods that can  
 175 approximate the EVD from a limited number of sample data. The fractional moment, as a generalization of the  
 176 traditional integer moment, has received a growing interest to characterize a positive random variable in many fields.  
 177 More recently, it has also been introduced to the area of EVD characterization [31–33, 35].

178 *3.1.1. Concept and properties of fractional moments*

179 The  $r$ -th fractional moment of the positive random variable  $\mathcal{Z}$  is defined as [33]

$$M_{\mathcal{Z}}^r = E[\mathcal{Z}^r] = \int_0^{+\infty} z^r f_{\mathcal{Z}}(z) dz, \quad (9)$$

180 where  $r$  can be any real number and  $E[\cdot]$  denotes the expectation operator. Note that when  $r$  takes an integer value,  
 181 Eq. (9) yields the  $r$ -th integer moment of  $\mathcal{Z}$ . Therefore, for any positive random variable, the integer moment of the  
 182 variable is a special case of its fractional moment.

183 If one expands  $\mathcal{Z}^r$  around its mean value  $\mu_{\mathcal{Z}} = M_{\mathcal{Z}}^1$  using the Taylor series expansion, we have

$$\mathcal{Z}^r = \sum_{k=0}^{\infty} \binom{r}{k} \mu_{\mathcal{Z}}^{r-k} (z - \mu_{\mathcal{Z}})^k, \quad (10)$$

184 where the fractional binomial coefficient  $\binom{r}{k}$  can be computed as  $\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k(k-1)\cdots 1}$ , and  $k$  can be any non-negative  
 185 integer. Taking the expectation of both sides of Eq. (10) yields:

$$E[\mathcal{Z}^r] = \sum_{k=0}^{\infty} \binom{r}{k} \mu_{\mathcal{Z}}^{r-k} E[(z - \mu_{\mathcal{Z}})^k]. \quad (11)$$

186 It can be seen that the right-hand side of Eq. (11) contains an infinite number of integer moments, i.e.,  $E[(z - \mu_{\mathcal{Z}})^k]$ ,  
 187 and the left-hand side of Eq. (11) is exactly the  $r$ -th fractional moment. Hence, Eq. (11) implies that a single  $r$ -order  
 188 fractional moment can embody statistical information of numerous integer moments. Further, as observed from Eq.  
 189 (11), when  $r$  is fixed, the value of coefficient  $\binom{r}{k} \mu_{\mathcal{Z}}^{r-k}$  decreases as  $k$  increases; when  $k$  is fixed,  $\binom{r}{k} \mu_{\mathcal{Z}}^{r-k}$  increases  
 190 as  $r$  increases. This indicates that the higher the fractional order, the greater the contribution of higher-order integer  
 191 moments. Since higher-order integer moments can provide more information about the shape of EVD, higher-order  
 192 fractional moments reflect more statistical features of EVD than lower-order fractional moments. In addition, it  
 193 should be mentioned that higher-order fractional moments have higher variability and are more difficult to obtain than  
 194 lower-order fractional moments [28, 33]. Note that one is able to generate any number of fractional moments given the  
 195 range of fractional orders. However, one can only generate a fixed number of integer moments if the maximum integer  
 196 order is given. As a compromise, a set of fractional moments up to second order, as adopted in Ref. [33], is used in this  
 197 work.

198 *3.1.2. Parallel adaptive estimation of fractional moments*

199 According to the principle of probability conservation, Eq. (9) can be rewritten in the random variable space of  $\mathbf{U}$ :

$$M_{\mathcal{Z}}^r = \int_{\Omega_{\mathbf{U}}} G^r(\mathbf{u}) p_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}, \quad (12)$$

200 where  $\Omega_{\mathbf{U}}$  denotes the random variable space of  $\mathbf{U}$ . For a general stochastic dynamic system, a considerably large  
 201 number of random variables are collected in  $\mathbf{U}$ , and strong nonlinearity exists in  $G(\mathbf{U})$ . In addition, the expression of  
 202  $G(\mathbf{U})$  cannot be explicitly given. Hence, a high-dimensional integral with a complex and implicit integrand is involved  
 203 in Eq. (12), which is impossible to solve analytically.

204 Alternatively, we can resort to the sampling methods to approximate the high-dimensional integral involved in Eq.  
 205 (12). In the literature, various variance reduction sampling methods with fixed sample sizes are employed to facilitate  
 206 the estimation of fractional moments. Under this setting,  $M_{\mathcal{Z}}^r$  can be approximated as:

$$\hat{M}_{\mathcal{Z}}^r = \sum_{k=1}^N \varpi_k \cdot G^r(\mathbf{u}_k), \quad (13)$$

207 where  $N$  denotes the sample size;  $\varpi_k$  represents the  $k$ -th sample weight,  $k = 1, \dots, N$ ;  $\mathbf{u}_k$  is the  $k$ -th sample of random  
 208 variables  $\mathbf{U}$ . Note that most variance reduction sampling methods do not allow sample size extension, and thus require  
 209  $N$  to be specified in advance from experience. However, for estimating fractional moments, an ‘‘optimal sample size’’  
 210 is desired, which is problem-dependent, and cannot be known in advance for a specified first-passage problem. The  
 211 optimal sample size enables the estimation to strike a balance between accuracy and computational efficiency. However,  
 212 with a predefined sample size, the fractional moment estimation may lose such balance, and may be trapped into  
 213 over-sampling or under-sampling situations. Specifically, if an overly conservative sample size is pre-specified, i.e., too  
 214 many samples are taken, oversampling occurs and leads to unnecessary computational waste. On the other hand, if the  
 215 predefined sample size is too small, under-sampling takes place, resulting in inaccurate evaluation of the fractional  
 216 moments.

217 To tackle with such dilemma, an adaptive sampling scheme should be developed for estimating fractional moments.  
 218 One feasible strategy is to generate samples one at a time or several at a time, and enrich the sample size progressively  
 219 until a specified convergence criterion is satisfied. In this manner, sample size extension is allowed, and the sample size  
 220 can be obtained adapted to different problems, which enables the estimated fractional moments to achieve both the  
 221 desired accuracy and computational efficiency. In addition, parallel computing technique can be equipped to further  
 222 accelerate the computational speed of such process. As such, we shall name this sampling scheme as parallel adaptive  
 223 sampling scheme. To illustrate the advantages of proposed scheme, Fig. 1 shows the comparison between traditional  
 224 sampling scheme and proposed parallel adaptive sampling scheme. In this figure,  $l$  denotes the  $l$ -th time of sample  
 225 size extension, and  $l \in \mathbb{Z}^+$ . As seen, by the proposed sampling scheme, the sample size for a given first-passage  
 226 problem can be determined in an adaptive way, where fractional moments can be approximated with a desired accuracy.  
 227 In addition, it is quite time-saving to evaluate additional samples of  $\mathcal{Z}$  only when it is required. In the process of  
 228 estimating the additional samples of  $\mathcal{Z}$ , the analysis time can be further decreased by adopting parallel computing  
 229 technique.

230 By employing the proposed parallel adaptive sampling scheme,  $\hat{M}_{\mathcal{Z}}^r$  after the  $l$ -th sample size extension can be  
 231 computed as follows:

$$\hat{M}_{\mathcal{Z}}^r = \sum_{k=1}^{(l-1)\hbar} \varpi^{(k)} \cdot G^r(\mathbf{u}^{(k)}) + \sum_{k=(l-1)\hbar+1}^{l\hbar} \varpi^{(k)} \cdot G^r(\mathbf{u}^{(k)}), \quad (14)$$

232 where the number of samples added in each time of sample size extension is denoted as  $\hbar$  and  $\hbar \in \mathbb{Z}^+$ ; the cur-  
 233 rent sample size is  $l\hbar$ ; the weight is reallocated in the  $l$ -th sample size extension and satisfies  $\sum_{k=1}^{l\hbar} \varpi^{(k)} = 1$ ;  
 234  $\{\mathbf{u}^{((l-1)\hbar+1)}, \dots, \mathbf{u}^{(l\hbar)}\}$  are the newly added samples in the  $l$ -th sample size extension, while  $\{\mathbf{u}^{(1)}, \dots, \mathbf{u}^{((l-1)\hbar)}\}$  are

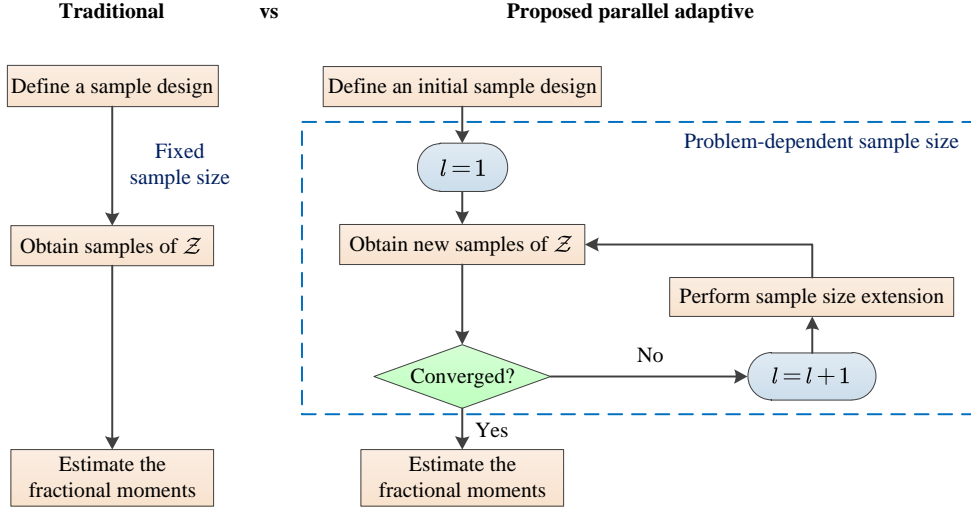


Figure 1: Comparison of traditional sampling scheme and proposed parallel adaptive scheme

235 samples generated in the previous  $(l - 1)$  sample size extensions. Note that when  $l = 1$ , initial samples of  $\mathcal{Z}$ , i.e.,  
 236  $\{G(\mathbf{u}^{(k)})\}_{k=1}^{\bar{h}}$  are evaluated. Since  $\{G(\mathbf{u}^{(k)})\}_{k=1}^{(l-1)\bar{h}}$  have been already obtained in the previous  $(l - 1)$  sample size  
 237 extensions, one only needs to evaluate  $\{G(\mathbf{u}^{(k)})\}_{k=(l-1)\bar{h}+1}^{l\bar{h}}$  in the  $l$ -th sample size extension.

238 In order to achieve the proposed parallel adaptive sampling scheme, the key is to employ a sampling strategy that  
 239 allows sequential sample size extension. Simple random sampling method, i.e., Monte Carlo simulation (MCS), can  
 240 naturally meet such aim. To obtain a better precision of fractional moments with fewer computational efforts, one  
 241 can apply a variance reduction sampling method to the proposed sampling scheme. In addition, sampling methods  
 242 that are applicable to high-dimensional problems are also desired. In fact, one recently developed sequential stratified  
 243 sampling technique, termed refined Latinized stratified sampling (RLSS) [37], is suitable for our purposes. On one  
 244 hand, RLSS is advantageous as it owns the ability to achieve effective variance reduction in terms of both main/additive  
 245 effects and variable interaction that appear in  $G(\mathbf{U})$ . On the other hand, RLSS is applicable to problems involving low-  
 246 and high-dimensional input random variables. By using the RLSS technique, we can evaluate  $\hat{M}_{\mathcal{Z}}^r$  according to Eq.  
 247 (14). Since the samples of RLSS are generated in the  $[0, 1]^{n_s}$  hyper-rectangular space, we need to transform the RLSS  
 248 sample points to the original distribution domain of random variables  $\mathbf{U}$ . Denote  $\hat{\varphi}^{(k)}$  and  $\varpi^{(k)}$  to be the  $k$ -th sample  
 249 point and corresponding weight obtained by RLSS and  $\Gamma$  to be the transformation operator,  $\hat{M}_{\mathcal{Z}}^r$  by RLSS at the  $l$ -th  
 250 sample size extension can be evaluated as:

$$\hat{M}_{\mathcal{Z}}^r = \sum_{k=1}^{(l-1)\bar{h}} \varpi^{(k)} \cdot G^r(\Gamma(\hat{\varphi}^{(k)})) + \sum_{k=(l-1)\bar{h}+1}^{l\bar{h}} \varpi^{(k)} \cdot G^r(\Gamma(\hat{\varphi}^{(k)})). \quad (15)$$

251 A brief illustration of the RLSS technique is discussed in the following. For more details, the interested readership  
 252 can refer to Appendix A or Ref. [37].

253 The first step of RLSS is generating  $\mathcal{N} \geq 1$  samples that follow a so-called Latinized stratified sampling (LSS)  
 254 scheme [39], which implies that these samples fulfill both the properties of Latin hypercube sampling (LHS) and  
 255 stratified sampling (SS). An schematic diagram of a LSS design is shown in Fig. 2(a), considering  $\mathcal{N} = 4$  and  $n_s = 2$ .  
 256 In this figure, the strata associated with LHS are shown with dashed black line, the strata associated with SS are marked



257 with solid green line, the samples per each dimension of analysis are marked with blue cross marks and the actual  
 258 samples are marked with blue dots. It is readily observed that the strata associated with SS possess the same area, and  
 259 boundaries of the strata associated with LHS match those associated with SS, which are the key properties of LSS.

260 The second step of RLSS consists of applying a Hierarchical Latin hypercube sampling (HLHS) design [37] over  
 261 the existing LHS design. This implies applying a refinement of each LHS strata by subdividing it  $\delta$  times, which is  
 262 illustrated schematically in Fig. 2(b), where  $\delta = 1$ . The new strata associated with LHS are shown with red dashed line  
 263 and the new candidate samples per each dimension on those strata are marked with orange cross marks. Note that up to  
 264 this point, no new actual samples have been generated. In addition, one identifies *candidate strata* for refining the SS  
 265 design by dividing the existing strata, which is shown schematically in Fig. 2(b) with blue solid lines.

266 The third step involves generating new *candidate samples* for RLSS. In this sense, candidate samples are those that  
 267 may include the already existing  $\mathcal{N}$  samples. These candidate samples must be identified following a special procedure  
 268 such that the properties of LSS continue being fulfilled. For materializing this third step, one must identify the strata  
 269 which must contain candidate samples in order to enforce the LSS condition, and the strata where candidate samples  
 270 can be generated randomly. This is illustrated schematically in Fig. 2(c). The pink color indicates those strata that must  
 271 contain candidate samples, while the yellow color shows those strata where a candidate sample may be generated at  
 272 random. With all these considerations, one can generate  $\mathcal{N}\delta$  candidate designs, as shown schematically in Fig. 2(c)  
 273 with 4 orange dots.

274 The fourth step of RLSS is to incorporate a batch of  $\bar{h}$  samples to the existing set of  $\mathcal{N}$  samples. This is performed  
 275 by selecting at random from the existing  $\mathcal{N}\delta$  candidate samples. Note that in this process, it is necessary to update the  
 276 strata associated with SS taking into account the candidate strata defined in the second step. Clearly, in such update,  
 277 one must also update the weights (areas) of the selected strata. Fig. 2(d) illustrates the case where  $\bar{h} = 4$  and also  
 278 shows the updated strata with green solid line.

279 It should be mentioned that the fourth step described above can be repeated as many times as necessary to select  
 280 many batches of  $\bar{h}$  samples as long as there are candidate samples left. In case one runs out of candidate samples,  
 281 it is necessary to return to the second step and perform a new run of HLHS, which implies subdividing the strata  
 282 associated with LHS. Furthermore, after each sample size extension, generated RLSS samples contain not only batches  
 283 of additional samples, but also samples from the initial LSS design. In this work, we take  $\bar{h} \geq \mathcal{N}$  in order to include  
 284 the initial LSS design in the initial RLSS samples when  $l = 1$  in Eq. (15).

285 In the proposed sampling scheme, a proper convergence criterion should be developed to determine the desired  
 286 number of sample size extensions. It is found that higher-order fractional moment always exhibits larger variability  
 287 than its lower-order counterpart. Accordingly, if the variability of maximum order fractional moment is controlled,  
 288 the variability of the lower-order ones will be automatically below a desired level. Note that the maximum order of  
 289 fractional moments is set to be 2 in this work, as mentioned in Section 3.1.1. Therefore, a convergence criterion is  
 290 proposed by judging the variability of the second-order fractional moment  $\hat{M}_{\mathcal{Z}}^2$  evaluated by RLSS. Specifically, the  
 291 coefficient of variation (COV) of the  $\hat{M}_{\mathcal{Z}}^2$  is compared with a user-defined small value  $\mathcal{E}$  (e.g.,  $\mathcal{E} = 0.02$ ) to determine  
 292 when to stop the sample size extension. The stopping criterion is defined as:

$$\text{COV} \left\{ \hat{M}_{\mathcal{Z}}^2 \right\} < \mathcal{E}. \quad (16)$$

293 Although the expression of  $\text{COV} \left\{ \hat{M}_{\mathcal{Z}}^2 \right\}$  is not available for RLSS, the bootstrap resampling technique [40] can be  
 294 alternatively implemented here to estimate it. Note that traditional bootstrap method generates samples with equal

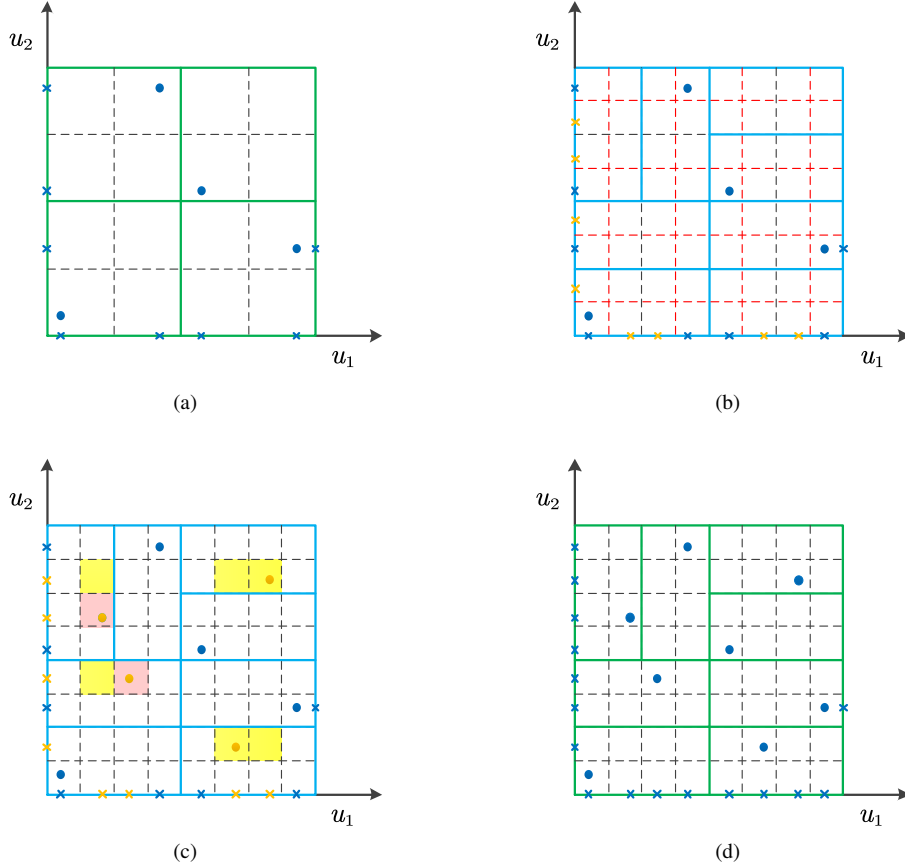


Figure 2: Schematic description of the RLSS technique for generating 8 samples in two dimensions

295 probability of occurrence, which is not the case for RLSS samples. To consider the unequal weight property of  
 296 RLSS samples, the approach proposed in Ref. [41] is adopted here, such that samples with higher weights have more  
 297 probability of being chosen for bootstrap. For more details on this approach, it is referred to Ref. [41].

298 With such parallel adaptive scheme above, once the samples of  $\mathcal{Z}$  that meet the convergence condition are obtained,  
 299 a set of lower-order (only up to 2) fractional moments can be estimated according to Eq. (15), which are then used to  
 300 represent the EVD.

### 301 3.2. Representing EVD by a mixture distribution with fractional moments

302 After obtaining the fractional moments of  $\mathcal{Z}$ , an adequate probability distribution model should be employed for the  
 303 EVD estimation. Generally, the state-of-art distribution models represent the EVD by adopting either maximum entropy  
 304 density [31, 32] or positively skewed distributions such as shifted generalized lognormal distribution [35] and a mixture  
 305 of lognormal distribution and inverse Gaussian distribution [33]. However, their flexibility is still limited for the EVDs  
 306 with heavy tails, leading to the inaccuracy of EVD reconstruction for some first-passage problems. To increase the  
 307 flexibility and enlarge the application scope, we first extend the traditional inverse Gaussian distributions by introducing  
 308 an exponential transformation with an additional shape parameter. Then, we introduce the log transformation to  
 309 the extended skew-normal distribution, to enhance its ability to accommodate fat tails. Further, these two improved

310 distributions are mixed together to produce a more flexible mixture distribution model, whose involved parameters can  
 311 be estimated from the estimated fractional moments.

### 312 3.2.1. Proposed extended inverse Gaussian distribution

313 The inverse Gaussian distribution (IGD) is a two-parameter skewed unimodal distribution and applies for positive  
 314 real values [42]. It is a first-passage time distribution for the Brownian motion with positive drift [43]. The PDF of the  
 315 IGD is:

$$f_{\text{IGD}}(z; a, b) = \sqrt{\frac{b}{2\pi z^3}} \exp\left[-\frac{b(z-a)^2}{2za^2}\right], \quad \text{with } z > 0, \quad (17)$$

316 where  $a > 0$  is the location parameter;  $b > 0$  is the shape parameter.

317 Denote the random variable which follows an IGD as  $\mathcal{Z}_{\text{IGD}}$ . The  $r$ -th fractional moment of  $\mathcal{Z}_{\text{IGD}}$  is given as:

$$M_{\mathcal{Z}_{\text{IGD}}}^r = E[\mathcal{Z}_{\text{IGD}}^r] = \int_0^{+\infty} z^r f_{\text{IGD}}(z) dz = \exp\left[\frac{b}{a}\right] \sqrt{\frac{2b}{\pi}} a^{r-1/2} K_{1/2-r}\left(\frac{b}{a}\right), \quad (18)$$

318 where  $K_\alpha(\beta)$  is the modified Bessel function of the second kind.

319 In fact, the IGD can be extended to obtain higher flexibility in its shape. Here, we introduce a transformation  
 320  $X = \mathcal{Z}^{1/\eta}$  to extend the original distribution, where  $\eta > 0$  is a shape parameter. The resulting distribution is called  
 321 extended inverse Gaussian distribution (EIGD). To obtain the PDF and fractional moments of the EIGD, the following  
 322 theorem is first given:

323 **Theorem 1.** Assume  $X$  and  $\mathcal{Z}$  are two continuous and positive real-valued random variables, and  $f_{\mathcal{Z}}(z)$  is already  
 324 available. Let  $X = \mathcal{Z}^{1/\eta}$  where  $\eta > 0$ , then we have  $f_X(x) = f_{\mathcal{Z}}(x^\eta) \cdot \eta \cdot x^{\eta-1}$ . Additionally, the  $r$ -th fractional  
 325 moment of  $X$  is  $E[X^r] = E[\mathcal{Z}^{r/\eta}]$ .

326 **Proof.** Since  $X = \mathcal{Z}^{1/\eta}$ , according to the principle of conservation of probability, it is straightforward to derive  
 327  $f_{\mathcal{Z}}(z) dz = f_X(x) dx$ . Thus, the PDF of  $X$  can be derived as  $f_X(x) = f_{\mathcal{Z}}(z) \frac{dz}{dx} = f_{\mathcal{Z}}(x^\eta) \cdot \eta \cdot x^{\eta-1}$ . We may also  
 328 derive the relationship between the  $r$ -th fractional moment of  $X$  and that of  $\mathcal{Z}$  as  $E[X^r] = E[(\mathcal{Z}^{1/\eta})^r] = E[\mathcal{Z}^{r/\eta}]$ .

329 Therefore, the PDF of EIGD reads:

$$f_{\text{EIGD}}(x; \eta, a, b) = \eta \sqrt{\frac{b}{2\pi}} x^{-\eta/2-1} \exp\left[-\frac{b(x^\eta - a)^2}{2x^\eta a^2}\right], \quad \text{with } x > 0. \quad (19)$$

330 Denote the random variable which follows the EIGD as  $X_{\text{EIGD}}$ . According to Eq. (18) and **Theorem 1**, the  $r$ -th  
 331 fractional moment of  $X_{\text{EIGD}}$  can be derived in analytic form:

$$M_{X_{\text{EIGD}}}^r = \exp\left[\frac{b}{a}\right] \sqrt{\frac{2b}{\pi}} a^{r/\eta-1/2} K_{1/2-r/\eta}\left(\frac{b}{a}\right). \quad (20)$$

332 Note that when  $\eta = 1$ , the EIGD reduces to the IGD according to Eq. (19). The limit or special cases of IGD also  
 333 belong to the EIGD, such as the chi-square distribution with single degree of freedom, normal distribution and Lévy  
 334 distribution. Besides, the shape flexibility of the EIGD is illustrated by Fig. 3 under four different sets of parameters.  
 335 In this figure, we make a comparison between the original IGD and the proposed EIGD by changing parameter  $\eta$  and  
 336 fixing  $a = 1, b = 1$  of the EIGD. It can be observed that, the proposed EIGD possesses much more flexibility in shape  
 337 of PDF than the original IGD.

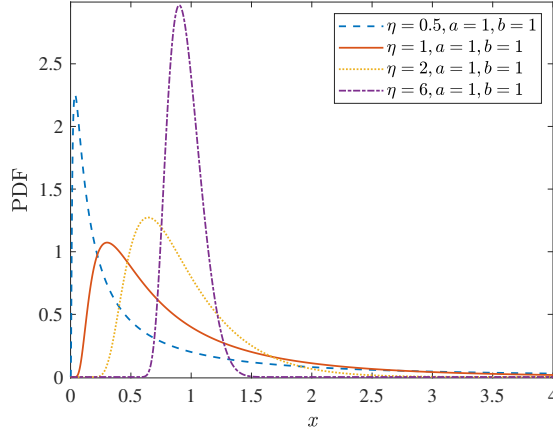


Figure 3: PDFs of extended inverse Gaussian distribution under four different sets of parameters

### 3.2.2. Proposed log extended skew-normal distribution

The extended skew-normal distribution (ESND) was first introduced by Azzalini [44]. This distribution is a four-parameter unimodal asymmetric distribution with support on  $(-\infty, +\infty)$ , which generalizes the traditional skew-normal distribution and normal distribution. The statistical properties of the ESND are discussed in detail in Ref. [45]. The PDF of the ESND of a real random variable  $\tilde{X} \in \mathbb{R}$  is:

$$f_{\text{ESND}}(\tilde{x}; c, d, \theta, \tau) = \frac{1}{d} \phi\left(\frac{\tilde{x} - c}{d}\right) \frac{\Phi\left(\tau\sqrt{1 + \theta^2} + \theta\frac{\tilde{x} - c}{d}\right)}{\Phi(\tau)}, \quad \text{with } \tilde{x} \in \mathbb{R}, \quad (21)$$

where  $c \in \mathbb{R}$  is the location parameter;  $d > 0$  is the scale parameter;  $\theta \in \mathbb{R}$  is the shape parameter;  $\tau \in \mathbb{R}$  is the truncation parameter;  $\phi(\cdot)$  and  $\Phi(\cdot)$  denote the PDF and CDF of the standard normal distribution.

The moment-generating function (MGF) of the ESND is:

$$M_{\tilde{X}}(\tilde{t}) = E\left[\exp(\tilde{t}\tilde{X})\right] = \exp\left(c\tilde{t} + \frac{1}{2}d^2\tilde{t}^2\right) \frac{\Phi\left(\tau + \frac{\theta d\tilde{t}}{\sqrt{1 + \theta^2}}\right)}{\Phi(\tau)}, \quad \text{with } \tilde{t} \in \mathbb{R}. \quad (22)$$

Although the ESND enables to accommodate asymmetry characteristics, its ability to fit heavier tails can be further improved by introducing a log transformation to the ESND. We shall refer the newly generated distribution as log extended skew-normal distribution (LESND). Denote the random variable which follows a LESND as  $X_{\text{LESND}}$ . Then, we have the relationship between  $X_{\text{LESND}}$  and  $\tilde{X}$  as  $X_{\text{LESND}} = \exp(\tilde{X})$ . That is, the logarithm of  $X_{\text{LESND}}$  follows the original ESND. Hence, we can get the PDF of the LESND as:

$$f_{\text{LESND}}(x; c, d, \theta, \tau) = \frac{1}{dx} \phi\left(\frac{\log(x) - c}{d}\right) \frac{\Phi\left(\tau\sqrt{1 + \theta^2} + \theta\frac{\log(x) - c}{d}\right)}{\Phi(\tau)}, \quad \text{with } x > 0. \quad (23)$$

From the relationship between the fractional moment of the LESND and the MGF of the ESND, it is easy to derive  $M_{X_{\text{LESND}}}^r = E[X_{\text{LESND}}^r] = E\left[\left(\exp(\tilde{X})\right)^r\right] = M_{\tilde{X}}(r)$ . Hence, the  $r$ -th fractional moment of  $X_{\text{LESND}}$  can be given in analytic form as:

$$M_{X_{\text{LESND}}}^r = \exp\left(cr + \frac{1}{2}d^2r^2\right) \frac{\Phi\left(\tau + \frac{\theta dr}{\sqrt{1 + \theta^2}}\right)}{\Phi(\tau)}. \quad (24)$$

354 Note that according to Eq. (23), when  $\tau = 0$ , the LESND reduces to the log skew-normal distribution [46]; and  
 355 when  $\theta = 0$ , the LESND reduces to the traditional lognormal distribution. It should be mentioned that if  $\theta = 0$ , the  
 356 shape of LESND will not be affected by changing the value of parameter  $\tau$ . Besides, to illustrate the flexibility of the  
 357 LESND, Fig. 4 depicts the LESND with four sets of parameters. In this figure, the log skew-normal distribution is  
 358 given for comparison by setting the parameters of LESND as  $c = 0, d = 1, \theta = 3, \tau = 0$ . As can be seen, the LESND  
 359 provides richer distribution shapes compared to the log skew-normal distribution, showing the increased flexibility of  
 360 LESND.

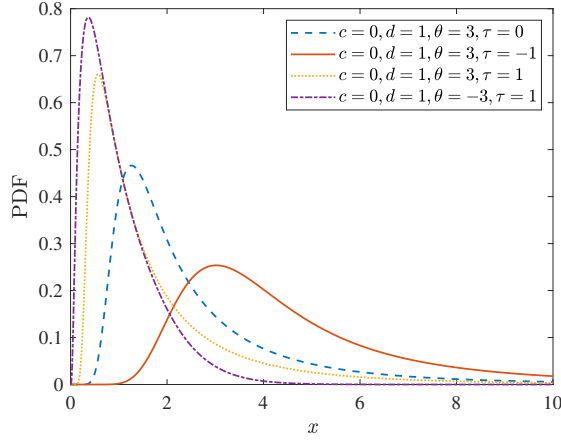


Figure 4: PDFs of log extended skew-normal distribution under four different sets of parameters

### 361 3.2.3. Proposed mixture distribution

362 It is worth mentioning that the first-passage probability estimation is closely associated to the distribution tail  
 363 of EVD. Besides, the EVD is usually asymmetric and possesses heavy tail in many cases. Hence, a highly flexible  
 364 distribution model is needed, which is suitable for fitting distributions with various tail properties, especially the  
 365 heavy-tailed distributions. For accurate EVD estimation, two single-component skewed distributions proposed above,  
 366 i.e., the EIGD and LESND, may still not be flexible enough and their applicability to various first-passage problems is  
 367 limited. To further improve the flexibility, one potential way is to mix the proposed single-component distributions  
 368 together by introducing a weight parameter. Such distribution model enables to incorporate both characteristics of two  
 369 single-component distributions, and can accommodate asymmetry in a more flexible way so as to properly estimate  
 370 the EVD. Therefore, motivated by the above, a novel mixture of the extended inverse Gaussian and log extended  
 371 skew-normal distributions (M-EIGD-LESND) is developed herein.

372 The PDF of M-EIGD-LESND is given as:

$$\begin{aligned}
 f_{\text{M-EIGD-LESND}}(x; \mathcal{Y}) &= w f_{\text{EIGD}}(x; \eta, a, b) + (1 - w) f_{\text{LESND}}(x; c, d, \theta, \tau) \\
 &= w \eta \sqrt{\frac{b}{2\pi}} x^{-\eta/2-1} \exp\left[-\frac{b(x^\eta - a)^2}{2x^\eta a^2}\right] + (1 - w) \frac{1}{dx} \phi\left(\frac{\log(x) - c}{d}\right) \frac{\Phi(\tau\sqrt{1+\theta^2} + \theta\frac{\log(x) - c}{d})}{\Phi(\tau)}, \quad \text{with } x > 0,
 \end{aligned} \tag{25}$$

373 where  $\mathcal{Y} = [w, \eta, a, b, c, d, \theta, \tau]$  is the set of eight unknown parameters and  $w \in [0, 1]$  is the weight parameter of  
 374 M-EIGD-LESND.

375 According to Eqs. (20) and (24), the  $r$ -th fractional moment of M-EIGD-LESND can be given in analytic form:

$$\begin{aligned}
 M_{X_{\text{M-EIGD-LESND}}}^r &= E[X_{\text{M-EIGD-LESND}}^r; \mathcal{X}] = wE[X_{\text{EIGD}}^r] + (1-w)E[X_{\text{LESND}}^r] \\
 &= w \exp\left[\frac{b}{a}\right] \sqrt{\frac{2b}{\pi}} a^{r/\eta-1/2} K_{1/2-r/\eta}\left(\frac{b}{a}\right) + (1-w) \exp\left(cr + \frac{1}{2}d^2r^2\right) \frac{\Phi\left(\frac{\tau + \frac{\theta dr}{\sqrt{1+\theta^2}}}{\Phi(\tau)}\right)}{\Phi(\tau)}.
 \end{aligned} \tag{26}$$

376 Note that the proposed M-EIGD-LESND can reduce to the mixture of lognormal and inverse Gaussian distributions  
 377 [33] if set  $\eta = 1$  and  $\theta = 0$ . To illustrate the flexibility of the proposed mixture distribution model, Fig. 5 shows the  
 378 plot of the PDFs associated with M-EIGD-LESND with different parameters. It can be seen that the proposed mixture  
 379 distribution model is highly flexible with rich shapes and enables to accommodate various heavy tails. In addition, the  
 380 M-EIGD-LESND is able to represent not only unimodal distributions but also bimodal distributions.

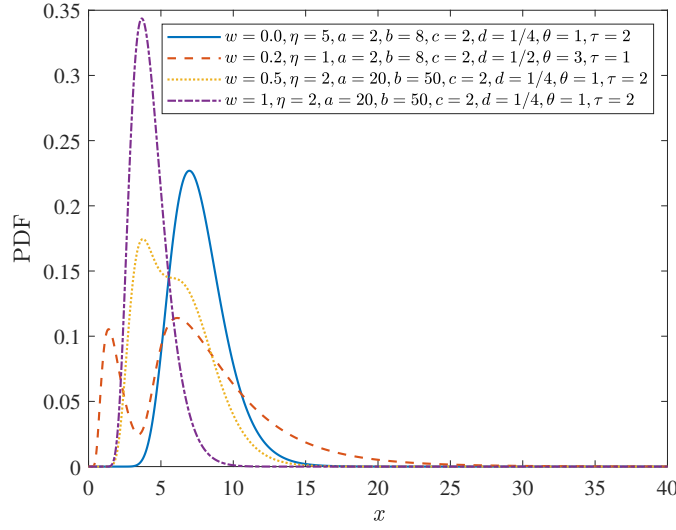


Figure 5: PDFs of the proposed mixture distribution under four different sets of parameters

### 381 3.2.4. Parameter estimation

382 The proposed mixture distribution model has the potential to characterize the EVD. Hence, in order to recover the  
 383 EVD of  $\mathcal{Z}$ , we assume that the EVD follows the proposed mixture distribution model, and determine the free parameters  
 384 of this model in an appropriate way. Note that the proposed distribution contains a set of eight free parameters. To  
 385 estimate these unknown distribution parameters, a natural way is to match the fractional moments of the proposed  
 386 mixture distribution model with the estimated fractional moments of the corresponding orders (hereafter referred to as  
 387 the fractional moment matching technique). Accordingly, the following nonlinear system of equations requires to be  
 388 solved:

$$\begin{cases} \hat{M}_{\mathcal{Z}}^{r_1} = M_{X_{\text{M-EIGD-LESND}}}^{r_1} \\ \hat{M}_{\mathcal{Z}}^{r_2} = M_{X_{\text{M-EIGD-LESND}}}^{r_2} \\ \dots \\ \hat{M}_{\mathcal{Z}}^{r_8} = M_{X_{\text{M-EIGD-LESND}}}^{r_8} \end{cases}, \tag{27}$$

389 where  $\hat{M}_{\mathcal{Z}}^{r_i}, i = 1, 2, \dots, 8$  are the  $r_i$ -th fractional moments estimated by RLSS;  $M_{X_{\text{M-EIGD-LESND}}}^{r_i}$  can be obtained by  
 390 Eq. (26); and the fractional order  $r_i$  takes  $[r_1, r_2, \dots, r_8] = \frac{2}{8} \times [1, 2, \dots, 8]$ . Here, the equally spaced fractional orders

are introduced for convenience, since it is straightforward to take such value without any prior knowledge of fractional orders. Besides, as adopted in Ref. [33], the maximum fractional order is set to be 2, since the second-order fractional moment can be estimated efficiently from a small number of samples, and it reflects more shape information of EVD than lower-order fractional moments, as discussed in Section 3.1.1.

Solution to Eq. (27) can be obtained in seconds by any appropriate nonlinear solver, such as *lsqnonlin* in Matlab. To facilitate the solving process, initial values for the free parameters are required. Denote the initial values of Eq. (27) as  $w_0, \hat{\eta}_0, \hat{a}_0, \hat{b}_0, \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0$ .  $w_0$  is set to be 0.5 to assign an equal initial weights to the two single-component functions. The other initial values, i.e.,  $\hat{\eta}_0, \hat{a}_0, \hat{b}_0, \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0$ , can be obtained by another low-order fractional moment matching technique, where a nonlinear system of equations is involved:

$$\begin{cases} \hat{M}_{\mathcal{Z}}^{1/2} = E \left[ X_{\text{EIGD}}^{1/2}; \hat{\eta}_0, \hat{a}_0, \hat{b}_0 \right] \\ \hat{M}_{\mathcal{Z}}^1 = E \left[ X_{\text{EIGD}}^1; \hat{\eta}_0, \hat{a}_0, \hat{b}_0 \right] \\ \hat{M}_{\mathcal{Z}}^{3/2} = E \left[ X_{\text{EIGD}}^{3/2}; \hat{\eta}_0, \hat{a}_0, \hat{b}_0 \right], \end{cases} \quad (28)$$

where  $\hat{\eta}_0 > 0, \hat{a}_0 > 0, \hat{b}_0 > 0$ ; and

$$\begin{cases} \hat{M}_{\mathcal{Z}}^{1/2} = E \left[ X_{\text{LESND}}^{1/2}; \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0 \right] \\ \hat{M}_{\mathcal{Z}}^1 = E \left[ X_{\text{LESND}}^1; \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0 \right] \\ \hat{M}_{\mathcal{Z}}^{3/2} = E \left[ X_{\text{LESND}}^{3/2}; \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0 \right], \\ \hat{M}_{\mathcal{Z}}^2 = E \left[ X_{\text{LESND}}^2; \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0 \right], \end{cases} \quad (29)$$

where  $\hat{c}_0 \in \mathbb{R}, \hat{d}_0 > 0, \hat{\eta}_0 \in \mathbb{R}, \hat{\tau}_0 \in \mathbb{R}$ . Note that the M-EIGD-LESND can reduce to the inverse Gaussian distribution (if set  $w = 0, \eta = 1$ ) or the lognormal distribution (if set  $w = 1, \theta = 0$ ), and the relationships between the parameters and the first two central moments of each reduced distribution are easy to be obtained. Besides, as discussed earlier, the value of parameter  $\tau$  will be irrelevant if  $\theta = 0$ . Hence, the initial values for Eqs (28) and (29) can be determined as:  $a_0 = \hat{\mu}_{\mathcal{Z}}, b_0 = \hat{\mu}_{\mathcal{Z}}^3 / \hat{\sigma}_{\mathcal{Z}}^2, \eta_0 = 1, c_0 = \log \left( \hat{\mu}_{\mathcal{Z}}^2 / \sqrt{\hat{\sigma}_{\mathcal{Z}}^2 + \hat{\mu}_{\mathcal{Z}}^2} \right), d_0 = \sqrt{\log \left( \hat{\sigma}_{\mathcal{Z}}^2 / \hat{\mu}_{\mathcal{Z}}^2 + 1 \right)}, \theta_0 = 0, \tau_0 = 0$ , where  $\hat{\mu}_{\mathcal{Z}} = \hat{M}_{\mathcal{Z}}^1$  and  $\hat{\sigma}_{\mathcal{Z}} = \sqrt{\hat{M}_{\mathcal{Z}}^2 - \left( \hat{M}_{\mathcal{Z}}^1 \right)^2}$ . The parameter estimation process of proposed M-EIGD-LESND is briefly summarized in Algorithm 1.

---

**Algorithm 1** Parameter estimation for M-EIGD-LESND using the fractional moment matching technique

---

**Input:** central moments  $\hat{\mu}_{\mathcal{Z}}, \hat{\sigma}_{\mathcal{Z}}$ , and fractional moments  $\hat{M}_{\mathcal{Z}}^{\mathbf{r}}$  ( $\mathbf{r} = [\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}, 2]$ ).

**Output:** estimated distribution parameters  $\mathcal{T} = [w, \eta, a, b, c, d, \theta, \tau]$ .

- 1: Use  $\hat{\mu}_{\mathcal{Z}}$  and  $\hat{\sigma}_{\mathcal{Z}}$  to evaluate  $\eta_0, a_0, b_0, c_0, d_0, \theta_0, \tau_0$  as the initial values of Eqs. (28) and (29);
  - 2: Solve Eqs. (28) and (29) with  $\eta_0, a_0, b_0, c_0, d_0, \theta_0, \tau_0$  to estimate the initial values  $\hat{\eta}_0, \hat{a}_0, \hat{b}_0, \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0$  of Eq. (27).
  - 3: Solve the fractional moment matching equations (Eq. (27)) by means of any appropriate nonlinear solver with  $\hat{\eta}_0, \hat{a}_0, \hat{b}_0, \hat{c}_0, \hat{d}_0, \hat{\theta}_0, \hat{\tau}_0$  and  $w_0 = 0.5$ , and then obtain the estimated distribution parameters  $\mathcal{T} = [w, \eta, a, b, c, d, \theta, \tau]$  of M-EIGD-LESND.
-

### 408 3.3. Procedure of the proposed method

409 Once the EVD is reconstructed by the proposed probability distribution model, the first-passage probability can  
 410 be evaluated by Eq. (8) for a given threshold. A flowchart of the proposed method is shown in Fig. 6, and a brief  
 411 procedure is summarized as follows:

412  
 413 **Step 1:** Initialization. Set the initial sample size  $\mathcal{N}$  of LSS, the refinement factor  $\delta$  of HLHS, the number of samples  
 414  $h$  added in each sample size extension and the value of tolerance  $\mathcal{E}$ . Determine the threshold  $b_{\text{lim}}$ .

415 **Step 2:** Generate  $h$  new samples by RLSS. Produce  $h$  new samples and update the weights by RLSS method  
 416 according to Algorithm 2 in Appendix A, and then compute the new samples of  $\mathcal{Z}$ .

417 **Step 3:** Judge the convergence criterion. Evaluate the COV of  $\hat{M}_{\mathcal{Z}}^2$  by using bootstrap technique. If Eq. (16) is  
 418 satisfied, then turn to step 4; otherwise, return to step 2.

419 **Step 4:** Moment evaluation. Calculate a set of fractional moments  $\hat{M}_{\mathcal{Z}}^{\mathbf{r}}$  ( $\mathbf{r} = [\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}, 2]$ ) according to  
 420 Eq. (15), and then compute the first-two central moments  $\hat{\mu}_{\mathcal{Z}}$  and  $\hat{\sigma}_{\mathcal{Z}}$  by  $\hat{\mu}_{\mathcal{Z}} = \hat{M}_{\mathcal{Z}}^1$  and  $\hat{\sigma}_{\mathcal{Z}} = \sqrt{\hat{M}_{\mathcal{Z}}^2 - (\hat{M}_{\mathcal{Z}}^1)^2}$ .

421 **Step 5:** EVD representation. Represent the EVD by using the proposed distribution model, i.e., M-EIGD-LESND,  
 422 where the involved free distribution parameters are estimated by the low-order fractional moment matching technique  
 423 described in Algorithm 1.

424 **Step 6:** First-passage probability estimation. Evaluate the first-passage probability  $P_f = \Pr\{\mathcal{Z} > b_{\text{lim}}\}$  via  
 425 obtained EVD and Eq. (8).

426

## 427 4. Numerical examples

428 In this section, three examples, including two test examples and one practical engineering example, will be  
 429 investigated to verify the efficacy of the proposed method. In all examples, the parameters of the proposed method  
 430 are set as  $\mathcal{N} = 1$ ,  $\delta = 1$ ,  $h = 8$  and  $\mathcal{E} = 0.015$ . The computational efficiency and accuracy of proposed methods  
 431 for first-passage probability estimation are compared with MCS, Subset simulation (SS) [21, 23] and two state-of-art  
 432 mixture distribution methods presented in Ref. [33] and [34]. Note that in SS, the number of samples per layer is  
 433 1000 and the conditional probability is 0.1. Both the existing mixture distribution methods for comparison employ the  
 434 Latinized partially stratified sampling (LPSS) to evaluate fractional moments of  $\mathcal{Z}$ . The mixture distribution method  
 435 in Ref. [33] develops a mixture distribution combining conventional inverse Gaussian and lognormal distributions  
 436 (MIGLD), and thus this method is referred as LPSS+MIGLD in the following examples. Another existing mixture  
 437 distribution method [34] develops a mixture of two generalized inverse Gaussian distributions (MTGIG), and this  
 438 method is denoted as LPSS+MTGIG in the following examples.

### 439 4.1. Example 1: a Duffing oscillator under Gaussian white noise

440 The first example considers a Duffing oscillator with uncertain parameters under Gaussian white noise, which is  
 441 governed by

$$\ddot{Y}(t) + \gamma \dot{Y}(t) + Y(t) + \varepsilon Y^3(t) = \mathcal{G}(t), \quad (30)$$

442 where  $\ddot{Y}$ ,  $\dot{Y}$  and  $Y$  are the acceleration, velocity and displacement at time  $t$ ;  $\gamma$  denotes the damping control coefficient;  
 443  $\varepsilon$  is the parameter controlling the degree of nonlinearity in the restoring force; and  $\mathcal{G}(t)$  is the Gaussian white noise.



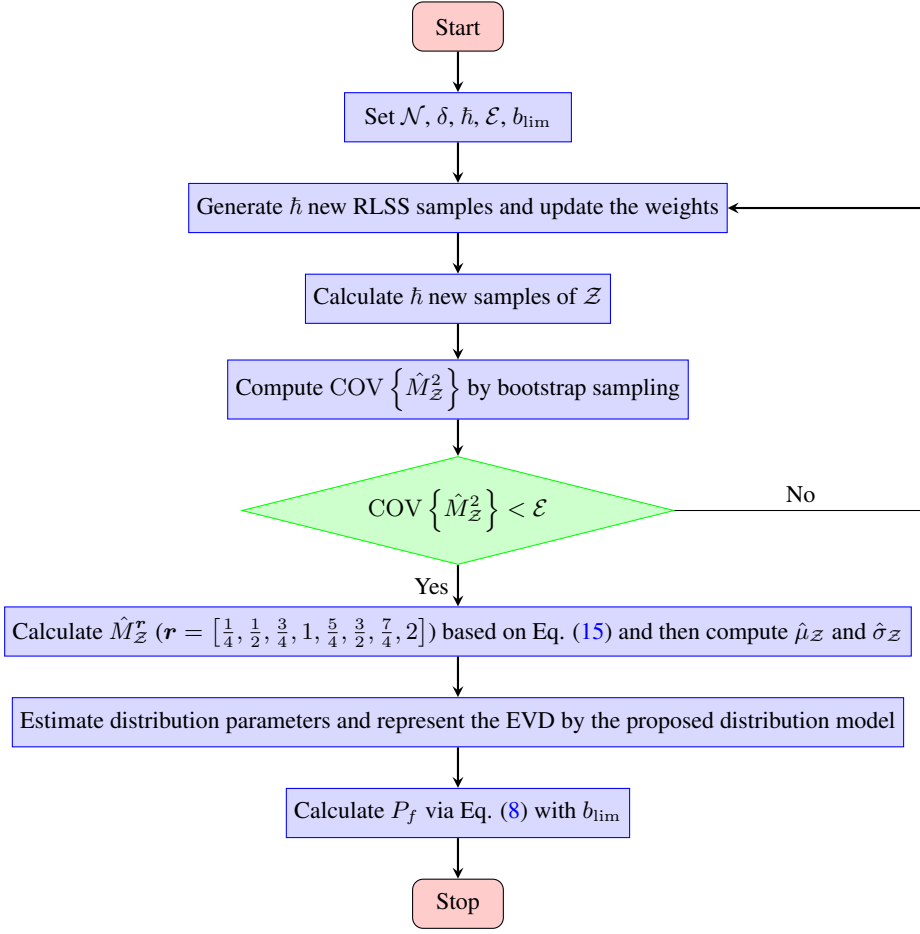


Figure 6: Flowchart of the proposed method

444 Differential equation solver *Ode45* in Matlab is utilized to solve Eq. (30). Both  $\gamma$  and  $\varepsilon$  follow the lognormal  
 445 distributions with mean values as 0.5 and 0.3, and standard deviation values as 0.2 and 0.1, respectively. The Gaussian  
 446 white noise is expressed as

$$\mathcal{G}(t_k) = \theta(t_k) \sqrt{2\pi S/\Delta t}, \quad (31)$$

447 where  $S = 1$  is the spectral intensity;  $\Delta t = 0.01$  s is the time interval;  $T = 30$  s is the time period;  $t_k = k\Delta t, k =$   
 448  $0, 1, \dots, n_t$  is the discrete time; and here we consider  $n_t = T/\Delta t + 1 = 3001$  random variables  $\theta(t_k)$  in the Gaussian  
 449 white noise following the standard normal distributions. Therefore, a total number of  $2 + n_t = 3003$  random variables  
 450 are involved in the present example.

451 The maximum absolute extreme value of displacement over time  $t \in [0, T]$ , i.e.,  $\mathcal{Z} = \max_{t \in [0, T]} \{|Y(t)|\}$ , is of  
 452 interest in this example. First, the proposed parallel adaptive sampling scheme is implemented for fractional moment  
 453 estimation. The proposed scheme performs sample size extension successively until the convergence criterion in  
 454 Eq. (16) is satisfied. In each sample size extension,  $\hbar = 8$  new RLSS samples are firstly generated for deterministic  
 455 dynamic analysis. Then, 8 new samples of  $\mathcal{Z}$  are produced at a time using parallel computing technique with 8 CPU  
 456 processors. After that, the RLSS weights are redistributed so that the weights produced by all performed sample size

457 extensions sum to 1. Subsequently, Eq. (16) is checked to determine whether to perform a new round of sample size  
458 extension. Accordingly, a total of  $\hat{\mathcal{N}} = 520$  samples of  $\mathcal{Z}$  are produced that satisfy the convergence criterion, where the  
459 corresponding  $\hat{M}_{\mathcal{Z}}^r$  ( $\mathbf{r} = [\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}, 2]$ ) can be obtained by Eq. (15). Table 1 compares the first-two central  
460 moments ( $\hat{\mu}_{\mathcal{Z}} = \hat{M}_{\mathcal{Z}}^1$  and  $\hat{\sigma}_{\mathcal{Z}} = \sqrt{\hat{M}_{\mathcal{Z}}^2 - (\hat{M}_{\mathcal{Z}}^1)^2}$ ) with the benchmark results given by MCS with  $10^6$  runs. In this  
461 table, relative errors of the first-two moments between proposed method and MCS are also given, i.e., 0.5656% and  
462 0.7195%, which indicate that proposed parallel adaptive scheme using RLSS enables to obtain accurate low-order  
463 central moments.

Table 1: Comparison of first-two central moments by the proposed method and MCS (Example 1)

Method( $\hat{\mathcal{N}}$ )	$\hat{\mu}_{\mathcal{Z}}$	$\hat{\sigma}_{\mathcal{Z}}$
Proposed(520)	3.6570	0.6623
MCS( $10^6$ )	3.6778	0.6671
R.E.	0.5656%	0.7195%

Note: R.E. = Relative error with reference to MCS.

464 Once the required fractional moments are obtained, eight unknown free parameters involved in the proposed mixture  
465 distribution (i.e., M-EIGD-LESND) can be determined by the fractional moment matching technique. Specifically, the  
466 nonlinear system of equations in Eq. (27) is solved according to Algorithm 1, where initial values of free parameters are  
467 given to speed up the solving process. Afterwards, the EVD could be approximated by the proposed mixture distribution  
468 model, where the PDF, CDF and probability of exceedance (POE) curves are all plotted in Fig. 7. For comparison,  
469 the benchmark results by MCS and the results from LPSS+MIGLD and LPSS+MTGIG are also depicted in Fig. 7.  
470 It can be found that both the PDF and POE curves obtained from the proposed method accord well with the MCS  
471 results. Although there is almost no difference between the CDF curves obtained by proposed method and those by  
472 existing mixture distribution methods, larger deviations appear in the POE curves obtained by the LPSS+MIGLD and  
473 LPSS+MTGIG. Moreover, both of the LPSS+MIGLD and LPSS+MTGIG require 625 LPSS samples to estimate the  
474 fractional moments used for distribution parameter evaluation, where the number of samples is empirically determined  
475 in advance and is larger than that required by the proposed method. In this regard, the proposed method shows a  
476 considerable improvement in both efficiency and accuracy to recover the EVD in this example.

477 After obtaining the reconstructed EVD, the first-passage probability can be evaluated by Eq. (8), where the  
478 safe threshold of this example is set to be  $b_{\text{lim}} = 7$ . Table 2 lists the first-passage probabilities estimated by the  
479 proposed method, LPSS+MIGLD, LPSS+MTGIG, SS and MCS. In this table, the estimated first-passage probabilities  
480 are denoted as  $\hat{P}_f$ . With reference to  $\hat{P}_f$  obtained by the MCS, i.e.,  $1.2200 \times 10^{-4}$ , the first-passage probability  
481 evaluated by the proposed method has acceptable accuracy, which reads  $1.2245 \times 10^{-4}$ . Unfortunately, the first-passage  
482 probabilities by SS, LPSS+MIGLD and LPSS+MTGIG largely deviate from the reference  $\hat{P}_f$  by the MCS.

Table 2: Comparison of first-passage probabilities by different methods (Example 1)

Method	MCS	SS	LPSS+MIGLD	LPSS+MTGIG	Proposed
$\hat{\mathcal{N}}$	$10^6$	4600	625	625	520
$\hat{P}_f$	$1.2200 \times 10^{-4}$	$8.3100 \times 10^{-5}$	$4.7154 \times 10^{-5}$	$4.5286 \times 10^{-5}$	$1.2245 \times 10^{-4}$

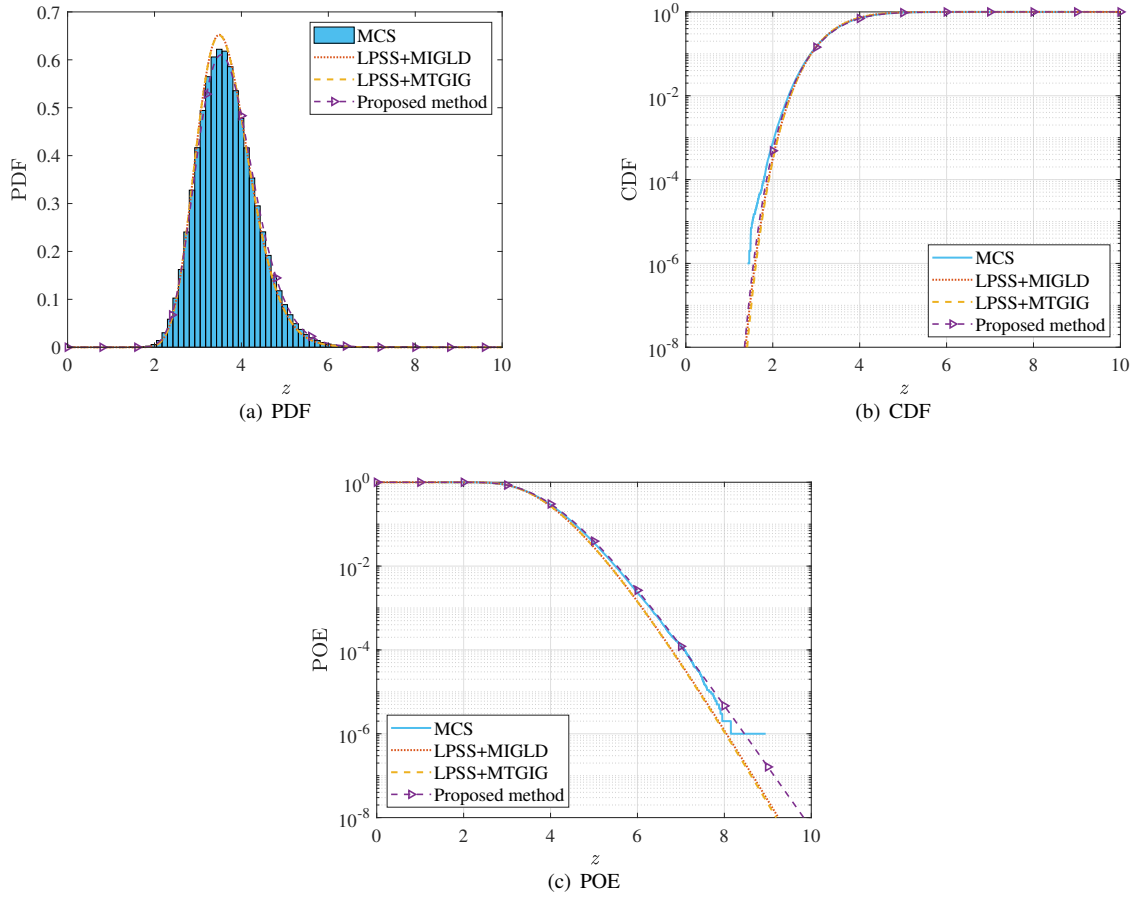


Figure 7: PDF, CDF and POE of  $Z$  in Example 1

483 4.2. Example 2: a 15-storey shear frame structure under fully nonstationary stochastic ground motion

484 A 15-storey nonlinear shear frame structure with uncertain structural properties under fully nonstationary stochastic  
 485 ground motion is investigated in this example, shown in Fig. 8. The equation of motion of this structure reads:

$$\mathbf{M}(\mathbf{U}_{\text{str}}) \ddot{\mathbf{Y}} + \mathbf{C}(\mathbf{U}_{\text{str}}) \dot{\mathbf{Y}} + \mathbf{K}(\mathbf{U}_{\text{str}}) [\tilde{a}\mathbf{Y} + (1 - \tilde{a}) \mathbf{V}] = -\mathbf{M}(\mathbf{U}_{\text{str}}) \mathbf{I} \ddot{\mathbf{x}}_g(\mathbf{U}_{\text{exl}}, t), \quad (32)$$

486 where  $\ddot{\mathbf{Y}}$ ,  $\dot{\mathbf{Y}}$  and  $\mathbf{Y}$  are the lateral acceleration, velocity and displacement matrices of the structure with respect to  
 487 the ground;  $\mathbf{M}$ ,  $\mathbf{C}$  and  $\mathbf{K}$  denote the mass, damping and stiffness matrices, respectively; Term  $\mathbf{I}$  denotes the unit  
 488 matrix. All of the lumped masses and the corresponding stiffnesses from bottom to top of the structure are assumed to  
 489 be independent random variables, following the lognormal distributions with same coefficients of variation 0.1 and  
 490 different mean values  $6 \times 10^4$  kg and  $7 \times 10^7$  N/m, respectively. Hence,  $n_{s_1} = 30$  random variables are involved  
 491 in the system properties, which are denoted as  $\mathbf{U}_{\text{str}}$ . The floor slabs are assumed to be rigid. Rayleigh damping is  
 492 implemented as  $\mathbf{C} = \hat{\alpha}\mathbf{M} + \hat{\beta}\mathbf{K}$ , where  $\hat{\alpha}$  and  $\hat{\beta}$  are obtained by taking both the damping ratios of the first and second  
 493 modes as 0.05. The Bouc-Wen resilience model [47] is adopted to describe the nonlinear behavior of the structure,  
 494 where the hysteretic displacement  $\mathbf{V}$  satisfies:

$$\dot{\mathbf{V}} = \mathcal{A}(\Delta\dot{\mathbf{Y}}) - \mathcal{B}|\Delta\dot{\mathbf{Y}}| |\mathbf{V}|^{\rho-1} \mathbf{V} - \xi(\Delta\dot{\mathbf{Y}}) |\mathbf{V}|^{\rho}, \quad (33)$$

495 in which  $\Delta\dot{\mathbf{Y}}$  is the relative velocity between two neighboring floors,  $\tilde{a} = 0.1$ ,  $\mathcal{A} = 1$ ,  $\mathcal{B} = \xi = 50$  and  $\rho = 1$  are  
 496 the dimensionless parameters controlling the hysteretic performance of Bouc-Wen model. The fully nonstationary  
 497 stochastic ground motion  $\ddot{\mathbf{x}}_g(\mathbf{U}_{\text{exl}}, t)$  is modeled by the second family of spectral representation method (SRM) [48]:

$$\ddot{\mathbf{x}}_g(\mathbf{U}_{\text{exl}}, t) = \sqrt{2} \sum_{j=0}^{n_{s_2}-1} \sqrt{2S_{\ddot{x}_g}(\omega_j, t)} \Delta\omega \cos(\omega_j t + U_{\text{exl},j}), \quad (34)$$

498 where  $\mathbf{U}_{\text{exl}} = [U_{\text{exl},1}, U_{\text{exl},2}, \dots, U_{\text{exl},n_{s_2}}]$  denotes the random vector with  $n_{s_2} = 1600$  independent random variables  
 499 uniformly distributed in  $[0, 2\pi]^{n_{s_2}}$ ;  $\omega_j = j\Delta\omega$ ,  $j = 1, 2, \dots, n_{s_2}$  is the discrete frequency and  $\Delta\omega = \omega_{\text{up}}/n_{s_2}$  denotes  
 500 the frequency interval with upper cut frequency  $\omega_{\text{up}} = 240$  rad/s;  $S_{\ddot{x}_g}(\omega_j, t)$  is the double-sided evolutionary power  
 501 spectrum density (EPSD) function:

$$S_{\ddot{x}_g}(\omega, t) = |\mathcal{A}(\omega, t)|^2 S(\omega), \quad (35)$$

502 in which  $\mathcal{A}(\omega, t)$  is the time-frequency modulation function and  $S(\omega)$  is the power spectrum density represented by  
 503 Clough-Penzien spectrum [49], which are given as

$$\mathcal{A}(\omega, t) = e^{-\chi_0 \frac{\omega t}{\omega_g T}} \cdot \left[ \frac{t}{\mathcal{C}_0} \cdot e^{\left(1 - \frac{t}{\mathcal{C}_0}\right)} \right]^\kappa, \quad (36)$$

$$S(\omega) = \frac{[\omega_g^4 + 4\zeta_g^2 \omega_g^2 \omega^2] \omega^4}{\left[ (\omega_g^2 - \omega^2)^2 + 4\zeta_g^2 \omega_g^2 \omega^2 \right] \left[ (\omega_f^2 - \omega^2)^2 + 4\zeta_f^2 \omega_f^2 \omega^2 \right]} \frac{\bar{a}_{\text{max}}^2}{\gamma_0^2 \left[ \pi \omega_g \left( 2\zeta_g + \frac{1}{2\zeta_g} \right) \right]}, \quad (37)$$

505 where  $\chi_0$  is the frequency modulation factor;  $\mathcal{C}_0$  is the approximate arrive time of peak ground acceleration (PGA);  $\kappa$   
 506 is the shape control coefficient;  $\omega_g$  and  $\zeta_g$  are the parameters describing the dominant frequency and damping ratio of  
 507 site soil;  $\omega_f$  and  $\zeta_f$  are similar parameters for the second filter that hinders the low-frequency component;  $\gamma_0$  is the  
 508 peak factor;  $T$  is the time duration; and  $\bar{a}_{\text{max}}$  denotes the PGA. Values of these involved parameters in EPSD take  
 509  $\chi_0 = 0.15$ ,  $\mathcal{C}_0 = 9$  s,  $\kappa = 2$ ,  $\omega_f = 0.1\omega_g = \frac{4}{7}\pi$ ,  $\zeta_f = \zeta_g = 0.64$ ,  $\gamma_0 = 2.85$ ,  $T = 20$  s,  $\bar{a}_{\text{max}} = 400$  cm/s<sup>2</sup>. Note  
 510 that a total number of  $n_{s_1} + n_{s_2} = 1630$  random variables are involved in this example.

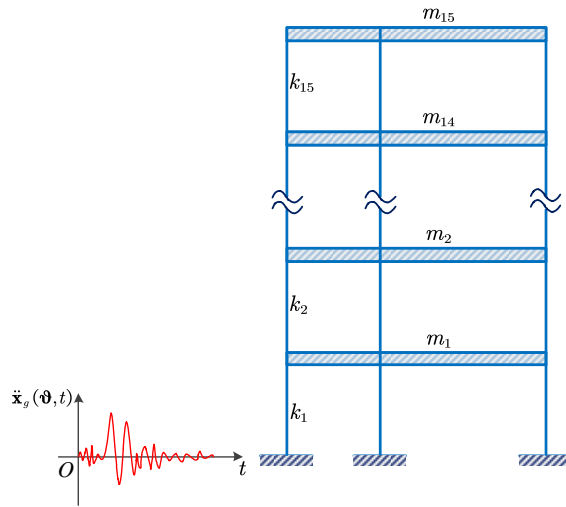


Figure 8: A 15-storey nonlinear shear frame structure

511 The maximum absolute extreme value of inter-storey drift on each storey over the time duration is considered as  
512 the response of interest in this example, which is denoted as  $\mathcal{Z}_i, i = 1, 2, \dots, 15$ . Function solver *Ode45* in Matlab is  
513 employed to perform deterministic dynamic analysis. Here, the second-order fractional moment of the maximum value  
514 of  $\mathcal{Z}$  of all layers, i.e.,  $\hat{M}_{\mathcal{Z}_{\max}}^2, \mathcal{Z}_{\max} = \max_{1 \leq i \leq 15} \{\mathcal{Z}_i\}$ , is considered in the convergence criterion (Eq. (16)). Accordingly,  
515 a total of  $\hat{N} = 520$  samples of  $\mathcal{Z}_i, i = 1, 2, \dots, 15$  are generated, and the required fractional moments are obtained  
516 according to Eq. (15). Besides, the speed up factor between the total computing time by using one CPU processor  
517  $T(1)$  and that by using 8 CPU processors  $T(8)$  is computed, which is  $S_p = T(1)/T(8) = 661 \text{ s}/246 \text{ s} = 2.7$ . This  
518 shows the benefit of using the parallel computing technique in the proposed parallel adaptive scheme.

519 Once the fractional moments are available, the EVDs of  $\mathcal{Z}_i, i = 1, 2, \dots, 15$  are then reconstructed by the proposed  
520 M-EIGD-LESND. Figs. 9-11 depict the PDFs and POEs of  $\mathcal{Z}_1$  on the 1st storey,  $\mathcal{Z}_7$  on the 7th storey and  $\mathcal{Z}_{15}$  on  
521 the 15th storey, respectively. As seen, the proposed mixture distribution model well captures the main parts and tail  
522 information of the EVDs for selected storeys. Specifically, for all the selected storeys, the proposed method gives  
523 almost same accurate results of PDF and POE compared to the reference results from MCS. Besides, to further illustrate  
524 the advantages of the proposed method, a comparison of the PDF and POE curves of  $\mathcal{Z}_1$  is depicted in Fig. 12, where  
525 results by LPSS+MIGLD and LPSS+MTGIG and those by the proposed method are given. As observed, with smaller  
526 sample size, the proposed method is able to capture the tail information more accurately than LPSS+MIGLD and  
527 LPSS+MTGIG, both of which require 625 samples.

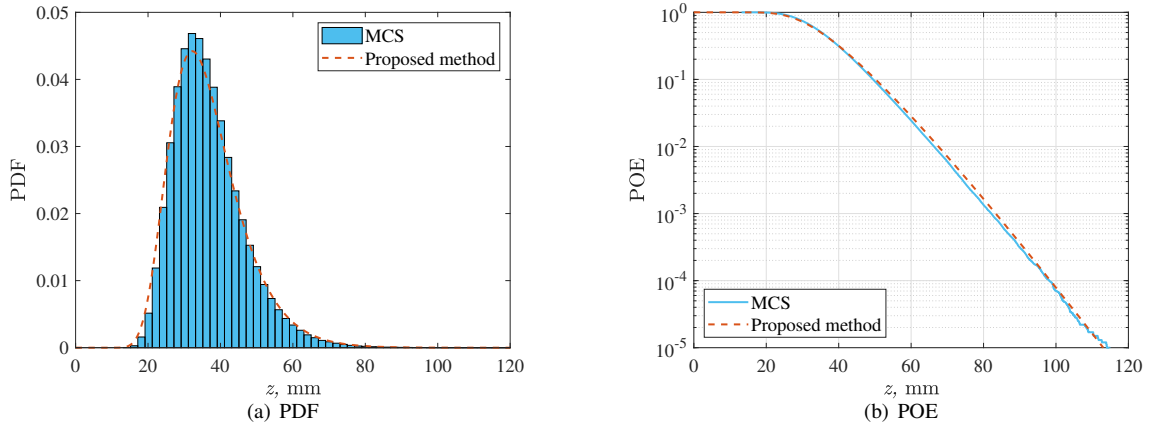


Figure 9: PDF and POE of  $\mathcal{Z}_1$  in Example 2

528 Further, we estimate the first-passage probabilities of the 1st, 7th and 15th storey of this example by Eq. (8), by  
529 setting three different thresholds as  $b_{\text{lim},1\text{st}} = 95 \text{ mm}$ ,  $b_{\text{lim},7\text{th}} = 80 \text{ mm}$  and  $b_{\text{lim},15\text{th}} = 67 \text{ mm}$ . Table 3 gives the  
530 comparison results of proposed method, SS and MCS. As seen, with only 520 samples involved, all three first-passage  
531 probabilities by the proposed method have better accuracy than probabilities by SS.

#### 532 4.3. Example 3: a spatial steel frame structure with viscous dampers under fully nonstationary stochastic ground 533 motion

534 To illustrate the practical applicability of the proposed method, a two-bay four-storey nonlinear spatial steel frame  
535 structure with three viscous dampers under fully nonstationary ground motion is considered in this example, as shown

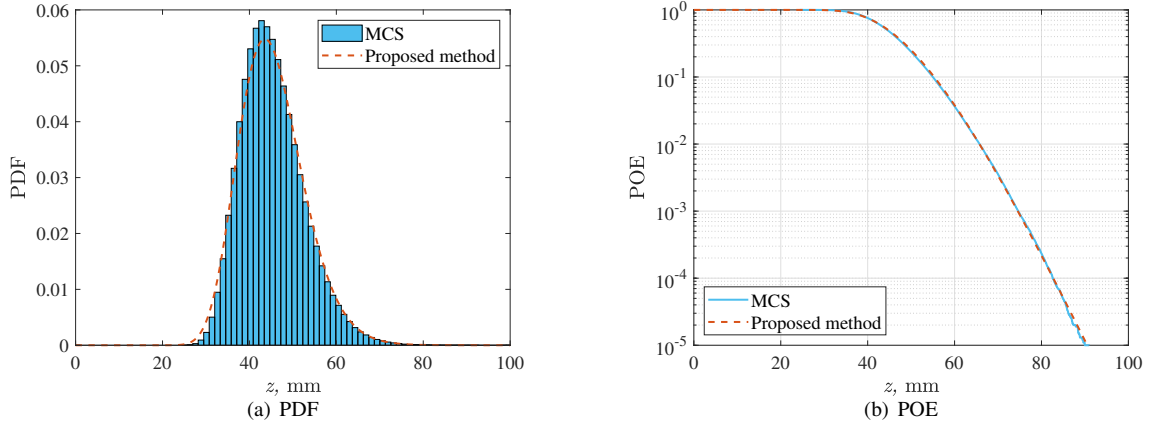


Figure 10: PDF and POE of  $Z_7$  in Example 2

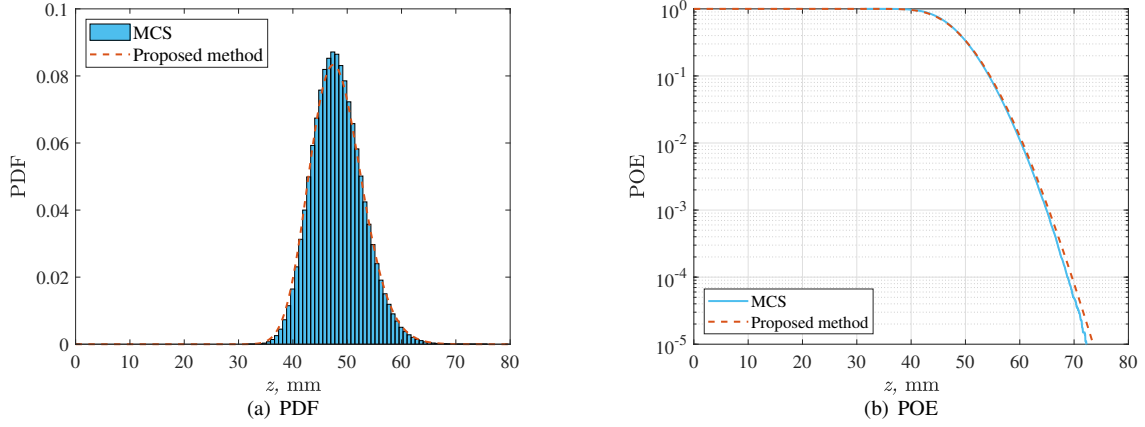


Figure 11: PDF and POE of  $Z_{15}$  in Example 2

Table 3: Comparison of first-passage probabilities by proposed method and MCS (Example 2)

Method( $\hat{N}$ )	1st storey		7th storey		15th storey	
	$b_{lim}(mm)$	$\hat{P}_f$	$b_{lim}(mm)$	$\hat{P}_f$	$b_{lim}(mm)$	$\hat{P}_f$
M-EIGD-LESND(520)	95	$1.5075 \times 10^{-4}$	80	$2.1708 \times 10^{-4}$	67	$4.2208 \times 10^{-4}$
SS(3700)	95	$1.9300 \times 10^{-4}$	80	$4.3600 \times 10^{-4}$	67	$4.5300 \times 10^{-4}$
MCS( $10^6$ )	95	$1.6300 \times 10^{-4}$	80	$2.3300 \times 10^{-4}$	67	$3.0000 \times 10^{-4}$

536 in Fig. 13. The whole structure is modeled and analyzed by the OpenSees software, where the *Steel01* model shown in  
 537 Fig. 14 is used to model the nonlinear stress-strain relationship of steel materials. The slab of each floor is supposed to  
 538 be rigid. The IPE270 beam and IPB300 column are adopted, where the column mass takes its self weight, while the  
 539 beam mass is defined by “self weight of beam + dead loads  $D_L + 0.2 \times$  live loads  $L_L$ ”. The viscous dampers are all  
 540 represented by the Maxwell model which includes a linear spring and nonlinear dashpot in series. Three coefficients

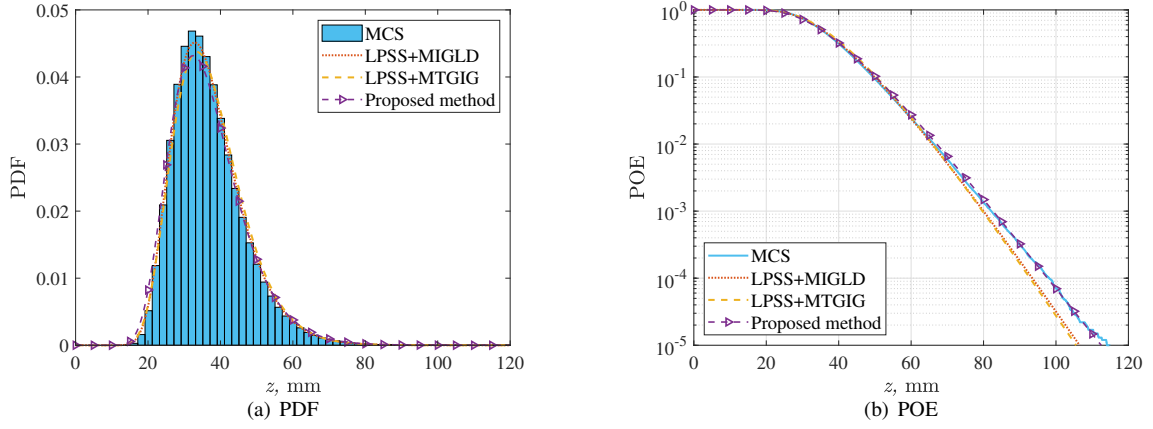


Figure 12: A comparison between the PDF and POE of  $Z_1$  in Example 2

541 are involved in these viscous dampers, i.e., axial elastic stiffness of linear spring  $K_d$ , damping coefficient  $C_d$ , and  
 542 velocity exponent  $\alpha_d$ . The Rayleigh damping is also employed here, where the damping ratios for both the first and  
 543 second modes are taken as 0.03. The fully nonstationary stochastic ground motion takes the same form and parameters  
 544 as employed in Example 2. It should be mentioned that the randomness of this structure comes from its external loads  
 545 (i.e., dead loads, live loads and ground motion) and its structural properties. The statistical information of uncertain  
 546 structural properties is collected in Table 4. In total, 1608 random variables are involved in this example.

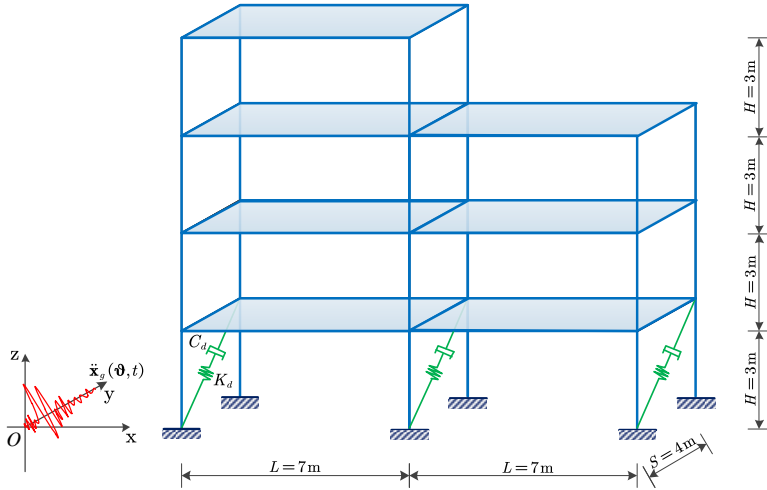


Figure 13: A two-bay nonlinear spatial steel frame structure with viscous dampers

547 We consider the maximum absolute inter-storey drift of the whole structure as the quantity of interest, denoted  
 548 by  $Z$ . By adopting the proposed parallel adaptive scheme,  $\hat{N} = 1032$  samples of  $Z$  are generated, where a set of up  
 549 to second order fractional moments can be estimated by Eq. (15). From the knowledge of the estimated fractional  
 550 moments, the EVD is represented by the proposed mixture distribution model, where the corresponding PDF and POE  
 551 curves are depicted in Fig. 15. For comparison, the results by LPSS+MIGLD and LPSS+MTGIG are also provided,

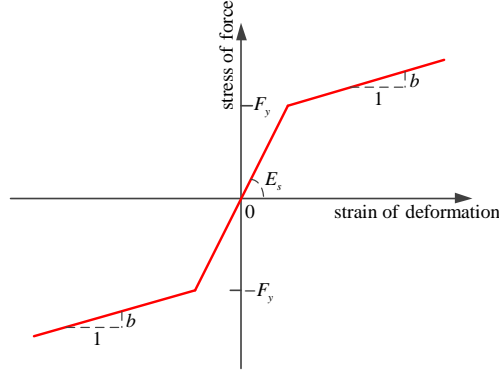


Figure 14: Constitutive law of material: Steel01

Table 4: Statistical information of the uncertain structural properties in Example 3

Parameter	Description	Distribution	Mean	Standard variation
$D_L$	Dead load	Lognormal	10 N/m <sup>2</sup>	0.5 N/m <sup>2</sup>
$L_L$	Live load	Lognormal	10 N/m <sup>2</sup>	1 N/m <sup>2</sup>
$F_y$	Yield strength of the steel	Normal	$250 \times 10^6$ Pa	$375 \times 10^5$ Pa
$E_s$	Young's modulus of the steel	Normal	$2 \times 10^{11}$ Pa	$3 \times 10^{10}$ Pa
$b$	Strain-hardening ratio	Normal	$10^{-3}$	$5 \times 10^{-5}$
$K_d$	Axial stiffness of linear spring	Normal	25 Pa	2.5 Pa
$C_d$	Damping coefficient	Normal	20.7452	2.07452
$\alpha_d$	Velocity exponent	Normal	0.35	0.0175

552 together with the benchmark results from MCS. Good accordance between results by proposed method and MCS  
553 is readily observed. Admittedly, LPSS+MIGLD and LPSS+MTGIG are more computationally efficient since only  
554 625 LPSS samples are employed. However, the tail distributions captured by the LPSS+MIGLD and LPSS+MTGIG  
555 unfortunately deviate from the benchmark results to a large extent. Moreover, we calculate the first-passage probability  
556 of this example by setting the threshold of  $\mathcal{Z}$  as 38 mm. The first-passage probabilities by the MCS, SS, LPSS+MIGLD,  
557 LPSS+MTGIG and proposed method are listed in Table 5. Remarkably, the proposed method yields a probability  
558 that is quite close to what MCS gives, i.e.,  $2.2439 \times 10^{-4}$  by the proposed method, and  $2.3600 \times 10^{-4}$  by MCS.  
559 The probability by LPSS+MIGLD and LPSS+MTGIG notably deviate from the probability by the MCS, reading  
560  $5.0859 \times 10^{-5}$  and  $5.0677 \times 10^{-5}$ , respectively. In addition, the first-passage probability by SS is also less accurate,  
561 reading  $2.0400 \times 10^{-4}$ , but requires much more model evaluations.

Table 5: Comparison of first-passage probabilities by different methods (Example 3)

Method	MCS	SS	LPSS+MIGLD	LPSS+MTGIG	Proposed
$\tilde{\mathcal{N}}$	$10^6$	3700	625	625	1032
$\hat{P}_f$	$2.3600 \times 10^{-4}$	$2.0400 \times 10^{-4}$	$5.0859 \times 10^{-5}$	$5.0677 \times 10^{-5}$	$2.2439 \times 10^{-4}$



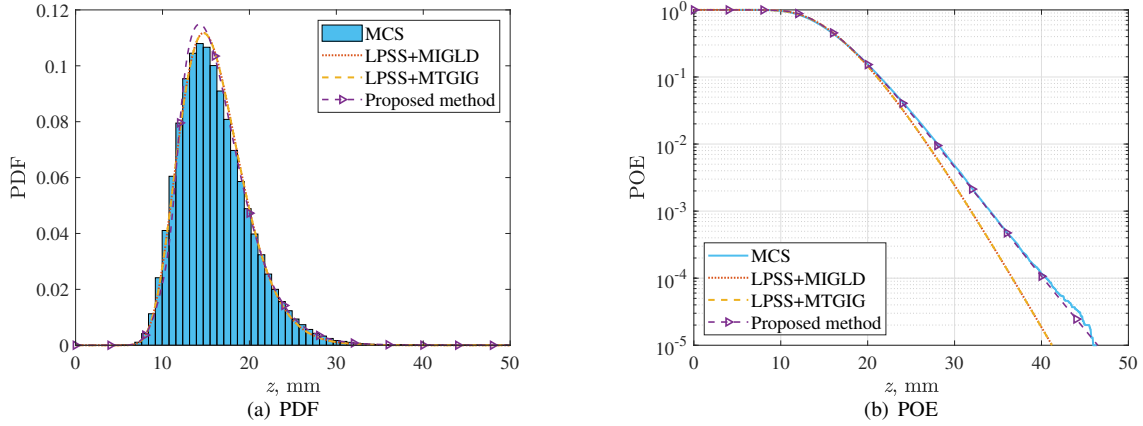


Figure 15: PDF and POE of  $Z$  in Example 3

## 5. Concluding remarks

This paper proposes a novel fractional moments-based mixture distribution method to estimate the EVD and the first-passage probabilities of high-dimensional nonlinear stochastic dynamic systems. Unlike the existing methods, a parallel adaptive sampling scheme that allows for sample size extension is first proposed for estimating fractional moments. By doing so, the sample size can be determined problem-dependently in conjunction with a proposed convergence criterion. Such scheme is realized by a sequential sampling method, i.e., refined latinized stratified sampling (RLSS), which also enables to achieve variance reduction in high dimensions. One versatile mixture distribution model, namely, M-EIGD-LESND, is proposed to represent the EVD with enhanced flexibility, whose free parameters are evaluated from obtained fractional moments. Three examples involving high-dimensional and strong-nonlinear stochastic dynamic systems are investigated to demonstrate the efficacy of the proposed method. The main conclusions are summarized as follows:

(1) The studied examples indicate that the proposed method is able to tackle with high-dimensional and strongly nonlinear stochastic dynamic systems, where the uncertainties in both internal structural properties and external excitations are considered. In addition, the proposed method is capable of accurately estimating small first-passage probabilities in the order of  $10^{-4}$ .

(2) Several byproducts can be obtained by adopting the proposed method, i.e., fractional moments (including integer moments such as mean and standard deviation) and EVD. Furthermore, for a general stochastic dynamic system, multiple EVDs and first-passage probabilities under different thresholds can be estimated from only a single run of the proposed method.

(3) The proposed method is computational efficient since the proposed parallel adaptive scheme allows to determine an optimal sample size for a particular problem at hand. In addition, only additional samples of extreme value need to be evaluated in each sample size extension, where parallel computing technique can be adopted to further improve the efficiency.

(4) The proposed eight-parameter mixture distribution model is highly flexible and can adapt to different levels of distribution asymmetry. This model generalizes several single-component distributions, such as the lognormal, skew-normal, log skew-normal, and inverse Gaussian distribution. In addition, the mixture of lognormal and inverse

588 Gaussian distributions is a special case of the proposed model. As a result, this model enables the proposed method to  
 589 accurately recover a wide variety of EVDs.

#### 590 **CRedit authorship contribution statement**

591 **Chen Ding:** Methodology, Software, Validation, Investigation, Writing - Original Draft, Writing - Revised draft;  
 592 **Chao Dang:** Conceptualization, Methodology, Investigation, Visualization, Writing - Original Draft, Writing - Revised  
 593 draft, Funding acquisition; **Marcos Valdebenito:** Validation, Writing- Reviewing and Editing, Funding acquisition;  
 594 **Matthias Faes:** Validation, Writing- Reviewing and Editing; **Matteo Broggi:** Validation, Supervision, Writing-  
 595 Reviewing and Editing; **Michael Beer:** Supervision, Project administration.

#### 596 **Declaration of Competing Interest**

597 The authors declare that they have no known competing financial interests or personal relationships that could have  
 598 appeared to influence the work reported in this paper.

#### 599 **Acknowledgement**

600 Chen Ding is grateful for the support by the European Union’s Horizon 2020 research and innovation programme  
 601 under Marie Skłodowska-Curie project GREYDIENT – Grant Agreement n°955393. Chao Dang is mainly supported  
 602 by the China Scholarship Council (CSC). Marcos Valdebenito acknowledges the support of ANID (National Agency  
 603 for Research and Development, Chile) under its program FONDECYT, grant number 1180271.

#### 604 **Appendix A. Refined Latinized stratified sampling**

605 To generate samples and weights by the Refined Latinized stratified sampling (RLSS) [37], first we need to generate  
 606 candidate samples and candidate strata by the combination of hierarchical Latin hypercube sampling (HLHS) [37] and  
 607 Latinized stratified sampling (LSS) [39].

608 Begin with a LSS design with  $\mathcal{N}$  samples in  $n_s$  dimensions. First, generate a  $n_s$ -dimensional Latin hypercube  
 609 sampling (LHS) design with  $\mathcal{N}$  one-dimensional LHS strata  $\Omega_{ij}$  and samples in each stratum  $\varphi_{ij}$ ,  $i = 1, \dots, n_s$ ;  $j =$   
 610  $1, \dots, \mathcal{N}$ . Denote  $\mathcal{S}$  as the  $[0, 1]^{n_s}$  space. Divide  $\mathcal{S}$  equally into  $\mathcal{N}$  mutually exclusive and collectively exhaustive strata  
 611  $\Omega^{(k)}$ ,  $k = 1, \dots, \mathcal{N}$ , where  $\Omega^{(k)} \cap \Omega^{(q)} = \emptyset$ ,  $k \neq q$  and  $\bigcup_{k=1}^{\mathcal{N}} \Omega^{(k)} = \mathcal{S}$ . Note that each  $\Omega^{(k)}$  is an equal-weighted  
 612 hyper-rectangle and its boundary coincides with the boundary of  $\Omega_{ij}$ . Each  $\Omega^{(k)}$  can be described by its starting  
 613 coordinate near the origin  $\mathbf{A}^{(k)} = \{A_1^{(k)}, \dots, A_{n_s}^{(k)}\}$  and its side length  $\lambda^{(k)} = \{\lambda_1^{(k)}, \dots, \lambda_{n_s}^{(k)}\}$ . The weight of each  
 614  $\Omega^{(k)}$  can be calculated as [37]:

$$\varpi^{(k)} = \prod_{i=1}^{n_s} \lambda_i^{(k)}, \quad (\text{A.1})$$

615 where  $\sum_{k=1}^{\mathcal{N}} \varpi^{(k)} = 1$ . For each  $\Omega^{(k)}$ , randomly pair each  $\varphi_{ij}$  without replacement to produce the  $k$ -th LSS sample  
 616  $\varphi^{(k)} = [\varphi_1^{(k)}, \dots, \varphi_{n_s}^{(k)}]$ ,  $k = 1, \dots, \mathcal{N}$ .

617 Afterwards, apply a  $\delta$ -level refinement of each  $\Omega_{ij}$  based on the idea of HLHS, where  $\delta \in \mathbb{Z}^+$  is the refinement  
 618 factor. Specifically, along each dimension, divide  $\Omega_{ij}$   $\delta$  times equally to obtain a total of  $\tilde{\mathcal{N}} = \mathcal{N}(\delta + 1)$  strata

619  $\Omega_{ijh}, h = 1, \dots, \tilde{\mathcal{N}}$ . Produce new candidate samples per each dimension by uniform sampling inside every empty newly  
620 produced stratum  $\Omega_{ijh}$ . Subsequently, generate the candidate strata of RLSS, denoted as  $\tilde{\Omega}^{(k^*)}, k^* = 1, \dots, \tilde{\mathcal{N}}$ , by  
621 dividing all the  $\Omega^{(k)}$   $\delta$  times along the LHS stratum boundaries in the dimension of largest side length  $\lambda^* = \max_i \{\lambda_i^{(k)}\}$ .  
622 Then, identify the candidate stratum  $\Xi_i^{(k^*)} = \left\{ \Omega_{ij} \in \left[ A_i^{(k^*)}, A_i^{(k^*)} + \lambda_i^{(k^*)} \right] \right\}$  which intersects with  $\tilde{\Omega}^{(k^*)}$  in each  
623  $i$ -th dimension. Count the number of  $\Xi_i^{(k^*)}$  as  $\varepsilon_i^{(k^*)}, i = 1, \dots, n_s$ , and then determine the minimum number of  $\Xi_{i,(k^*)}$   
624 as  $\varepsilon_i^* = \min_{k^*} \left\{ \varepsilon_i^{(k^*)} \right\}$ . The candidate samples of RLSS, denoted as  $\tilde{\varphi}^{(k^*)}, k^* = 1, \dots, \tilde{\mathcal{N}}$ , are generated by drawing  
625 samples to the stratum  $\Omega^{(k^*)}$  satisfying  $\varepsilon_i^{(k^*)} = \varepsilon_i^*$ : if  $\varepsilon_i^* = 1$ ,  $\Omega^{(k^*)}$  contains only one single candidate LHS stratum,  
626 one must draw a sample from it; if  $\varepsilon_i^* > 1$ , one can draw samples from  $\Xi_i^{(k^*)}$  at random without replacement. Repeat  
627 the sample adding process until all the dimensions of  $\Omega^{(k^*)}$  have one related sample.

628 Once the candidate samples  $\tilde{\varphi}^{(k^*)}$  and strata  $\tilde{\Omega}^{(k^*)}$  of RLSS are obtained, we can generate  $\hbar$  RLSS samples at  
629 a time. First, randomly select  $\hbar$  RLSS strata  $\hat{\Omega}^{(l)}, l = 1, \dots, \hbar$  from the candidate strata  $\tilde{\Omega}^{(k^*)}$ . Then form RLSS  
630 samples  $\hat{\varphi}^{(l)}, l = 1, \dots, \hbar$  by drawing corresponding samples from  $\tilde{\varphi}^{(k^*)}$  to  $\hat{\Omega}^{(l)}$ . Update the stratum weight according  
631 to Eq. (A.1) by specifying the side length of  $\hat{\Omega}^{(l)}$ . Repeat several times to add  $\hbar$  RLSS samples continuously until a  
632 user-defined convergence criterion is met or the number of remaining candidate samples  $\tilde{\varphi}^{(k^*)}$  of RLSS is less than  $\hbar$ .  
633 Note that if the number of candidate samples is insufficient, a new extension of the sample candidate pool is required.  
634 If  $\varsigma > 1$  extensions of the candidate sample pool can finally produce enough samples and weights of RLSS that meet  
635 the convergence criterion, then the total number of  $\tilde{\varphi}^{(k^*)}$  and  $\tilde{\Omega}^{(k^*)}$  at this time will be  $\tilde{\mathcal{N}} = \mathcal{N}(\delta + 1)^\varsigma$ . Briefly, the  
636 procedure of RLSS scheme is summarized in Algorithm 2, where  $\hat{\mathcal{N}}$  denotes the obtained optimal sample size.

## 637 References

- 638 [1] H. Risken, The Fokker-Planck Equation: Methods of Solution and Applications, 2nd Edition, Springer Series in Synergetics 18, Springer-Verlag  
639 Berlin Heidelberg, 1989.
- 640 [2] M.-Z. Lyu, J.-B. Chen, First-passage reliability of high-dimensional nonlinear systems under additive excitation by the ensemble-evolving-based  
641 generalized density evolution equation, Probabilistic Engineering Mechanics 63 (2021) 103119.
- 642 [3] S. O. Rice, Mathematical analysis of random noise, The Bell System Technical Journal 24 (1) (1945) 46–156.
- 643 [4] A. Naess, O. Gaidai, Monte carlo methods for estimating the extreme response of dynamical systems, Journal of Engineering Mechanics  
644 134 (8) (2008) 628–636.
- 645 [5] J. He, An approximation of the first passage probability of systems under nonstationary random excitation, Applied Mathematics and Mechanics  
646 30 (2) (2009) 255–262.
- 647 [6] L. Lutes, Amplitude correlation in first-passage problems, Journal of Engineering Mechanics 138 (9) (2012) 1205–1213.
- 648 [7] J. Li, J. Chen, Stochastic Dynamics of Structures, John Wiley & Sons, 2009.
- 649 [8] H. Vanvinckenroye, I. Kougoumtzoglou, V. Denoël, Reliability function determination of nonlinear oscillators under evolutionary stochastic  
650 excitation via a galerkin projection technique, Nonlinear Dynamics 95 (1) (2019) 293–308.
- 651 [9] A. Kovaleva, An exact solution of the first-exit time problem for a class of structural systems, Probabilistic Engineering Mechanics 24 (3)  
652 (2009) 463–466.
- 653 [10] D. Iourtchenko, E. Mo, A. Naess, Reliability of strongly nonlinear single degree of freedom dynamic systems by the path integration method,  
654 Journal of Applied Mechanics 75 (6) (2008) 061016.
- 655 [11] I. A. Kougoumtzoglou, P. D. Spanos, Response and first-passage statistics of nonlinear oscillators via a numerical path integral approach,  
656 Journal of Engineering Mechanics 139 (9) (2013) 1207–1217.
- 657 [12] C. Bucher, A. Di Matteo, M. Di Paola, A. Pirrotta, First-passage problem for nonlinear systems under lévy white noise through path integral  
658 method, Nonlinear Dynamics 85 (3) (2016) 1445–1456.
- 659 [13] P. D. Spanos, I. A. Kougoumtzoglou, Galerkin scheme based determination of first-passage probability of nonlinear system response, Structure  
660 and Infrastructure Engineering 10 (10) (2014) 1285–1294.

---

**Algorithm 2** Refined Latinized stratified sampling approach [37]

---

**Input:** Dimension  $n_s$  of the random parameter vector  $\mathbf{U}$ , LSS size  $\mathcal{N}$ , refinement factor  $\delta$  and number of samples for each new sample size extension  $\hat{h}$ .

**Output:** RLSS samples  $\hat{\varphi} = \{\hat{\varphi}^{(1)}, \dots, \hat{\varphi}^{(\mathcal{N})}\}$  and corresponding weights  $\varpi = \{\varpi^{(1)}, \dots, \varpi^{(\mathcal{N})}\}$ .

- 1: Initialize with  $\varsigma = 1$ . Define a LHS design with  $\mathcal{N}$  ungrouped LHS sample components  $\varphi_{ij}$  and corresponding one dimension LHS strata  $\Omega_{ij}, i = 1, \dots, n_s; j = 1, \dots, \mathcal{N}$ .
  - 2: Establish a  $n_s$ -dimensional stratification  $\Omega^{(k)}, k = 1, \dots, \mathcal{N}$  to form LSS strata such that each stratum is an equal-weighted hyper-rectangle and its boundary coincides with the boundary of  $\Omega_{ij}$ . Calculate the stratum weight of  $\Omega^{(k)}$  according to Eq. (A.1).
  - 3: Generate LSS samples  $\varphi^{(k)} = [\varphi_1^{(k)}, \dots, \varphi_{n_s}^{(k)}], k = 1, \dots, \mathcal{N}$  by randomly drawing  $\varphi_{ij}$  to its related LSS stratum without replacement.
  - 4: Produce candidate samples per each dimension by applying a  $\delta$ -level refinement of each  $\varphi_{ij}$  inherent in  $\Omega^{(k)}$  according to HLHS design.
  - 5: Generate candidate strata of RLSS  $\tilde{\Omega}^{(k^*)}, k^* = 1, \dots, \mathcal{N}(\delta + 1)^\varsigma$  by dividing all the strata  $\Omega^{(k)}$  equally  $\delta$  times along every dimension with largest side length  $\lambda_i^*$ .
  - 6: Identify the strata  $\Xi_i^{(k^*)} = \left\{ \Omega_{ij} \in [A_i^{(k^*)}, A_i^{(k^*)} + \lambda_i^{(k^*)}] \right\}, k^* = 1, \dots, \mathcal{N}(\delta + 1)^\varsigma$  which intersect with  $\tilde{\Omega}^{(k^*)}$  in each  $i$ -th dimension. Count the number of  $\Xi_i^{(k^*)}$  in the  $i$ -th dimension as  $\varepsilon_i^{(k^*)}$ , and then calculate  $\epsilon_i^* = \min_{k^*} \left\{ \varepsilon_i^{(k^*)} \right\}$ .
  - 7: Generate candidate samples of RLSS  $\tilde{\varphi}^{(k^*)}, k^* = 1, \dots, \mathcal{N}(\delta + 1)^\varsigma$  inside the stratum  $\tilde{\Omega}^{(k^*)}$  satisfying  $\varepsilon_i^{(k^*)} = \epsilon_i^*$ : if  $\epsilon_i^* = 1$ , draw samples from  $\Omega_{ij}$ ; if  $\epsilon_i^* > 1$ , draw samples from  $\Xi_i^{(k^*)}$  at random; repeat sample selection until all the dimensions are filled.
  - 8: Select  $\hat{h}$  RLSS strata  $\hat{\Omega}^{(k)}, k = 1, \dots, \hat{h}$  randomly from candidate  $\tilde{\Omega}^{(k^*)}$  and generate  $\hat{h}$  RLSS samples  $\hat{\varphi}^{(k)}, k = 1, \dots, \hat{h}$  by drawing corresponding samples from candidate  $\tilde{\varphi}^{(k^*)}$  to  $\hat{\Omega}^{(k)}$ . Calculate the stratum weight according to Eq. (A.1) by specifying the side length of  $\hat{\Omega}^{(k)}$ .
  - 9: Repeat step 8 to add samples continuously until Eq. (16) is satisfied or an enlargement of the pool of candidate samples  $\tilde{\varphi}^{(k^*)}$  is required. Then return to step 4 with  $\varsigma = \varsigma + 1$  and  $\Omega^{(k)} = \tilde{\Omega}^{(k^*)}$ .
- 

- 661 [14] K. R. dos Santos, I. A. Kougioumtzoglou, P. D. Spanos, Hilbert transform-based stochastic averaging technique for determining the survival  
662 probability of nonlinear oscillators, *Journal of Engineering Mechanics* 145 (10) (2019) 04019079.
- 663 [15] J. Li, J. Chen, W. Sun, Y. Peng, Advances of the probability density evolution method for nonlinear stochastic systems, *Probabilistic Engineering*  
664 *Mechanics* 28 (2012) 132–142.
- 665 [16] M. Shinozuka, Monte carlo solution of structural dynamics, *Computers & Structures* 2 (5-6) (1972) 855–874.
- 666 [17] S.-K. Au, J. Beck, Important sampling in high dimensions, *Structural Safety* 25 (2) (2003) 139–163.
- 667 [18] Z. Wang, J. Song, Cross-entropy-based adaptive importance sampling using von mises-fisher mixture for high dimensional reliability analysis,  
668 *Structural Safety* 59 (2016) 42–52.
- 669 [19] O. Kanjilal, I. Papaioannou, D. Straub, Cross entropy-based importance sampling for first-passage probability estimation of randomly excited  
670 linear structures with parameter uncertainty, *Structural Safety* 91 (2021) 102090.
- 671 [20] Z. Zhao, Z.-H. Lu, C.-Q. Li, Y.-G. Zhao, Efficient simulation method for first passage problem of linear systems subjected to non-gaussian  
672 excitations, *Journal of Engineering Mechanics* 148 (1) (2022) 04021128.
- 673 [21] S.-K. Au, J. L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, *Probabilistic Engineering Mechanics*  
674 16 (4) (2001) 263–277.
- 675 [22] J. Ching, J. L. Beck, S. Au, Hybrid subset simulation method for reliability estimation of dynamical systems subject to stochastic excitation,  
676 *Probabilistic Engineering Mechanics* 20 (3) (2005) 199–214.

- 677 [23] S. K. Au, J. Ching, J. Beck, Application of subset simulation methods to reliability benchmark problems, *Structural Safety* 29 (3) (2007)  
678 183–193.
- 679 [24] J.-B. Chen, J. Li, The extreme value distribution and dynamic reliability analysis of nonlinear structures with uncertain parameters, *Structural*  
680 *Safety* 29 (2) (2007) 77–93.
- 681 [25] G. Chen, D. Yang, A unified analysis framework of static and dynamic structural reliabilities based on direct probability integral method,  
682 *Mechanical Systems and Signal Processing* 158 (2021) 107783.
- 683 [26] J. He, J. Gong, Estimate of small first passage probabilities of nonlinear random vibration systems by using tail approximation of extreme  
684 distributions, *Structural Safety* 60 (2016) 28–36.
- 685 [27] Z. Zhao, Y.-G. Zhao, P.-P. Li, Efficient approach for dynamic reliability analysis based on uniform design method and box-cox transformation,  
686 *Mechanical Systems and Signal Processing* 172 (2022) 108967.
- 687 [28] G. Li, W. He, Y. Zeng, An improved maximum entropy method via fractional moments with laplace transform for reliability analysis, *Structural*  
688 *and Multidisciplinary Optimization* 59 (4) (2019) 1301–1320.
- 689 [29] J. Xu, W. Zhang, R. Sun, Efficient reliability assessment of structural dynamic systems with unequal weighted quasi-monte carlo simulation,  
690 *Computers & Structures* 175 (2016) 37–51.
- 691 [30] J. Xu, D. Wang, A two-step methodology to apply low-discrepancy sequences in reliability assessment of structural dynamic systems, *Structural*  
692 *and Multidisciplinary Optimization* 57 (4) (2018) 1643–1662.
- 693 [31] J. Xu, C. Dang, A novel fractional moments-based maximum entropy method for high-dimensional reliability analysis, *Applied Mathematical*  
694 *Modelling* 75 (2019) 749–768.
- 695 [32] Z.-Q. Chen, S.-X. Zheng, J. Zhang, C. Dang, K. Wei, X. Li, Efficient seismic reliability analysis of non-linear structures under non-stationary  
696 ground motions, *Soil Dynamics and Earthquake Engineering* 139 (2020) 106385.
- 697 [33] C. Dang, J. Xu, A mixture distribution with fractional moments for efficient seismic reliability analysis of nonlinear structures, *Engineering*  
698 *Structures* 208 (2020) 109912.
- 699 [34] C. Dang, P. Wei, M. Beer, An approach to evaluation of evd and small failure probabilities of uncertain nonlinear structures under stochastic  
700 seismic excitations, *Mechanical Systems and Signal Processing* 152 (2021) 107468.
- 701 [35] Z.-Q. Chen, S.-X. Zheng, Z.-H. Ding, J. Zhang, Y.-J. Tai, Seismic reliability evaluation of bridges under spatially varying ground motions  
702 using a four-parameter distribution, *Engineering Structures* 247 (2021) 113157.
- 703 [36] L.-W. Zhang, Z.-H. Lu, Y.-G. Zhao, Dynamic reliability assessment of nonlinear structures using extreme value distribution based on l-moments,  
704 *Mechanical Systems and Signal Processing* 159 (2021) 107832.
- 705 [37] M. D. Shields, Refined latinized stratified sampling: A robust sequential sample size extension methodology for high-dimensional latin  
706 hypercube and stratified designs, *International Journal for Uncertainty Quantification* 6 (1) (2016) 79–97.
- 707 [38] J. Li, J.-b. Chen, W.-l. Fan, The equivalent extreme-value event and evaluation of the structural system reliability, *Structural Safety* 29 (2)  
708 (2007) 112–131.
- 709 [39] M. D. Shields, J. Zhang, The generalization of latin hypercube sampling, *Reliability Engineering & System Safety* 148 (2016) 96–108.
- 710 [40] B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap*, CRC press, 1994.
- 711 [41] M. D. Shields, K. Teferra, A. Hapij, R. P. Daddazio, Refined stratified sampling for efficient monte carlo based uncertainty quantification,  
712 *Reliability Engineering & System Safety* 142 (2015) 310–325.
- 713 [42] J. L. Folks, R. S. Chhikara, The inverse gaussian distribution and its statistical application—a review, *Journal of the Royal Statistical Society:*  
714 *Series B (Methodological)* 40 (3) (1978) 263–275.
- 715 [43] R. Chhikara, *The Inverse Gaussian Distribution: Theory: Methodology, and Applications*, Vol. 95, CRC Press, 1988.
- 716 [44] A. Azzalini, *The Skew-normal and Related Families*, Vol. 3, Cambridge University Press, 2013.
- 717 [45] A. Canale, Statistical aspects of the scalar extended skew-normal distribution, *Metron* 69 (3) (2011) 279–295.
- 718 [46] K. Wang, S. Yu, W. Peng, A novel moment method using the log skew normal distribution for particle coagulation, *Journal of Aerosol Science*  
719 134 (2019) 95–108.
- 720 [47] Y.-K. Wen, Method for random vibration of hysteretic systems, *Journal of the Engineering Mechanics Division* 102 (2) (1976) 249–263.
- 721 [48] Z. Liu, W. Liu, Y. Peng, Random function based spectral representation of stationary and non-stationary stochastic processes, *Probabilistic*  
722 *Engineering Mechanics* 45 (2016) 115–126.
- 723 [49] R. Clough, J. Penzien, *Dynamics of Structures*, New York:McGraw-Hill, 1975.