https://doi.org/10.1016/j.ymssp.2021.108762

Efficient subset simulation for rare-event integrating point-evolution kernel density and adaptive polynomial chaos kriging

Hongyuan Guo[1], You Dong[1], and Paolo Gardoni[2]

1. Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, China

2. Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Abstract: Rare-event probability estimation has a wide range of applications, including the design and manufacture of precision equipment, aerospace systems, and critical industrial and civil structures. However, traditional simulation-based reliability calculation methods, such as brute Monte Carlo simulation (MCS) and subset simulation (SS), face challenges in efficiently evaluating small-failure probabilities due to the need for a large number of simulations, especially for non-linear and complex scenarios. Thus, to efficiently assess the probability of rare failure events in structural engineering, this paper develops a novel method for assessing the small-failure probability by integrating the point-evolution kernel density (PKDE), SS, and polynomial chaos kriging (PCK). The proposed **P**KDE-**A**daptive **P**CK-based **S**S (PAPS) method aims to reduce the implementation of the original performance function by PCK and enrich the training set using an adaptive strategy. Moreover, the initial cumulative density function (CDF) of the performance function estimated by PKDE is modified gradually to facilitate the estimation of small-failure probability. Four numerical examples of small-failure probability estimation involving classical analytical cases, time-variant cases, and non-linear stochastic structures are used to illustrate the accuracy and efficiency of the proposed method. The computational results show that the proposed method can provide accurate computational results with a smaller computational burden than traditional methods (e.g., MCS, SS, LHS-PCK-SS).

Keywords: Point-evolution kernel density estimation; surrogate model; reliability analysis; rare-event probability; subset simulation; adaptive Monte Carlo Simulation.

[11] **Corresponding Author**
E-mail address: you.dong@polyu.edu.hk
Tel.: +852-3400 8818

1. Introduction

The uncertainty and stochasticity of the environment and structural performance bring difficulties and challenges to structural design and performance evaluation. To overcome the drawbacks of traditional deterministic methods in structural design and assessment, probabilistic methods have gained the attention of many scholars and engineers, where reliability-informed methods have become one of the most widely used methods in modern structural engineering [1–3]. In the reliability-informed method, the main task is to quantify and evaluate the failure probability of structures, i.e., to calculate the occurrence probability of a member or system exceeding a limit state during its service life [4]. It is of great importance to quantify failure risks accurately and design conservatively for significant industrial and civil structures such as sea-crossing bridges, large dams, and nuclear power stations. However, their failure probabilities are usually extremely small and challenging to be evaluated accurately, especially for highly complex and non-linear systems. Therefore, it is crucial to develop an efficient and highly accurate method to assess the failure probability of engineering systems.

To date, many reliability analysis methods have been invented and applied, such as the first-order reliability method (FORM) [5,6], the second-order reliability method (SORM) [3,7], and Monte Carlo simulation (MCS) [8]. FORM and SORM are applied after evaluating the design point using a numerical search algorithm, while MCS estimates the failure probability through a large number of samples. The application of the first two methods may be limited by the considered scenarios and expression of performance function (e.g., limited to the case of component reliability), while MCS can avoid mathematical problems and achieve high accuracy in solving the failure probability. Therefore, MCS has been widely used in classical reliability calculation methods. However, concerning the scenarios of small-failure probabilities, the classical MCS requires many samples to reach the desired accuracy in estimating the failure probability, which may cause a significant computational burden and limit its application. Thus, some strategies to reduce the computational cost, e.g., importance sampling method (IS) [9], subset simulation (SS) [10], asymptotic sampling [11], etc., have been proposed. Although these methods are well established and more efficient than the

classical MCS, they still require high computational cost, especially for complex performance functions and small-failure probabilities. Therefore, there is still a need for a reliability analysis with high precision and low computational cost.

Currently, a popular path for reducing the computational cost of reliability analysis is to build surrogate models, such as response surface models [12,13], kriging models [14,15], support vector machine models [16], and polynomial chaos kriging (PCK) models [17,18]. The main merit of surrogate models is that a large number of calls to the original performance function can be avoided, which reduces the computational cost. The prediction accuracy of a surrogate model depends on its training set, affecting the reliability calculation results. Therefore, Echard *et al.* [19] proposed a strategy to identify updating points and enrich the training set of the Kriging model, which is called the adaptive Kriging-MCS (AK-MCS) method of reliability analysis. Schöbi *et al.* [18] also applied the PCK model to achieve adaptive MCS to estimate the small-failure probability. Recently, the idea of adaptive MCS has been widely developed. Some studies proposed learning functions in adaptive MCS [20,21], while others focused on reducing the variation of failure probability and improving the robustness of rare-event probability estimation with IS and SS. For example, Echard *et al.* [22] and Cadini *et al.* [23] suggested combining AK-MCS with IS. Besides, Huang *et al.* [24] and Ling *et al.* [25] proposed adaptive SS by integrating the Kriging model and SS to estimate small failure probability. Wei *et al.* [26] implemented global sensitivity and reliability analysis for rare events by employing adaptive SS. To further improve the performance of adaptive SS, Xu *et al.* [27] proposed a modified SS-based adaptive MCS method, and Tong *et al.* [28] developed an adaptive MCS method by integrating SS and IS. The adaptive MCS method can be implemented in both the time-invariant and time-variant cases [29,30]. Although existing adaptive MCS-based methods have been successfully applied in different applications, they still depend on training sets and candidate sample collection. Regarding the rare-event probability estimation and non-linear scenarios, many samples might be required to build an accurate surrogate model for reliability estimation. It is not easy to achieve a tradeoff between the accuracy and efficiency of reliability analysis, even for the adaptive MCS-based approach.

On the other hand, a probability density function (PDF)-oriented method has been recently developed to reduce the computational cost of reliability analysis. The PDF-oriented method aims to calculate reliability utilizing the PDF of a performance function. For instance, Chen and Li [31] developed the probabilistic density evolution method (PDEM) to perform reliability analysis of non-linear stochastic structures by directly solving the generalized density evolution equation (GDEE) to obtain the PDF of the structural response. Also, Guo *et al.* [32] proposed some complementary strategies to extend the application scale of PDEM to the reliability analysis of degraded structures. Although PDEM performed well in previous studies, its focus on solving GDEE may limit its application. A more flexible strategy is to evaluate the PDF of performance function through nonparametric evaluation ways such as kernel density estimation (KDE) [33–35]. For instance, based on the maximum entropy principle, Alibrandi and Mosalam [36] proposed a KDE-maximum entropy method (MEM) to estimate the probability distribution with limited samples. Also, Xu and Kong [37] combined the point evolution method and used KDE in reliability analysis. Similarly, Guo *et al.* [38] proposed a point-evolution kernel density estimation (PKDE) based time-dependent reliability analysis framework for engineering structures subject to multiple deterioration mechanisms. Although existing studies confirmed the excellent performance of KDE in general reliability analysis, the method has not been validated within rare events of small probabilities. Therefore, it is still necessary to further investigate and improve the KDE-based methods (KDEM) with more cases and scenarios.

The purpose of this paper is to develop an adaptive estimation method for the small-failure probability in terms of some advanced probabilistic techniques such as PKDE, PCK, and SS. The **P**KDE-**A**daptive **P**CK-based **S**S (PAPS) is developed. Section 2 presents the definition of the small-failure probability and the basic idea of the proposed method. The main technical details of the proposed method are described in Section 2, where the algorithm of PKDE is reviewed, followed by a detailed description of complementary strategies such as SS and PCK. Next, in Section 3, four benchmark cases demonstrate the computational effectiveness of the proposed PAPS method, including a simple two-dimensional case, a classical static rod truss, a time-variant steel beam, and a non-linear random structure. Meanwhile, the reliability analysis

results are compared with other existing methods. Finally, conclusions are drawn, and directions for future work are indicated.

2. Estimation of rare-event probability

In this study, the performance function g and failure probability p_f are denoted as Eqs. (1) and (2), respectively

$$g = G(\boldsymbol{X}) \tag{1}$$

$$p_f = \int_{\Omega_f} f_G(g)\,\mathrm{d}g \tag{2}$$

in which $\boldsymbol{X}$ denotes the input vector containing d random variables $[X_1,X_2,\ldots,X_d]^T$; $f_G(g)$ is the probability density function (PDF) of performance function; and Ω_f is the failure zone.

Besides, the small failure or rare-event probability is defined as the occurrence probability with an order not more than 10^{-4} [39]. Traditional estimation methods of small-failure probability focus on using techniques to precisely calculate the failure probability, such as IS or SS. Both two methods focus on reducing the variations involved in the failure probability estimation. The former IS works on sampling through an importance sampling density (ISD), and the latter SS splits the failure event into a series of events. The accuracy and efficiency of IS depend on the choice of ISD and the dimension of the investigated case [10], while those of SS relate to the separation level and the number of samples within each subset [40].

Due to the small sample size required and the high accuracy in practical applications, this paper adopts the PDF-oriented idea of capturing the PDF of $G(\boldsymbol{X})$, i.e., $f_G(g)$, and then integrating the $f_G(g)$ within the failure zone to obtain the failure probability p_f as

$$p_f = \int_{g<0} f_G(g)\mathrm{d}g = F_G(0) \tag{3}$$

where $F_G(\cdot)$ is the cumulative distribution function (CDF) of the performance function.

This paper presents an estimation method for the small-failure probability that requires small inputs while maintaining high accuracy. To obtain $f_G(g)$ by a limited number of samples, there generally exist two popular paths: maximum entropy-based method [41,42] and point-evolution-based method [31,38]. For the former, the main idea is to capture the integer or

fractional moments of the performance function and then optimize the distribution of the performance function, i.e., f_G (g). However, regarding the discontinuous and non-convex objection function, such a strategy might not be efficient [42]. For the latter, a limited number of representative points are chosen to compute their assigned probabilities and the values of the performance function. Then, $f_G(g)$ is captured by solving GDEE [31] or PKDE [38]. PKDE is employed to capture the initial $f_G(g)$ in this study due to its convenience and accuracy. However, existing point-evolution methods, including PDEM and PKDE, are not developed for small probability estimation. Thus, corresponding improvements are proposed to improve the accuracy of PKDE for small probability estimation.

2.1. A review of point-evolution based kernel density estimation (PKDE) and its limitation for rare events

To perform PKDE, the first step is to select a point set $\boldsymbol{x}$ of n_{sel} representative points [38]. Existing studies proved that the uniformity of representative point sets significantly influences the accuracy of point-evolution methods [43,44]. Thus, this study employed a generalized F-discrepancy (GF-discrepancy) based point selection method (GFM) [44], and the information of the algorithms refers to Appendix A.1.

After capturing n_{sel} representative points and their assigned probability $p_{a,i}$ (i=1,…,n_{sel}) (i.e., Eq.(A1.4)), the estimation of the target PDF $\hat{f}_G(g)$ can be expressed as follows:

$$\hat{f}_G(g) = \sum_{i=1}^{n_{\text{sel}}} p_{a,i} \cdot \delta(g - g_i) \tag{4}$$

where g_i is the value of the i-th representative point calculated by Eq.(1); and δ (·) is Dirac's delta function.

In PKDE, $\delta(\cdot)$ of Eq.(4) is represented through a kernel density estimator, and then the target PDF can be evaluated as follows:

$$\hat{f}_G(g,h) = \sum_{i=1}^{n_{\text{sel}}} p_{a,i} \cdot K(g, g_i, h) \tag{5}$$

where $K(\cdot)$ is the kernel function that could be selected based on the distribution region and

characteristics of representative points [45,46]; and h is the bandwidth.

Supposing the target distribution is an infinite region, $K(\cdot)$ in Eq.(5) could be the PDF of a Gaussian distribution whose mean is g_i and the standard deviation is h

$$K\left(g, g_i, h\right) = \frac{1}{h\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{g - g_i}{h} \right)^2 \right] \tag{6}$$

For most scenarios, applying many samples is costly due to experimental conditions and computational burden limitations. Thus, selecting an appropriate bandwidth h becomes significant due to the limited number of samples. To obtain the optimal bandwidth h, a common strategy is to find the h satisfying the minimum mean integrated square error (MISE). Such a strategy is hard to be implemented directly since $f_G(g)$ is unknown. Thus, a plug-in bandwidth selection method is employed based on a diffusion partial differential equation [47]

$$\frac{\partial}{\partial h} \hat{f}_G(g,h) = \frac{1}{2} \frac{\partial^2}{\partial g^2} \hat{f}_G(g,h), g \in \chi, h > 0 \tag{7}$$

where $\chi \equiv R$.

Besides, the initial condition $\lim_{h \to 0} \hat{f}_G(g,h)$ can be approximately calculated by a histogram with assigned probability $p_{\mathrm{a},i}$, and a small bin Δg

$$\lim_{h \to 0} \hat{f}_G(g,h) \approx \sum_{i=1}^{n_{\mathrm{sel}}} p_{\mathrm{a},i} \cdot H\left(\Delta g - \left| g - g_i \right| \right) \tag{8}$$

where $H(\cdot)$ is the Heaviside function, and it is equal to 1 when the value in the bracket $\geqslant 0$ and 0 otherwise.

For simplicity, the values of n_{sel} representative points can be linearly mapped to [0, 1] and the mapped value of the i-th representative point u_i by Eq. (9).

$$u_i = \frac{1}{g_{\mathrm{rg}}} \cdot \left[g_i - \min\left(g_i \right) \right], i = 1, ..., n_{\mathrm{sel}} \tag{9}$$

in which g_{rg} is the distribution range of n_{sel} representative points which could be calculated by

$$g_{\mathrm{rg}} = \max\left(g_i \right) - \min\left(g_i \right), i = 1, ..., n_{\mathrm{sel}} \tag{10}$$

Furthermore, the optimal square bandwidth t can be obtained by employing the algorithms in Appendix A.2, and the optimal bandwidth h can be expressed by

$$h = g_{rg}\sqrt{t} \tag{11}$$

In terms of the values of representative points x_i and the optimal bandwidth h, the estimation of $\hat{f}(g,h)$ could be obtained. The principal algorithms of PKDE are summarized in Algorithm 1.

Algorithm 1. Computation of PKDE

1: Generate n_{sel} representative points via the method in Appendix A.1
2: Set initial error variable $\varepsilon_0 = 10^3$, precision ε, initial $z_0 = \varepsilon$, $n = 0$, $k=10^4$ and g_{rg} by Eq.(10)
3: Set the initial PDF by Eq.(8) with $\Delta g = g_{rg}/k$
4: **while** $\varepsilon_0 > \varepsilon$ **do**
5: Set $z_{n+1}=\xi\gamma^{[l]}(z_n)$ through Eq. (A2.6)
6: $\varepsilon 0=| z_{n+1}- z_n|$ and n:= n+1
7: **end do**
8: Obtain the optimal bandwidth $h = g_{rg}\sqrt{z_{n+1}}$
9: Calculate $\hat{f}_G(g,h)$ by Eq.(5)

Although PKDE in Section 2.1 could accurately estimate the PDF of performance for most scenarios, additional improvements might be required under the scenario of rare-event estimation. To demonstrate this statement, a simple case is taken for illustrative purpose, and its performance function g is written as [38]

$$g = G(x_1, x_2) = g_{cr} + x_1 \cdot H(|x_1| - |x_2|) + x_2 \cdot \left[1 - H(|x_1| - |x_2|)\right] \tag{12}$$

in which g_{cr} is the threshold value of 8; and x_1 and x_2 are Gaussian variables following N(-2,1) and N(2,3), respectively.

As illustrated in Fig. 1a, 54 samples and representative points are generated by Latin Hypercube Sampling (LHS) [48] and GF-discrepancy based point selection method (GF), respectively. As shown, the representative points look more uniform than LHS samples. Then, LHS samples and GF points are used to capture the PDF of g, i.e., $f_G(g)$, by classical KDEM [47] and PKDE, respectively, as presented in Fig. 1b. Compared with 10^7 samples of classical KDEM, it can be noticed that PKDE performs better than KDEM in describing the PDF of Eq. (12) through a limited number of samples. However, it also can be observed that both the tails of PDFs by 54 samples of KDEM and PKDE are inaccurate on the logarithmic scale. Due to the inaccurate PDF, the CDF $F_G(g)$ evaluated by the small samples of KDEM and PKDE is

incorrect. In terms of Eq.(3), the p_f obtained by 54 samples of KDEM and PKDE are 1.61×10^{-6} and 4.09×10^{-6}, respectively, which deviate much from 10^7 samples of classical KDEM (4.28×10^{-4}). Based on the definition of rare-event probability, such a phenomenon might frequently occur in the existing PDF-oriented methods [32,38]. Therefore, it is crucial to overcome this shortcoming and improve the accuracy in the tails of PDF and CDF by PDF-oriented methods. Therefore, based on the results of PDF-oriented methods, the idea of using other techniques to refine the PDF/CDF of its tail is needed and conducted in this study.

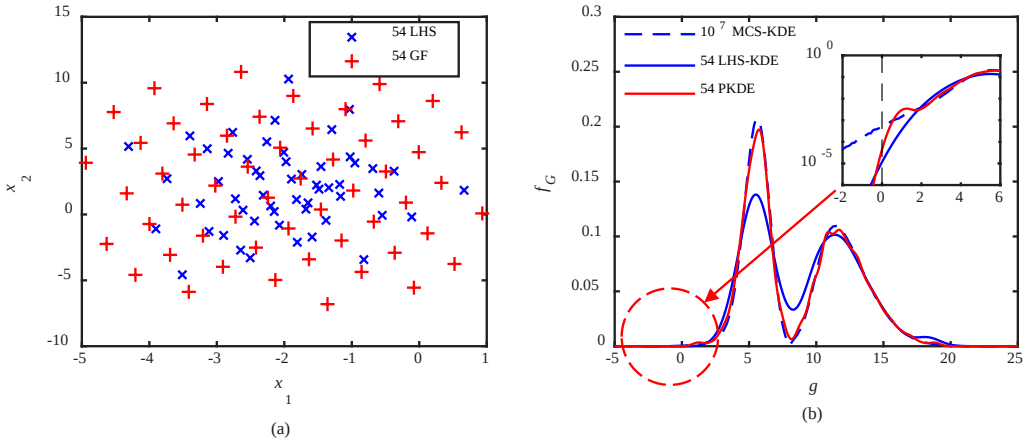


Fig. 1. Schematic of point sets and PDFs of Eq.(12): (a) point sets by LHS and GF-discrepancy based point selection method; and (2) PDF by classical KDE and PKDE

2.2. Proposed method for efficient reliability method for small probability failure

2.2.1. Basic concept of PAPS

Since the failure probability calculation depends directly on the CDF of the performance function $F_G(g)$, the main goal is to improve the accuracy of the CDF calculation. This paper also applies the concept of SS to conservatively evaluate the tail of the CDF by approaching the limit state step-by-step. The primary thought of SS is to split a failure event $F \equiv \{G < 0\}$ into a series of intermediate events F_j [10]

$$F_j = \left\{G \le b_j\right\}, F_1 \supset F_2 \supset \ldots \supset F_m = F, b_1 > b_2 > \ldots > b_m = 0, \left(j = 1,\ 2,\ \ldots, m\right) \tag{13}$$

where b_j is the threshold value at the j-th level of SS.

Then, failure probability p_f can be expressed as the product of $P(F_1)$ and a series of conditional probabilities $P(F_j|F_{j-1})$ obtained through the j-th level SS,

$$p_{\mathrm{f}} = P(F) = P(F_1) \cdot \prod_{j=2}^{m} P(F_j \mid F_{j-1}),\ j = 2,\ 3,\ \ldots, m \tag{14}$$

where $P(F_1)$ is the probability that the performance function is below b_1. Being different from the traditional SS, $P(F_1)$ can be obtained through the PKDE results of $F_G(\cdot)$, i.e., integrating the PDF obtained in Section 2.1:

$$P(F_1) = F_G(b_1) = \int_{g<b_1} f_G(g)\mathrm{d}g \tag{15}$$

Besides, $P(F_j|F_{j\text{-}1})$ in Eq.(14) can be estimated through Markov chain Monte Carlo (MCMC) as follows [49]:

$$P(F_j \mid F_{j-1}) \approx \frac{1}{N_j} \sum_{k=1}^{N_j} I_{F_j}(\boldsymbol{x}_k^j),\ j = 2, \ldots, m \tag{16}$$

in which N_j is the number of samples at the j-th level of SS; $\boldsymbol{x}_k^j$ is the k-th sample at the j-th level of SS generated by MCMC and an implicit conditional PDF $q(\boldsymbol{x}|F_{j\text{-}1})$ expressed by Eq.(17); and I_{Fj} $(\boldsymbol{x}_k{}^j)$ is an indicator function written by Eq.(18).

$$q(\boldsymbol{x} \mid F_{j-1}) = \frac{I_{Fj-1}(\boldsymbol{x}) \cdot f_{\boldsymbol{X}}(\boldsymbol{x})}{P(F_{j-1})},\ j = 2, 3, \ldots, m \tag{17}$$

where $f_{\boldsymbol{X}}(\boldsymbol{x})$ is the joint PDF of random variables in $\boldsymbol{X}$.

$$I_{F_j}(\boldsymbol{x}_k^j) = \begin{cases} 1, G(\boldsymbol{x}_k^j) \le b_j \\ 0, G(\boldsymbol{x}_k^j) > b_j \end{cases} \tag{18}$$

In addition, during the implementation of SS, the values of $P(F_1)$ and $P(F_j|F_{j\text{-}1})$ (j = 2, 3, …, m-1) are generally fixed to a constant value $p_0 \in [0.1, 0.3]$ [50]. Given $P(F_1)$ of p_0, b_1 can be determined by the inverse function of CDF of G, $F^{-1}{}_G(p_0)$. Other thresholds of b_j are determined by the p_0 quantile of N_j response values of the performance function at the j-th level of SS. All response values at the j-th level of SS are ordered in ascending order, i.e., $G(\boldsymbol{x}_1^j) \le G(\boldsymbol{x}_2^j) \le \ldots \le G(\boldsymbol{x}_{N_j}^j)$, and b_i can be expressed as $G(\boldsymbol{x}_{[N_j p_0]}^j)$, in which '[]' refers to the round of the argument. Furthermore, p_{f} in Eq.(14) can be written as:

$$p_{\mathrm{f}} = p_0^{m-1} \cdot \frac{1}{N_m} \sum_{k=1}^{N_m} I_{F_m}(\boldsymbol{x}_k^m) \tag{19}$$

Besides, based on [10], the coefficient of variation (COV) of p_{f} could be evaluated through Eq.(20).

$$\delta_{\mathrm{f}} = Cov\left(p_{\mathrm{f}}\right) \approx \sqrt{\sum_{j=1}^{m} Cov^2\left[P\left(F_j\right)\right]} \tag{20}$$

where Cov[$P(F_1)$] could be calculated using Eq.(21) for sampling-based methods but is zero for PDKE since the results of PKDE are determined given representative points, and Cov[$P(F_j)$] (j>1) can be evaluated by Eq.(22).

$$Cov\left[P\left(F_1\right)\right] = \sqrt{\frac{1-P\left(F_1\right)}{N_1 P\left(F_1\right)}} \tag{21}$$

$$Cov\left[P\left(F_j\right)\right] = \sqrt{\frac{1-P\left(F_j\right)}{N_j P\left(F_j\right)}\left(1+\gamma_j\right)},\ j = 2,....,m \tag{22}$$

in which γ_j can be expressed as

$$\gamma_j = 2\sum_{k=1}^{N_j/N_c^{(j)}-1}\left[\left(1-\frac{k\cdot N_c^{(j)}}{N_j}\right)\rho_j\left(k\right)\right] \tag{23}$$

where $N_{\mathrm{c}}^{(j)}$ is the number of Markov chains at the j-th level of SS; and $\rho_j(k)$ is the correlation coefficient which could be calculated as

$$\rho_j\left(k\right) = \frac{\dfrac{1}{N_j - k\cdot N_c^{(j)}}\displaystyle\sum_{u=1}^{N_c^{(j)}}\sum_{v=1}^{N_j/N_c^{(j)}-j} I_{F_j}^{(u,v)}\cdot I_{F_j}^{(u,v+k)} - p_{\mathrm{f}}^2}{p_{\mathrm{f}}\left(1-p_{\mathrm{f}}\right)} \tag{24}$$

where $I_{F_j}^{(u,v)}$ and $I_{F_j}^{(u,v+k)}$ denote the values of Eq.(18) associated with the v-th and (v+k)-th sample in the u-th Markov chain [10,25].

Accordingly, the failure probability for the rare events can be calculated, and corresponding COV could be evaluated. The empirical CDF of performance can be captured by summarizing the samples for each level of SS and the initial PDF obtained through PKDE. However, if the traditional SS is directly employed, each level of SS still needs many samples, bringing a high computational burden. Therefore, it is necessary to build surrogate models based on the samples involved in PKDE to avoid the mass invocations of the original performance function.

With respect to surrogate models, Gaussian Kriging models are generally applied in conventional adaptive reliability analysis [19,24,27]. However, existing studies proved that Gaussian Kriging models might not be the optimal surrogate model, especially for high-dimensional cases [18,51,52]. Thus, in this study, surrogate modeling is implemented by PCK since PCK integrates the merits of the Kriging model and polynomial chaos expansion (PCE) [17]. In PCK, the Kriging model is used to interpolate the local variations of the target variables as a function of input points [17,53]. Meanwhile, PCE is applied to approximately evaluate the global behavior of the target variable through a set of orthogonal polynomials [54]. Thus, PCK has its theoretical advantage compared with the conventional Gaussian Kriging (GK) model and has been widely applied in adaptive reliability analysis, especially for small failure probability estimation [18,52]. The establishment approaches of PCK are introduced in Appendix A.3.

Furthermore, to improve the probability estimation accuracy, at each level of the SS, a limited number of additional samples is adaptively selected by employing learning functions, and the corresponding computational result is revised. Finally, the process of SS stops once the limit state is reached, and the final CDF is obtained. Meanwhile, p_f can be computed through Eq.(3).

2.2.2. Integration between PKDE and adaptive PCK-based SS (PAPS)

Fig. 2 shows the basic flow of the proposed PKDE-Adaptive PCK-based SS (PAPS) method for calculating the probability of small-failure events. As shown in Fig. 2, the initial PCK model G^{PCK} can be built based on the initial point selection in PKDE and the procedures of PCK in Appendix A.3. Then, by replacing the original performance function with G^{PCK}, SS can be performed using the algorithm in Section 2.2.1. Since the original performance functions are not used extensively, the computational burden can be significantly reduced. However, such a PCK model based on the initial point selection cannot guarantee the computational accuracy of the CDF and failure probability. As a result, based on the thought of PAPS described in Section 2.2.1, it is necessary to update the PCK model by adaptively adding new point selections during

the computation.

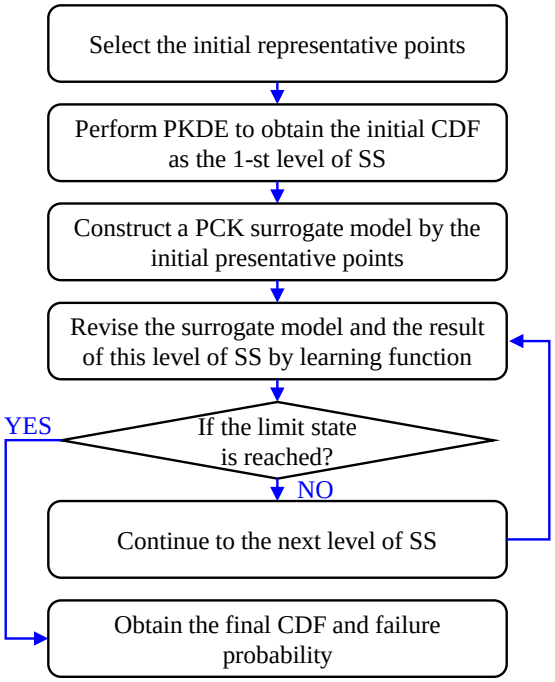


Fig. 2. Procedures underlying the proposed PAPS method of small-failure probability

In the proposed method, initial point sets are regarded as training sets. At each level of SS, update points could be found by the learning function and then added to the training sets. For adaptive reliability analysis, the most common learning function are U function [19] ($U(\boldsymbol{x})$, Eq.(25)) and the expected feasibility function (EFF) [55] (EFF($\boldsymbol{x}$)).

$$U\left(\boldsymbol{x}_i^j\right)=\frac{\left|G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)-b_j\right|}{\sigma^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)},\ \boldsymbol{x}_i^j\in\boldsymbol{X}^j, i=1,2,\ldots,N_j,\ j=1,2,....,m \tag{25}$$

where N_j is the number of samples in the j-th level of SS; $\boldsymbol{X}^j$ is the set of samples at the j-th level of SS; and $G^{\mathrm{PCK}}(\boldsymbol{x}^j{}_i)$ and $\sigma^{\mathrm{PCK}}(\boldsymbol{x}^j{}_i)$ are the predicted mean value and standard deviation (STD) of PCK model at point $\boldsymbol{x}^j{}_i$.

$$
\begin{aligned}
&\mathrm{EFF}\left(\boldsymbol{x}_i^j\right)=\left(G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)-b_j\right)\cdot \\
&\left[2\Phi\left(\frac{b_j-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)-\Phi\left(\frac{\left(b_j-\varepsilon\right)-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)-\Phi\left(\frac{\left(b_j+\varepsilon\right)-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)\right] \\
&-\sigma^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)\cdot \\
&\left[2\phi\left(\frac{b_j-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)-\phi\left(\frac{\left(b_j-\varepsilon\right)-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)-\phi\left(\frac{\left(b_j+\varepsilon\right)-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)\right] \\
&+\left[\Phi\left(\frac{\left(b_j+\varepsilon\right)-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)-\Phi\left(\frac{\left(b_j-\varepsilon\right)-G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right)}\right)\right], \\
&\boldsymbol{x}_i^j\in\boldsymbol{X}^j, i=1,2,\ldots,N_j,\ j=1,2,\ldots,m
\end{aligned}
\tag{26}
$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the CDF and PDF of standard Gaussian distribution, respectively; and ε is a small value could be selected as $2\cdot[\sigma^{\mathrm{PCK}}(\boldsymbol{x}_i^j)]^2$ [19].

If the minimum value of U(x) is lower than 2 or the maximum value of EFF($\boldsymbol{x}$) is higher than 0.001[19], the corresponding point is picked up, and the actual value of its performance function is calculated; otherwise, no update occurs, and the next level of SS is executed. The detailed procedures of the proposed PAPS are as follows:

(1) Initialize j = 1 and denote the initial representative point selection as a training set for PCK $\mathbf{X}^{\mathrm{train}}=\left\{\boldsymbol{x}_i^{\mathrm{train}}, i=1,2,\ldots,n_{\mathrm{sel}}\right\}$. A set of performance functions $G\left(\mathbf{X}^{\mathrm{train}}\right)=\left\{G\left(\boldsymbol{x}_i^{\mathrm{train}}, i=1,2,\ldots,n_{\mathrm{sel}}\right)\right\}$ could be obtained and used to conduct PKDE to get PDF and CDF $F_{\mathrm{G}}(\cdot)$ and build a PCK model G^{PCK};

(2) Determine the parameter p_0 (p_0 is suggested to be ranging from 0.1 to 0.3[50]), b_j by the inverse function of CDF of G, i.e., $F^{-1}{}_G(p_0)$, and a seed set of points $\boldsymbol{X}_{\mathrm{seed}}^j$ in terms of $\boldsymbol{x}^{\mathrm{train}}$, to satisfy the condition that for arbitrary $\boldsymbol{x}_{\mathrm{seed},i}^j\in\boldsymbol{X}_{\mathrm{seed}}^j, G(\boldsymbol{x}_{\mathrm{seed},i}^j)\le b_j$ at the j level of SS;

(3) Generate a set of N_j random samples $\boldsymbol{X}^j=\left\{\boldsymbol{x}_i^j, i=1,2,\ldots,N_j\right\}$ ($\boldsymbol{X}^j$ is the candidate sample pool and its size N_j relies on several parameters (e.g., the target coefficient of variation, p_0 value, the performance of the computing platform)), a set of PCK values $\boldsymbol{G}^{\mathrm{PCK}}(\boldsymbol{X}^j)$ and their standard deviations (STD) $\boldsymbol{\sigma}^{\mathrm{PCK}}(\boldsymbol{X}^j)$,

$$
\boldsymbol{G}^{\mathrm{PCK}}\left(\boldsymbol{X}^j\right)=\left\{G^{\mathrm{PCK}}\left(\boldsymbol{x}_i^j\right), i=1,2,\ldots,N_j\right\},
\tag{27}
$$

$$\boldsymbol{\sigma}^{\mathrm{PCK}}\left(\boldsymbol{X}^{j}\right)=\left\{\sigma^{\mathrm{PCK}}\left(\boldsymbol{x}_i^{j}\right), i=1,2,\ldots,N_j\right\}, \tag{28}$$

based on b_j, $\boldsymbol{X}_{\mathrm{seed}}^{j}$, G^{PCK} and MCMC technique, while $U(\boldsymbol{x}_i^j)$ are computed through Eq.(25) for each sample $\boldsymbol{x}_i^j$ in $\boldsymbol{X}^j$;

(4) Skip this step to Step (5), if $\min\left[U\left(\boldsymbol{x}_i^{j}\right)\right]\geq 2$ or $\max\left[\mathrm{EFF}\left(\boldsymbol{x}_i^{j}\right)\right]\leq 10^{-3}$. Otherwise, pick up $\boldsymbol{x}_{\mathrm{up}}^{j}$, i.e.,

$$\boldsymbol{x}_{\mathrm{up}}^{j}=\arg\min\left[U\left(\boldsymbol{x}_i^{j}\right)\right] \text{ or } \arg\max\left[\mathrm{EFF}\left(\boldsymbol{x}_i^{j}\right)\right], i=1,2,\ldots,N_j, \tag{29}$$

to update the point set $\boldsymbol{X}^{\mathrm{train}}=\boldsymbol{X}^{\mathrm{train}}\bigcup\boldsymbol{x}_{\mathrm{up}}^{j}$ and then update G^{PCK} based on $\boldsymbol{X}^{\mathrm{train}}$ and $\boldsymbol{G}(\boldsymbol{X}^{\mathrm{train}})$. This step would be repeated until $\min\left[U\left(\boldsymbol{x}_i^{j}\right)\right]\geq 2$ or $\max\left[\mathrm{EFF}\left(\boldsymbol{x}_i^{j}\right)\right]\leq 10^{-3}$ is satisfied;

(5) Regenerate $\boldsymbol{X}^j$ by MCMC in terms of b_j, $\boldsymbol{X}_{\mathrm{seed}}^{j}$ and updated G^{PCK}, and then resort PCK values $\boldsymbol{G}^{\mathrm{PCK}}(\boldsymbol{X}^j)$ through its ascending order;

(6) Let $j=j+1$ and obtain b_j by the p_0 quantile of $\boldsymbol{G}^{\mathrm{PCK}}(\boldsymbol{X}^{j-1})$, i.e., $G^{\mathrm{PCK}}(\boldsymbol{X}_{[N_{j-1}p_0]}^{j-1})$. If $b_j \leq 0$, estimate $P\left(F_j \mid F_{j-1}\right)$ by Eq.(16); otherwise, $P\left(F_j \mid F_{j-1}\right)=p_0$;

(7) Determine a new seed set of points $\boldsymbol{X}_{\mathrm{seed}}^{j}$ according to $\boldsymbol{X}^{j-1}$ to satisfy the condition that for arbitrary $\boldsymbol{x}_{\mathrm{seed},i}^{j}\in\boldsymbol{X}_{\mathrm{seed}}^{j}, G(\boldsymbol{x}_{\mathrm{seed},i}^{j})\leq b_j$. Next, repeat Steps (3)-(5) to generate a new $\boldsymbol{X}^j$ and update G^{PCK}; and

(8) Revise b_j by calculating $\boldsymbol{G}^{\mathrm{PCK}}(\boldsymbol{X}^j)$, and regenerate $\boldsymbol{X}^j$ through the revised b_j, updated G^{PCK} and $\boldsymbol{X}_{\mathrm{seed}}^{j}$. Let $m=j$, calculate $P\left(F_j \mid F_{j-1}\right)$ and p_f by Eqs.(16) and (19), respectively, and obtain the final CDF of the performance function based on $F_G(\cdot)$ and the empirical CDF of all $\boldsymbol{G}^{\mathrm{PCK}}(\boldsymbol{X}^j)$, if $b_j \leq 0$. Otherwise, go to step (6).

The above procedure of the proposed PAPS method is summarized in Algorithm 2.

Algorithm 2. Computation of PAPS

Generate N_{sel} representative points as initial point set $\boldsymbol{X}^0$
Evaluate initial PDF and CDF $F_{\text{G}}(\cdot)$ by Algorithm 1
Build a PCK model G^{PCK} through initial representative points
Determine $p_0 \in [0.1, 0.3]$ and Let $j = 1$ and $b_j = F_G^{-1}\left(p_0\right)$
while $b_{\text{j}} > 0$
  if $j >= 2$
    Let $b_j = G^{\text{PCK}}(\boldsymbol{X}^{j-1}_{[N_{j-1}p_0]})$
    if $b_j \le 0$
      Calculate $P(F_j|F_{j\text{-}1})$ by Eq.(16)
    else:
      $P(F_j|F_{j\text{-}1}) = p_0$
    end if
  end if
  Collect a seed set of points $\boldsymbol{X}^j_{\text{seed}} = \left\{G(\boldsymbol{x}^j_{\text{seed},i}) \le b_j, \boldsymbol{x}^j_{\text{seed},i} \in \boldsymbol{X}^j_{\text{seed}} \subset \boldsymbol{X}^{j-1}\right\}$
  Generate a candidate sample pool with N_j samples $\boldsymbol{X}^j$ by MCMC based on b_j, $\boldsymbol{X}^j_{\text{seed}}$, and G^{PCK}
  Compute $U\left(\boldsymbol{x}_i^j\right)$ by Eq.(25) or $\text{EFF}\left(\boldsymbol{x}_i^j\right)$ by Eq.(26) for each sample
  while $\min\left[U\left(\boldsymbol{x}_i^j\right)\right] < 2$ or $\max\left[\text{EFF}\left(\boldsymbol{x}_i^j\right)\right] \le 10^{-3}$ **do**
    Pick up the $\boldsymbol{x}^j_{\text{up}} = \arg\min\left[U\left(\boldsymbol{x}_i^j\right)\right]$ or $\arg\max\left[\text{EFF}\left(\boldsymbol{x}_i^j\right)\right], i = 1, 2, \ldots, N_j$
    Update point set $\boldsymbol{X}^{\text{train}} = \boldsymbol{X}^{\text{train}} \bigcup \boldsymbol{x}^j_{\text{up}}$
    Rebuild G^{PCK} based on X^{train}
    Compute $U\left(\boldsymbol{x}_i^j\right)$ by Eq.(25) or $\text{EFF}\left(\boldsymbol{x}_i^j\right)$ by Eq.(26) for each sample
  end do
  if $j >= 2$
    Revise $b_j = G^{\text{PCK}}(\boldsymbol{X}^j_{[N_j p_0]})$
    if $b_j \le 0$
      Let $m=j$
      Calculate p_{f} by Eq.(19)
    end if
  end if
  Regenerate N_j samples $\boldsymbol{X}^j$ by MCMC based on b_j, $\boldsymbol{X}^j_{\text{seed}}$, and G^{PCK}
  Resort and store $\boldsymbol{G}^{\text{PCK}}(\boldsymbol{X}^j)$ by an ascending order
  Let j:= j +1
end do
Calculate the refined CDF in terms of $F_{\text{G}}(\cdot)$ and empirical CDF of all stored $\boldsymbol{G}^{\text{PCK}}(\boldsymbol{X}^j)$

3. Illustrative examples

In this Section, four numerical cases are presented to validate the accuracy and efficiency of the proposed small-probability analysis method. Some conventional reliability estimation methods, e.g., brute MCS, brute SS, and other adaptive SS methods, are also compared with the proposed

PAPS method. For the SS methods, p_0 is chosen as 0.1, and the candidate sample size of each level is 100,000 for comparing with brute SS. Under such a sample size, the COV of the brute SS is relatively low under the scenarios associated with the small failure probability.

3.1. Case 1: A simple 2D case

A classical 2D case is applied to illustrate the proposed method. To demonstrate the efficiency and accuracy of PAPS, conventional LHS [48] and Gaussian Kriging (GK) modeling [53] are also employed. The performance function g can be written as [27]:

$$g = 0.5\left(x_1 - 2\right)^2 - 1.5\left(x_2 - 5\right)^3 - g_u \tag{30}$$

where g_u is a constant value of three, and x_1 and x_2 are two independent standard Gaussian variables.

To begin with, 12 representative points are selected through the GFM in Appendix A.1. By applying the diffusion-based bandwidth selector in Section 2.1, the bandwidth is chosen as 45.86, and then the PDF and CDF of g can be obtained, as shown in Fig. 3. The CDF of PKDE matches the CDF of 10^8 trials of MCS better than the empirical CDF from 12 samples of LHS. However, the tails of both PDKE and LHS-based KDEM are inaccurate, as shown in the subplot in Fig. 3.

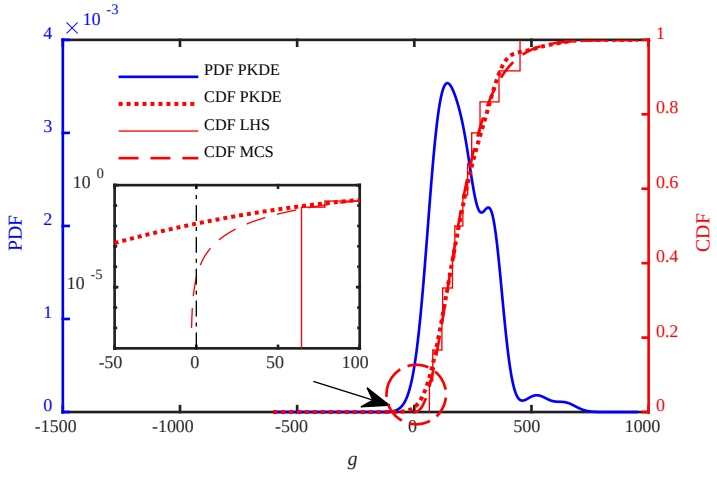


Fig. 3. PDF and CDF estimation by PKDE, LHS, and MCS in Case 1

Furthermore, eight types of approaches, i.e., PAPS-U, PKDE-GK-SS-U, LHS-PCK-SS-U, LHS-GK-SS-U, PAPS-EFF, PKDE-GK-SS-EFF, LHS-PCK-SS-EFF, and LHS-GK-SS-EFF,

are applied to refine the CDF of g. PAPS, including PAPS-U and PAPS-EFF, refers to the algorithms introduced in Section 2.2. PKDE-GK-SS-U and PKDE-GK-SS-EFF refer to replacing PCK in PAPS-U and PAPS-EFF with GK, respectively. Also, LHS-PCK-SS-U and LHS-PCK-SS-EFF refer to replacing PKDE with LHS-KDEM. In addition, LHS-GK-SS-U and LHS-GK-SS-EFF are replacing PCK and PKDE with GK and LHS-KDEM simultaneously.

One run of reliability analysis is taken as an example to compare the performance of different reliability analysis methods. By comparing Fig. 4a, b, c, and d with Fig. 4e, f, g, and h, it can be concluded that the representative points are more uniform than LHS samples. The initial representative points or LHS samples are used to build the surrogate model and are considered the samples at the first level of SS. Based on the built surrogate models, the second, third, fourth, and fifth levels of SS are performed, where 100,000 samples are generated from the MCMC and surrogate models at each level of SS except for the first level of SS.

In addition, Fig. 4a and e show that under the scenarios of PCK, two samples are picked up by the U-learning function and then updated at the third and fourth level of SS, respectively, while in Fig. 4b and f, applying the EFF-learning function does not capture extra samples. Also, in Fig. 4c, d, g, and h, the GK is updated at each level of SS, and the total numbers of extra samples for PKDE-GK-SS-U, PKDE-GK-SS-EFF, LHS-GK-SS-U, and LHS-GK-SS-EFF are 421, 170, 868, and 158, respectively. It can be noticed that both PCK and EFF-learning function could significantly reduce the number of extra samples.

Additionally, Fig. 5 compares the CDF obtained by running all methods in Case 1 once, including brute MCS, original SS, PAPS, PAPS-U, PKDE-GK-SS-U, LHS-PCK-SS-U, LHS-GK-SS-U, PAPS-EFF, PKDE-GK-SS-EFF, LHS-PCK-SS-EFF, and LHS-GK-SS-EFF. As shown, except for LHS-GK-SS-EFF, most CDFs are close to each other. The CDF curves of the failure zone are relatively small or even disappeared for the scenarios using SS methods because the computational process of SS methods stops once b_j is below or equal to zero. In the original SS, 500,000 samples (five levels of SS and 100,000 samples at each level) are generated by the original performance function Eq.(30). Besides, 433, 182, 880, 170 samples (12 initial samples and 421, 170, 868, or 158 extra samples) are used in PKDE-GK-SS-U, PKDE-GK-SS-EFF,

LHS-GK-SS-U, and LHS-GK-SS-EFF. In PAPS-U, PAPS-EFF, LHS-PCK-SS-U, and LHS-PCK-SS-EFF, only 14, 12, 14, and 12 samples (12 initial samples and 2, 0, 2, or 0 extra samples) are needed.

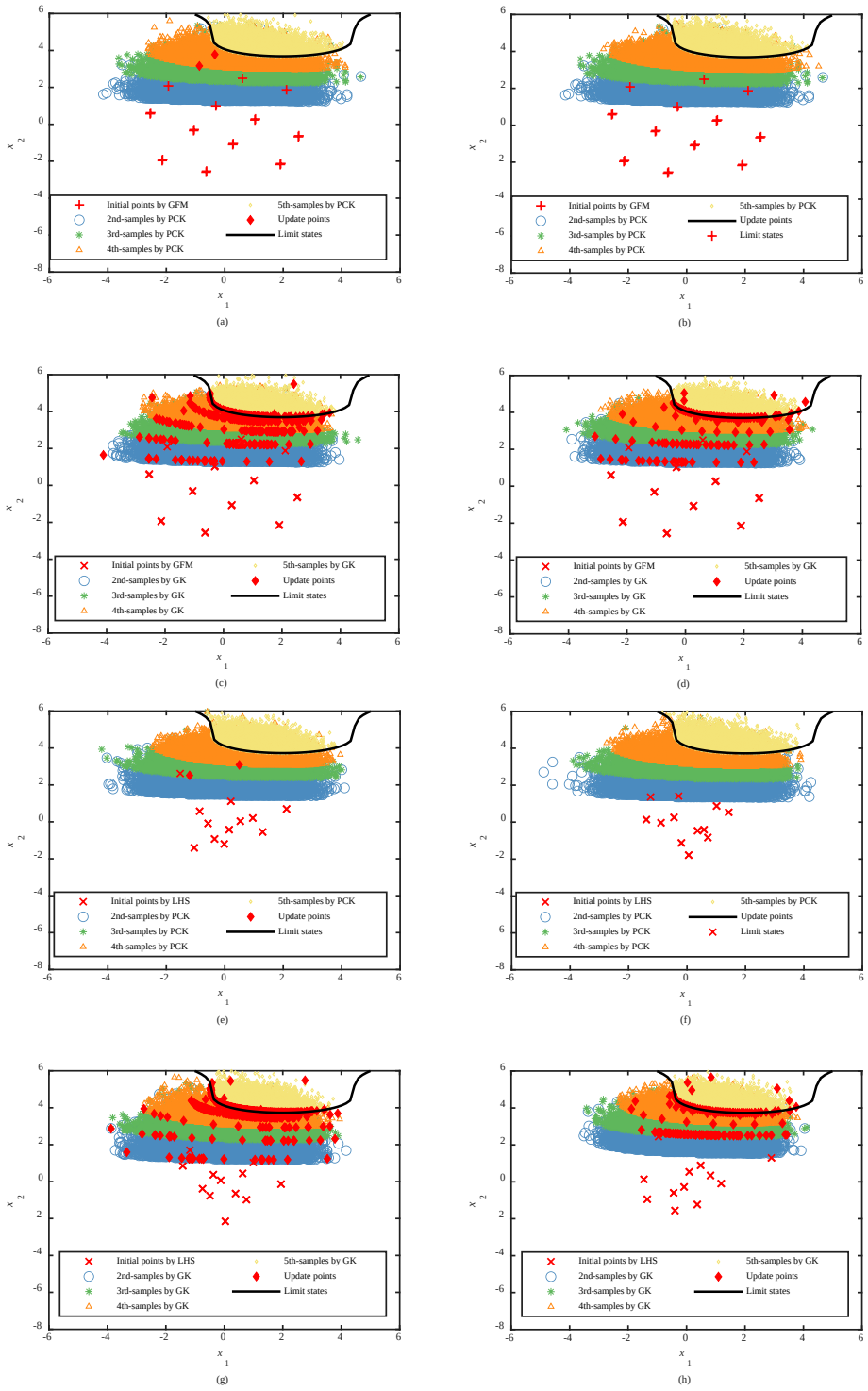


Fig. 4. Classifications of different methods for Case 1: (a) PAPS-U; (b) PKDE-GK-SS-U; (c) LHS-PCK-SS-U; (d) LHS-GK-SS-U; (e) PAPS-EFF; (f) PKDE-GK-SS-EFF; (g) LHS-PCK-SS-EFF; and (h) LHS-GK-SS-EFF

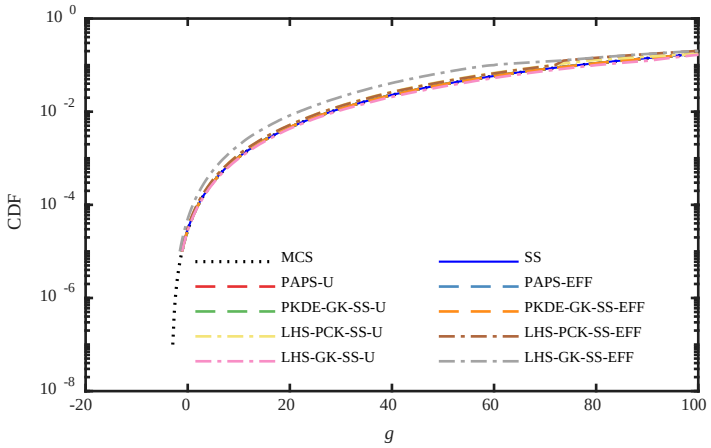


Fig. 5. Comparison of CDF by one run of different methods for Case 1

To further evaluate the performance of different methods in computing failure probabilities, all methods are run 20 times to obtain the average failure probability p_f and COV δ_f (calculated by Eqs.(20)-(24)), as listed in Table 1. As indicated, compared with 10^8 samples of MCS, 12 samples of PKDE might overestimate the failure probability. Among those SS-based methods, the proposed PAPS-U performs the best (the relative error ε_f of 0.10% is the smallest with COV δ_f of 2.80%). In addition, the average numbers of runs of the original performance function, N_{call}, for PAPS-EFF and LHS-PCK-SS-EFF are the smallest (no extra samples are identified).

Table 1 Reliability analysis results associated with different methods for Case 1

Method	N_{call}	p_f	δ_f (%)	ε_f (%)
PKDE	12	0.0122	-	42453.19%
MCS	10^8	2.867×10^{-5}	1.87%	-
SS	500,000	2.876×10^{-5}	2.95%	0.30%
PAPS-U	**12.8**	$\mathbf{2.864\times10^{-5}}$	**2.80%**	**0.10%**
PKDE-GK-SS-U	382.8	2.874×10^{-5}	2.80%	0.26%
LHS-PCK-SS-U	12.7	2.892×10^{-5}	2.94%	0.88%
LHS-GK-SS-U	536.4	2.877×10^{-5}	2.93%	0.34%
PAPS-EFF	12.0	2.874×10^{-5}	2.80%	0.27%
PKDE-GK-SS-EFF	171.4	2.875×10^{-5}	2.80%	0.29%
LHS-PCK-SS-EFF	12.0	2.878×10^{-5}	2.93%	0.38%
LHS-GK-SS-EFF	169.5	2.900×10^{-5}	2.94%	1.13%

Note: N_{call} and ε_f are the average number of runs of the original performance function and the relative error of the failure probability compared with MCS ($|p_f-p_{f,mcs}|/p_{f,mcs}\times100\%$), respectively.

Besides, all adaptive SS methods using PCK have smaller N_{call} (less than 13) than those using GK; the methods using PKDE have smaller δ_{f} and ε_{f} than those using LHS; and the methods with EFF-learning function have smaller N_{call} than those with *U*-learning function. Thus, the calculation results of Case 1 can display the accuracy and efficiency of the PAPS method, where PDKE and PCK outperform LHS and GK. In addition, although the EFF-learning function can reduce the number of calls to the performance function compared with the U-learning function, the accuracy of reliability analysis with the EFF-learning function might not be as good as the U-learning function.

3.2. Case 2: Classical roof truss

As shown in Fig. 6, a classical roof truss structure is employed to illustrate the proposed method [15,20]. This roof truss is assumed to carry a uniformly distributed load q (N/m), which can be transformed into three-point loads $P = ql/4$. More details refer to [15,20].

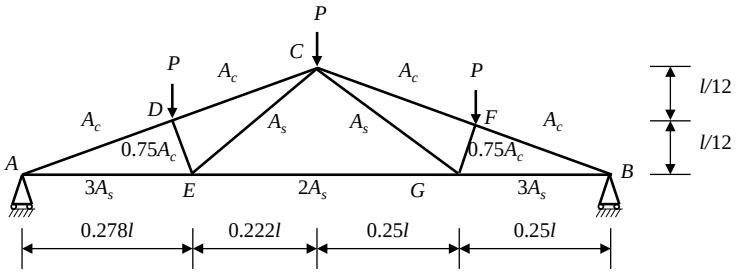


Fig. 6. Schematic of the roof truss

Concerning the mid-span deflection of this roof truss, the performance function g can be defined as:

$$g = \delta_u - \left(\frac{ql^2}{2}\right)\cdot\left(\frac{3.81}{E_c A_c} + \frac{1.13}{E_s A_s}\right) \tag{31}$$

where $\delta_u = 0.04\ m$ is the permissive deflection; l is the span of the roof; A_{c} and E_{c} are the cross-sectional area and Young's modulus of concrete members, respectively; and A_s and E_s are the cross-sectional area and Young's modulus of reinforcement, respectively. All parameters are listed in Table 2.

In this case, 53 initial points are selected through the GFM in Appendix A.1. In terms of

diffusion-based methods, PKDE is achieved to calculate the PDF of the performance function, and the optimal bandwidth is 0.0016. Fig. 7 compares the results of PDF and CDF for different methods. The CDF of PKDE is closer to the one from 10^{10} samples of MCS than the empirical CDF from 53 samples of LHS. Also, the subplot of Fig. 7 shows that the CDF differences among PKDE, LHS-based KDEM, and MCS increase as g tends to negative infinity. Thus, it shows the limitations of PKDE in estimating the PDF of the performance function, which might underestimate the occurrence probabilities of different limit states.

Table 2 Parameters of variables in Case 2 [15,20]

Variable	Distribution	Mean	STD	Variable	Distribution	Mean	STD
δ_u (m)	-	0.04	-	E_s (N/m^2)	Gaussian	10^{11}	6×10^9
q (N/m)	Gaussian	2×10^4	1.4×10^3	A_c (m^2)	Gaussian	0.04	0.0048
l (m)	Gaussian	12	0.12	E_c (N/m^2)	Gaussian	2×10^{10}	1.2×10^9
A_s (m^2)	Gaussian	9.82×10^{-4}	5.98×10^{-5}				

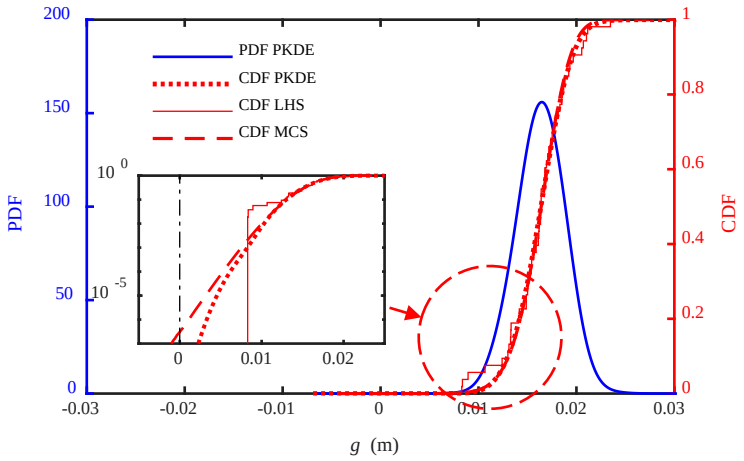


Fig. 7. PDF and CDF estimation by PKDE and MCS in Case 2

Furthermore, brute and adaptive SS methods, including the proposed PAPS, are utilized to calculate the CDF and failure probability. Fig. 8 illustrates the CDFs by one run of all methods. As shown, the CDFs of most methods agree well with MCS except for PAPS-EFF, PKDE-GK-SS-EFF, and LHS-PCK-SS-EFF. The smoothness of CDF curves is not only influenced by the chosen adaptive strategy but also the SS method itself because the tails using SS are derived from the empirical CDF (similar to the stair function) of the candidate samples at each stage [10].

To further assess the performance of different methods in computing failure probabilities, all methods are run 20 times to capture the average p_f and COV δ_f, as summarized in Table 3. The failure probabilities of all methods show that 53 samples of PKDE are infeasible in the failure probability estimation. Also, with respect to SS-based methods, the relative deviation and COV of PAPS-U are the smallest (ε_f of 1.06% and δ_f of 3.57%) and even better 700,000 trials of SS (seven levels of SS with 100,000 samples for each level and its ε_f and δ_f are 1.76% and 3.69%, respectively). The results verify the accuracy and efficiency of the proposed PAPS method.

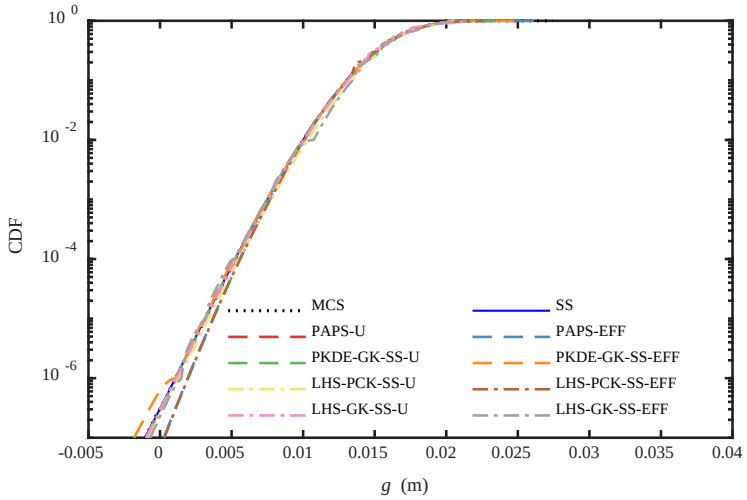


Fig. 8. Comparison of CDF among different methods

Table 3 Reliability analysis results associated with different methods for Case 2

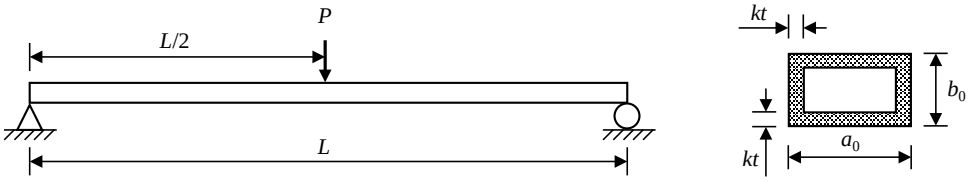
Method	N_{call}	p_f	δ_f (%)	ε_f (%)
PKDE	53	1.641×10^{-10}	-	99.95%
MCS	10^{10}	3.070×10^{-7}	1.80%	-
SS	700,000	2.985×10^{-7}	3.69%	1.76%
PAPS-U	**403.10**	**3.038×10^{-7}**	**3.57%**	**1.06%**
PKDE-GK-SS-U	1000.75	2.940×10^{-7}	3.58%	4.22%
LHS-PCK-SS-U	397.50	2.978×10^{-7}	3.70%	3.00%
LHS-GK-SS-U	918.20	2.764×10^{-7}	3.71%	9.98%
PAPS-EFF	53.50	7.083×10^{-8}	3.81%	76.20%
PKDE-GK-SS-EFF	79.98	3.198×10^{-7}	3.65%	4.99%
LHS-PCK-SS-EFF	54.48	1.194×10^{-7}	3.91%	67.89%
LHS-GK-SS-EFF	80.82	3.651×10^{-7}	3.81%	18.94%

Note: N_{call} and ε_f are the average number of runs of the original performance function and the relative error of the failure probability compared with MCS ($|p_f-p_{f,mcs}|/p_{f,mcs}\times100\%$), respectively.

Besides, similar to Case 1, all methods using the U-learning function and GK need more N_{call} than those using the EFF-learning function and PCK. Although PAPS-EFF has the lowest N_{call}, its accuracy is unsatisfying as well as LHS-PCK-SS-EFF. In addition, although the CDF by a single run of PKDE-GK-SS-EFF is not well consistent with MCS (Fig. 8), its average p_f is close to MCS (δ_f of 4.99%) with N_{call} of 79.98. These results also suggest that it might be more appropriate to apply the EFF learning function combining GK and PKDE to obtain the average failure probability with a relatively small number of samples through multiple runs.

3.3. Case 3: Simply supported steel beam with stochastic external force

In this case, a corroded steel beam is assumed to suffer from steel corrosion and a stochastic external load, as illustrated in Fig. 9 [56].



(a) Corroded steel beam (b) Cross-section

Fig. 9. Schematic of a corroded steel beam

Supposing the failure occurs when the stress of the beam exceeds the ultimate strength of the steel σ_u, the time-variant performance function $g_v(t)$ could be written as:

$$g_v(t) = \sigma_u / 4 \cdot \left(a_0 - 2 \cdot k_t \cdot t\right) \cdot \left(b_0 - 2 \cdot k_t \cdot t\right)^2 - \left(F\left(t\right) \cdot L / 4 + \rho_{st} \cdot a_0 \cdot b_0 \cdot L^2 / 8\right) \tag{32}$$

where t (year) is the time parameter; a_0 and b_0 are the initial dimensions of cross-section (Fig. 9b); L (m) is beam length (Fig. 9b); k_t (m/year) is the corrosion rate; ρ_{st} (N) is the mass density of steel beam; and $F(t)$ is a stochastic process as follows:

$$F\left(t\right) = 6500 + \sum_{i=1}^{7} \xi_i \left(\sum_{j=1}^{7} a_{ij} \sin\left(b_{ij} t + c_{ij}\right) \right) \tag{33}$$

where ξ_i are independent random variables (i = 1, 2, …, 7); a_{ij}, b_{ij} and c_{ij} (j = 1, 2, ..., 7) are the coefficients of the 7×7 matrixes in Eq.(34)-(36); and the parameters of other variables in Eqs.(32) and (33) are listed in Table 4.

$$a=\begin{bmatrix} 0.13 & 0.36 & 0.14 & 3.07 & 0.17 & 0.13 & 0.12 \\ 0.02 & 0.18 & 0.09 & 0.13 & 0.69 & 0.04 & 0.27 \\ 0.08 & 0.29 & 0.14 & 3.09 & 0.05 & 0.37 & 0.13 \\ 0.03 & 0.06 & 0.01 & 0.04 & 0.63 & 0.30 & 0.06 \\ 0.03 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix} \tag{34}$$

$$b=\begin{bmatrix} 0.06 & 0.31 & 0.15 & 0.28 & 0.24 & 0.44 & 0.48 \\ 0.38 & 0.15 & 0.40 & 0.06 & 0.42 & 0.09 & 0.01 \\ 0.10 & 0.33 & 0.03 & 0.29 & 0.11 & 0.26 & 0.38 \\ 0.28 & 0.07 & 0.59 & 0.55 & 0.42 & 0.23 & 0.29 \\ 0.52 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.77 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.91 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix} \tag{35}$$

$$c=\begin{bmatrix} 2.91 & -2.34 & -2.43 & -2.82 & -2.15 & 0.47 & 2.90 \\ -2.91 & 2.21 & -0.97 & 0.98 & -1.03 & -3.81 & -0.35 \\ 1.25 & 0.52 & 2.62 & 0.23 & 0.91 & -1.39 & -2.45 \\ 0.73 & 0.00 & -0.45 & -0.50 & 1.93 & -3.64 & -3.00 \\ 0.18 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -1.71 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -2.46 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix} \tag{36}$$

Table 4 Parameters of variables in Case 3

Variable	Distribution	Mean	STD	Variable	Distribution	Mean	STD
L (m)	-	3	-	ξ_2	Gaussian	0	50
k_t (m/year)	-	5×10^{-5}	-	ξ_3	Gaussian	0	98
ρ_{st} (N)	-	7.85×10^{4}	-	ξ_4	Gaussian	0	121
σ_u (Pa)	Gaussian	2.4×10^{8}	2.4×10^{7}	ξ_5	Gaussian	0	227
a_0 (m)	Gaussian	0.2	0.01	ξ_6	Gaussian	0	98
b_0 (m)	Gaussian	0.04	0.004	ξ_7	Gaussian	0	121
ξ_1	Gaussian	0	100				

Considering the first passage problem, the time-dependent performance function $g(t)$ is written as:

$$g(t)=\min_{\tau\in[0,t]} g_v(\tau) \tag{37}$$

To begin with, 139 initial points are selected by GFM, and eight-time instances from 1 to 35 years (1, 5, 15, …, 35 years) are chosen for this case. By employing PKDE, the time-

dependent optimal bandwidth $h(t)$ is captured, as shown in Fig. 10a. It is found that $h(t)$ gradually decreases versus t. Next, the PDF and CDF of g are obtained, and $p_f(t)$ is further computed and compared with 10^8 trials of MCS. Fig. 10b compares the PDFs and CDFs of different methods at three typical time instants: 1, 15, and 35 years. As in Cases 1 and 2, Fig. 10b shows that the CDFs of PKDE are close to brute MCS. However, the subplot in Fig. 10b shows that the tails of CDFs are quite different from MCS, especially for $g \leq 10^3$. These results also suggest that PKDE may not be feasible in the failure probability estimation of rate events and supplement modifications are necessary.

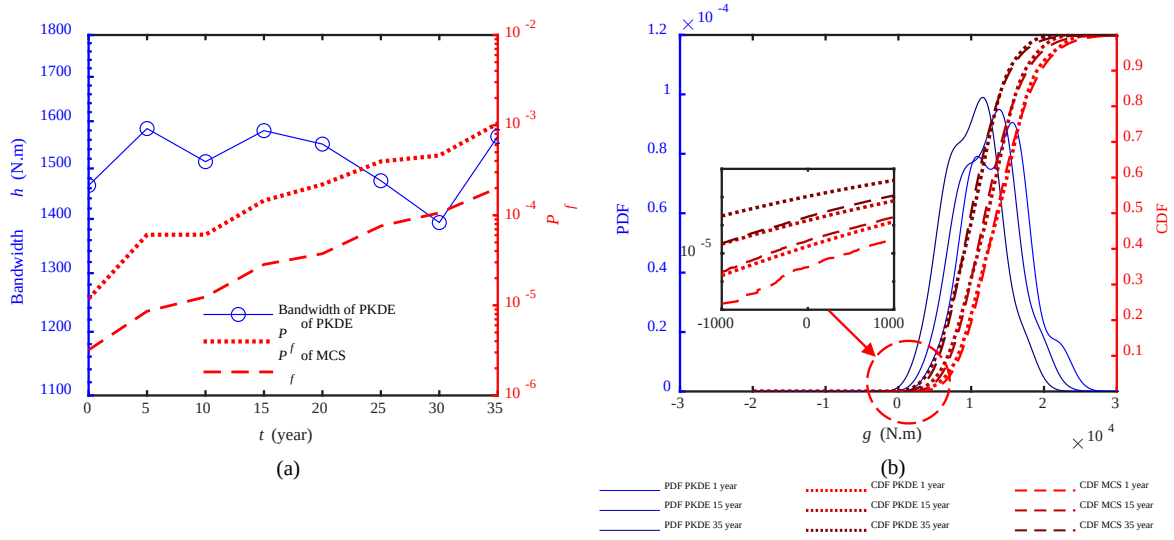


Fig. 10. (a) $h(t)$ and P_f of PKDE and MCS in Case 3; and (b) PDF and CDF estimation by PKDE and MCS in Case 3

By applying brute and adaptive SS methods, the CDFs at different time instants are computed once and summarized in Fig. 11. It can be noticed that most CDFs are consistent with 10^8 trials of MCS except for PKDE-GK-SS-EFF and LHS-GK-SS-EFF, which demonstrates the accuracy of SS methods in CDF estimation of rare time-dependent events. Besides, for the results using SS methods, the CDF curves of the failure zone are relatively small since the computation process stops once b_j is below or equal to zero.

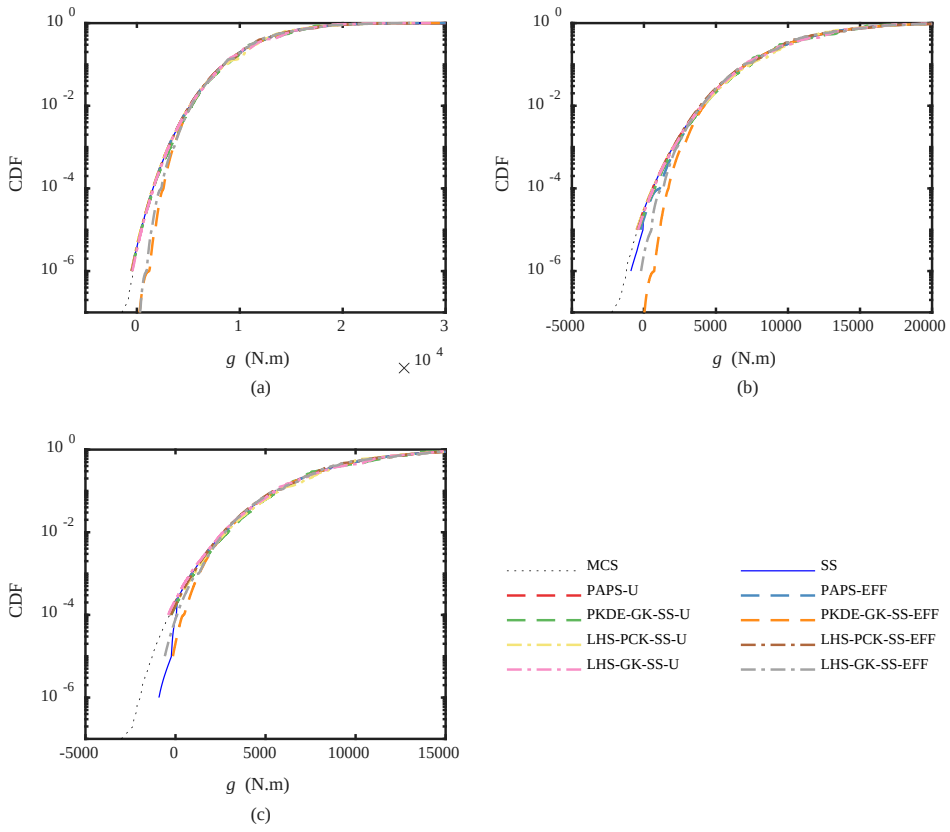


Fig. 11. Comparison of CDF among different methods under different time instants: (a) 1st year; (b) 15th year; and (c) 35th year

Furthermore, all methods are run 20 times to capture the average p_f and COV δ_f, at different time instants, as shown in Fig. 12. As indicated in Fig. 12a, $P_f(t)$ increases by around 77 times from t =1 to 35 years. Besides, after 20 times of runs, $P_f(t)$ of most methods agree with MCS except for PKDE-GK-SS-EFF and LHS-GK-SS-EFF. In addition, Fig. 12b shows that the N_{call} of all methods decrease versus time, where N_{call} of brute SS is the largest (ranging from 400,000 to 600,000, i.e., four to six levels of SS), and N_{call} of methods using EFF-learning function (ranging from 139 to 193) is much smaller than those using U-learning function (ranging from 282 to 1010). Also, in Fig. 12c, δ_f of LHS-GK-SS-U and LHS-GK-SS-EFF is the largest, while δ_f of the proposed PAPS (including PAPS-U and PAPS-EFF) is the smallest. Moreover, Fig. 12d illustrates that, except for the first and tenth years, the ε_f of PAPS-U is the smallest in most years (ranging from 0.09 to 1.8%). These results show that the proposed PAPS could capture more accurate time-dependent small failure probabilities than conventional SS methods with fewer samples. In addition, the robustness of PAPS using the U-learning function

is better than that with the EFF-learning function in this case.

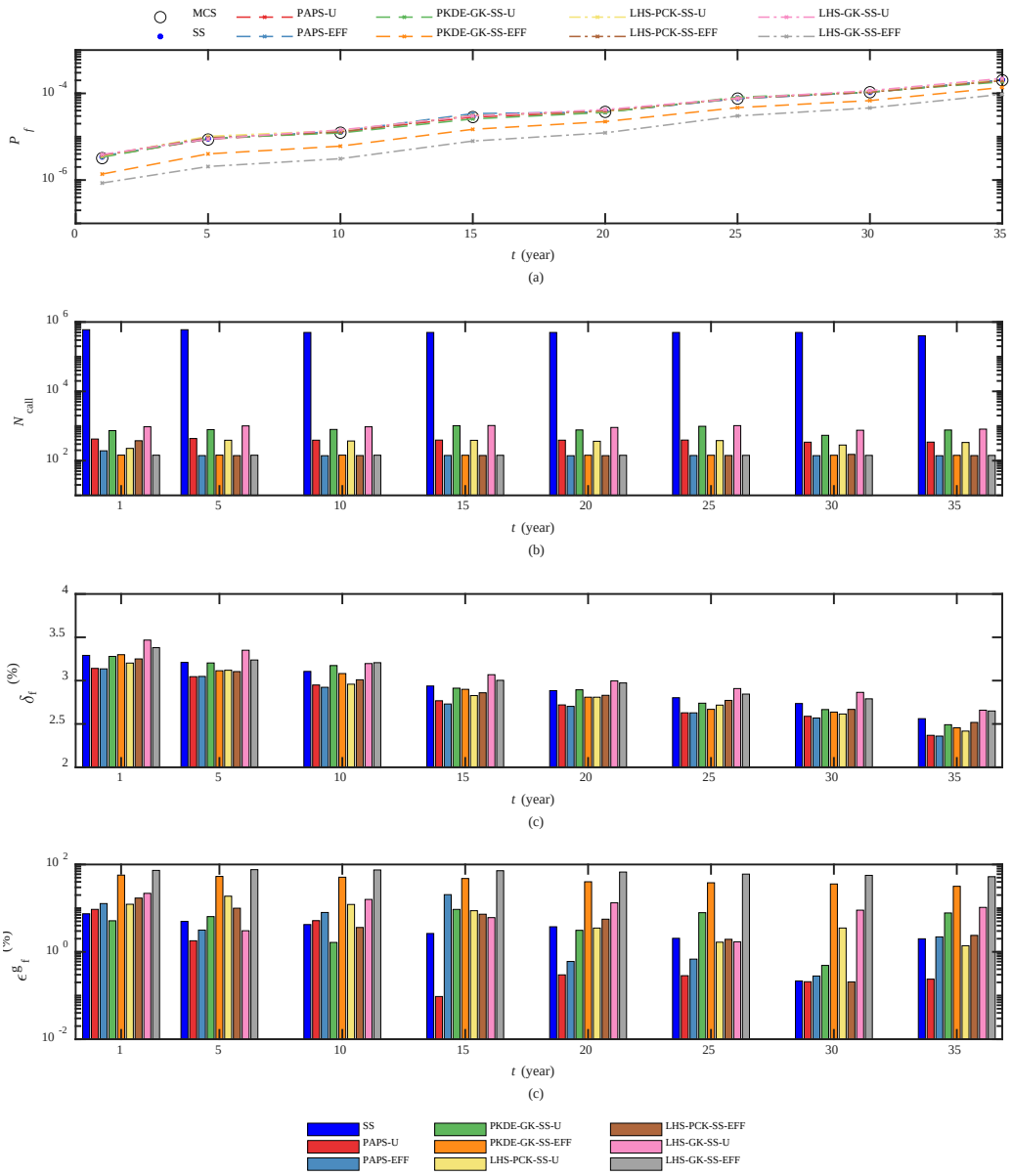


Fig. 12 Reliability analysis results associated with different methods for Case 3

Note: N_{call} and ε_f are the average number of runs of the original performance function and the relative error of the failure probability compared with MCS ($|p_f-p_{f,mcs}|/p_{f,mcs}\times100\%$), respectively.

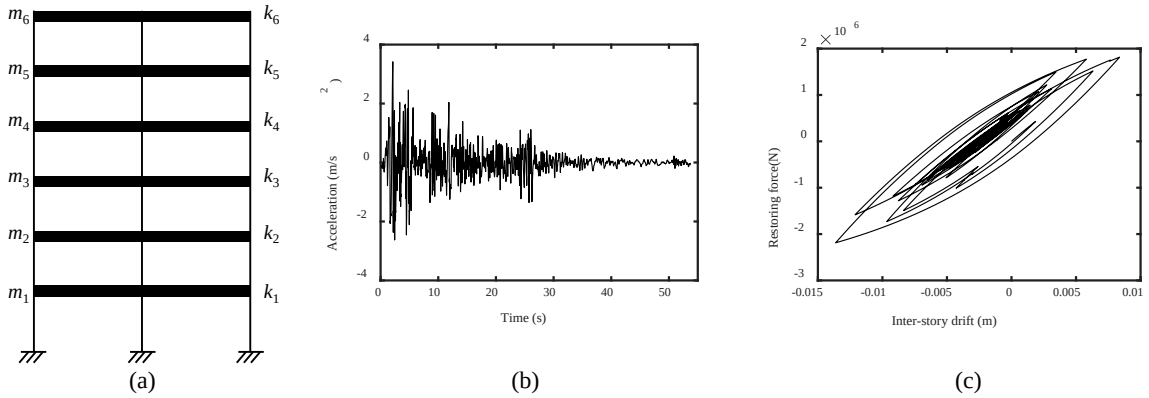
3.4. Case 4: A MDOF non-linear structure subjected to seismic excitation

As shown in Fig. 13a, a 6-story shear frame structure subject to the El Centro ground acceleration in the N-S direction (Fig. 13b) is considered in this case. The masses from bottom to top are denoted as m_1, m_2, ..., and m_6, and the initial lateral inter-story stiffnesses from bottom to top are k_1, k_2, …, k_6, where m_i and k_i (i=1, 2, …, 6) are regarded as independent random

variables, as listed in Table 5. To determine the seismic response of such a frame structure, the motion equation of a non-linear structure with multiple degrees of freedom (MDOF) is employed:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{g}\left(\mathbf{u}, \dot{\mathbf{u}}\right) = -\mathbf{M}\mathbf{I}\ddot{\mathbf{u}}_g \tag{38}$$

where **M** and **C** are the 6×6 matrices of mass and damping, respectively; **g** is a 6×1 vector of restoring force; and $\mathbf{u}, \dot{\mathbf{u}}, \text{and } \ddot{\mathbf{u}}$ are 6×1 vectors of displacement, velocity, and acceleration, respectively.



(a) (b) (c)

Fig. 13. Shear frame structure model: (a) schematic of the 6-story building structure; (b) acceleration record of EI Centro in N-S direction; and (c) typical restoring force versus inter-story drift

By employing Rayleigh damping, **C** is written as:

$$\mathbf{C} = a\mathbf{M} + b\mathbf{K} \tag{39}$$

where a = 0.2643 s^{-1}, b = 0.0071 s [57] and **K** is the 6×6 stiffness matrix. Also, the stiffness of each floor is modeled by the Bouc-Wen model [58], whose main 13 parameters are set as α = 0.01, A = 1, n = 1, q = 0, p = 600, ψ = 0.2, d_ψ = 0, λ = 0.5, ζ_s = 0.95, β = 60, γ = 10, d_v = 200, and d_η = 200 [59].

Table 5 Distribution parameters of different random variables

Variables	m_1	m_2	m_3	m_4	m_5	m_6
Distribution			Lognormal			
Mean (×10^5kg)	3.50	3.30	3.30	3.00	3.00	2.70
STD (×10^4kg)	7.00	6.60	6.60	6.00	6.00	5.40
Variables	k_1	k_2	k_3	k_4	k_5	k_6
Distribution			Lognormal			
Mean (×10^8kN/m)	3.24	4.71	4.71	4.71	4.71	5.15
STD (×10^7kN/m)	6.48	9.42	9.42	9.42	9.42	10.29

In this case, the main interesting output is the inter-story drift of the first floor u_1 whose typical hysteretic curve is illustrated in Fig. 13c. To begin with, 367 initial points are chosen through GFM. Then, based on the PKDE method, PDFs of u_1 at different time instant, i.e., $u_1(t)$, could be captured. Fig. 14 shows the obtained PDFs of $u_1(t)$ for a typical period of 5s to 7s. It can be seen that $u_1(t)$ is highly non-linear and stochastic and varies with time.

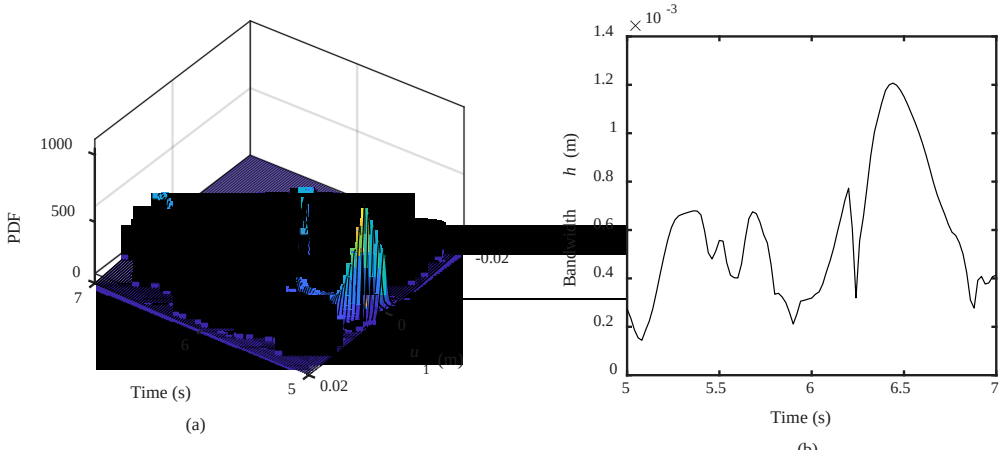


Fig. 14. PKDE results at typical time instants: (a) PDF surface; and (b) optimal bandwidths

Fig. 15 compares the mean and standard deviation of $u_1(t)$ based on 367 samples of PKDE and 100,000,000 trials of MCS, indicating that the PKDE has a good performance in evaluating the non-linear response.

Given the limit drift u_d=0.024 m, the performance function g is written as:

$$g = u_d - \max_{t\in[0,t_u]}\left[u_1(t)\right] \tag{40}$$

where t_u is the maximum time 53.74s.

By applying the PKDE and LHS-based KDEM, the PDF and CDF of g can be drawn, as illustrated in Fig. 16, which also plots a subgraph to magnify the tails of two CDFs. The CDF of g from PKDE is closer to MCS than LHS-based KDEM, but the subgraph of Fig. 16 shows the tails of the PDFs by PDKE and KDEM are not accurate.

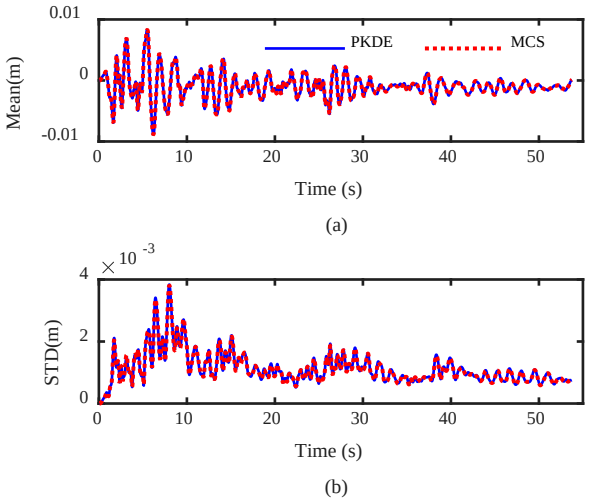


Fig. 15. Comparisons of the mean and standard deviation of $u_1(t)$ between PKDE and MCS

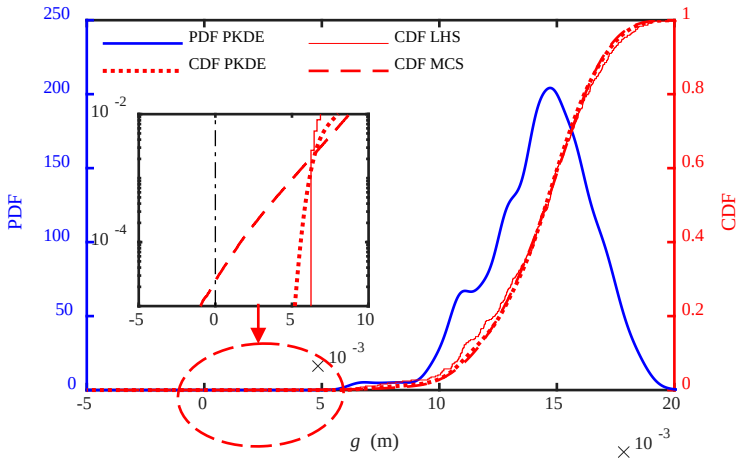


Fig. 16. PDF and CDF estimation by PKDE and MCS

Therefore, all SS-based methods are used to calculate the CDF of g, as illustrated in Fig. 17. It can be found that most CDFs are relatively close to each other in which LHS-PCK-SS-EFF deviates from MCS. Furthermore, Table 6 summarizes the reliability analysis results after 20 times of runs by all methods. In Table 6, 500,000 N_{call} of brute SS (five levels of SS with 100,000 samples for each level) is the closest to 10^8 trials of MCS (ε_{f} of 2.38%), followed by PAPS-U (ε_{f} of 3.68%) and PAPS-EFF (ε_{f} of 9.01%). Besides, all methods using PKDE have similar δ_{f} (from 2.84% to 2.86%). The above results demonstrate the accuracy of the proposed PAPS for small failure probability estimation of the non-linear dynamic structural response with a little computational cost.

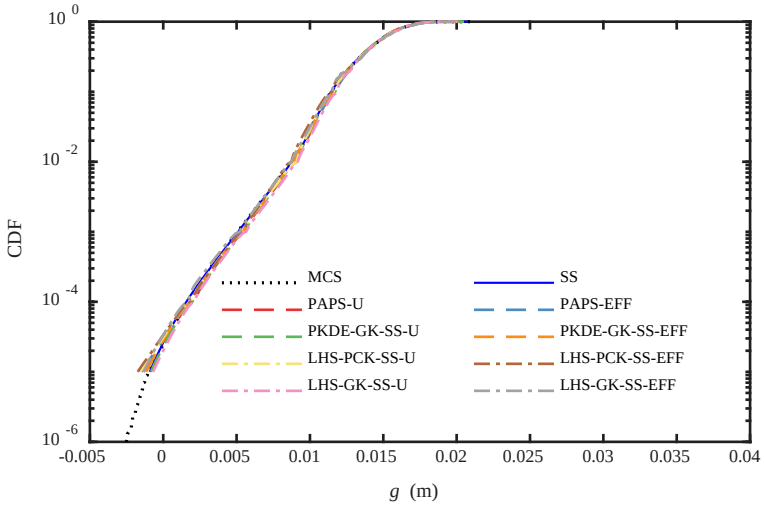


Fig. 17. Comparison of CDF among different methods

Table 6 Reliability analysis results associated with different methods for Case 4

Method	N_{call}	p_f	δ_f (%)	ε_f (%)
PKDE	367	1.868×10^{-10}	-	100.00%
MCS	10^8	2.533×10^{-5}	1.80%	-
SS	500,000	2.593×10^{-5}	3.02%	2.38%
PAPS-U	**867.2**	$\mathbf{2.626\times10^{-5}}$	**2.84%**	**3.68%**
PKDE-GK-SS-U	1380.7	2.292×10^{-5}	2.86%	9.50%
LHS-PCK-SS-U	868.5	2.839×10^{-5}	3.00%	12.08%
LHS-GK-SS-U	1421.7	2.005×10^{-5}	3.04%	20.85%
PAPS-EFF	876.0	2.761×10^{-5}	2.84%	9.01%
PKDE-GK-SS-EFF	1373.8	2.814×10^{-5}	2.84%	11.11%
LHS-PCK-SS-EFF	871.3	3.401×10^{-5}	2.96%	34.29%
LHS-GK-SS-EFF	1382.0	3.244×10^{-5}	2.98%	28.06%

Note: N_{call} and ε_f are the average number of runs of the original performance function and the relative error of the failure probability compared with MCS ($|p_f-p_{f,mcs}|/p_{f,mcs}\times100\%$), respectively.

4. Conclusions

This paper develops a novel adaptive method for the small-failure probability analysis that aims at computing the PDF and CDF of the performance function. Such a method integrates PKDE, PCK, SS, and the thought of adaptive MCS. The main idea of the proposed method is to use the CDF and representative points from PKDE as the initial CDF and initial training set, respectively, and then perform adaptive SS step-by-step to refine the CDF of performance. Four different cases of estimating small-failure probabilities are considered to show the benefits of

654 the proposed PAPS method. The four cases are two classical and straightforward analytical
655 cases, a time-variant case and a non-linear seismic response case. The following conclusions
656 could be drawn:
657 (1) The results of illustrative cases show that compared with conventional sampling strategy,
658 e.g., LHS, PKDE could generate uniform representative points and evaluate PDF. However,
659 existing PKDE may not be appropriate in PDF and CDF estimation under the scenarios of
660 rare failure events. Based on the results of PKDE, the proposed PAPS successfully refines
661 the CDF of performance function and provides accurate evaluation results with fewer
662 samples and smaller COV under most scenarios, compared with traditional methods, e.g.,
663 original SS, brute MCS, LHS, and GK based adaptive SS, verifying the feasibility and
664 accuracy of the proposed method; and
665 (2) The proposed PAPS requires only a small number of additional samples to refine the CDF
666 of the performance function. The number of additional samples relies on the complexity of
667 the studied case and chosen learning function. For simple and static cases, additional
668 samples are quite limited when the PCK model and EFF-learning function are adopted. For
669 the time-variant case, the number of additional samples depends on the investigated periods.
670 For the case of dynamic structural behavior, the number of additional samples becomes
671 more extensive due to the complexity of the case. Moreover, although the case studies show
672 that using the EFF-learning function might need fewer samples than using the U-learning
673 function, relative errors suggest that the U-learning function seems to be more robust than
674 the EFF-learning function.
675 In general, the proposed method efficiently estimates small-failure probabilities. In the
676 future, several directions need further investigation:
677 (1) Although the proposed PAPS is validated effectively and accurately in the presented
678 cases, the performance of PAPS is promising on high-dimensional issues, especially for
679 problems with thousands of random variables. Thus, in the future, to conduct high-dimensional
680 reliability analysis, a more advanced point selection method for the PKDE under high
681 dimensional issues should be developed, and dimensionality reduction-based advanced

surrogate models need to be proposed; and

(2) In the existing PAPS, the adaptive learning processes are driven by traditional learning functions (i.e., U-learning function and EFF-learning function) and surrogate models (i.e., PCK and GK model). To improve the performance of PAPS, more advanced learning functions and surrogate models should be developed and deeply integrated with PKDE.

(3) The existing PAPS may not be the optimal method to handle the small probability estimation considering multiple failure modes. Regarding the multi-failure modes and non-convex failure domains, it is still necessary to improve the efficiency of PKDE and PAPS in the future.

Acknowledgments

Funding: This study has been supported by the National Natural Science Foundation of China (Grant No. 52078448) and the Research Grants Council of the Hong Kong Special Administrative Region, China (No. PolyU 15219819).

References:

[1] Schuëller GI, Pradlwarter HJ, Koutsourelakis PS. A critical appraisal of reliability estimation procedures for high dimensions. Probabilistic Eng Mech 2004;19:463–74. https://doi.org/10.1016/j.probengmech.2004.05.004.

[2] Barone G, Frangopol DM. Reliability, risk and lifetime distributions as performance indicators for life-cycle maintenance of deteriorating structures. Reliab Eng Syst Saf 2014;123:21–37. https://doi.org/10.1016/j.ress.2013.09.013.

[3] Gardoni P. Risk and Reliability Analysis: Theory and Applications. Springer; 2017.

[4] Melchers RE, Beck AT. Structural reliability analysis and prediction. John Wiley & Sons; 2018.

[5] Frangopol DM, Lin K-Y, Estes AC. Reliability of reinforced concrete girders under corrosion attack. J Struct Eng 1997;123:286–97. https://doi.org/10.1061/(asce)0733-9445(1997)123:3(286).

[6] Bjerager P. Methods for structural reliability computations. Reliab. Probl. Gen. Princ. Appl. Mech. solids Struct., Springer; 1991, p. 89–135.

[7] Der Kiureghian A, Lin H, Hwang S. Second-Order Reliability Approximations. J Eng Mech 1987;113:1208–25. https://doi.org/10.1061/(asce)0733-9399(1987)113:8(1208).

[8] Ditlevsen O, Madsen HO. Structural reliability methods. vol. 178. New York: Wiley New York; 1996.

[9] Melchers RE. Importance sampling in structural systems. Struct Saf 1989;6:3–10.

https://doi.org/10.1016/0167-4730(89)90003-9.
[10] Au SK, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Eng Mech 2001;16:263–77. https://doi.org/10.1016/S0266-8920(01)00019-4.
[11] Bucher C. Asymptotic sampling for high-dimensional reliability analysis. Probabilistic Eng Mech 2009;24:504–10. https://doi.org/10.1016/j.probengmech.2009.03.002.
[12] Faravelli L. Response-surface approach for reliability analysis. J Eng Mech 1989;115:2763–81. https://doi.org/10.1061/(asce)0733-9399(1989)115:12(2763).
[13] Kioumarsi MM, Hendriks MAN, Geiker M. Interference of localised corrosion on adjacent reinforcement bars of a beam in bending. Concr Innov Conf 2014:Oslo, Norw.
[14] Kaymaz I. Application of kriging method to structural reliability problems. Struct Saf 2005;27:133–51. https://doi.org/10.1016/j.strusafe.2004.09.001.
[15] Zhu ZF, Du XP. Reliability analysis with monte carlo simulation and dependent kriging predictions. J Mech Des 2016;138:121403. https://doi.org/10.1115/1.4034219.
[16] Hurtado JE. An examination of methods for approximating implicit limit state functions from the viewpoint of statistical learning theory. Struct Saf 2004;26:271–93. https://doi.org/10.1016/j.strusafe.2003.05.002.
[17] Schöbi R, Sudret B, Wiart J. Polynomial-chaos-based Kriging. Int J Uncertain Quantif 2015;5:171–93. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2015012467.
[18] Schöbi R, Sudret B, Marelli S. Rare Event Estimation Using Polynomial-chaos Kriging. ASCE-ASME J Risk Uncertain Eng Syst Part A Civ Eng 2017;3:1–12. https://doi.org/10.1061/ajrua6.0000870.
[19] Echard B, Gayton N, Lemaire M. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Struct Saf 2011;33:145–54. https://doi.org/10.1016/j.strusafe.2011.01.002.
[20] Meng Z, Zhang Z, Li G, Zhang D. An active weight learning method for efficient reliability assessment with small failure probability. Struct Multidiscip Optim 2020;61:1157–70. https://doi.org/10.1007/s00158-019-02419-z.
[21] Shi Y, Lu ZZ, He RY, Zhou YC, Chen SY. A novel learning function based on Kriging for reliability analysis. Reliab Eng Syst Saf 2020;198:106857. https://doi.org/10.1016/j.ress.2020.106857.
[22] Echard B, Gayton N, Lemaire M, Relun N. A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. Reliab Eng Syst Saf 2013;111:232–40. https://doi.org/10.1016/j.ress.2012.10.008.
[23] Cadini F, Santos F, Zio E. An improved adaptive kriging-based importance technique for sampling multiple failure regions of low probability. Reliab Eng Syst Saf 2014;131:109–17. https://doi.org/10.1016/j.ress.2014.06.023.
[24] Huang XX, Chen JQ, Zhu HP. Assessing small failure probabilities by AK-SS: An active learning method combining Kriging and Subset Simulation. Struct Saf 2016;59:86–95. https://doi.org/10.1016/j.strusafe.2015.12.003.
[25] Ling C, Lu Z, Feng K, Zhang X. A coupled subset simulation and active learning kriging reliability analysis method for rare failure events. Struct Multidiscip Optim

2019;60:2325–41. https://doi.org/10.1007/s00158-019-02326-3.
[26] Wei P, Tang C, Yang Y. Structural reliability and reliability sensitivity analysis of extremely rare failure events by combining sampling and surrogate model methods. Proc Inst Mech Eng Part O J Risk Reliab 2019;233:943–57. https://doi.org/10.1177/1748006X19844666.
[27] Xu CL, Chen WD, Ma JX, Shi YQ, Lu SZ. AK-MSS: An adaptation of the AK-MCS method for small failure probabilities. Struct Saf 2020;86:101971. https://doi.org/10.1016/j.strusafe.2020.101971.
[28] Tong C, Sun Z, Zhao Q, Wang Q, Wang S. A hybrid algorithm for reliability analysis combining Kriging and subset simulation importance sampling. J Mech Sci Technol 2015;29:3183–93. https://doi.org/10.1007/s12206-015-0717-6.
[29] Hu YS, Lu ZZ, Wei N, Zhou CC. A single-loop Kriging surrogate model method by considering the first failure instant for time-dependent reliability analysis and safety lifetime analysis. Mech Syst Signal Process 2020;145:106963. https://doi.org/10.1016/j.ymssp.2020.106963.
[30] Shi Y, Lu ZZ, He RY. Advanced time-dependent reliability analysis based on adaptive sampling region with Kriging model. Proc Inst Mech Eng Part O J Risk Reliab 2020;234:588–600. https://doi.org/10.1177/1748006X20901981.
[31] Chen JB, Li J. Dynamic response and reliability analysis of non-linear stochastic structures. Probabilistic Eng Mech 2005;20:33–44. https://doi.org/10.1016/j.probengmech.2004.05.006.
[32] Guo HY, Dong Y, Gu XL. Two-step translation method for time-dependent reliability of structures subject to both continuous deterioration and sudden events. Eng Struct 2020;225:111291. https://doi.org/10.1016/j.engstruct.2020.111291.
[33] Silverman BW. Density estimation for statistics and data analysis. vol. 26. CRC press; 1986.
[34] Colbrook MJ, Botev ZI, Kuritz K, MacNamara S. Kernel density estimation with linked boundary conditions. Stud Appl Math 2020:1–40. https://doi.org/10.1111/sapm.12322.
[35] Jia G, Taflanidis AA, Beck JL. A new adaptive rejection sampling method using kernel density approximations and its application to subset simulation. ASCE-ASME J Risk Uncertain Eng Syst Part A Civ Eng 2017;3:D4015001. https://doi.org/10.1061/AJRUA6.0000841.
[36] Alibrandi U, Mosalam KM. Kernel density maximum entropy method with generalized moments for evaluating probability distributions, including tails, from a small sample of data. Int J Numer Methods Eng 2018;113:1904–28. https://doi.org/10.1002/nme.5725.
[37] Xu J, Kong F. A new unequal-weighted sampling method for efficient reliability analysis. Reliab Eng Syst Saf 2018;172:94–102. https://doi.org/10.1016/j.ress.2017.12.007.
[38] Guo HY, Dong Y, Gardoni P, Gu XL. Time-dependent reliability analysis based on point-evolution kernel density estimation: comprehensive approach with continuous and shock deterioration and maintenance. ASCE-ASME J Risk Uncertain Eng Syst

Part A Civ Eng 2021;7:04021032. https://doi.org/10.1061/ajrua6.0001153.
[39] Huang A, Botev ZI. Rare-event probability estimation via empirical likelihood maximization. ArXiv Prepr ArXiv13123027 2013.
[40] Li HS, Cao ZJ. MATLAB codes of Subset Simulation for reliability analysis and structural optimization. Struct Multidiscip Optim 2016;54:391–410. https://doi.org/10.1007/s00158-016-1414-5.
[41] Xu J, Dang C. A novel fractional moments-based maximum entropy method for high-dimensional reliability analysis. Appl Math Model 2019;75:749–68. https://doi.org/10.1016/j.apm.2019.06.037.
[42] Zhang X, Low YM, Koh CG. Maximum entropy distribution with fractional moments for reliability analysis. Struct Saf 2020;83:101904. https://doi.org/10.1016/j.strusafe.2019.101904.
[43] Chen JB, Zhang SH. Improving point selection in cubature by a new discrepancy. SIAM J Sci Comput 2013;35:A2121–49. https://doi.org/10.1137/12089377X.
[44] Chen JB, Yang JY, Li J. A GF-discrepancy for point selection in stochastic seismic response analysis of structures with uncertain parameters. Struct Saf 2016;59:20–31. https://doi.org/10.1016/j.strusafe.2015.11.001.
[45] Alibrandi U, Ricciardi G. Efficient evaluation of the pdf of a random variable through the kernel density maximum entropy approach. Int J Numer Methods Eng 2008;75:1511–48. https://doi.org/10.1002/nme.
[46] Rawa MJH, Thomas DWP, Sumner M. Kernel density estimation and its application. Proceeding Int Conf Electr Power Qual Util EPQU 2011;00037:102–7. https://doi.org/10.1109/EPQU.2011.6128915.
[47] Botev ZI, Grotowski JF, Kroese DP. Kernel density estimation via diffusion. Ann Stat 2010;38:2916–57. https://doi.org/10.1214/10-AOS799.
[48] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Saf 2003;81:23–69. https://doi.org/10.1016/S0951-8320(03)00058-9.
[49] Papaioannou I, Betz W, Zwirglmaier K, Straub D. MCMC algorithms for Subset Simulation. Probabilistic Eng Mech 2015;41:89–103. https://doi.org/10.1016/j.probengmech.2015.06.006.
[50] Zuev KM, Beck JL, Au SK, Katafygiotis LS. Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions. Comput Struct 2012;92–93:283–96. https://doi.org/10.1016/j.compstruc.2011.10.017.
[51] Yu Z, Sun Z, Cao R, Wang J, Yan Y. RCA-PCK: A new structural reliability analysis method based on PC-Kriging and radial centralized adaptive sampling strategy. Proc Inst Mech Eng Part C J Mech Eng Sci 2021;235:3424–38. https://doi.org/10.1177/0954406220957711.
[52] Moustapha M, Marelli S, Sudret B. A generalized framework for active learning reliability : survey and benchmark n.d.:1–50.
[53] Rasmussen CE, Williams CK. Gaussian processes for machine learning. vol. 2. MIT press Cambridge, MA; 2006.

[54] Marelli S, Sudret B. UQLab user manual – Polynomial Chaos Expansions. Rep UQLab-V09-104, Chair Risk, Saf Uncertain Quantif ETH Zurich, 2015:1–52.
[55] Bichon BJ, Eldred MS, Swiler LP, Mahadevan S, McFarland JM. Efficient global reliability analysis for nonlinear implicit performance functions. AIAA J 2008;46:2459–68. https://doi.org/10.2514/1.34321.
[56] Hu Z, Du X. Mixed efficient global optimization for time-dependent reliability analysis. J Mech Des Trans ASME 2015;137:1–9. https://doi.org/10.1115/1.4029520.
[57] Xu J, Chen J, Li J. Probability density evolution analysis of engineering structures via cubature points. Comput Mech 2012;50:135–56. https://doi.org/10.1007/s00466-011-0678-2.
[58] Ma F, Zhang H, Bockstedte A, Foliente GC, Paevere P. Parameter analysis of the differential model of hysteresis. J Appl Mech Trans ASME 2004;71:342–9. https://doi.org/10.1115/1.1668082.
[59] Li J, Chen JB. Probability density evolution method in stochastic dynamics. Encycl Earthq Eng 2014:1–14. https://doi.org/10.1007/978-3-642-36197-5_333-1.
[60] Li J, Chen JB. The number theoretical method in response analysis of nonlinear stochastic structures. Comput Mech 2007;39:693–708. https://doi.org/10.1007/s00466-006-0054-9.
[61] Radović I, Sobol IM, Tichy RF. Quasi-Monte Carlo methods for numerical integration: Comparison of different low discrepancy sequences. Monte Carlo Methods Appl 1996;2:1–14. https://doi.org/10.1515/mcma.1996.2.1.1.
[62] Conway JH, Sloane NJA. Sphere Packings, Lattices and Groups. vol. 290. 1999. https://doi.org/10.1007/978-1-4757-6568-7.
[63] Jiang Z, Li J. High dimensional structural reliability with dimension reduction. Struct Saf 2017;69:35–46. https://doi.org/10.1016/j.strusafe.2017.07.007.

Appendix:

A.1. GF-discrepancy based point selection method (GFM)

In this section, the approach of obtaining uniform point set $\boldsymbol{x}$ including n_{sel} d-dimensional representative points $x_{i,j}$ (i = 1, 2, … , n_{sel}; j = 1, 2, … , d) is introduced. To begin with, a uniformly distributed point set $\boldsymbol{\theta}\in[0, 1]$ is captured by the dimension of $\boldsymbol{x}$ applying a low-discrepancy sequence e.g., number-theoretical method (NTM) based [60] or Sobol based sequence [61]. In terms of the distribution type of each random variable in the vector $\mathbf{X}$, $x_{i,j}$ can be computed by

$$x_{i,j} = F_j^{-1}\left(\theta_{i,j}\right), i = 1,2,...,n_{\text{sel}},\ j = 1,2,...,d \tag{A1.1}$$

where $F_j^{-1}(\cdot)$ is the inverse cumulative distribution function (CDF) of the j-th random variable.

However, the uniformity of the point set $\boldsymbol{x}$ obtained by Eq.(A1.1) is generally unsatisfying except for a huge n_{sel} [44]. To reduce the uniformity of $\boldsymbol{x}$, a GF-discrepancy (i.e., Eq.(A1.2)) based refine strategy is employed, and then the refined representative point $x'_{i,j}$ can be obtained byEq.(A1.3) [44].

$$D_{\text{GF}}(\boldsymbol{x}) = \max_{1\le j\le d}\left\{\sup\left|F_{n_{sel},j}(\boldsymbol{x}) - F_j(\boldsymbol{x})\right|\right\} \tag{A1.2}$$

where $F_{n_{sel},j}(\cdot)$ and $F_j(\cdot)$ are the empirical and actual CDF of the j-th random variable, respectively.

$$x'_{i,j} = F_j^{-1}\left\{\sum_{k=1}^{n_{\text{sel}}} p_{\text{a},k}\cdot I\left\{\theta_{k,j} < \theta_{i,j}\right\} + 0.5\,p_{\text{a},i}\right\} \tag{A1.3}$$

where $I\{\cdot\}$ is an indicator function equals one once the condition in the bracket is satisfied; $p_{\text{a},I}$ is the assigned probability of the i-th representative point, which is written as Eq. (A1.4).

$$p_{\text{a},i} = P\left\{\boldsymbol{X}\in\Omega_i\right\} = \int_{V_i} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, i = 1,2,...,n_{\text{sel}} \tag{A1.4}$$

in which $f_{\boldsymbol{X}}(\boldsymbol{x})$ is the joint PDF of random variables in $\boldsymbol{X}$; and V_i is the Voronoi volume of the i-th sub-domain [62].

Existing studies proved that by employing Eqs. (A1.1)-(A1.4), a uniform and low GF-discrepancy point set could be captured and used in point-evolution-based PDFM [32,44,63].

A.2. Diffusion-based bandwidth selection method

In the diffusion-based bandwidth selection method, the second derivative function of $f_G(g)$, which is denoted as f'', is supposed to be a continuous square-integrable function and the square of bandwidth $t = h^2$, the optimal t is the minimum value of the first-order asymptotic approximation of MISE:

$$_*t = \left(2M\sqrt{\pi}\left\|f''\right\|^2\right)^{-0.4} \tag{A2.1}$$

Considering the $\left\|f^j\right\|^2$ for arbitrary integer $j \geq 1$, there exist two estimators $(-1)^j \mathbb{E}_f\left[f^{2j}(g)\right]$ and $\left\|\widehat{f^j}\right\|^2$:

$$(-1)^j \mathbb{E}_f\left[f^{2j}(g)\right] = \frac{(-1)^j}{M^2}\sum_{k=1}^{M}\sum_{n=1}^{M} K^{2j}(g_k, g_m, \sqrt{t_j}) \tag{A2.2}$$

$$\left\|\widehat{f^j}\right\|^2 = \frac{(-1)^j}{M^2}\sum_{k=1}^{M}\sum_{n=1}^{M} K^{2j}(g_k, g_m, \sqrt{2t_j}) \tag{A2.3}$$

Let estimator $(-1)^j \mathbb{E}_f\left[f^{2j}(g)\right]$ and $\left\|\widehat{f^j}\right\|^2$ have the same asymptotic mean square error, the square of bandwidth $_*t_j$ can be evaluated as [47]

$$_*t_j = \left(\frac{1+1/2^{j+1/2}}{3}\frac{1\times 3\times 5\cdots\times(2j-1)}{N\sqrt{\pi/2}\left\|f^{j+1}\right\|^2}\right)^{2/(3+2j)} \tag{A2.4}$$

where $\left\|f^{j+1}\right\|^2$ is unknown and can be estimated by $\left\|\widehat{f^{j+1}}\right\|^2$ which needs the evaluation of $_*\hat{t}_{j+1}$ (Eq.(A2.3)). Furthermore, $_*t$ requires estimating a sequence $\left\{_*\hat{t}_{j+k}, k \geq 1\right\}$. Denoting the relationship between $_*\hat{t}_j$ and $_*\hat{t}_{j+1}$ as $_*\hat{t}_j = \gamma_j\left(_*\hat{t}_{j+1}\right)$, $_*t$ can be evaluated by:

$$_*\hat{t} = \xi\gamma^{[l]}\left(_*\hat{t}_{l+1}\right), \xi = \left(\frac{6\sqrt{2}-3}{7}\right)^{2/5} \approx 0.90, l > 0 \tag{A2.5}$$

$$\gamma^{[l]}(g) = \gamma_1\left(\gamma_2\left(\cdots\gamma_{l-1}\left(\gamma_l(g)\right)\right)\right), l \geq 1 \tag{A2.6}$$

Moreover, the optimal t can be calculated by solving the non-linear equation Eq.(A2.5) through fixed-point iteration or the Newtonian method [38,47].

A.3. Polynomial chaos kriging (PCK)

The idea of PCK is to model a set of orthogonal polynomials and stochastic processes related to the input variables [17]:

$$G(\boldsymbol{X}) \approx G^{\mathrm{PCK}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in \mathrm{A}} \beta_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) + \sigma^2 \cdot Z(\boldsymbol{X}, \omega) \tag{A3.1}$$

in which $\psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$ are multivariate orthonormal polynomials $\psi_{\boldsymbol{\tau}}(\boldsymbol{X})$ indexed by a multi-indices $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_{N_{dim}}\}$; $\beta_{\boldsymbol{\alpha}}$ are the corresponding coefficients; and σ^2 and $Z(\boldsymbol{X},\omega)$ are the variances of $G^{\mathrm{PCK}}(\boldsymbol{X})$ stationary Gaussian process of zero-mean and unit variance described through a correlation function and hyperparameters $\boldsymbol{\theta}$, respectively.

The $\psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$ in Eq.(A3.1) consists of the tensor product of univariate orthogonal polynomials

$$\psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) = \prod_{i=1}^{N_{\mathrm{dim}}} \psi_{\alpha_i}^{i}(x_i) \tag{A3.2}$$

where $\psi_{\alpha i}{}^{i}(x_i)$ is the univariate polynomial of degree α_i in the i-th variable.

For each variable X_i, $i = 1, 2, \ldots, d$, an orthonormal polynomial basis can be written as:

$$\left\langle \psi_j^i, \psi_k^i \right\rangle = \int_{D_i} \psi_j^i(x) \psi_k^i(x)\, f_{X_i}(x)\, dx = \delta_{jk} \tag{A3.3}$$

where $\psi_j{}^i$ and $\psi_k{}^i$ are two candidate univariate polynomials for the i-th variable; D_i is the support of the distribution of X_i; f_{Xi} is marginal PDF; and δ_{jk} is the Kronecker symbol ($\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ if $j \neq k$).

To build the PCK model, the main problem is to determine the optimal polynomial set in the trend and the calibration of the Kriging model, which can be achieved through the methods of sequential PCK [17] and optimal PCK [18]. In the sequential PCK, the optimal polynomial set is determined by least-angle regression selection (LARS), whereas in the optimal PCK, the PCK model is iteratively computed, and the optimal PCK is selected by minimizing the leave-one-out (LOO) error. In this study, the sequential PCK is employed for the consideration of computational efficiency.