Parallel Bayesian probabilistic integration for structural reliability analysis with small failure probabilities

Zhuo Hu[a,b], Chao Dang[b,*], Lei Wang[a,*], Michael Beer[b,c,d]

[a]School of Civil Engineering, Changsha University of Science & Technology, Changsha 410114, PR China
[b]Institute for Risk and Reliability, Leibniz University Hannover, Callinstr. 34, Hannover 30167, Germany
[c]Institute for Risk and Uncertainty, University of Liverpool, Liverpool L69 7ZF, United Kingdom
[d]Department of Civil Engineering, Tsinghua University, Beijing 100084, PR China

Abstract

Bayesian active learning methods have emerged for structural reliability analysis and shown more attractive features than existing active learning methods. However, it remains a challenge to actively learn the failure probability by fully exploiting its posterior statistics. In this study, a novel Bayesian active learning method termed 'Parallel Bayesian Probabilistic Integration' (PBPI) is proposed for structural reliability analysis, especially when involving small failure probabilities. A pseudo posterior variance of the failure probability is first heuristically proposed for providing a pragmatic uncertainty measure over the failure probability. The variance amplified importance sampling is modified in a sequential manner to allow the estimations of posterior mean and pseudo posterior variance with a large sample population. A learning function derived from the pseudo posterior variance and a stopping criterion associated with the pseudo posterior coefficient of variance of the failure probability are then presented to enable active learning. In addition, a new adaptive multi-point selection method is developed to identify multiple sample points at each iteration without the need to predefine the number, thereby allowing parallel computing. The effectiveness of the proposed PBPI method is verified by investigating four numerical examples, including a turbine blade structural model and a transmission tower structure. Results indicate that the proposed method is capable of estimating small failure probabilities with superior accuracy and efficiency over several other existing active learning reliability methods.

Keywords: Bayesian probabilistic integration; Small failure probability; Gaussian process; Bayesian active learning; Importance sampling; Parallel computing

1. Introduction

A major task of structural reliability analysis is to compute the failure probability in the presence of various uncertainties, which may arise from external loads, material properties, and environmental factors, etc. The uncertainties are represented by a d-dimensional random vector $\boldsymbol{X} = [X_1, X_2, ..., X_d]$ with known joint probability density function (PDF) $f_{\boldsymbol{X}}(\boldsymbol{x})$. The failure probability is generally formulated as the d-dimensional integral:

$$P_f = \Pr\{g(\boldsymbol{X}) < 0\} = \int_{\mathbb{R}^d} I(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{1}$$

where $\Pr\{\cdot\}$ is the probability operation; $g(\boldsymbol{X})$ is the performance function (a.k.a. limit state function); $\boldsymbol{x}$ denotes a realization of $\boldsymbol{X}$; $I(\boldsymbol{x})$ is the indicator function: $I(\boldsymbol{x}) = 1$ if $g(\boldsymbol{x}) < 0$ and $I(\boldsymbol{x}) = 0$ otherwise.

In the past decades, various methods have been developed to approximate the intractable integral in Eq. (1), which can be roughly divided into four categories. The first category is the simulation methods, including the Monte Carlo simulation (MCS) and its variants, e.g., importance sampling (IS) [1, 2], subset simulation [3], line sampling [4] and directional sampling [5], etc. The second category is the analytical approximation methods such as the well known first-order and second-order reliability methods (FORM and SORM) [6, 7]. The third category consist of the methods of moments, for instance, integer moments-based methods [8, 9] and fractional moments-based methods [10, 11]. The fourth category is the surrogate assisted methods. Some commonly used surrogate models in reliability analysis include response surface methods [12, 13], polynomial chaos expansion [14, 15], support vector machines [16, 17], artificial neural networks [18, 19], and Kriging (a.k.a. Gaussian process regression, i.e., GPR) [20, 21].

Among the various surrogates, Kriging-based methods have received much attention due to its interpolative and probabilistic properties. The desirable features promote the development of Kriging in combination with the active learning strategies. The earlier proposed efficient global reliability analysis (EGRA) [22] and the adaptive Kriging Monte Carlo simulation (AK-MCS) [23] are two prominent examples. These active learning Kriging methods start from constructing an initial Design of Experiment (DoE), and then pro-

*Corresponding authors
Email addresses: chao.dang@irz.uni-hannover.de (Chao Dang), leiwang@csust.edu.cn (Lei Wang)

gressively add new sample points into the initial DoE until a predefined stopping criterion is fulfilled. An essential component in this respect is the so-called learning function, which provides a meaningful guidance for selecting the best points to evaluate the performance function. Various learning functions have been developed from different perspectives. Except for U function in AK-MCS and expected feasibility function (EFF) in EGRA, some other representative learning functions consist of least improvement function (LIF) [24], reliability-based expected improvement function (REIF) [25], folded normal based expected improvement function (FNEIF) [26], H function [27], potential risk function (PRF) [28], reliability-based lower confidence bounding (RLCB) function [29], expected integrated error reduction (EIER) [30] and so forth. Another critical component while designing an active learning algorithm is the stopping criterion, which is used to terminate the learning process at an appropriate stage. Many existing researches directly prescribe a threshold on the learning function as the stopping criterion, e.g., $\min(\mathrm{U}) > 2$ in AK-MCS. Some other stopping criteria have also been developed by judging the accuracy of failure probability, such as the error-based stopping criterion (ESC) [31], and cumulative confidence level (CCL) measure [32], etc. One can refer to [33, 34] for a comprehensive literature review. Despite great efforts, the performance of the existing active learning algorithms can still be further improved in terms of accuracy, efficiency, and applicability.

More recently, the failure probability integral (i.e., Eq. (1)) has been interpreted from a Bayesian probabilistic integration perspective [35–37]. In a Bayesian viewpoint, the numerical uncertainty induced from limited observations on performance function is regarded as a kind of epistemic uncertainty. This uncertainty propagates through the indicator function and in turn propagates into the failure probability. A probabilistic uncertainty measure over failure probability can thus be derived, which allows to develop the two critical components of active learning algorithm (i.e., learning function and stopping criterion). The resulting methodology is the so-called Bayesian active leaning method. The ability for providing the uncertainty measure over the failure probability makes the Bayesian active learning method more advantageous than the existing active learning reliability methods. In [35], a method, called active learning probabilistic integration (ALPI), is proposed, where an upper-bound posterior variance of the failure probability is given. Based on the conceptual framework of ALPI, a parallel adaptive Bayesian quadrature (PABQ) method is developed, which allows to estimate small failure probabilities and support parallel computing [36]. The exact expression of posterior variance of failure probability is then derived in a Bayesian failure probability inference framework [37]. A parallel adaptive-Bayesian failure probability learning (PA-BFPL) method is developed within this framework [37]. These methods can provide uncertainty measure over failure probability, among which, the two parallel methods (i.e., PABQ and PA-BFPL) can be applied to small failure probability problems. However, they still possess respective limitations. In ALPI and PABQ, the strict upper-bound posterior variance largely overestimates the true posterior variance of failure probability, making it difficult to prescribe a reasonable threshold in the stopping criterion to truly reflect the uncertainty level of failure probability. As for PA-BFPL, the numerical computation of exact posterior variance is very time-consuming when entailing a large amount of samples. In addition, both PABQ and PA-BFPL identify multiple points using weighted k-means clustering algorithm, in which an important parameter, i.e., the number of added points at each iteration, needs to be empirically specified. This algorithm would decrease the number of iterations but sacrifice the number of performance function calls if one specifies a large k.

In order to overcome the issues above, a new Bayesian active learning method, termed 'Parallel Bayesian Probabilistic Integration' (PBPI) is developed in this study for efficient structural reliability analysis, especially for estimating small failure probabilities. Specifically, a pseudo posterior variance (PPV) of the failure probability is first heuristically proposed under a Gaussian process prior over the performance function, thereby providing a simple and pragmatic uncertainty measure over the failure probability. The variance amplified importance sampling (VAIS) developed in [37] is modified as sequential sampling to estimate the posterior mean and PPV of failure probability. A learning function derived from the PPV and a stopping criterion associated with the pseudo posterior coefficient of variation (COV) of failure probability are then presented to enable active learning. Moreover, a novel adaptive multi-point selection method is proposed based on the identification of local maxima of learning function, which allows parallel computing without the need to predefine the number of added points as required by k-means algorithm.

The remaining of this work is organized as follows. Section 2 briefly reviews two existing methods (i.e., PABQ and PA-BFPL) that are closely related to our development. Section 3 presents the proposed PBPI method in detail. Four numerical examples characterized by small failure probabilities are then investigated in Section 4 to demonstrate the performance of the proposed method. Conclusions are finally drawn in Section 5.

2. Brief review of related methods

This section briefly reviews the two adaptive Bayesian quadrature methods (i.e., PABQ and PA-BFPL), which are closely related to our development. Besides, some discussions about the PABQ and PA-BFPL are given.

2.1. Adaptive Bayesian quadrature for failure probability estimation

The problem of estimating the intractable failure probability integral in Eq. (1) is interpreted from a perspective of Bayesian quadrature (a.k.a. integration or cubature) in PA-BFPL [37] and PABQ [36]. Specifically, the performance function $g(\cdot)$ is regarded as random, that is, the value $g(\boldsymbol{x})$ at a given site $\boldsymbol{x}$ is uncertain before it is evaluated. The discretization error as an epistemic uncertainty arises herein since evaluating $g(\cdot)$ at every point is impractical. Following a standard Bayesian approach, both methods thus start at putting a prior on the performance function $g(\cdot)$ and combining it with a dataset $\boldsymbol{\mathcal{D}}$ that consists of some observations of the g-function. The posterior mean and variance of the failure probability are then derived. Additional informative observations are identified using a so-called learning function to enrich the dataset $\boldsymbol{\mathcal{D}}$ through successive iterations until a stopping criterion is fulfilled.

2.1.1. Bayesian inference of failure probability

A Gaussian process (GP) prior is first placed over the performance function $g(\cdot)$, which is written as:

$$g_0 \sim \mathcal{GP}\left(m_{g_0}(\boldsymbol{x}), k_{g_0}\left(\boldsymbol{x}, \boldsymbol{x}'\right)\right) \tag{2}$$

where g_0 denotes the prior distribution of $g(\cdot)$ prior to seeing any observations; $m_{g_0}(\boldsymbol{x})$ and $k_{g_0}\left(\boldsymbol{x}, \boldsymbol{x}'\right)$ are the prior mean and covariance functions respectively. The mean function $m_{g_0}(\boldsymbol{x})$ takes the constant type (i.e., $m_{g_0}(\boldsymbol{x})=\beta$), and the covariance function adopts the widely used squared exponential kernel function:

$$k_{g_0}\left(\boldsymbol{x}, \boldsymbol{x}^{\prime}\right)=\sigma^2 \exp \left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}^{\prime}\right) \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{x}-\boldsymbol{x}^{\prime}\right)^{\top}\right) \tag{3}$$

where σ^2 is the process variance; $\boldsymbol{\Sigma}=\operatorname{diag}\left(l_1^2, l_1^2, \ldots, l_d^2\right)$ is a diagonal matrix with $l_i>0$ being the length scale in the i-dimension.

Conditioning on n observations to constitute the dataset $\boldsymbol{\mathcal{D}}=\{\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}\}$ ($\boldsymbol{\mathcal{X}}$ is a $d \times n$ matrix with i-th column being observation $\boldsymbol{x}^{(i)}$ and $\boldsymbol{\mathcal{Y}}$ is a $n \times 1$ vector with i-th row being $g(\boldsymbol{x}^{(i)})$), the $d+2$ hyper-parameters $\boldsymbol{\theta}=[\beta, \sigma, l_1, l_2, \ldots l_d]$ can be estimated by minimizing the negative log marginal likelihood:

$$L(\boldsymbol{\theta})=-\log [p(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{X}}, \boldsymbol{\theta})]=\frac{1}{2}(\boldsymbol{\mathcal{Y}}-\beta)^{\top} \boldsymbol{K}_{g_0}^{-1}(\boldsymbol{\mathcal{Y}}-\beta)+\frac{1}{2} \log \left|\boldsymbol{K}_{g_0}^{-1}\right|+\frac{n}{2} \log (2 \pi) \tag{4}$$

where $\boldsymbol{K}_{g_0}$ is the $n \times n$ covariance matrix with (i, j)-th element $\left[\boldsymbol{K}_{g_0}\right]_{i, j}=k_{g_0}\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right)$.

Then, the posterior distribution of $g(\cdot)$ can be obtained as:

$$g_n \sim \mathcal{GP}\left(m_{g_n}(\boldsymbol{x}), k_{g_n}\left(\boldsymbol{x}, \boldsymbol{x}^{\prime}\right)\right) \tag{5}$$

where g_n denotes the posterior distribution of $g(\cdot)$ conditional on n observations; $m_{g_n}(\boldsymbol{x})$ and $k_{g_n}\left(\boldsymbol{x}, \boldsymbol{x}^{\prime}\right)$ are the posterior mean and covariance functions respectively, which can be analytically derived as:

$$m_{g_n}(\boldsymbol{x})=m_{g_0}(\boldsymbol{x})+\boldsymbol{k}_{g_0}(\boldsymbol{x}, \boldsymbol{\mathcal{X}})^{\top} \boldsymbol{K}_{g_0}^{-1}\left(\boldsymbol{\mathcal{Y}}-\boldsymbol{m}_{g_0}(\boldsymbol{\mathcal{X}})\right) \tag{6}$$

$$k_{g_n}\left(\boldsymbol{x}, \boldsymbol{x}^{\prime}\right)=k_{g_0}\left(\boldsymbol{x}, \boldsymbol{x}^{\prime}\right)-\boldsymbol{k}_{g_0}(\boldsymbol{x}, \boldsymbol{\mathcal{X}})^{\top} \boldsymbol{K}_{g_0}^{-1} \boldsymbol{k}_{g_0}\left(\boldsymbol{x}^{\prime}, \boldsymbol{\mathcal{X}}\right) \tag{7}$$

where $\boldsymbol{k}_{g_0}(\boldsymbol{x}, \boldsymbol{\mathcal{X}})$ and $\boldsymbol{k}_{g_0}(\boldsymbol{x}^{\prime}, \boldsymbol{\mathcal{X}})$ are two $n \times 1$ covariance vectors with i-th element being $k_{g_0}\left(\boldsymbol{x}, \boldsymbol{x}^{(i)}\right)$ and $k_{g_0}\left(\boldsymbol{x}^{\prime}, \boldsymbol{x}^{(i)}\right)$, respectively; $\boldsymbol{m}_{g_0}(\boldsymbol{\mathcal{X}})$ is an $n \times 1$ mean vector with i-th element being $m_{g_0}(\boldsymbol{x}^{(i)})$.

The posterior distribution $g_n(\boldsymbol{x})$ will imply the posterior distribution of indicator function I (denoted as I_n), and then imply the posterior distribution of failure probability P_f (denoted as $P_{f,n}$). The posterior mean and exact posterior variance of P_f adopted in PA-BFPL are written as [37]:

$$m_{P_{f,n}} = \int_{\mathcal{X}} \Phi\left(-\frac{m_{g_n}(\boldsymbol{x})}{\sigma_{g_n}(\boldsymbol{x})}\right) f_{\boldsymbol{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{8}$$

$$\sigma^2_{P_{f,n}} = \int_{\mathcal{X}}\int_{\mathcal{X}} F\left([0\ 0];\left[m_{g_n}(\boldsymbol{x}), m_{g_n}(\boldsymbol{x}')\right], \boldsymbol{K}_{g_n}(\boldsymbol{x},\boldsymbol{x}')\right) f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}' - m^2_{P_{f,n}} \tag{9}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of standard normal variable; $\sigma_{g_n}(\boldsymbol{x}) = \sqrt{k_{g_n}(\boldsymbol{x},\boldsymbol{x})}$ is the posterior standard deviation of $g(\cdot)$; $f_{\boldsymbol{X}}(\boldsymbol{x})$ and $f_{\boldsymbol{X}}(\boldsymbol{x}')$ are the joint PDF of $\boldsymbol{X}$ and $\boldsymbol{X}'$, respectively; F is the joint CDF of a bivariate normal distribution; $\boldsymbol{K}_{g_n}(\boldsymbol{x},\boldsymbol{x}') = \begin{bmatrix} \sigma^2_{g_n}(\boldsymbol{x}) & k_{g_n}(\boldsymbol{x},\boldsymbol{x}') \\ k_{g_n}(\boldsymbol{x}',\boldsymbol{x}) & \sigma^2_{g_n}(\boldsymbol{x}') \end{bmatrix}$ is the posterior covariance matrix of $g(\cdot)$.

An upper-bound posterior variance (UPV) is derived in PABQ according to Cauchy-Schwarz inequality, which is expressed as [36]:

$$\sigma^2_{P_{f,n}} \leq \bar{\sigma}^2_{P_{f,n}} = \left(\int_{\mathcal{X}} \sqrt{\Phi\left(-\frac{m_{g_n}(\boldsymbol{x})}{\sigma_{g_n}(\boldsymbol{x})}\right)\Phi\left(\frac{m_{g_n}(\boldsymbol{x})}{\sigma_{g_n}(\boldsymbol{x})}\right)} f_{\boldsymbol{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right)^2 \tag{10}$$

where $\bar{\sigma}_{P_{f,n}}$ is the upper-bound posterior standard deviation.

Note that numerical integration techniques are necessary to estimate $m_{P_{f,n}}$, $\sigma_{P_{f,n}}$ and $\bar{\sigma}_{P_{f,n}}$ due to the analytical intractability. The VAIS method and a importance ball sampling method are respectively developed in PA-BFPL [37] and PABQ [36] to approximate the integrals.

2.1.2. Multi-point selection strategy and stopping criterion

In PA-BFPL, an expected misclassification probability contribution (EMPC) function is developed to identify new points and enrich the dataset $\boldsymbol{\mathcal{D}}$. As for PABQ, the aforementioned UPV $\bar{\sigma}^2_{P_{f,n}}$ is utilized to derive a learning function called upper-bound posterior variance contribution (UPVC), which is defined as:

$$\mathrm{UPVC}(\boldsymbol{x}) = \sqrt{\Phi\left(-\frac{m_{g_n}(\boldsymbol{x})}{\sigma_{g_n}(\boldsymbol{x})}\right)\Phi\left(\frac{m_{g_n}(\boldsymbol{x})}{\sigma_{g_n}(\boldsymbol{x})}\right)} \times f_{\boldsymbol{X}}(\boldsymbol{x}) \tag{11}$$

The EMPC function and UPVC function are respectively combined with k-means clustering algorithm in PA-BFPL and PABQ to select multiple points at each iteration, thereby enabling parallel computing. Note that the number of clusters k, which corresponds the number of identified points at each iteration, should be predefined.

In order to terminate the active learning process, the stopping criteria in PA-BFPL and PABQ are constructed by judging the exact posterior COV and upper-bound posterior COV of failure probability, respectively. The two stopping criteria are given as:

$$\text{PA-BFPL}: \quad \text{COV}_{P_{f,n}} = \frac{\sigma_{P_{f,n}}}{m_{P_{f,n}}} < \epsilon_T \tag{12}$$

$$\text{PABQ}: \quad \overline{\text{COV}}_{P_{f,n}} = \frac{\bar{\sigma}_{P_{f,n}}}{m_{P_{f,n}}} < \epsilon_U \tag{13}$$

where ϵ_T and ϵ_U are user-specified thresholds (0.05 and 0.1 suggested in PA-BFPL and PABQ, respectively).

2.2. Discussions on PABQ and PA-BFPL

The two adaptive Bayesian quadrature methods, PABQ and PA-BFPL, adopt the posterior mean and their respective posterior variance expressions (i.e., exact posterior variance and UPV) to represent the numerical uncertainty of failure probability arising from limited observations on g-function. Both methods can assess small failure probabilities without excessively large amount of samples and allow parallel computing to decrease the number of iterations. However, several drawbacks still exist in both methods, mainly lying in their respective posterior variance expressions and multi-point selection strategies.

The posterior variance promotes the development of learning function and stopping criterion. However, the numerical estimation of exact posterior variance $\sigma^2_{P_{f,n}}$ is very time-consuming in PA-BFPL, especially when involving large amount of samples. This is mainly due to the fact that the costly computation of bivariate normal CDF should be performed at each iteration. The expensive computation is also the reason why EMPC function is utilized in PA-BFPL instead of that directly based on the exact posterior variance contribution. Although the upper-bound posterior variance $\bar{\sigma}^2_{P_{f,n}}$ in PABQ can alleviate the computational difficulties, it still has two main drawbacks. First, the equality in Inequality (10) holds only when $I_n(\boldsymbol{x})$ and

$I_n(\boldsymbol{x}')$ are perfectly positively correlated for any $\boldsymbol{x}, \boldsymbol{x}' \in \boldsymbol{\mathcal{X}}$, which is hardly impractical. Second, the strict upper bound considerably overestimate the posterior variance, making it difficult to specify a reasonable threshold ϵ_U in the stopping criterion (i.e., Eq. (13)) to truly reflect the uncertainty level of posterior failure probability.

When it comes to the multi-point selection strategy, the learning function weighted k-means clustering algorithms presented in PABQ and PA-BFPL require to specify the number of added points at each iteration. With such a technique, some not necessarily optimal points, which contribute little to the convergence of active learning, are also evaluated, leading to the increase of number of performance function calls.

3. Parallel Bayesian probabilistic integration

This section presents a novel method termed PBPI for small failure probability estimation. Specifically, a PPV of the failure probability is first heuristically proposed to approximate the true posterior variance. The VAIS is then introduced and modified in a sequential way to numerically approximate the posterior mean and PPV with a large sample population. A stopping criterion and a learning function are presented according to the posterior statistics of failure probability. Finally, an adaptive multi-point selection method is proposed by identifying local maximum points of learning function. A set of points can thus be selected to enable parallel distributed processing, eliminating the the necessity of predefining the number of points as required by the k-means algorithm in PABQ and PA-BFPL.

The proposed PBPI method is defined in the standard normal space ($\boldsymbol{U}$ space), which can be formulated through an isoprobabilistic transformation $\boldsymbol{u} = T(\boldsymbol{x})$ (e.g., Nataf or Rosenblatt transformation). The transformed performance function is written as $\mathcal{G}(\boldsymbol{u}) = g(T^{-1}(\boldsymbol{x}))$.

3.1. Proposed pseudo posterior variance

The posterior variance of failure probability is significantly meaningful for constructing efficient learning function and stopping criterion. The exact posterior variance in Eq. (9), however, is very expensive to evaluate when involving a large number of samples. The UPV $\bar{\sigma}^2_{P_{f,n}}$ in Eq. (10), as a strict upper bound, greatly overestimates the true posterior variance. Note that $\Phi\left(-\frac{m_{\mathcal{G}_n}(\boldsymbol{u})}{\sigma_{\mathcal{G}_n}(\boldsymbol{u})}\right)$ in Eq. (10) represents the probability (denoted as p) of GP prediction less than zero at the point $\boldsymbol{u}$, i.e.,

$$p(\boldsymbol{u}) = \Phi\left(-\frac{m_{\mathcal{G}_n}(\boldsymbol{u})}{\sigma_{\mathcal{G}_n}(\boldsymbol{u})}\right) = 1 - \Phi\left(\frac{m_{\mathcal{G}_n}(\boldsymbol{u})}{\sigma_{\mathcal{G}_n}(\boldsymbol{u})}\right) \leq 1 \tag{14}$$

Hence $[p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))]^{\frac{1}{2}} \leq 1$ holds. Inspired by this aspect, a PPV is heuristically proposed by introducing a parameter α greater than 1 into the UPV(i.e., $[p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))]^{\frac{\alpha}{2}}$), thereby narrowing the UPV and further approximating the true posterior variance. The PPV is expressed as:

$$\hat{\sigma}^2_{P_{f,n}} = \left(\int_{\mathcal{U}} [p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))]^{\frac{\alpha}{2}} f_{\boldsymbol{U}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}\right)^2 \tag{15}$$

where $\hat{\sigma}_{P_{f,n}}$ is the pseudo posterior standard deviation.

Obviously, the PPV $\hat{\sigma}^2_{P_{f,n}}$ is smaller than or equal to UPV $\bar{\sigma}^2_{P_{f,n}}$, which is expressed as:

$$\hat{\sigma}^2_{P_{f,n}} = \left(\int_{\mathcal{U}} [p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))]^{\frac{\alpha}{2}} f_{\boldsymbol{U}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}\right)^2 \leq \bar{\sigma}^2_{P_{f,n}} = \left(\int_{\mathcal{U}} [p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))]^{\frac{1}{2}} f_{\boldsymbol{U}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}\right)^2 \tag{16}$$

The PPV $\hat{\sigma}^2_{P_{f,n}}$ decreases with the increase of α. If one specifies a very large value for α, the PPV would approach to zero and greatly underestimate the true posterior variance. In contrast, the PPV would degenerate to the UPV when α approaches to 1. The optimal value of α is therefore between 1 and positive infinity and can be calculated by setting the PPV equal to the exact posterior variance, i.e.,

$$\begin{aligned}
\text{find} \quad \alpha \quad \text{s.t.} \quad \hat{\sigma}^2_{P_{f,n}} &= \left(\int_{\mathcal{U}} [p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))]^{\frac{\alpha}{2}} f_{\boldsymbol{U}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}\right)^2 = \sigma^2_{P_{f,n}} \\
&= \int_{\mathcal{U}} \int_{\mathcal{U}} F\left([0\ 0]; [m_{\mathcal{G}_n}(\boldsymbol{u}), m_{\mathcal{G}_n}(\boldsymbol{u}')], \boldsymbol{K}_{\mathcal{G}_n}(\boldsymbol{u}, \boldsymbol{u}')\right) f_{\boldsymbol{U}}(\boldsymbol{u}) f_{\boldsymbol{U}}(\boldsymbol{u}') \mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{u}' - m^2_{P_{f,n}}
\end{aligned} \tag{17}$$

However, it is quite hard to theoretically derive an optimal value for α in PPV due to the difficulty in solving the exact posterior variance $\sigma^2_{P_{f,n}}$ that involves the expensive computation of the bivariate normal CDF. Compared with $\sigma^2_{P_{f,n}}$, the proposed PPV $\hat{\sigma}^2_{P_{f,n}}$ greatly simplifies the expression and computation of the exact posterior variance, thereby providing a simple and pragmatic uncertainty measure of failure probability. As an alternative to the theoretical derivation, a parameter analysis will be conducted by specifying different values for α in the numerical examples, and then a reasonable value will be suggested (see Section 4 for details). Note that although we have not theoretically derived an optimal value for α, the proposed PPV is also very important for the development of Bayesian active learning method as it facilitates us to construct the simple and efficient learning function and stopping criterion.

As a compromise between the exact posterior variance $\sigma^2_{P_{f,n}}$ and the UPV $\bar{\sigma}^2_{P_{f,n}}$, the proposed PPV $\hat{\sigma}^2_{P_{f,n}}$ with a reasonable setting of α mainly has two advantages. First, the PPV with a reasonable α allows to more realistically represent the true posterior variance of failure probability, as compared with the UPV. Second, the proposed PPV avoids the cumbersome evaluation of bivariate normal CDF in exact posterior variance (i.e., Eq. (9)), thereby significantly saving the computational time especially when large amount of samples are needed to evaluate $\sigma^2_{P_{f,n}}$.

3.2. Sequential variance-amplified importance sampling

In order to numerically approximate the analytically intractable integrals (i.e., posterior mean $m_{P_{f,n}}$ in Eq. (8) and PPV $\hat{\sigma}^2_{P_{f,n}}$ in Eq. (15)), MCS is the most straightforward method. However, most of the generated samples are distributed near the peak of the joint PDF $f_{\boldsymbol{U}}(\boldsymbol{u})$. Excessively large amount of samples are required for accurate estimation of the integrals in some cases (e.g., small failure probability). With regard to importance sampling, the optimal sampling density is infeasible in practice as it involves the quantity to be computed. As a simple but efficient alternative, the VAIS technique developed in PA-BFPL [37] can produce more dispersedly distributed samples than MCS and facilitate the small failure probability estimation. However, the original VAIS faces a problem of computer memory when involving a very large sample population to accurately approximate the integrals. In this paper, the VAIS is modified in a sequential manner, forming the sequential VAIS technique. This modification not only avoids the memory problem caused by one-shot GP prediction of a large sample population, but also eliminates the need to pre-specify the total sample size and improves the computational efficiency, as compared with original VAIS.

The $m_{P_{f,n}}$ in Eq. (8) and $\hat{\sigma}_{P_{f,n}}$ in Eq. (15) are first rewritten as:

$$m_{P_{f,n}} = \int_{\mathcal{U}} p(\boldsymbol{u}) \frac{f_{\boldsymbol{U}}(\boldsymbol{u})}{h_{\boldsymbol{U}}(\boldsymbol{u})} h_{\boldsymbol{U}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} \tag{18}$$

$$\hat{\sigma}_{P_{f,n}} = \int_{\mathcal{U}} \left[p(\boldsymbol{u}) \cdot (1 - p(\boldsymbol{u}))\right]^{\frac{\alpha}{2}} \frac{f_{\boldsymbol{U}}(\boldsymbol{u})}{h_{\boldsymbol{U}}(\boldsymbol{u})} h_{\boldsymbol{U}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} \tag{19}$$

where $h_{\boldsymbol{U}}(\boldsymbol{u})$ is the importance sampling density (ISD). The ISD $h_{\boldsymbol{U}}(\boldsymbol{u})$ is constructed by amplifying the standard deviation $\sigma_{\boldsymbol{U}}$ (or equivalently amplifying the variance $\sigma_{\boldsymbol{U}}^2$) of $f_{\boldsymbol{U}}(\boldsymbol{u})$. The constructed ISD is thus formulated as $h_{\boldsymbol{U}}(\boldsymbol{u}) = f_{\boldsymbol{U}}\left(\boldsymbol{u}; \boldsymbol{0}, \gamma \cdot \boldsymbol{I}\right)$, where $\gamma > 1$ is the amplification coefficient of standard deviation; $\boldsymbol{I}$ is a $d \times d$ identity matrix.

The estimators of $m_{P_{f,n}}$ and $\hat{\sigma}_{P_{f,n}}$ with N_{vas} samples generated from $h_{\boldsymbol{U}}(\boldsymbol{u})$ are expressed as:

$$\tilde{m}_{P_{f,n}} = \frac{1}{N_{vas}} \sum_{i=1}^{N_{vas}} \left[p(\boldsymbol{u}^{(i)}) \frac{f_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)}{h_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)} \right] \tag{20}$$

$$\tilde{\hat{\sigma}}_{P_{f,n}} = \frac{1}{N_{vas}} \sum_{i=1}^{N_{vas}} \left[p(\boldsymbol{u}^{(i)}) \cdot \left(1 - p(\boldsymbol{u}^{(i)})\right) \right]^{\frac{\alpha}{2}} \frac{f_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)}{h_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)} \tag{21}$$

The variances of the above estimators are given as:

$$\mathbb{V}\left[\tilde{m}_{P_{f,n}}\right] = \frac{1}{N_{vas} - 1} \left(\frac{1}{N_{vas}} \sum_{i=1}^{N_{vas}} \left[p(\boldsymbol{u}^{(i)}) \frac{f_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)}{h_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)} \right]^2 - \tilde{m}_{P_{f,n}}^2 \right) \tag{22}$$

$$\mathbb{V}\left[\tilde{\hat{\sigma}}_{P_{f,n}}\right] = \frac{1}{N_{vas} - 1} \left(\frac{1}{N_{vas}} \sum_{i=1}^{N_{vas}} \left[p(\boldsymbol{u}^{(i)}) \cdot \left(1 - p(\boldsymbol{u}^{(i)})\right) \right]^{\alpha} \left[\frac{f_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)}{h_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)} \right]^2 - \tilde{\hat{\sigma}}_{P_{f,n}}^2 \right) \tag{23}$$

Note that in Eqs. (20)-(23) $p(\boldsymbol{u}^{(i)}) = \Phi\left(-\frac{m_{\mathcal{G}_n}(\boldsymbol{u}^{(i)})}{\sigma_{\mathcal{G}_n}(\boldsymbol{u}^{(i)})}\right)$ are simultaneously utilized for calculating $m_{P_{f,n}}$, $\hat{\sigma}_{P_{f,n}}$ and their variances, thus one only need to calculate the $p(\boldsymbol{u}^{(i)})$ once, avoiding the time-consuming recalculations of $\Phi(\cdot)$ for N_{vas} samples.

The samples are sequentially generated from the ISD $h_{\boldsymbol{U}}(\boldsymbol{u})$ and predicted using the GP model to further save computational time and facilitate the GP prediction with a large sample population. First,

N_{vas} samples are generated and let $j = 1$. The posterior mean and pseudo posterior standard deviation are then estimated by Eqs. (20) and (21), denoted as $m^{(j)}$ and $\sigma^{(j)}$, respectively. To reserve the GP prediction information for calculating the variances of estimators, let $s_1^{(j)}$ and $s_2^{(j)}$ be respectively expressed as:

$$s_1^{(j)} = \sum_{i=1}^{N_{vas}} \left[p(\boldsymbol{u}^{(i)}) \frac{f_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)}{h_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)} \right]^2 \tag{24}$$

$$s_2^{(j)} = \sum_{i=1}^{N_{vas}} \left[p(\boldsymbol{u}^{(i)}) \cdot \left(1 - p(\boldsymbol{u}^{(i)})\right) \right]^\alpha \left[\frac{f_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)}{h_{\boldsymbol{U}}\left(\boldsymbol{u}^{(i)}\right)} \right]^2 \tag{25}$$

Additional N_{vas} samples are generated from $h_{\boldsymbol{U}}(\boldsymbol{u})$ and let $j = j + 1$. The $m^{(j)}$, $\sigma^{(j)}$, $s_1^{(j)}$ and $s_2^{(j)}$ are calculated with the new N_{vas} generated samples by Eqs. (20)-(21) and Eqs. (24)-(25). The estimators of $m_{P_{f,n}}$ and $\hat{\sigma}_{P_{f,n}}$ and the corresponding variances in Eqs. (20)-(23) are reformulated as:

$$\tilde{m}_{P_{f,n}} = \frac{1}{j} \sum_{i=1}^{j} m^{(i)} \tag{26}$$

$$\tilde{\hat{\sigma}}_{P_{f,n}} = \frac{1}{j} \sum_{i=1}^{j} \sigma^{(i)} \tag{27}$$

$$\mathbb{V}\left[\tilde{m}_{P_{f,n}}\right] = \frac{1}{j \cdot N_{vas} - 1} \left(\frac{1}{j \cdot N_{vas}} \sum_{i=1}^{j} s_1^{(i)} - \tilde{m}_{P_{f,n}}^2 \right) \tag{28}$$

$$\mathbb{V}\left[\tilde{\hat{\sigma}}_{P_{f,n}}\right] = \frac{1}{j \cdot N_{vas} - 1} \left(\frac{1}{j \cdot N_{vas}} \sum_{i=1}^{j} s_2^{(i)} - \tilde{\hat{\sigma}}_{P_{f,n}}^2 \right) \tag{29}$$

The sequential sampling process is repeated until the target COVs of the posterior mean $\tilde{m}_{P_{f,n}}$ and pseudo posterior standard deviation $\tilde{\hat{\sigma}}_{P_{f,n}}$ are below the corresponding specified thresholds, that is, $\mathrm{COV}(\tilde{m}_{P_{f,n}}) < \epsilon_\mu$ and $\mathrm{COV}(\tilde{\hat{\sigma}}_{P_{f,n}}) < \epsilon_{\hat{\sigma}}$.

3.3. Stopping criterion and adaptive multi-point selection

Once the posterior mean $\tilde{m}_{P_{f,n}}$ and pseudo posterior standard deviation $\tilde{\hat{\sigma}}_{P_{f,n}}$ of the failure probability are numerically estimated, a reasonable stopping criterion is needed to judge whether $\tilde{m}_{P_{f,n}}$ is accurate enough as the finial failure probability. The stopping criterion can be naturally defined based on the judgment of the pseudo posterior COV of failure probability:

$$\widehat{\text{COV}} = \frac{\tilde{\hat{\sigma}}_{P_{f,n}}}{\tilde{m}_{P_{f,n}}} < \epsilon_p \tag{30}$$

where ϵ_p is a user-specified threshold. In order to avoid the possible fake convergence in the first few iterations, the active learning is terminated only when Eq. (30) is satisfied twice in succession.

If the stopping criterion is not satisfied, additional informative observations should be identified to enrich the dataset $\boldsymbol{\mathcal{D}}$. The key to achieve this aim is to develop a suitable learning function that provides a useful guidance to add new points. Based on the PPV in Eq. (15), a learning function called pseudo posterior variance contribution (PPVC) is defined as:

$$\text{PPVC}(\boldsymbol{u}) = \left[\Phi\left(-\frac{m_{\mathcal{G}_n}(\boldsymbol{u})}{\sigma_{\mathcal{G}_n}(\boldsymbol{u})}\right) \cdot \Phi\left(\frac{m_{\mathcal{G}_n}(\boldsymbol{u})}{\sigma_{\mathcal{G}_n}(\boldsymbol{u})}\right)\right]^{\frac{\alpha}{2}} \times f_{\boldsymbol{U}}(\boldsymbol{u}) \tag{31}$$

where $\text{PPVC}(\boldsymbol{u})$ measures the contribution of numerical uncertainty at point $\boldsymbol{u}$ to the PPV that equals to $\hat{\sigma}^2_{P_{f,n}} = \left(\int_{\mathcal{U}} \text{PPVC}(\boldsymbol{u})\mathrm{d}\boldsymbol{u}\right)^2$.

As the most convenient way, the best next point can be identified by maximizing the PPVC function. However, selecting a single point at each iteration would result in the underuse of the information provided by learning function and hinder the use of parallel computing facilities. In order to identify multiple points at each iteration, two aspects need to be considered. First, the points should be selected based on the learning function value, e.g., the points with large PPVC value. Second, the selected points in a certain iteration cannot be too clustered. Considering these aspects, some existing researches adopt learning function-based k-means clustering algorithm to identify a batch of points [36–40]. Nevertheless, a main limitation of the k-means algorithm is the need to specify the number of clusters k which corresponds to the number of selected points at each iteration. Although a large k can reduce the number of iterations, more performance function calls are required, causing the unnecessary waste of computing resource.

The learning function PPVC is generally multi-modal during the iteration process. The local peaks of PPVC function correspond to the points with local maximum contribution of the numerical uncertainty to the PPV of failure probability. These points possess relatively large PPVC values and are not too close, which simultaneously satisfy the two previously mentioned aspects. In order to identify the local maxima of PPVC function, a novel adaptive multi-point selection method is thus developed to enable parallel computing without the need to predefine the number of added points at each iteration.

Quasi-Newton method, as an alternative to Newton method, can approximate the computationally costly Hessian matrix and efficiently search the local minima (or maxima) of functions. The widely used Broyden Fletcher Goldfarb Shanno (BFGS) quasi-Newton algorithm is employed in the present study. Note that one of the main limitations of quasi-Newton method is its sensitivity to initial point. Thus, we first generate n_q uniform points $\boldsymbol{U}_q = \{\boldsymbol{u}^{(i)}\}_1^{n_q}$ within a d-ball as the initial points to identify multiple local peaks of PPVC function. Given an arbitrary point $\boldsymbol{u}_k \in \boldsymbol{U}_q$, the BFGS quasi-Newton algorithm is executed. The $fminunc$ function in Optimization Toolbox of Matlab is utilized for implementing the BFGS algorithm, with which we start at the point $\boldsymbol{u}_k$ and attempt to find a local minima of the objective function $\mathcal{F}(\boldsymbol{u})$. The objective function $\mathcal{F}(\boldsymbol{u})$ is defined as the negative learning function, i.e., $-\text{PPVC}(\boldsymbol{u})$. It should be noted that the learning function values in most regions are extremely small (typically, $10^{-3} \sim 10^{-10}$ or smaller) and the corresponding gradients are very close to 0. In order to efficiently find the local maxima of PPVC function, the objective function $\mathcal{F}(\boldsymbol{u})$ is re-formulated as $\mathcal{F}(\boldsymbol{u}) = -log\left(\text{PPVC}(\boldsymbol{u}) + eps\right)$. eps is a very small value and introduced herein to avoid the antilogarithm being zero.

The initial points in $\boldsymbol{U}_q$ gradually converge to the local peaks of PPVC function using BFGS quasi-Newton algorithm. After all initial points are converged, two additional aspects should be considered. First, some local maximum points with relatively low PPVC values are also identified, which contribute little to the convergence of active learning and are undesired to be selected. Second, some initial points would converge to the same local maximum point of the PPVC function. These converged points are very close but different in numerical values due to the recursive approximation nature of BFGS algorithm. Among the converged points near a true local peak, only one point is desired to be selected to enrich the dataset $\boldsymbol{\mathcal{D}}$ and evaluate on performance function. For the former aspect, we first eliminate those points with extremely small PPVC values after all initial points are converged. For the latter one, a DBSCAN clustering algorithm [41], which does not require to predefine the number of clusters, is introduced to cluster the points near one true local peak. A batch of points consisting of the point with the largest PPVC value in each cluster are selected as the final points to evaluate the $\mathcal{G}$-function and update the GP model.

The procedures for implementing the proposed adaptive multi-point selection method are summarized as below and schematically illustrated in Fig. 1.

Step I: Generate uniform initial points $\boldsymbol{U}_q = \{\boldsymbol{u}^{(i)}\}_1^{n_q}$ within a d-ball of radius R for searching local peaks. The radius R is determined as $R = \sqrt{\chi_d^{-2}(1 - p_{f,0})}$. The number of initial points $n_q = 400$ and the parameter $p_{f,0} = 10^{-8}$ are adopted in the present study.

Step II: Search the local peaks of the learning function PPVC($\boldsymbol{u}$) with BFGS quasi-Newton algorithm.

Step III: Obtain the maximum PPVC value M_p in all local peaks, and then eliminate the points with their PPVC values less than $\rho \cdot M_p$ ($\rho = 0.01$ is adopted).

Step IV: Divide the remaining points into q_m clusters with DBSCAN algorithm, identify q_m points which consists of the point with largest PPVC value at each cluster, enrich the dataset $\boldsymbol{\mathcal{D}}$ with the q_m points and their corresponding $\mathcal{G}$-function evaluations.

Note that two parameters are involved in the DBSCAN clustering algorithm, i.e., measure of distance ξ and minimum number of points in a cluster $minPts$. $minPts$ is generally determined by adopting a rule of thumb. As for ξ, an adaptive scheme is developed in [42] by selecting ξ as the minimum value that minimises the number of outliers and does not compromise the definition of separate clusters. However, this scheme would increase the total computational time and are therefore not used herein. Through some numerical tests, $\xi = 0.2\sqrt{d}$ and $minPts = 3$ are adopted for convenience in this study.

3.4. Implementation of the proposed PBPI method

The implementation procedure of the proposed PBPI method is summarized as follows (see Fig. 2 for the flowchart):

Step 1: Generate the initial observations

Generate N_0 samples $\boldsymbol{\mathcal{U}} = \{\boldsymbol{u}^{(i)}\}_{i=1}^{N_0}$ within a d-ball of radius R as the initial observations. These observations are evaluated on the performance function $\mathcal{G}(\cdot)$ to obtain the corresponding responses $\boldsymbol{\mathcal{Y}} =$

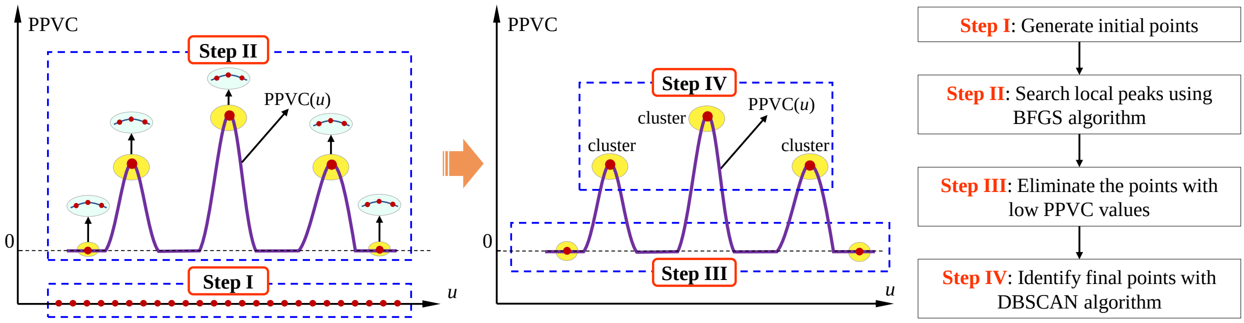


Figure 1: A schematic illustration of the proposed adaptive multi-point selection method.

$\{\boldsymbol{y}^{(i)}\}_{i=1}^{N_0}$. Construct the initial dataset $\boldsymbol{\mathcal{D}} = \{\boldsymbol{\mathcal{U}}, \boldsymbol{\mathcal{Y}}\}$. Let the number of $\mathcal{G}$-function calls $N_{call} = N_0$ and $m = 1$.

Step 2: Make Bayesian inference about the $\mathcal{G}$-function

By assigning a GP prior for the $\mathcal{G}$-function, the posterior distribution of $\mathcal{G}$-function is inferred based on dataset $\boldsymbol{\mathcal{D}}$. The prior mean and covariance function are assumed to be a constant type and squared exponential kernel, respectively (see Section 2.1 for details). In this paper, the *fitrgp* function in Statistics and Machine Learning Toolbox of Matlab is utilized for this purpose.

Step 3: Sequential VAIS for estimating posterior mean and PPV of failure probability

Initialize the parameter $j = 1$;

Step 3.1: Generate N_{vas} samples from the ISD $h_{\boldsymbol{U}}(\boldsymbol{u})$, and compute the corresponding GP predictions.

Step 3.2: Estimate the $m^{(j)}$, $\sigma^{(j)}$, $s_1^{(j)}$ and $s_2^{(j)}$ based on Eqs. (20)-(21) and Eqs. (24)-(25) respectively.

Step 3.3: Compute the COVs of posterior mean and pseudo posterior standard deviation based on Eqs. (26)-(29). If $\mathrm{COV}(\tilde{m}_{P_{f,n}}) < \epsilon_\mu$ and $\mathrm{COV}(\tilde{\tilde{\sigma}}_{P_{f,n}}) < \epsilon_{\tilde{\sigma}}$ are fulfilled, then the sequential sampling process is finished; else, return to Step 3.1 and let $j = j + 1$.

Step 4: Check the stopping criterion

If $\widehat{\mathrm{COV}} = \frac{\tilde{\tilde{\sigma}}_{P_{f,n}}}{\tilde{m}_{P_{f,n}}} < \epsilon_p$ is satisfied twice in succession, go to Step 6; else, go to Step 5.

Step 5: Adaptively identify multiple points and enrich the dataset

Identify q_m points $\boldsymbol{\mathcal{U}}_+ = \{\boldsymbol{u}_+^{(i)}\}_{i=1}^{q_m}$ using the proposed adaptive multi-point selection method (see Section 3.3). Evaluate the $\mathcal{G}$-function on the q_m points and obtain the corresponding responses $\boldsymbol{\mathcal{Y}}_+ = \left\{y_+^{(i)}\right\}_{i=1}^{q_m}$.

Let $\mathcal{D} = \{\mathcal{U} \cup \mathcal{U}_+, \mathcal{Y} \cup \mathcal{Y}_+\}$, $N_{call} = N_{call} + q_m$ and $m = m + 1$ and go to Step 2.

Step 6: End of PBPI

Return the estimated failure probability $\tilde{m}_{P_{f,n}}$ in Eq. (26).

4. Numerical examples

In this section, four numerical examples characterized by small failure probabilities are presented to demonstrate the performance of the proposed method. Several different values are considered for the parameter α in PPV to study its effects on the results. The efficiency, accuracy and robustness are compared with several other non-parallel methods (e.g., ALK-KDE-IS [43], AK-SDMCS [44] and AK-MCMC [45], etc.) and parallel methods (e.g., PABQ [36] and ALR in UQLab [46], etc.) in terms of the average number of iterations N_{iter}, the average number of $\mathcal{G}$-function calls N_{call}, the average failure probability P_f, the relative error of failure probability ϵ_{P_f} and the coefficient of variation $COV[P_f]$. Except for MCS and IS, the reported results are averaged over 20 repeated runs unless otherwise specified. It should be noted that the two parallel methods, PABQ and ALR in UQLab, need to predefine the number of added points at each iteration, i.e., k. Specifying a large k would decrease the number of iterations but increase the number of $\mathcal{G}$-function calls [36, 37]. In the following four examples, the values of k in PABQ and ALR are specified according to the average number of added points per iteration using the proposed PBPI method, in order to compare the performance between different methods more fairly.

In the proposed method, the number of initial observations is set to $N_0 = 10$. Specially, the threshold ϵ_p in the stopping criterion is set to $\epsilon_p = 5\%$. The variance amplification factor and initial sample size for sequential VAIS are set to $\gamma = 2.0$ and $N_{vas} = 10^6$, respectively. The thresholds ϵ_μ and $\epsilon_{\hat{\sigma}}$ are set to 2% and 10%, respectively. For ALR in UQLab, the Kriging model is adopted as the surrogate, in which the Gaussian function is employed as the correlation function. The stopping criterion is also modified to be that the beta bounds and stability are less than the default threshold (i.e., 0.01) within three consecutive iterations (refer to [46] for more details).

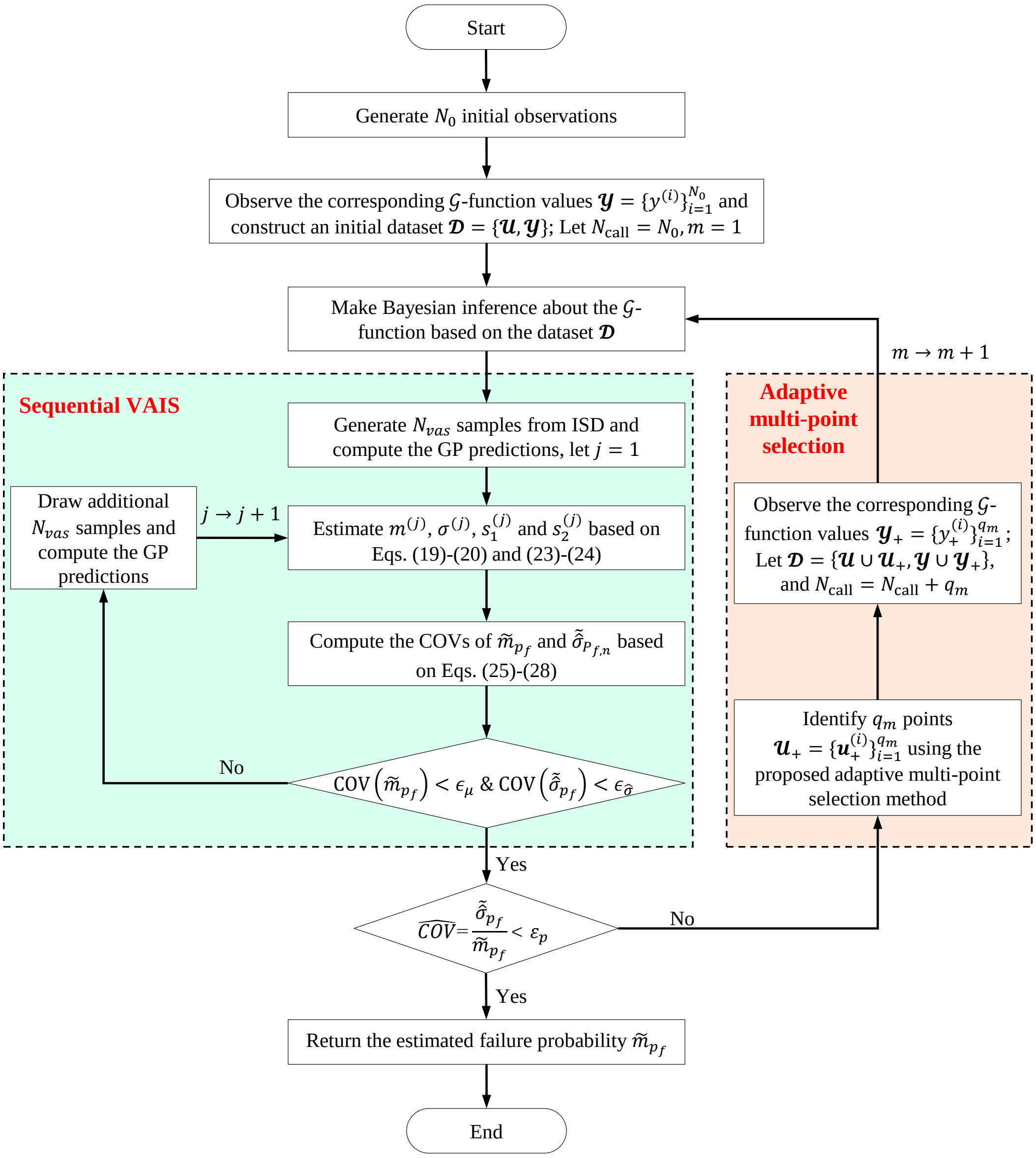


Figure 2: Flowchart of the proposed PBPI method.

4.1. Example 1: Series system with four branches

The first example considers the reliability analysis problem of a series system with four branches [23, 47]. The performance function is given as:

$$g(\boldsymbol{X}) = \min \begin{cases} a + \frac{(X_1 - X_2)^2}{10} - \frac{(X_1 + X_2)}{\sqrt{2}} \\ a + \frac{(X_1 - X_2)^2}{10} + \frac{(X_1 + X_2)}{\sqrt{2}} \\ (X_1 - X_2) + \frac{b}{\sqrt{2}} \\ (X_2 - X_1) + \frac{b}{\sqrt{2}} \end{cases} \tag{32}$$

where X_1 and X_2 are two independent standard normal variables; a and b are two constant parameters which affect the failure probability of the series system. Two cases are considered in this example: $a = 4$, $b = 7$ for the first case, and $a = 5.5$, $b = 11$ for the second case.

Case 1: $a = 4$ and $b = 7$

Table 1 shows the detailed results given by the proposed method and several other parallel (e.g., ALR [46], PABQ [36], and ds-AKP [42]) and non-parallel methods (e.g., AK-MCS [23], AK-SDMCS [44], ALK-KDE-IS [43]). The failure probability $P_f = 4.93 \times 10^{-4}$ provided by MCS is considered as the reference result. It can be observed that the proposed method ($\alpha = 1.5, 2.0$ or 2.25) and other active learning methods can provide accurate average failure probability estimates with the COVs less than 5% and the relative errors lower than 2%, except for ALR that produces biased results with relative errors greater than 6%. When it comes to the efficiency, the average number of iterations N_{iter} and the average number $\mathcal{G}$-function calls N_{call} of the proposed PBPI method are comparable to those of PABQ, which are obviously less than those of the other parallel methods, i.e., ALR and ds-AKP. Compared with the non-parallel counterparts (i.e., AK-MCS, AK-SDMCS and ALK-KDE-IS), the proposed method also exhibits computational advantages in terms of N_{iter} and N_{call}, though the comparable number of $\mathcal{G}$-function calls as AK-SDMCS.

Fig. 3 shows the identified points at each iteration of the proposed method ($\alpha = 2.25$). It can be observed that the number of added points at each iteration changes during the active learning process. These identified points are typically local optimal points of learning function (i.e., PPVC) and have relatively large contribution of uncertainty to the PPV of the failure probability. The points selected at different stages and the final experimental designs are shown in Fig. 4(a). The true limit state surface (black solid line) and the final predicted limit state surface (red solid line) are also plotted. Most of the identified points are found to locate in the vicinity of true limit state and distributed in the critical regions with major contributions to the failure probability. It can be observed that the predicted limit state surface fits well in the critical regions, though weakly approximating at regions with small probability densities that have negligible effects on failure probability. The results indicate that the proposed method can estimate the failure probability efficiently and accurately.

Case 2: $a = 5.5$ *and* $b = 11$

The failure probability is very small in the second case (in the order of 10^{-8}). Table 2 presents the reliability analysis results by the proposed method and other compared methods. The proposed method and the two non-parallel methods (i.e., AK-SDMCS and ALK-KDE-IS) can produce satisfactory average failure probability estimates with their COVs less than 5%. Although ALR provides a close average failure probability to the reference result provided MCS when $k = 4$, the COVs of ALR are larger than 30% for $k = 4$ or 5. Meanwhile, another parallel method PABQ yields inaccurate average failure probability with the relative errors of 13.68% for $k = 4$ or 5. The biased results mainly result from the fixed sampling region of importance ball sampling adopted in PABQ. As for the efficiency, the proposed method requires slightly less iterations and $\mathcal{G}$-function calls than PABQ when α is large (e.g., $\alpha = 2.5$). In addition, the proposed method greatly outperforms the ALR and ALK-KDE-IS in terms of N_{iter} and N_{call}. This case demonstrates the superior performance of the proposed method compared with several other methods in terms of accuracy and efficiency.

The identified points at each iteration of the proposed method ($\alpha = 2.25$) are depicted in Fig. 5. It can be seen that the local peaks of PPVC function are almost identified and the number of identified points is not fixed but changes during the iteration process. Fig. 4(b) shows the selected points at different stages and final experimental designs. The predicted and true limit state surfaces are depicted with red and black

Table 1: Reliability analysis results of Example 1 (Case 1).

Method		N_{iter}	N_{call}	P_f	$COV[P_f]$	ϵ_{P_f}
MCS		-	10^7	4.93×10^{-4}	1.42%	-
AK-MCS-U		63.15	74.15	4.99×10^{-4}	4.74%	1.01%
AK-SDMCS		33.55	44.55	4.95×10^{-4}	3.14%	0.41%
ALK-KDE-IS		65.35	76.35	5.02×10^{-4}	2.10%	1.83%
ALR in UQLab	$k = 4$	17.70	76.80	5.26×10^{-4}	3.60%	6.69%
	$k = 5$	15.15	80.75	5.27×10^{-4}	4.03%	6.90%
PABQ	$k = 4$	8.75	41.00	4.96×10^{-4}	2.39%	0.61%
	$k = 5$	7.00	40.00	4.98×10^{-4}	1.96%	1.01%
ds-AKP[1]		29.93	59.50	4.98×10^{-4}	4%	1.01%
Proposed method	$\alpha = 1.5$	8.00	47.00	4.98×10^{-4}	1.54%	1.01%
	$\alpha = 2.0$	7.30	41.00	4.96×10^{-4}	1.71%	0.61%
	$\alpha = 2.25$	7.15	39.70	4.94×10^{-4}	3.19%	0.20%
	$\alpha = 2.5$	6.60	38.55	4.93×10^{-4}	6.51%	0

[1] The results are taken from research [42] based on 30 independent runs.

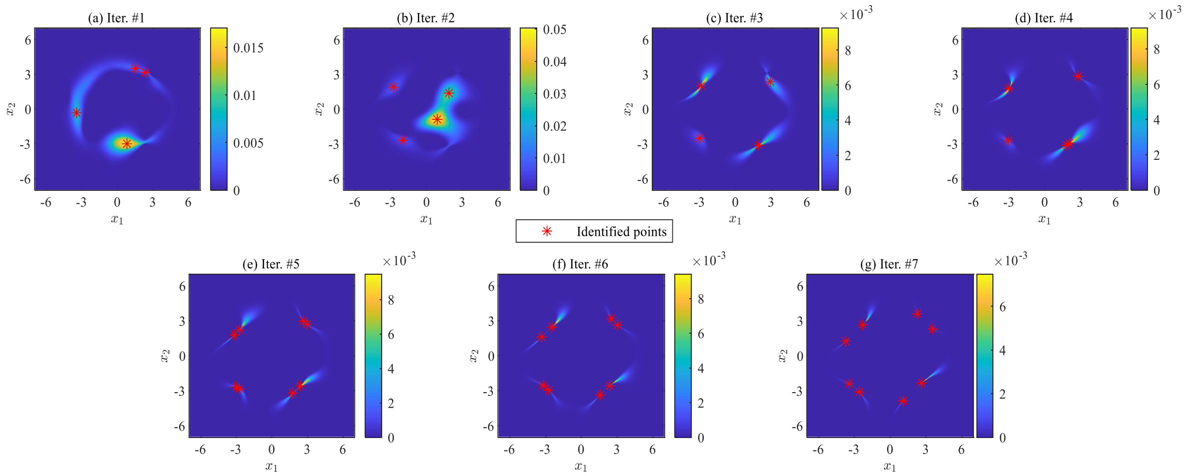
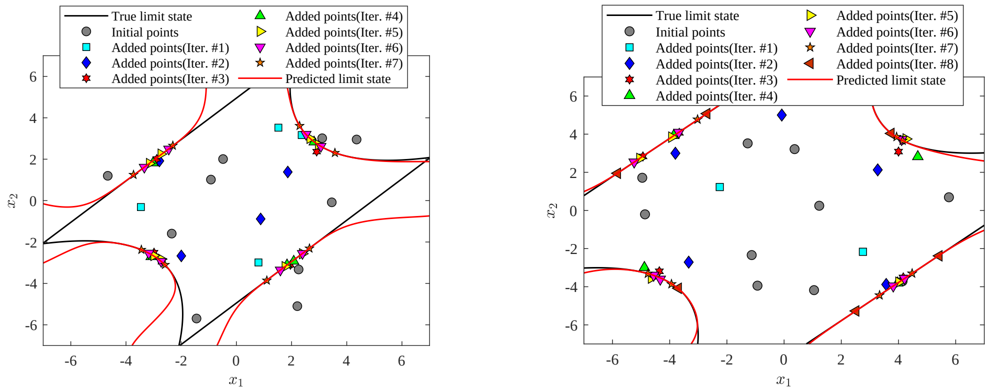


Figure 3: The points identified at each iteration for Example 1 (Case 1).

solid lines, respectively. The identified points are found to gradually move to the critical regions on the limit state surface. Most of the identified points reside in the critical regions and the predicted limit state surface is generally consistent with the true limit state surface. The results verify that the proposed method can estimate the small failure probability accurately.



(a) Case 1 (b) Case 2

Figure 4: Selected points by the proposed method for Example 1.

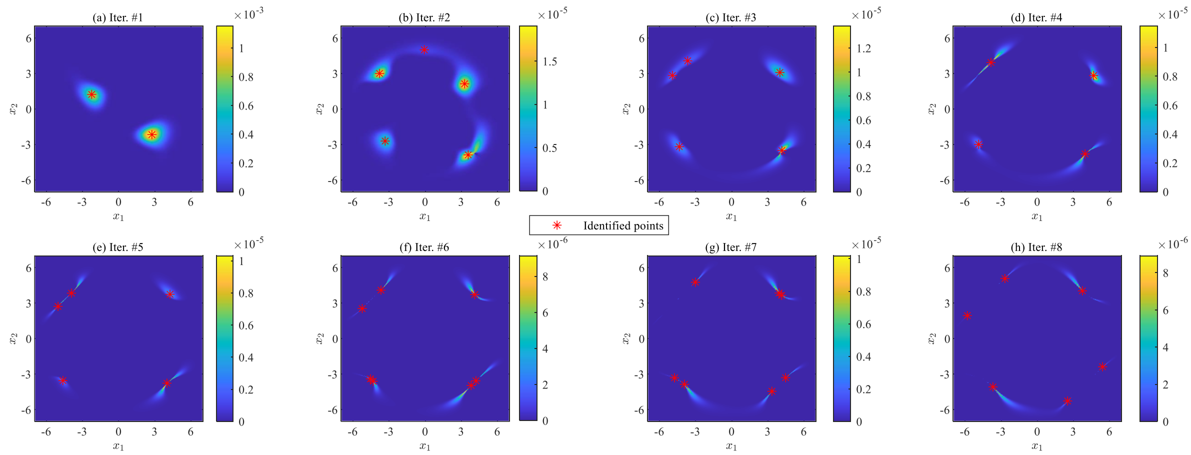


Figure 5: The points identified at each iteration for Example 1 (Case 2).

4.2. Example 2: Nonlinear oscillator

In this example, a nonlinear undamped single degree of freedom oscillator subjected to rectangular pulse load (Fig. 6) is investigated [23, 36]. The performance function is defined as:

$$g\left(c_1, c_2, m, r, t_1, F_1\right) = 3r - \left|\frac{2F_1}{m\omega_0^2}\sin\left(\frac{\omega_0 t_1}{2}\right)\right| \tag{33}$$

where $\omega_0 = \sqrt{(c_1 + c_2)}/m$. The six random variables are listed in Table 3.

Table 4 lists the numerical results of the proposed method and other compared methods. The failure probability $P_f = 1.50 \times 10^{-8}$ with COV of 2.58% provided by MCS is considered as the reference result. It

Table 2: Reliability analysis results of Example 1 (Case 2).

Method		N_{iter}	N_{call}	P_f	$COV[P_f]$	ϵ_{P_f}
MCS		-	3×10^{10}	5.92×10^{-8}	2.37%	-
AK-SDMCS[1]		35.5	46.5	5.79×10^{-8}	3.99%	2.20%
ALK-KDE-IS[2]		65.6	76.6	5.86×10^{-8}	2.80%	1.01%
ALR in UQLab	$k = 4$	21.45	91.80	5.96×10^{-8}	45.22%	0.68%
	$k = 5$	19.40	104.00	6.41×10^{-8}	31.01%	8.28%
PABQ	$k = 4$	13.35	59.40	5.11×10^{-8}	1.87%	13.68%
	$k = 5$	9.50	52.50	5.11×10^{-8}	2.58%	13.68%
Proposed method	$\alpha = 1.5$	8.95	54.60	5.90×10^{-8}	1.60%	0.34%
	$\alpha = 2.0$	8.90	51.75	5.84×10^{-8}	2.32%	1.35%
	$\alpha = 2.25$	8.20	48.65	5.86×10^{-8}	3.17%	1.01%
	$\alpha = 2.5$	8.05	47.55	5.87×10^{-8}	2.34%	0.84%

[1] The results are taken from research [44] based on 100 independent runs.
[2] The results are taken from research [43].

Table 3: Details of random variables in Example 2.

Variable	Distribution	Mean	Standard deviation
m	Normal	1	0.05
c_1	Normal	1	0.1
c_2	Normal	0.1	0.01
r	Normal	0.5	0.05
t_1	Normal	1	0.2
F_1	Normal	0.45	0.075

is found that the proposed method ($\alpha = 1.5, 2.0$ or 2.25), AK-SDMCS, ALK-KDE-IS and AK-MCMC [45] can provide accurate average failure probability estimates with the COVs around 5%, while the proposed method greatly outperforms its counterparts in terms of N_{iter} and N_{call}. The proposed method also exhibits computational advantages compared with ALR. In addition, both ALR and PABQ produce biased average failure probability estimates with relatively large COVs in this example. In particular, the relative errors of PABQ are greater than 60% though it costs similar N_{iter} and N_{call} to the proposed method. Overall, the results demonstrate the superior accuracy, robustness and efficiency of the proposed method over several others active learning methods.

4.3. Example 3: Turbine blade structural model

The third example considers a turbine blade structural model of jet engine available in the Matlab Partial Differential Equation (PDE) Toolbox, which is made of nickel-base alloy (NIMONIC90). The finite element (FE) model of the turbine blade with the maximum element size of 0.01 is depicted in the left of Fig. 7.

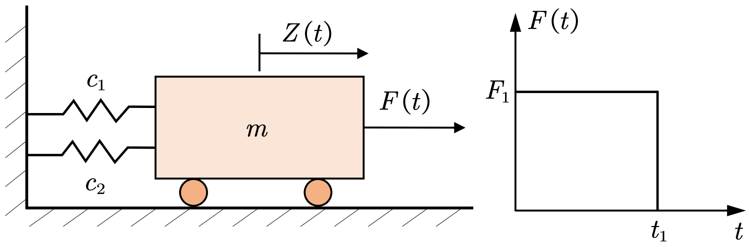


Figure 6: Nonlinear oscillator subjected to pulse load.

Table 4: Reliability analysis results of Example 2.

Method		N_{iter}	N_{call}	P_f	$COV[P_f]$	ϵ_{P_f}
MCS		-	10^{11}	1.50×10^{-8}	2.58%	-
AK-SDMCS[1]		53.6	64.6	1.46×10^{-8}	5.20%	2.67%
ALK-KDE-IS		47.25	58.25	1.51×10^{-8}	2.77%	0.67%
AK-MCMC		171.70	182.70	1.51×10^{-8}	1.19%	0.67%
ALR in UQLab	$k = 1$	42.00	53.00	1.70×10^{-8}	9.26%	13.33%
	$k = 2$	22.75	55.50	1.70×10^{-8}	8.96%	13.33%
PABQ	$k = 1$	13.45	22.45	5.97×10^{-9}	8.81%	60.20%
	$k = 2$	7.90	23.80	5.67×10^{-9}	9.28%	62.20%
Proposed method	$\alpha = 1.5$	10.90	25.35	1.49×10^{-8}	4.00%	0.67%
	$\alpha = 2.0$	9.25	22.50	1.47×10^{-8}	4.37%	2.00%
	$\alpha = 2.25$	9.15	22.00	1.51×10^{-8}	5.83%	0.67%
	$\alpha = 2.5$	8.10	20.95	1.46×10^{-8}	8.38%	2.67%

[1] The results are taken from research [44] based on 30 independent runs.

The von Mises stress distribution of the combined structural and thermal analysis is shown in the right of Fig. 7. Considering the uncertainties of material properties, pressure loads and temperature condition of the turbine blade structural model, the maximum von Mises stress should be less than a given allowable threshold. The limit state function is thus defined as:

$$g(\boldsymbol{X}) = \sigma_{th} - \sigma_{max}(E, CTE, \lambda, K_{app}, p_1, p_2, T_1, T_2) \tag{34}$$

where σ_{th} denotes the allowable threshold ($\sigma_{th} = 1.5$ GPa is adopted); σ_{max} denotes the maximum von Mises stress under the combined thermal and pressure effects; The Young's modulus E, coefficient of thermal expansion CTE, Poisson's ration λ, thermal conductivity K_{app}, pressure load on the pressure side p_1, pressure load on the suction side p_2, temperature of the interior cooling air T_1, and temperature on the pressure and suction sides T_2 are assumed to be independent random variables. The details of the eight

Table 5: Details of random variables in Example 3.

Variable	Description	Distribution	Parameter 1	Parameter 2
E(GPa)	Young's modulus	Normal	200	0.15
$CTE(1/K)$	Coefficient of thermal expansion	Normal	1.27×10^{-5}	0.1
λ	Poisson's ratio	Lognormal	0.27	0.1
$K_{app}(W/m/K)$	Thermal conductivity	Lognormal	11.5	0.1
p_1(kPa)	Pressure loads	Lognormal	500	0.20
p_2(kPa)	Pressure loads	Lognormal	450	0.20
T_1(°C)	Temperature	Uniform	130	170
T_2(°C)	Temperature	Uniform	950	1050

Note: Parameter 1 and 2 respectively denote the mean and coefficient of variation for normal and lognormal distribution, and the lower and upper-bounds for uniform distribution.

random variables are listed in Table 5.

Table 6 presents the numerical results provided by the proposed PBPI method and other compared methods. The failure probability $P_f = 4.19 \times 10^{-6}$ estimated by IS is regarded as the reference value. The AK-SDMCS and ALK-KDE-IS do not converge after multiple trials, hence the results are absent. As shown in Table 6, the proposed method provides fairly accurate average failure probability estimates with their COVs around 5% when the parameter $\alpha = 1.5, 2.0$ or 2.25. As for the efficiency, the proposed method costs significantly less iterations and $\mathcal{G}$-function calls than AK-MCMC and ALR. It is noted in this example that AK-MCMC fail to converge after 200 iterations in all 20 independent runs. Compared to PABQ, the proposed method ($\alpha = 2.0, 2.25$ or 2.5) also shows better efficiency in terms of N_{iter} and N_{call}. In addition, the COVs in PBPI are smaller than those in PABQ, indicating the better robustness of the proposed method. The results verify that the proposed PBPI can efficiently produce accurate and robust failure probability estimate for this turbine blade problem.

4.4. Example 4: A transmission tower

A transmission tower structure is studied in the last example to further illustrate the performance of the proposed method. The 19.3 m-tall tower structure is modified from [37]. Four forces with random direction in XOZ plane are applied on this structure, as depicted in Fig. 8(a) and (b). The tower structure is modeled as FE model constructed in OpenSees platform. The FE model consists of 53 nodes and 172 elements. The bilinear stress-strain curve is used, as shown in Fig. 8(c). Twelve independent random variables are considered in this example. Table 7 lists the details of these random variables. The performance function is

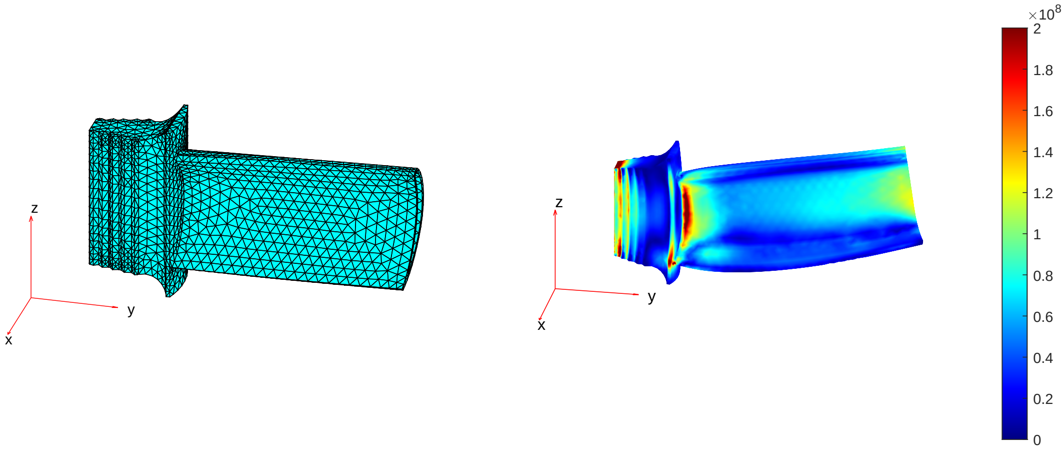


Figure 7: Finite element model (left) and von Mises stress distribution of combined structural and thermal analysis (right) of turbine blade.

Table 6: Reliability analysis results of Example 3.

Method		N_{iter}	N_{call}	P_f	$COV[P_f]$	ϵ_{P_f}
IS[1]		-	4101	4.19×10^{-6}	3.89%	-
AK-MCMC		200.00	211.00	4.49×10^{-6}	6.41%	7.16%
AK-SDMCS		-	-	-	-	-
ALK-KDE-IS		-	-	-	-	-
ALR in UQLab	$k = 2$	147.35	308.70	4.96×10^{-6}	5.17%	18.38%
	$k = 3$	93.65	293.95	4.93×10^{-6}	6.11%	17.66%
PABQ	$k = 2$	36.65	81.30	3.87×10^{-6}	12.99%	7.64%
	$k = 3$	19.95	66.85	3.97×10^{-6}	10.70%	5.25%
Proposed method	$\alpha = 1.5$	28.30	79.60	4.20×10^{-6}	3.55%	0.24%
	$\alpha = 2.0$	17.90	52.35	4.12×10^{-6}	5.29%	1.67%
	$\alpha = 2.25$	15.40	44.35	4.08×10^{-6}	5.73%	2.63%
	$\alpha = 2.5$	14.15	38.00	4.11×10^{-6}	8.77%	1.91%

[1] The results of IS are calculated using UQLab [46].

defined as follows:

$$g(\boldsymbol{X}) = \delta_{th} - D(P_1, P_2, P_3, P_4, \theta_1, \theta_2, \theta_3, \theta_4, A, F_y, E_0, b) \tag{35}$$

where δ_{th} denotes the specified threshold and $\delta_{th} = 15\,\mathrm{cm}$ is adopted in this example; $D(\cdot)$ denotes the horizontal displacement of the topmost node.

The results provided by different methods are summarized in Table 8. The failure probability $P_f = 6.06 \times 10^{-7}$ with the COV of 1.32% estimated by IS is adopted as the reference value. The results of AK-SDMCS and ALK-KDE-IS are not listed as they fail to converge after multiple trials. It is observed that the proposed method and AK-MCMC can provide fairly accurate average failure probability estimates

Table 7: Details of random variables in Example 4.

Variable	Description	Distribution	Parameter 1	Parameter 2
$P_1, P_2(kN)$	Load	Lognormal	60	0.2
$P_3, P_4(kN)$	Load	Lognormal	50	0.2
$\theta_1, \theta_2(°)$	Angle	Uniform	0	10
$\theta_3, \theta_4(°)$	Angle	Uniform	0	20
$A(\text{mm}^2)$	Cross-sectional area	Normal	5000	0.10
$F_y(\text{MPa})$	Yield strength	Normal	400	0.15
$E_0(\text{GPa})$	Young's modulus	Normal	200	0.10
b	Strain-hardening ratio	Normal	0.02	0.05

Note: Parameter 1 and 2 respectively denote the mean and coefficient of variation for normal and lognormal distribution, and the lower and upper-bounds for uniform distribution.

Table 8: Reliability analysis results of Example 4.

Method		N_{iter}	N_{call}	P_f	$COV[P_f]$	ϵ_{P_f}
IS[1]		-	50120	6.06×10^{-7}	1.32%	-
AK-MCMC		200.00	211.00	5.86×10^{-7}	5.63%	3.30%
AK-SDMCS		-	-	-	-	-
ALK-KDE-IS		-	-	-	-	-
ALR in UQLab	$k = 2$	147.45	316.90	6.71×10^{-7}	6.73%	10.73%
PABQ	$k = 2$	44.50	97.00	2.55×10^{-7}	30.25%	57.92%
	$\alpha = 2.0$	40.45	89.15	6.16×10^{-7}	5.49%	1.65%
Proposed method	$\alpha = 2.25$	28.65	69.00	6.05×10^{-7}	6.03%	0.17%
	$\alpha = 2.5$	24.65	60.00	5.89×10^{-7}	6.24%	2.81%

[1] The results of IS are calculated using UQLab [46].

with their COVs around 5% though AK-MCMC can not converge after 200 iterations in all 20 independent runs. ALR and PABQ yield biased results, especially for PABQ that the relative error is larger than 57%. As for the computational efficiency, the proposed method costs obviously less iterations and $\mathcal{G}$-function calls than AK-MCMC, ALR and PABQ, especially when α is large (e.g., $\alpha = 2.5$). In addition, the COVs in the proposed method (5.49% to 6.24%) are greatly less than that in PABQ (30.25%). These results demonstrates that the proposed method outperforms several other existing methods in terms of accuracy, robustness and efficiency.

4.5. Discussions on the parameter α

A parameter α is introduced in the proposed PPV (i.e., Eq. (15)) to approximate the true posterior variance. Parametric analysis is conducted in the four numerical examples to investigate its effect on the performance of the proposed method. It is observed that the average number of iterations N_{iter} and the average number of $\mathcal{G}$-function calls N_{call} generally decrease with the increase of the parameter α. Meanwhile,

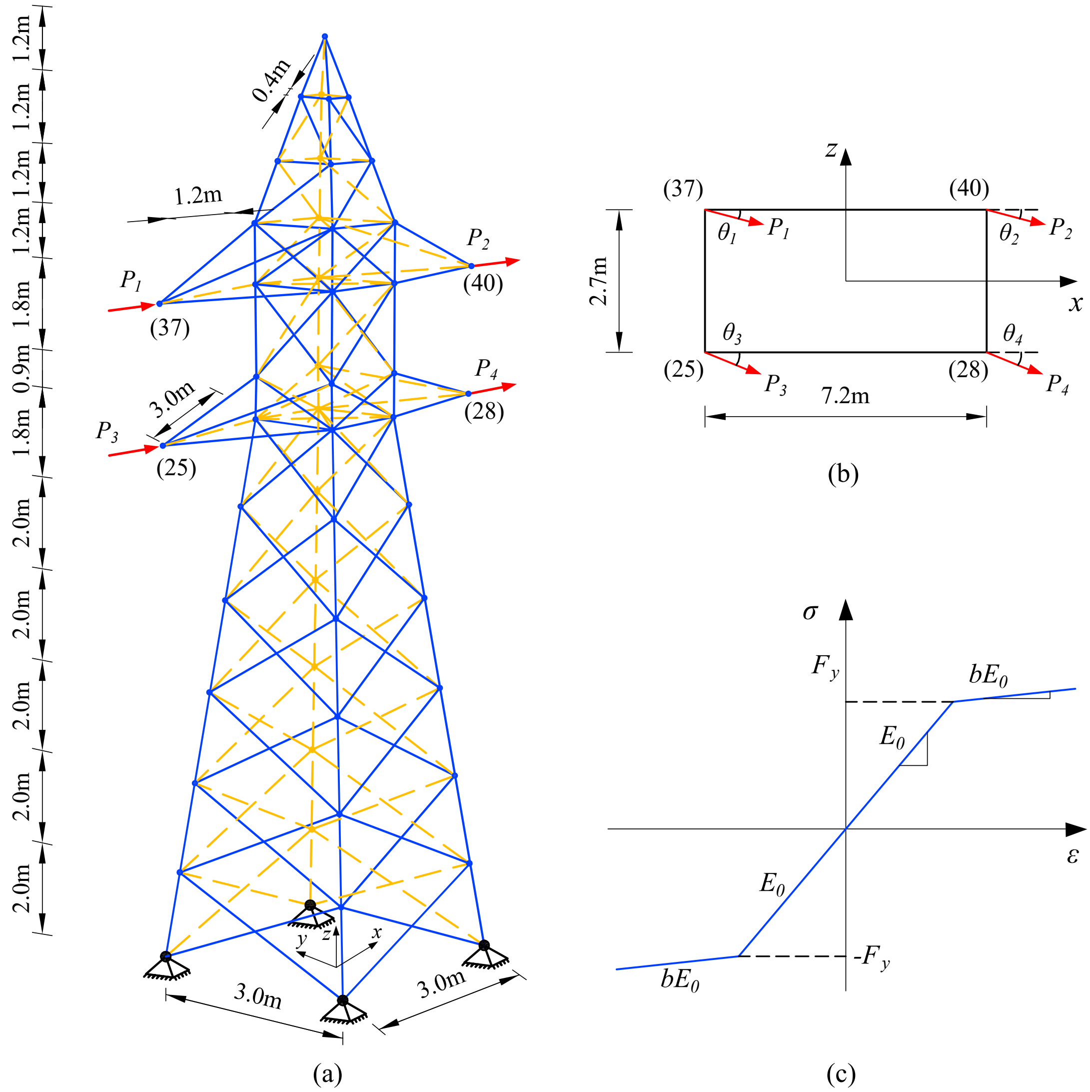


Figure 8: Finite element modeling for transmission tower: (a) 172-elements truss structures; (b) Schematic diagram of load directions; (c) Bilinear stress-strain curve for elements.

the COVs of failure probability estimate increase with α. The results are mainly due to the fact that the increase of α would decrease the PPV value and the corresponding pseudo posterior COV in Eq. (30), accelerating the convergence of active learning though sacrificing the robustness of the algorithm.

The setting of parameter α involves a trade-off between accuracy and efficiency. According to the four investigated examples, a parameter $\alpha = 2.25$ is suggested in PPV to construct the learning function and stopping criterion. Results show that the accuracy and efficiency can be guaranteed simultaneously under this setting. Specifically, the COVs of the failure probability estimates are around 5% and the relative errors are within 3% in the four investigated examples. Note that the threshold ϵ_p in the stopping criterion, constructed based on the judgment of the pseudo posterior COV, is specified as exactly 5%. In addition, the proposed method with $\alpha = 2.25$ typically shows better efficiency in terms of N_{iter} and N_{call} than several other existing active learning reliability methods. It is worth mentioning that the PPV with $\alpha = 2.25$ may not be able to accurately approximate the true posterior variance. However, as a compromise between the UPV and the true posterior variance, it can provide a simple but pragmatic uncertainty measure of failure probability, which contributes the development of Bayesian active learning methods.

5. Conclusions

This paper presents a novel Bayesian active learning method termed 'Parallel Bayesian Probabilistic Integration' (PBPI) for efficiently estimating small failure probabilities. Specifically, a pseudo posterior variance (PPV) of failure probability is first heuristically proposed for providing a pragmatic uncertainty measure over failure probability. The PPV with a reasonable setting of α can not only alleviate the expensive computational cost of exact posterior variance in PA-BFPL, but also more realistically reflect the true posterior variance compared with the upper-bound posterior variance in PABQ. Besides, the variance amplified importance sampling is modified in a sequential manner to allow the estimations of posterior mean and PPV of failure probability with large sample population. According to the posterior statistics of failure probability, a learning function and a stopping criterion are then presented to enable active learning. Finally, a novel adaptive multi-point selection method is developed to identify multiple points without the need to predefine the number of points added at each iteration, thereby supporting parallel computing more intelligently.

The effectiveness of the proposed PBPI method is demonstrated by investigating four numerical examples, including a turbine blade model and a transmission tower structure. According to the investigated numerical examples, a parameter $\alpha = 2.25$ is suggested in the proposed PPV of failure probability. Numerical results indicate that the proposed method is capable of providing accurate failure probability estimates with the COVs around 5% and the relative error less than 3% under this setting. In addition, the proposed method generally requires less performance function evaluations and iterations compared to several other state-of-the-art active learning methods. Overall, the proposed PBPI method can assess small failure probabilities (e.g., in the order of $10^{-4} \sim 10^{-8}$) with satisfactory accuracy, efficiency and robustness.

The proposed PBPI method is expected to perform well for linear, weakly nonlinear and moderately nonlinear problems in low to moderate dimensions. The performance of the proposed method may degrade for high dimensional and/or strong nonlinearity problems due to the limitations of the DBSCAN clustering algorithm and GP model. Additional research efforts are still needed to address the limitations.

Declaration of competing interest

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

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Grant No. 2021YFB2600900), and China Scholarship Council.

References

[1] R. Melchers, Importance sampling in structural systems, Structural Safety 6 (1) (1989) 3–10.

[2] A. Tabandeh, G. Jia, P. Gardoni, A review and assessment of importance sampling methods for reliability analysis, Structural Safety 97 (2022) 102216.

[3] S.-K. Au, J. L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, Probabilistic Engineering Mechanics 16 (4) (2001) 263–277.

[4] P.-S. Koutsourelakis, H. J. Pradlwarter, G. I. Schueller, Reliability of structures in high dimensions, part i: algorithms and applications, Probabilistic Engineering Mechanics 19 (4) (2004) 409–417.

[5] P. Bjerager, Probability integration by directional simulation, Journal of Engineering Mechanics 114 (8) (1988) 1285–1302.

[6] A. M. Hasofer, N. C. Lind, Exact and invariant second-moment code format, Journal of the Engineering Mechanics Division 100 (1) (1974) 111–121.

[7] A. Der Kiureghian, H.-Z. Lin, S.-J. Hwang, Second-order reliability approximations, Journal of Engineering Mechanics 113 (8) (1987) 1208–1225.

[8] Y.-G. Zhao, T. Ono, Moment methods for structural reliability, Structural safety 23 (1) (2001) 47–75.

[9] J. Xu, C. Dang, A new bivariate dimension reduction method for efficient structural reliability analysis, Mechanical Systems and Signal Processing 115 (2019) 281–300.

[10] X. Zhang, M. D. Pandey, Structural reliability analysis based on the concepts of entropy, fractional moment and dimensional reduction method, Structural Safety 43 (2013) 28–40.

[11] C. Dang, J. Xu, A mixture distribution with fractional moments for efficient seismic reliability analysis of nonlinear structures, Engineering Structures 208 (2020) 109912.

[12] L. Faravelli, Response-surface approach for reliability analysis, Journal of Engineering Mechanics 115 (12) (1989) 2763–2781.

[13] A. Hadidi, B. F. Azar, A. Rafiee, Efficient response surface method for high-dimensional structural reliability analysis, Structural Safety 68 (2017) 15–27.

[14] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering & System Safety 93 (7) (2008) 964–979.

[15] J. Zhang, W. Gong, X. Yue, M. Shi, L. Chen, Efficient reliability analysis using prediction-oriented active sparse polynomial chaos expansion, Reliability Engineering & System Safety 228 (2022) 108749.

[16] C. M. Rocco, J. A. Moreno, Fast Monte Carlo reliability evaluation using support vector machine, Reliability Engineering & System Safety 76 (3) (2002) 237–243.

[17] H. Li, Z. Lü, Z. Yue, Support vector machine for structural reliability analysis, Applied Mathematics and Mechanics 27 (10) (2006) 1295–1303.

[18] A. A. Chojaczyk, A. P. Teixeira, L. C. Neves, J. B. Cardoso, C. G. Soares, Review and application of Artificial Neural Networks models in reliability analysis of steel structures, Structural Safety 52 (2015) 78–89.

[19] W. J. de Santana Gomes, Structural reliability analysis using adaptive artificial neural networks, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering 5 (4) (2019) 041004.

[20] I. Kaymaz, Application of kriging method to structural reliability problems, Structural Safety 27 (2) (2005) 133–151.

[21] B. Gaspar, A. P. Teixeira, C. G. Soares, Assessment of the efficiency of Kriging surrogate models for structural reliability analysis, Probabilistic Engineering Mechanics 37 (2014) 24–34.

[22] B. J. Bichon, M. S. Eldred, L. P. Swiler, S. Mahadevan, J. M. McFarland, Efficient global reliability analysis for nonlinear implicit performance functions, AIAA Journal 46 (10) (2008) 2459–2468.

[23] B. Echard, N. Gayton, M. Lemaire, AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation, Structural Safety 33 (2) (2011) 145–154.

[24] Z. Sun, J. Wang, R. Li, C. Tong, LIF: A new kriging based learning function and its application to structural reliability analysis, Reliability Engineering & System Safety 157 (2017) 152–165.

[25] X. Zhang, L. Wang, J. D. Sørensen, REIF: a novel active-learning function toward adaptive kriging surrogate models for structural reliability analysis, Reliability Engineering & System Safety 185 (2019) 440–454.

[26] Y. Shi, Z. Lu, R. He, Y. Zhou, S. Chen, A novel learning function based on kriging for reliability analysis, Reliability Engineering & System Safety 198 (2020) 106857.

[27] Z. Lv, Z. Lu, P. Wang, A new learning function for kriging and its applications to solve reliability problems in engineering, Computers & Mathematics with Applications 70 (5) (2015) 1182–1197.

[28] L. Hong, H. Li, J. Fu, A novel surrogate-model based active learning method for structural reliability analysis, Computer Methods in Applied Mechanics and Engineering 394 (2022) 114835.

[29] J. Yi, Q. Zhou, Y. Cheng, J. Liu, Efficient adaptive Kriging-based reliability analysis combining new learning function and error-based stopping criterion, Structural and Multidisciplinary Optimization 62 (5) (2020) 2517–2536.

[30] P. Wei, Y. Zheng, J. Fu, Y. Xu, W. Gao, An expected integrated error reduction function for accelerating bayesian active learning of failure probability, Reliability Engineering & System Safety 231 (2023) 108971.

[31] Z. Wang, A. Shafieezadeh, ESC: an efficient error-based stopping criterion for kriging-based reliability analysis methods, Structural and Multidisciplinary Optimization 59 (5) (2019) 1621–1637.

[32] Z. Wang, P. Wang, A maximum confidence enhancement based sequential sampling scheme for simulation-based design, Journal of Mechanical Design 136 (2) (2014) 021006.

[33] M. Moustapha, S. Marelli, B. Sudret, Active learning for structural reliability: Survey, general framework and benchmark, Structural Safety 96 (2022) 102174.

[34] R. Teixeira, M. Nogal, A. O'Connor, Adaptive approaches in metamodel-based reliability analysis: A review, Structural Safety 89 (2021) 102019.

[35] C. Dang, P. Wei, J. Song, M. Beer, Estimation of failure probability function under imprecise probabilities by active learning–augmented probabilistic integration, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering 7 (4) (2021) 04021054–04021054.

[36] C. Dang, P. Wei, M. G. Faes, M. A. Valdebenito, M. Beer, Parallel adaptive Bayesian quadrature for rare event estimation, Reliability Engineering & System Safety 225 (2022) 108621.

[37] C. Dang, M. A. Valdebenito, M. G. Faes, P. Wei, M. Beer, Structural reliability analysis: A Bayesian perspective, Structural Safety 99 (2022) 102259.

[38] N. Lelièvre, P. Beaurepaire, C. Mattrand, N. Gayton, AK-MCSi: A Kriging-based method to deal with small failure probabilities and time-consuming models, Structural Safety 73 (2018) 1–11.

[39] H. Zhan, N.-C. Xiao, Y. Ji, An adaptive parallel learning dependent Kriging model for small failure probability problems, Reliability Engineering & System Safety 222 (2022) 108403.

[40] N. Razaaly, P. M. Congedo, Extension of AK-MCS for the efficient computation of very small failure probabilities, Reliability Engineering & System Safety 203 (2020) 107084.

[41] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining, Vol. 96, 1996, pp. 226–231.

[42] R. Teixeira, M. Nogal, A. O'Connor, B. Martinez-Pastor, Reliability assessment with density scanned adaptive kriging, Reliability Engineering & System Safety 199 (2020) 106908.

[43] X. Yang, Y. Liu, C. Mi, X. Wang, Active learning Kriging model combining with kernel-density-estimation-based importance sampling method for the estimation of low failure probability, Journal of Mechanical Design 140 (5) (2018) 051402.

[44] M. Su, G. Xue, D. Wang, Y. Zhang, Y. Zhu, A novel active learning reliability method combining adaptive Kriging and spherical decomposition-MCS (AK-SDMCS) for small failure probabilities, Structural and Multidisciplinary Optimization 62 (6) (2020) 3165–3187.

[45] P. Wei, C. Tang, Y. Yang, Structural reliability and reliability sensitivity analysis of extremely rare failure events by combining sampling and surrogate model methods, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 233 (6) (2019) 943–957.

[46] M. Moustapha, S. Marelli, B. Sudret, UQLab user manual – Active learning reliability, Tech. rep., Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Switzerland, report UQLab-V2.0-117 (2022).

[47] S. Katsuki, D. M. Frangopol, Hyperspace division method for structural reliability, Journal of Engineering Mechanics 120 (11) (1994) 2405–2427.