Efficient procedure for failure probability function estimation in augmented space

Xiukai Yuan[a,c,*], Shaolong Liu[a], M.A. Valdebenito[b], Jian Gu[a]， Michael Beer[c,d,e]
[a] School of Aerospace Engineering, Xiamen University, Xiamen 361005, R.P. China
[b] Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Av. Padre Hurtado 750, 2562340 Viña del Mar, Chile
[c] Institute for Risk and Reliability, Leibniz Universität Hannover, Callinstr. 34, Hannover, Germany
[d] Institute for Risk and Uncertainty, University of Liverpool, Peach Street, L69 7ZF Liverpool, United Kingdom
[e] International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, Shanghai 200092, China

Abstract

An efficient procedure is proposed to estimate the failure probability function (FPF) with respect to design variables, which correspond to distribution parameters of basic structural random variables. The proposed procedure is based on the concept of an augmented reliability problem, which assumes the design variables as uncertain by assigning a prior distribution, transforming the FPF into an expression that includes the posterior distribution of those design variables. The novel contribution of this work consists of expressing this target posterior distribution as an integral, allowing it to be estimated by means of sampling, and no distribution fitting is needed, leading to an efficient estimation of FPF. The proposed procedure is implemented within three different simulation strategies: Monte Carlo simulation, importance sampling and subset simulation; for each of these cases, expressions for the coefficient of variation of the FPF estimate are derived. Numerical examples illustrate performance of the proposed approaches.

Keywords:
Failure probability function; Bayesian theory; Reliability analysis; Reliability-based optimization

1. Introduction

Reliability and reliability-based design optimization (RBDO) provide useful tools for quantifying uncertainty and performing optimal design under uncertainty, respectively. The application of reliability methods has been widely accepted in structural design, since there are inherent sources of uncertainties affecting the performance of structural systems. RBDO attempts to determine optimal design solutions while explicitly taking into account the effects of uncertainty [1][2]. Note that the failure probability may be highly sensitive to the distribution parameters that characterize basic structural random variables, i.e., mean value or standard deviation. Thus, it is of paramount importance and of great interest to estimate the failure probability as a function of these parameters [3][4]. The latter is particularly true as in several cases, these distribution parameters can be actually interpreted as design variables, as they can represent, for example, the outcome of a fabrication process. Given that the failure probability becomes a function of these distribution parameters / design variables, it is termed within the context of this work as the *Failure Probability Function* (FPF). The FPF can be seen as “reliability sensitivity analysis” that indicates how the failure probability changes with respect to the design variables. Also, the FPF can be used within the context of RBDO, as it allows to decouple the problem into a traditional optimization task without implementing a nested double-loop [5][6].

In practice, it is difficult to obtain the FPF with regard to the design parameters, since it usually requires repeated reliability analyses executed at various design parameter values. Even though a variety of methods have been developed to estimate the failure probability of structural systems, performing repeated reliability analyses constitutes still a challenge. In this sense, well established methods for reliability based on approximate analytic representations (e.g., first/second order reliability method - FORM/SORM [7][8]) or simulation methods (such as Monte Carlo simulation [9], importance sampling [10], Subset simulation [11], and Line sampling [12] etc.) still demand computational and numerical efforts which may become remarkable for practical problems, especially when finite element models are involved. As a result, repeated reliability analyses render the direct estimation of FPF computationally intractable.

There are many contributions addressing the estimation of the FPF, and a number of methods have been developed, which can be classified into three classes. One class comprises surrogate models. For example, one can build an approximation of the FPF based on design of experiments (DOE). That is, to construct an approximation by selecting some predefined interpolation points in the space of the design parameters. For example, Gasser [13] adopts a pre-defined quadratic function with respect to the design parameters to approximate the logarithm of FPF. Then, a number of reliability analyses are carried out over some interpolation grid points of design parameters, finally to determine the coefficients of the predefined function by a least squares approach. Jensen [14] adopted a linear function to approximate the logarithm of FPF; in this approach, the number of reliability analyses equals that of the design variables. Note that there is a considerable number of surrogate model methods, for example, Kriging method [15][16], Support vector machine [17][18], etc. These methods are widely applied to reliability analysis to approximate the limit state function [19][20], which can be also used to approximate the FPF. However, their practical application may become involved due to the necessity of repeated evaluation of the structural model for training the surrogate. Another variant of surrogate models was proposed by Wei [21][22], that adopted High-dimensional model representation (HDMR) to approximate the FPF within the framework of imprecise stochastic simulation. The second class of approaches for approximating the FPF consists of a post-processing step of a standard reliability analysis. This class of methods focuses on obtaining the FPF with respect to the distribution parameters. For example, Zou [5] expressed the FPF as a linear function of the distribution parameters by applying first-order Taylor series expansion based on the reliability sensitivity information. As the reliability sensitivity is a byproduct of reliability analysis, the FPF can be built by means of a single reliability analysis. Yuan [23] proposed a weighted approach to obtain the FPF. By introducing an instrumental sampling function, the estimate of FPF can be expressed as a function of a set of samples which are generated in a single reliability analysis. Its efficiency depends on the simulation method used, such as, Monte Carlo simulation, importance sampling, Subset simulation. Further, an advanced Line sampling approach is proposed to solve the FPF [24], which is similar with the weighted approach as it only needs one simulation run of line sampling. The third and last class of strategies for estimating the FPF involves the formulation of the reliability problem in an augmented space. Au [25] first consider the design variables as uncertain with an assigned probability distribution. This leads to a reliability problem in an augmented space, which is composed of the basic structural variables and the design parameters. Then, the Bayesian rule is applied to express the FPF as the product of three terms. The key term out of these three is the posterior distribution of the design variables. In this way, the FPF is estimated by performing a single augmented reliability analysis. Ching [26][27] follows the augmented reliability idea and adopted the maximum entropy principle to estimate the posterior distribution of the design variables. Despite all these progresses, there is still a strong need to enhance our ability and efficiency for estimating the FPF for general problems.

This contribution proposes an efficient procedure called 'Augmented space integral' (ASI) for estimating the structural FPF. In particular, the proposed procedure develops the augmented reliability idea further to handle the FPF with respect to distribution design parameters. The features of the proposed approach are: (1) It is based on the augmented reliability idea [25][26], allowing the FPF to be obtained in a single simulation run and thus, repeated reliability analyses can be avoided; (2) It casts the posterior distribution appearing in the expression of the FPF as an integral, which allows its estimation through simulation by averaging over samples, without need of fitting prescribed probability density functions [25][26]; (3) The work reported here solves the FPF in augmented space, while the weighted approach proposed in [23] is solved in the original random variable space. Thus, the proposed approach can be interpreted as an extended version of the weighted approach.

This contribution is organized as follows. In Section 2, the FPF problem and the original augmented reliability method are first briefly given. Then, the mathematical formulation of the proposed procedure is derived, and the implementation with Monte Carlo simulation, importance sampling and Subset simulation are also developed in Section 3. At last, in Section 4, various examples are presented to show the performance of the proposed approach.

2. Failure probability function and its estimation in augmented space

2.1 Failure probability function definition

In this contribution, we focus on the estimation of the FPF with respect to the distribution parameters of basic random variables, such as mean and standard deviation. These distribution parameters could be interpreted as design variables. Note that the failure probability can be highly sensitive to the distribution parameters and thus, it is of great interest to know the relationship of the failure probability with respect to the distribution parameters. This type of problems is encountered, for example, when the mean of the geometrical dimension of a structural member such as thickness, length, height, etc. corresponds to the design variables in reliability-based design optimization.

The FPF considered in this contribution is given by

$$P_F(\boldsymbol{\theta}) = \int I_F(\boldsymbol{x}) f(\boldsymbol{x}|\boldsymbol{\theta}) d\boldsymbol{x} \tag{1}$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$ is the vector of basic random variables of the structure/system; $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_{n_\theta}]^T$ is the vector of distribution parameters related with $\boldsymbol{x}$; $f(\boldsymbol{x}|\boldsymbol{\theta})$ represents the joint conditional Probability Distribution Function (PDF) of $\boldsymbol{x}$ given $\boldsymbol{\theta}$; $I_F(\boldsymbol{x})$ is the indicator function, $I_F(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in F$ and $I_F(\boldsymbol{x}) = 0$, otherwise; $F = \{\boldsymbol{x}: g(\boldsymbol{x}) \leq 0\}$ is the failure region and $g(\cdot)$ is the limit state function.

Note that there are usually two different types of design variables in structural reliability-based design. One type refers to the variables that affect the limit state function and the other type refers to variables that affect in the distribution parameters of basic random variables [4]. In this contribution, it is found that, when the design variable corresponds to the distribution parameters, the estimator of FPF can be obtained in an efficient way by the proposed procedure. For the sake of simplicity and without loss of generality, it is also assumed that all the basic random variables are independent with respect to each other.

2.2 Failure probability function transformation in augmented space

The augmented reliability idea provides an efficient means for calculating FPF and was first proposed by Au [25]. In an augmented space, the design variable $\boldsymbol{\theta}$ is no longer a deterministic quantity but it is modeled as a random variable vector with an instrumental probability distribution $\varphi(\boldsymbol{\theta})$. Then, applying the Bayesian theory, the sought failure probability function in Eq. (1) can be transformed as [25]:

$$P_F(\boldsymbol{\theta}) = \frac{\varphi(\boldsymbol{\theta}|F)P(F)}{\varphi(\boldsymbol{\theta})} \tag{2}$$

where $\varphi(\boldsymbol{\theta}|F)$ is the posterior distribution of $\boldsymbol{\theta}$ conditioned on the occurrence of the failure event; and $P(F)$ is the failure probability of the augmented reliability problem:

$$P(F) = \iint I_F(\boldsymbol{x}) f(\boldsymbol{x}|\boldsymbol{\theta})\varphi(\boldsymbol{\theta})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{\theta} \tag{3}$$

By virtue of Eq. (2), the FPF is expressed in terms of three components, namely $\varphi(\boldsymbol{\theta})$, $P(F)$ and $\varphi(\boldsymbol{\theta}|F)$. Among them, $\varphi(\boldsymbol{\theta})$ can be selected arbitrarily (due to its instrumental nature) and it is important to note that different distributions for $\boldsymbol{\theta}$ do not affect the results of FPF from a theoretical viewpoint. For example, either Gaussian or Uniform distribution can be employed [26]. However, care should be taken whenever the Gaussian distribution is used, as it may associate negative values with quantities that are positive due to physical reasons. $P(F)$ can be estimated by performing reliability analysis in augmented space. The remaining part is to estimate $\varphi(\boldsymbol{\theta}|F)$, which is the most important and challenging term for the estimation of FPF by means of Eq. (2). In [25], Au used histograms to represent this term, and latter Ching adopted the maximum entropy method to approximate the posterior distribution, leading to an estimator for the FPF which is an explicit expression of $\boldsymbol{\theta}$ [26]-[27]. In this contribution, the proposed procedure develops the augmented reliability method further, such that there is no need to fit a density function to describe the posterior distribution $\varphi(\boldsymbol{\theta}|F)$.

3. Proposed approach for FPF estimation

As the key for solving the FPF according to Eq. (2) is the assessment of the term $\varphi(\boldsymbol{\theta}|F)$, the proposed procedure is further developed for handling a particular type of problem where the design variables are the distribution parameters of basic random variables. It is found that for this type of problem, the posterior distribution associated with the FPF can be expressed as an integral which can be estimated by means of samples. Thus, there is no need to fit a probability distribution. In the following, the proposed procedure, as well as the implementations with three different simulation approaches are presented.

3.1 Basic formulation of the proposed approach

The proposed procedure attempts to establish a relationship between the posterior distribution $\varphi(\boldsymbol{\theta}|F)$ and the variable $\boldsymbol{x}$. If such goal can be fulfilled, then there is no need to fit a probability distribution in order to estimate $\varphi(\boldsymbol{\theta}|F)$.

According to the formula of conditional probability, $\varphi(\boldsymbol{\theta}|F)$ can be developed as:

$$\varphi(\boldsymbol{\theta}|F) = \int \varphi(\boldsymbol{\theta}|\boldsymbol{x},F) f(\boldsymbol{x}|F)\mathrm{d}\boldsymbol{x} = E_{\boldsymbol{x}|\boldsymbol{F}}[\varphi(\boldsymbol{\theta}|\boldsymbol{x},F)] \tag{4}$$

where $E_{\boldsymbol{x}|\boldsymbol{F}}[\cdot]$ means the expectation under $f(\boldsymbol{x}|F)$. Eq. (4) reveals the relationship between the two posterior distributions, $\varphi(\boldsymbol{\theta}|F)$ and $f(\boldsymbol{x}|F)$. It provides an alternative way to obtain $\varphi(\boldsymbol{\theta}|F)$, instead of using density fitting methods. According to Eq. (4), if $\varphi(\boldsymbol{\theta}|\boldsymbol{x},F)$ is obtained beforehand, then $\varphi(\boldsymbol{\theta}|F)$ can be obtained using simulation, i.e., estimating expectation by the mean of samples.

According to Bayesian theory, $\varphi(\boldsymbol{\theta}|\boldsymbol{x},F)$ can be simplified to

$$\varphi(\boldsymbol{\theta}|\boldsymbol{x},F)=\frac{I_F(\boldsymbol{x})\varphi(\boldsymbol{\theta}|\boldsymbol{x})}{\int I_F(\boldsymbol{x})\varphi(\boldsymbol{\theta}|\boldsymbol{x})\,\mathrm{d}\,\boldsymbol{\theta}}=I_F(\boldsymbol{x})\varphi(\boldsymbol{\theta}|\boldsymbol{x}) \tag{5}$$

where the term $\varphi(\boldsymbol{\theta}|\boldsymbol{x})$ can also be rewritten as follows by using Bayesian theory

$$\varphi(\boldsymbol{\theta}|\boldsymbol{x})=\frac{f(\boldsymbol{x}|\boldsymbol{\theta})\varphi(\boldsymbol{\theta})}{f(\boldsymbol{x})} \tag{6}$$

where $f(\boldsymbol{x})$ is the marginal distribution of $\boldsymbol{x}$ in augmented space $(\boldsymbol{x},\boldsymbol{\theta})$ which is given by:

$$f(\boldsymbol{x})=\int f(\boldsymbol{x},\boldsymbol{\theta})\,\mathrm{d}\,\boldsymbol{\theta}=\int f(\boldsymbol{x}|\boldsymbol{\theta})\varphi(\boldsymbol{\theta})\,\mathrm{d}\,\boldsymbol{\theta} \tag{7}$$

It has already been stated in [25] that the role of the PDF $\varphi(\boldsymbol{\theta})$ is not to reflect the uncertainty of $\boldsymbol{\theta}$. Rather, it is a device to yield information about $P_F(\boldsymbol{\theta})$ versus $\boldsymbol{\theta}$. The choice of $\varphi(\boldsymbol{\theta})$ depends on the region in the design parameter space where $P_F(\boldsymbol{\theta})$ is to be studied. Without particular preference for the region to be emphasized, a uniform distribution may be chosen for convenience [25], i.e., $\boldsymbol{\theta}\sim\boldsymbol{U}[\underline{\boldsymbol{\theta}},\ \bar{\boldsymbol{\theta}}]$. In this context, $\varphi(\boldsymbol{\theta})$ is a constant within $\boldsymbol{\theta}\in[\underline{\boldsymbol{\theta}},\ \bar{\boldsymbol{\theta}}]$. Then, the marginal distribution $f(\boldsymbol{x})$ can be rewritten as:

$$f(\boldsymbol{x})=\int_{\underline{\boldsymbol{\theta}}}^{\bar{\boldsymbol{\theta}}} f(\boldsymbol{x}|\boldsymbol{\theta})\varphi(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}=\varphi(\boldsymbol{\theta})\Delta(\boldsymbol{x}) \tag{8}$$

where $\Delta(\boldsymbol{x})=\int_{\underline{\boldsymbol{\theta}}}^{\bar{\boldsymbol{\theta}}} f(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$ is an integral over the design region. Further details on the calculation of $\Delta(\boldsymbol{x})$ are presented in Appendix A.

Substitution of Eq. (8) into (6) allows determining the sought the posterior distribution $\varphi(\boldsymbol{\theta}|\boldsymbol{x})$, which is equal to:

$$\varphi(\boldsymbol{\theta}|\boldsymbol{x})=\frac{f(\boldsymbol{x}|\boldsymbol{\theta})\varphi(\boldsymbol{\theta})}{f(\boldsymbol{x})}=\frac{f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})} \tag{9}$$

Replacing Eqs. (5) and (9) into Eq. (4) leads to:

$$\varphi(\boldsymbol{\theta}|F)=\int I_F(\boldsymbol{x})\frac{f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})}f(\boldsymbol{x}|F)\mathrm{d}\boldsymbol{x}=E_{\boldsymbol{x}|\boldsymbol{F}}\left[\frac{f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})}\right] \tag{10}$$

And $f(\boldsymbol{x}|F)$ is the PDF of $\boldsymbol{x}$ conditional on F, which is given by:

$$f(\boldsymbol{x}|F)=\frac{I_F(\boldsymbol{x})f(\boldsymbol{x})}{\int I_F(\boldsymbol{x})f(\boldsymbol{x})\,\mathrm{d}\,\boldsymbol{x}}=\frac{I_F(\boldsymbol{x})f(\boldsymbol{x})}{\int I_F(\boldsymbol{x})\int f(\boldsymbol{x},\boldsymbol{\theta})\,\mathrm{d}\,\boldsymbol{\theta}\,\mathrm{d}\,\boldsymbol{x}}=\frac{I_F(\boldsymbol{x})f(\boldsymbol{x})}{P(F)} \tag{11}$$

Substitution of Eq. (11) into Eq. (10) allows rewriting $\varphi(\boldsymbol{\theta}|F)$ as:

$$\varphi(\boldsymbol{\theta}|F)=\frac{1}{P(F)}\int\frac{I_F(\boldsymbol{x})f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})}f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{12}$$

Finally, substitution of Eq. (12) into Eq. (2) leads to the final expression for the FPF, which can be expressed as:

$$P_F(\boldsymbol{\theta})=\frac{1}{\varphi(\boldsymbol{\theta})}\int\frac{I_F(\boldsymbol{x})f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})}f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=\frac{1}{\varphi(\boldsymbol{\theta})}E_{\boldsymbol{x}}\left[\frac{I_F(\boldsymbol{x})f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})}\right] \tag{13}$$

where $E_{\boldsymbol{x}}[\cdot]$ represents expectation under the marginal distribution $f(\boldsymbol{x})$.

3.2 Proposed procedure based on Monte Carlo simulation

After a general formula for the FPF has been obtained above, it can be implemented by means of a simulation-based method. The most direct approach is using Monte Carlo simulation. In the following, the proposed procedure based on Monte Carlo simulation is presented, which is denoted as 'ASI-MCS' for compactness.

Note that contrary to traditional reliability analysis, the augmented reliability problem must be solved in the augmented space $(\boldsymbol{x}, \boldsymbol{\theta})$. According to Eq. (13), if $P_F(\boldsymbol{\theta})$ is solved by using Monte Carlo Simulation, the key is to generate samples from $\boldsymbol{x}$ which follow the marginal distribution of $f(\boldsymbol{x})$**.** However, in a general case, one may not sample directly from $f(\boldsymbol{x})$, as it may not correspond to a known probability distribution. However, the following workaround can be implemented. First, generate samples $\{\boldsymbol{\theta}^{(j)}, j = 1, \dots, N\}$ that follow $\varphi(\boldsymbol{\theta})$. Then, for each of these samples, generate samples $\{\boldsymbol{x}^{(j)}, j = 1, \dots, N\}$, each of them distributed according to $f(\boldsymbol{x}|\boldsymbol{\theta}^{(j)})$. Thus, the set of samples $\{(\boldsymbol{x}^{(j)}, \boldsymbol{\theta}^{(j)}), j = 1, \dots, N\}$ follows $f(\boldsymbol{x}, \boldsymbol{\theta})$. Ignoring the samples associated with $\boldsymbol{\theta}$, then $\{\boldsymbol{x}^{(j)}, j = 1, \dots, N\}$ are distributed as $f(\boldsymbol{x})$.

According to Eq. (10), the estimator for the posterior distribution is given by:

$$\hat{\varphi}(\boldsymbol{\theta}|F) = \frac{1}{N_F} \sum_{j=1}^{N_F} \frac{f(\boldsymbol{x}^{(j)}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x}^{(j)})} \tag{14}$$

This implies that, instead of using distribution fitting approach to estimate $\varphi(\boldsymbol{\theta}|F)$ (as performed in [26]), the proposed approach can estimate directly the $\varphi(\boldsymbol{\theta}|F)$ by means of sampling.

According to Eq. (13), the FPF $P_F(\boldsymbol{\theta})$ is estimated as:

$$\hat{P}_F(\boldsymbol{\theta}) = \frac{1}{\varphi(\boldsymbol{\theta})} \frac{1}{N} \sum_{j=1}^{N} \frac{I_F(\boldsymbol{x}^{(j)}) f(\boldsymbol{x}^{(j)}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x}^{(j)})} \tag{15}$$

It is obvious that the estimator $\hat{P}_F(\boldsymbol{\theta})$ is unbiased, and its variance can be readily obtained as:

$$Var[\hat{P}_F(\boldsymbol{\theta})] \approx \frac{1}{N-1} \left\{ \frac{1}{N} \sum_{j=1}^{N} \left\{ \frac{I_F(\boldsymbol{x}^{(j)}) f(\boldsymbol{x}^{(j)}|\boldsymbol{\theta})}{\varphi(\boldsymbol{\theta})[\Delta(\boldsymbol{x}^{(j)})]} \right\}^2 - \hat{P}_F^2(\boldsymbol{\theta}) \right\} \tag{16}$$

And the Coefficient of variation (C.o.v.) of $\hat{P}_F(\boldsymbol{\theta})$ is given by:

$$Cov[\hat{P}_F(\boldsymbol{\theta})] = \frac{\sqrt{Var[\hat{P}_F(\boldsymbol{\theta})]}}{E[\hat{P}_F(\boldsymbol{\theta})]} \approx \frac{\sqrt{Var[\hat{P}_F(\boldsymbol{\theta})]}}{\hat{P}_F(\boldsymbol{\theta})} \tag{17}$$

3.3 Proposed procedure based on importance sampling

The proposed procedure can be also implemented with importance sampling, which is denoted as 'ASI-IS'. Introducing an appropriate importance sampling function $H(\boldsymbol{x})$ in augmented space, the FPF in Eq. (13) can be rewritten as:

$$P_F(\boldsymbol{\theta}) = \frac{1}{\varphi(\boldsymbol{\theta})} \int \frac{I_F(\boldsymbol{x}) f(\boldsymbol{x}|\boldsymbol{\theta})}{\Delta(\boldsymbol{x})} \frac{f(\boldsymbol{x})}{H(\boldsymbol{x})} H(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{18}$$

Substitution of Eq. (8) into (18) allows rewriting $P_F(\boldsymbol{\theta})$ as:

$$P_F(\boldsymbol{\theta}) = \int \frac{I_F(\boldsymbol{x}) f(\boldsymbol{x}|\boldsymbol{\theta})}{H(\boldsymbol{x})} H(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = E_H\left[\frac{I_F(\boldsymbol{x}) f(\boldsymbol{x}|\boldsymbol{\theta})}{H(\boldsymbol{x})}\right] \tag{19}$$

where $E_H[\cdot]$ denotes the expectation under $H(\boldsymbol{x})$.

It should be noted that the final expression in Eq. (19) is the same as the one associated with weighted importance sampling introduced in a previous work [23], but its meaning is quite different. In the previous work, everything is formulated in the space associated with $\boldsymbol{x}$, but in the present contribution, it is solved in the augmented space $(\boldsymbol{x}, \boldsymbol{\theta})$. Thus, in this sense, it can be seen as an extended version of the previous work.

The importance sampling density $H(\boldsymbol{x})$ should be selected properly. There are plenty of contributions addressing the way of determining importance sampling density. However, few of them address how to determine such density in an augmented space. Here, we present some suggestions. There are two ways for determining $H(\boldsymbol{x})$ in augmented space. One is based on a design point $\boldsymbol{x}^*$. The design point can be solved according to a nominal value of design parameter, say, $\boldsymbol{\theta_0}$, which can be simply set as the center of the domain associated with $\boldsymbol{\theta}$, i.e., $\boldsymbol{\theta_0} = (\underline{\boldsymbol{\theta}} + \bar{\boldsymbol{\theta}})/2$. Alternatively, the design point can be obtained by searching the point $\boldsymbol{x}^*$ in the augmented space which has the largest joint PDF value $f(\boldsymbol{x}, \boldsymbol{\theta})$ and that belongs to the failure domain F. For this purpose, the random variables can be classified into two types, that is, $\boldsymbol{x} = [\boldsymbol{x_\theta}, \boldsymbol{x_r}]$, where $\boldsymbol{x_\theta}$ is the variable vector related with design parameters $\boldsymbol{\theta}$ and $\boldsymbol{x_r}$ is the vector of the rest of random variables (whose distribution is not affected by $\boldsymbol{\theta}$). Then, based on the design point $\boldsymbol{x}^* = [\boldsymbol{x}_{\boldsymbol{\theta}}^*, \boldsymbol{x}_{\boldsymbol{r}}^*]$, $H(\boldsymbol{x})$ can be chosen as

$$H(\boldsymbol{x}) = f(\boldsymbol{x_\theta}) H(\boldsymbol{x_r}|\boldsymbol{x}_{\boldsymbol{r}}^*) \tag{20}$$

where $f(\boldsymbol{x_\theta})$ is the marginal distribution for $\boldsymbol{x_\theta}$ given in Eq. (8).

The other way to choose $H(\boldsymbol{x})$ is based on adaptive importance sampling density [29]. That is, to establish an approximate optimal density. For example, we can pre-sample in the failure region (in augmented space), and then, based on these samples, obtain an approximated sampling density.

Suppose $H(\boldsymbol{x})$ has been chosen, then samples can be generated according to $H(\boldsymbol{x})$. Suppose a total of N samples are generated, $\{\boldsymbol{x}^{(j)}, j = 1, \dots, N\}$. Then according to Eq. (19), $P_F(\boldsymbol{\theta})$ is estimated as:

$$\hat{P}_F(\boldsymbol{\theta}) = \frac{1}{N} \sum_{j=1}^{N} \frac{I_F(\boldsymbol{x}^{(j)}) f(\boldsymbol{x}^{(j)}|\boldsymbol{\theta})}{H(\boldsymbol{x}^{(j)})} \tag{21}$$

It is obvious that the estimator $\hat{P}_F(\boldsymbol{\theta})$ is unbiased, and its variance is obtained as

$$Var[\hat{P}_F(\boldsymbol{\theta})] \approx \frac{1}{N-1} \left\{ \frac{1}{N} \sum_{j=1}^{N} \left\{ \frac{I_F(\boldsymbol{x}^{(j)}) f(\boldsymbol{x}^{(j)}|\boldsymbol{\theta})}{H(\boldsymbol{x}^{(j)})} \right\}^2 - \hat{P}_F^2(\boldsymbol{\theta}) \right\} \tag{22}$$

And the C.o.v. of $\hat{P}_F(\boldsymbol{\theta})$ is given by:

$$Cov[\hat{P}_F(\boldsymbol{\theta})] = \frac{\sqrt{Var[\hat{P}_F(\boldsymbol{\theta})]}}{E[\hat{P}_F(\boldsymbol{\theta})]} \approx \frac{\sqrt{Var[\hat{P}_F(\boldsymbol{\theta})]}}{\hat{P}_F(\boldsymbol{\theta})} \tag{23}$$

3.4 Proposed procedure based on Subset simulation

In this section, Subset simulation is adopted and integrated within the proposed procedure, which is denoted as 'ASI-SS'. Subset simulation is an efficient approach for evaluating the failure probability of general reliability problems, and is especially suitable for high dimensional, low failure probability reliability problems [11]. The basic idea of Subset simulation is that a low failure probability event is expressed as the product of a series of conditional, larger probabilities, whose estimation is straightforward. The proposed ASI procedure can be implemented by using Subset Simulation, as described in the following.

Let $F_1 \supset F_2 \supset \ldots \supset F_m = F$ be a nested sequence of failure events in subset simulation in augmented space $(\boldsymbol{x}, \boldsymbol{\theta})$ where $F_i = \{g(\boldsymbol{x}) \leq b_i\}(i = 1,2,\ldots,m)$; then the failure probability can be expressed by

$$P(F) = P(F_1)\prod_{i=2}^{m} P(F_i|F_{i-1}) \tag{24}$$

Note that $b_1, b_2, \ldots, b_{m-1}$ are the intermediate threshold values which are adaptively determined, so that the corresponding probabilities $P(F_1)$, $P(F_2|F_1)$,..., $P(F_{m-1}|F_{m-2})$ can be all set to be p_0, e.g., $p_0 = 0.1$ for convenience. The final threshold $b_m = 0$ is not chosen adaptively.

Suppose there are N_s samples generated at (m-1)th level. Moreover, it is considered that there are a number of N_F failure samples located in the final level (target failure region F), $\{(\boldsymbol{x}^{(j)}, \boldsymbol{\theta}^{(j)}), j = 1, \ldots, N_F\}$, which are distributed as $f(\boldsymbol{x}, \boldsymbol{\theta}|F)$. Ignoring the $\boldsymbol{\theta}$ part, the samples $\{\boldsymbol{x}^{(j)}, j = 1, \ldots, N_F\}$ are distributed as $f(\boldsymbol{x}|F)$. Then according to Eq. (10), $\varphi(\boldsymbol{\theta}|F)$ can be estimated by

$$\hat{\varphi}(\boldsymbol{\theta}|F) = \frac{1}{N_F}\sum_{j=1}^{N_F}\frac{f\left(\boldsymbol{x}^{(j)}\middle|\boldsymbol{\theta}\right)}{\Delta(\boldsymbol{x}^{(j)})} \tag{25}$$

This means that the distribution $\varphi(\boldsymbol{\theta}|F)$ can be estimated using simulation, and no distribution fitting is required. The failure probability in Eq. (24) can be estimated by

$$\hat{P}(F) = p_0^{m-1}\frac{N_F}{N_s} \tag{26}$$

Finally, substitution of Eqs. (25) and (26) into (2) leads to the following expression for FPF $P_F(\boldsymbol{\theta})$:

$$\hat{P}_F(\boldsymbol{\theta}) = \frac{p_0^{m-1}}{\varphi(\boldsymbol{\theta})}\frac{1}{N_s}\sum_{j=1}^{N_s}\frac{I_F(\boldsymbol{x}^{(j)})f\left(\boldsymbol{x}^{(j)}\middle|\boldsymbol{\theta}\right)}{\Delta(\boldsymbol{x}^{(j)})} = \frac{p_0^{m-1}\hat{P}_m(\boldsymbol{\theta})}{\varphi(\boldsymbol{\theta})} \tag{27}$$

where

$$\hat{P}_m(\boldsymbol{\theta}) = \frac{1}{N_s}\sum_{j=1}^{N_s}\frac{I_F(\boldsymbol{x}^{(j)})f\left(\boldsymbol{x}^{(j)}\middle|\boldsymbol{\theta}\right)}{\Delta(\boldsymbol{x}^{(j)})} \tag{28}$$

And the C.o.v. of $\hat{P}_F(\boldsymbol{\theta})$ can be approximated by:

$$Cov\left[\hat{P}_F(\boldsymbol{\theta})\right] = \sqrt{\sum_{i=1}^{m}\frac{Var(\hat{P}_i)}{P_i^2}} \approx \sqrt{\sum_{i=1}^{m}\frac{Var(\hat{P}_i)}{\hat{P}_i^2}} \tag{29}$$

A detailed derivation of Eq. (29) is given in Appendix B.

3.5 Comparison of different approaches

Table 1 presents the comparison of different approaches that can produce an estimate of the FPF by means of a single reliability analysis.

In Table 1, 'ASI' refers to the proposed procedure herein; 'WA' refers to the Weighted Approaches previously proposed by the author, which includes Weighted Monte Carlo Simulation (WMCS), Weighted importance sampling (WIS) and Weighted Subset simulation (WSS); 'ALS' refers to Advanced Line Sampling method. 'Au' refers to the augmented reliability idea proposed by Au; 'Ching' refers to the advanced augmented reliability idea with maximum entropy; 'Wei' refers to non-intrusive imprecise stochastic simulation for uncertainty propagation.

Table 1. Comparison of different approaches

Methods	Space	Expression	Simulation	Design variable
ASI	Augmented space	Integral	MCS, IS, SS	Distribution parameter
WA[23]	Original space	Integral	MCS, IS, SS	Distribution parameter
ALS[24]	Original space	Integral	Line sampling	Distribution parameter
Au[25]	Augmented space	Histogram	SS	Distribution parameter or deterministic parameter
Ching[26]	Augmented space	Maximum entropy	SS	Distribution parameter or deterministic parameter
Wei[21]	Augmented space	Integral	MCS, SS	Distribution parameter

Among these methods, WA and ALS are carried out in original space of basic random variables, while the other methods are all formulated in augmented space. In Au and Ching, the estimator of FPF is finally expressed by using histogram or by using maximum entropy estimation. However, they can handle both the distribution parameters and design variables affecting the limit state function, whereas other approaches only handle the former one. The proposed approach preserves the advantage of using only a single reliability simulation in augmented space. Furthermore, it does not need to estimate the conditional distribution $\varphi(\boldsymbol{\theta}|F)$ by using density approximation (considering, e.g. maximum entropy principle) and in addition, there is no need to select the shape of the probability density function.

Thus, in summary, the proposed procedure can be seen as a particular version of the augmented reliability idea (Au and Ching), which improves the efficiency when handling the FPF with respect to distribution parameters, resulting in better performance. Also, the proposed procedure can be seen as a further advanced version of the weighted approach (Yuan), which extends the original space to the augmented space, resulting in a better improvement on accuracy. These characteristics will be shown in the examples given in Section 4.

3.6 Procedure of the proposed approach

The procedure of the proposed approach is summarized as follows.

1) Choose a distribution $\varphi(\boldsymbol{\theta})$.

   For general cases, uniform distribution may be chosen.

2) Carry out the simulation in augmented space.

   ASI-MCS, ASI-IS or ASI-SS can be selected to carry out reliability analysis in the augmented space $(\boldsymbol{x}, \boldsymbol{\theta})$, producing failure samples $\{(\boldsymbol{x}^{(j)}, \boldsymbol{\theta}^{(j)}): j = 1, \ldots, N_F\}$.

3) Obtain the FPF estimator

The FPF can be obtained according to Eq. (15) for ASI-MCS, Eq. (21) for ASI-IS or (27) for ASI-SS; their respective C.o.v.'s can be calculated as well according to Eq. (17) for ASI-MCS, Eq. (23) for ASI-IS or (29) for ASI-SS.

4. Examples

In order to verify the feasibility and accuracy of the proposed approaches (ASI-MCS, ASI-IS and ASI-SS), numerical and practical engineering examples are presented in this section. Meanwhile, different methods are also used for comparison purposes. 'Direct MCS' refers to the results obtained by Direct MCS, which can be seen as the reference results. 'WMCS' refers to the 'Weighted Monte Carlo simulation' method [23]; 'WIS' refers to the 'Weighted Importance sampling' method [23]. Note that in the proposed procedure, the prior distribution is selected as a uniform distribution for all the examples.

4.1 Example 1: A test example

The first example considers a simple limit state function, which is given by:

$$g(\boldsymbol{x}) = e^{0.04x_1+7} - e^{0.3{x_1}^2+5} * x_2 \tag{30}$$

where $\boldsymbol{x} = [x_1, x_2]$, x_1 and x_2 are independent, Gaussian distributed random variables, such that $x_1 \sim N(\theta, 0.5)$, $x_2 \sim N(0, 0.5)$; the mean of x_1 is taken as the design parameter, and the design domain is $\theta \in [-2, 2]$.

The proposed ASI approaches (ASI-MCS, ASI-IS and ASI-SS) are applied to this toy problem. For comparison, the weighted approaches (WMCS, WIS and WSS) [23] and GEMCS in [21] are also applied. The computational cost for these approaches is listed in Table 2. Note that for the application of all of these methods, a single reliability analysis is required to obtain the FPF and its corresponding C.o.v. In addition, the direct MCS is carried out in ten independent simulation runs to generate point wise failure probability estimates, which are regarded as the exact reference results. The obtained FPF results (according to Eq. (15) for ASI-MCS, Eq. (21) for ASI-IS or (27) for ASI-SS) and C.o.v. of the estimate (according to Eq. (17) for ASI-MCS, Eq. (23) for ASI-IS or (29) for ASI-SS.) are given in Fig. 1 and Fig. 2, respectively.

Note that for the implementation of ASI-MCS, a total of 10^5 samples are generated, and 143 failure samples fall within the failure domain; whereas WMCS was implemented considering 10^9 samples, with a nominal setting $\theta_0 = 0$, and only 50 failure samples are obtained. This is due to the fact that the proposed ASI approaches are implemented in augmented space where it is assumed $\theta \sim U[-2, 2]$, and the corresponding augmented failure probability is larger than the one where it is assumed that $\theta_0 = 0$. It can be seen from Fig. 1 that both ASI-MCS and WMCS obtained quite accurate FPF results (consistent with the reference results provided by Direct MCS). However, the samples used for WMCS are 1000 (=$10^9/10^5$) times larger than that of ASI-MCS. In this sense, the proposed ASI-MCS overperforms the WMCS in terms of efficiency. Compared with GEMCS, the proposed approach is more accurate (see Fig. 1) and more robust, as the C.o.v. of estimate by GEMCS vary drastically (see Fig. 2).

Table 2. Comparisons of different methods for Example 1

Methods	No. of samples	No. of failure samples
ASI-MCS	10^5	163
ASI-IS	2000	61
ASI-SS	200×4*	185
WMCS	10^9	50
WIS	2000	962
WSS	500×9	258
GEMCS	10^5	135
Direct MCS	10^8×10	

*: “200×4” means 200 samples are used for each level and a total of 4 levels are used.

ASI-IS is applied such that the sampling function is constructed based on the design point. The design point $\boldsymbol{x}^* = [-2.35, 1.26]$ is found when assuming $\theta_0 = 0$. Then, the sampling function setting for ASI-IS is $H(\boldsymbol{x}) = f(x_1)H(x_2|\boldsymbol{x}_r^* = 1.26)$. A number of N= 2000 samples are generated to estimate the FPF. Meanwhile, WIS is also applied. The instrumental sampling function for WIS is Gaussian distributed and centered on design point $\boldsymbol{x}^*$ [23]. Note that, the difference between ASI-IS and WIS is that they are carried out in different spaces, i.e., augmented space and original space, respectively, so they obtain different numbers of failure samples. From Fig. 1, it can be seen that the FPF obtained by ASI-IS matches quite well the reference values. However, the result by WIS possesses considerable error when $\theta \in [0.5,2]$. The reason is illustrated in Fig. 3 which shows the failure samples generated by the different approaches. It is clearly seen that the failure samples generated by WIS are located in one important region, while the region corresponding to $x_1 \in [2, 3]$ contains no samples. Note that this region will be the importance failure region when $\theta \in [0.5,2]$. Failure in exploring this region leads to underestimate the FPF over $\theta \in [0.5,2]$. The proposed ASI-IS can overcome this disadvantage, as it can generate failure samples in both important regions (see Fig. 3). A similar phenomenon can be seen in the figures related with ASI-SS and WSS. The reason behind this is that, the proposed approach can explore all the important failure regions (corresponding to different values of $\theta \in [\underline{\theta}, \bar{\theta}]$) by sampling in the augmented space, while the weighted approach may just concentrate in an important failure region corresponding to $\theta = \theta_0$. Thus, the superiority of accuracy and robustness of the proposed approach over the weighted approach are well demonstrated through this example.

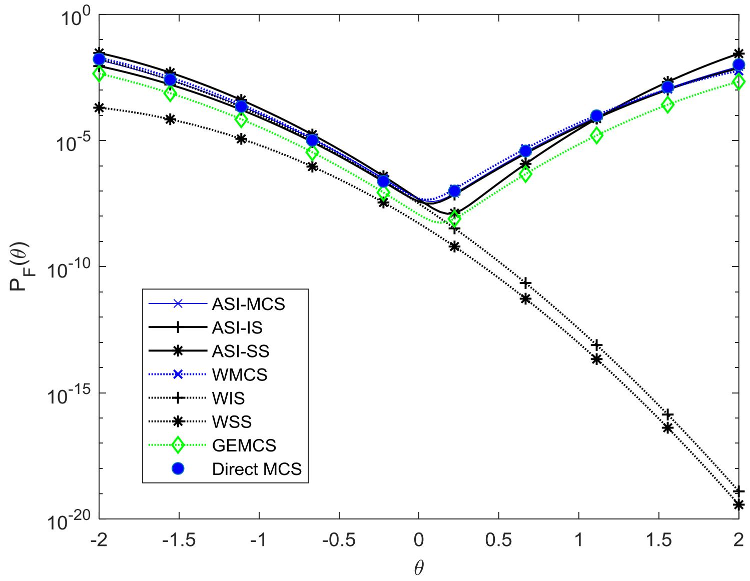

Fig. 1. FPF results obtained by different approaches for Example 1

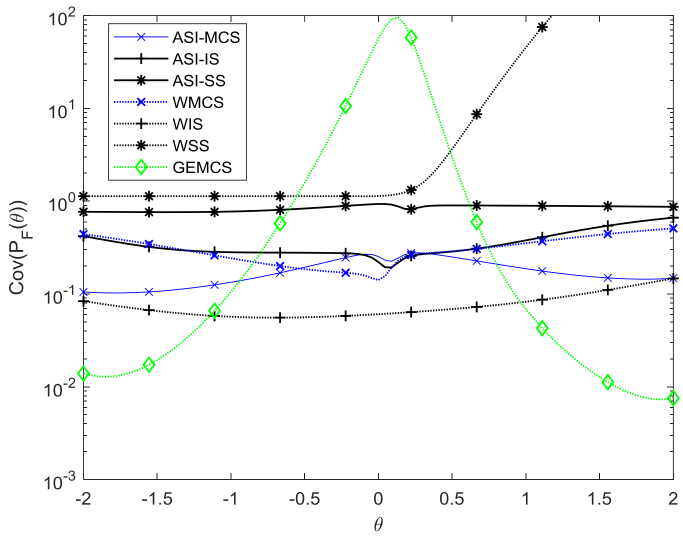

Fig. 2. The C.o.v.'s of FPF results obtained by different approaches for Example 1

Also, the proposed ASI procedure is compared with the maximum entropy estimation which is used in [26]. Fig. 4 shows the posterior distribution $\varphi(\boldsymbol{\theta}|F)$ obtained by these methods. 'MaxEnt-MCS' and 'MaxEnt-SS' refer to the approaches that use Maximum Entropy estimate to fit the target distribution based the failure samples generated from MCS and SS, respectively. Note that the same set of samples are used for different approaches, i.e., both ASI-MCS and MaxEnt-MCS use the same set of 163 failure samples, while both ASI-SS and MaxEnt-SS use the same set of 183 failure samples. It can be seen that the results by the proposed ASI-MCS and ASI-SS are consistent with the reference result. However, the results by the maximum entropy estimation have remarkable errors in this example. Note that as there are less than 200 failure samples, methods based on distribution fitting may lead to considerable errors. On the contrary, the proposed procedure actually calculates an integral by means of sampling and accurate estimates can be obtained. Therefore, the advantages on accuracy and efficiency of the proposed procedure have been clearly shown.

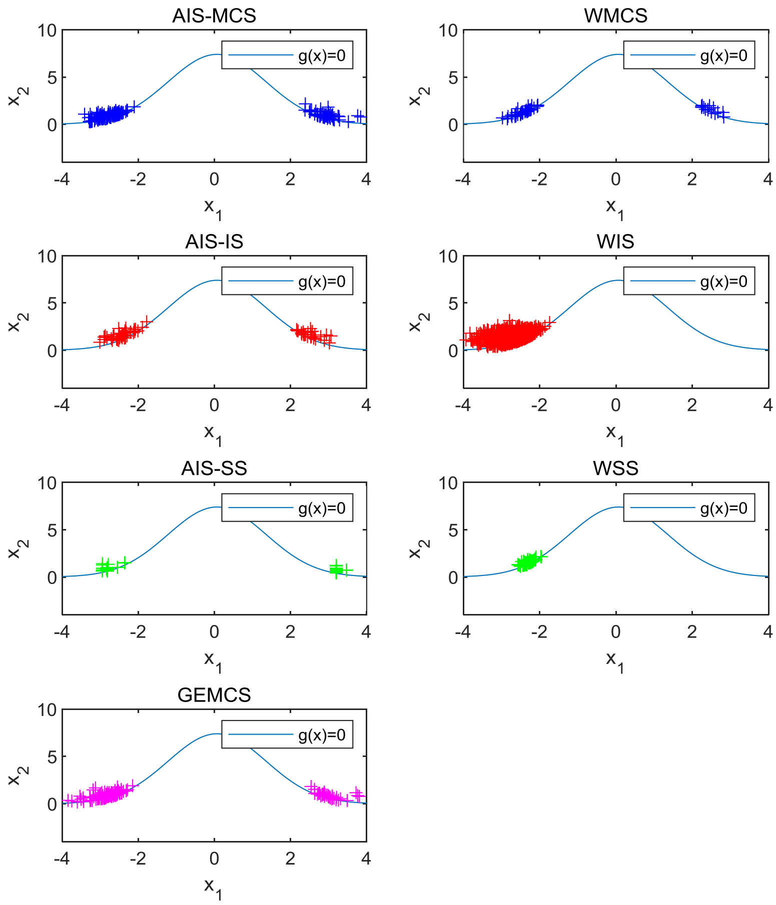

Fig. 3. The failure samples generated by different approaches for Example 1

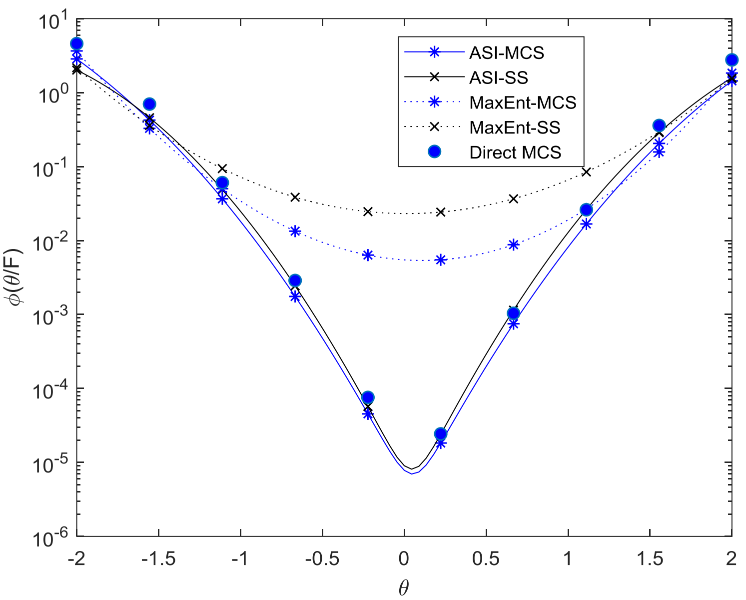

Fig. 4. The distribution $\varphi(\boldsymbol{\theta}|F)$ obtained by different approaches for Example 1

4.2 Example 2: Automobile front axle

Front axle is an important component of automobile that bears heavy loads [31] (Fig. 5). An I-beam is often used in the design of front axle due to its high bend strength and light weight. As shown in Fig. 5, a critical component of the axle is located in the I-beam part. To test the static strength of the front axle, the limit-state function can be expressed as

$$g(\boldsymbol{x}) = \sigma_s - \sqrt{\sigma^2 + 3\tau^2} \tag{31}$$

where σ_S is the limit-state stress associated with yielding. According to the material property of the front axle, the limit stress of yielding σ_S is 680 MPa. The maximum normal stress and shear stress are $\sigma = M/W_x$ and $\tau = T/W_\rho$, where M and T are bending moment and torque, respectively, W_x and W_ρ are section factor and polar section factor, respectively, which are given as [32]

$$W_x = \frac{a(h-2t)^3}{6h} + \frac{b}{6h}[h^3 - (h-2t)^3] \tag{32}$$

$$W_\rho = 0.8bt^2 + 0.4[a^3(h-2t)/t] \tag{33}$$

The geometry variables of I-beam a, b, t, h and the load M and T are independent variables with distribution parameters listed in Table 3. Note that all the variables are restricted to positive value due to physical reason, actually they are all truncated variables.

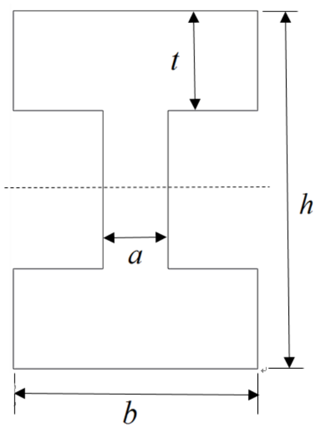

Fig. 5. Diagram of automobile front axle

Table 3. The distribution information of the random variables in Example 2

Random variable	a (mm)	t (mm)	b (mm)	h (mm)	M (KN·m)	T (KN·m)
Location parameter	θ_1=μ_a	θ_3=μ_t	65	85	3.65	θ_4=μ_T
Scale parameter	θ_2=σ_a	1.5	6.5	8.5	0.27	0.24
Distribution	Normal	Normal	Normal	Normal	Gumbel	Gumbel

The design parameters given in Table 3 include the mean value and standard deviation of the normal variable, and also the location parameter of the non-normal distributed variable. The design domains are $\theta_1 = \mu_a \in [11, 15]$mm, $\theta_2 = \mu_t \in [0.8, 1.6]$mm, $\theta_3 = \mu_t \in [12, 18]$ mm and $\theta_4 = \mu_T \in [2.8, 3.8]$ KN·m, respectively.

Table 4. Comparisons of different methods for Example 2

Methods	No. of samples	No. of failure samples
ASI-MCS	10^5	825
ASI-IS	3000	107
ASI-SS	1000×3	790
WMCS	10^5	38
WIS	3000	1371
WSS	1000×4	310
Direct MCS	10^6×10	-

The FPF is estimated by means of the proposed approaches, ASI-MCS, AIS-IS and AIS-SS. For comparison, the weighted approaches, WMCS, WIS and WSS are also applied. In addition, direct MCS is carried out considering ten independent simulation runs to generate point wise failure probability estimates which are regarded as the reference results.

Fig. 6 shows the FPF results obtained by different approaches. Note that the FPF is a four-dimensional function, and in the figure, the FPF with respect to each dimension is shown (while others are fixed at the center values of the design intervals). Information on the implementation details for each approach is listed in Table 4. It is seen from the figure that the results by the proposed approaches are in good agreement with the reference results (denoted by dots).

AIS-MCS is implemented considering a total of $N = 10^5$ samples, and $N_F = 825$ failure samples are obtained. Whereas WMCS also involves 10^5 samples and only $N_F = 38$ failure samples are obtained. The reason is that AIS-MCS is carried out in the augmented space which owns a bigger failure probability than that of WMCS. A similar situation happens when comparing AIS-SS and WSS. WSS is applied with $N = 1000×4$ samples (1000 for each of the 4 levels considered). While the proposed ASI-SS uses only $N = 1000×3$ samples to obtain a satisfactory estimate of FPF. Computation can be saved when considering the formulation in the augmented space by AIS-MCS and AIS-SS. The results by the weighted approaches possess remarkable errors in this example. For instance, a large error exists when $\theta_2 \in [1, 1.6]$ and also $\theta_4 \in [3.6, 3.8]$. The reason behind such error is that, since the sampling function of the weighted approach is centered in the midpoint of the design region, it cannot cover the important failure region in the original space sufficiently well as the proposed ASI approaches do. In conclusion, from the figure and table, the results obtained shows that the proposed approaches are applicable for multiply-dimension design parameters and for both the normal and non-normal variables, also showing more accuracy than other approaches.

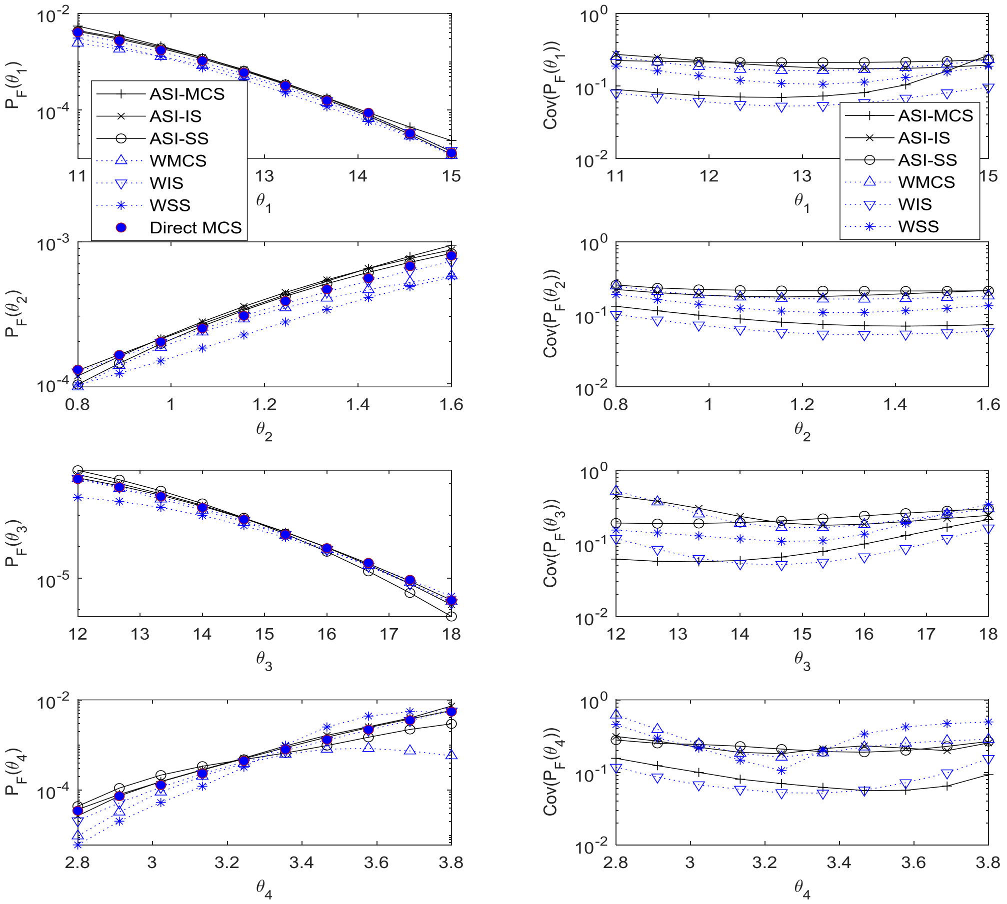

Fig. 6. The FPF estimates obtained by the different methods for Example 2

4.3 Example 3: Steel frame subject to stochastic acceleration

4.3.1 General model introduction

Consider a six-story steel frame as shown in Fig. 7, which has been previously investigated in [23]. This structure is assumed to have rigid floors and behave within the linear-elastic range, with classical damping. The damping ratio of the i-th mode is denoted as $\zeta_i(i = 1, \ldots, 6)$ Linear viscous damping elements, called passive dampers, are installed as diagonal bracings. The damping coefficient of the damper in the i-th $(i = 1, \ldots, 6)$ story is denoted as C_{di} and their values can be adjusted. In this example, the damping coefficient of each damper $C_{di}(i = 1, \ldots, 6)$, the original modal damping ratios of the building, $\zeta_i(i = 1, \ldots, 6)$, the stiffness of each story, $k_i(i = 1, \ldots, 6)$, and the mass of each story, $m_i(\mathrm{i} = 1, \ldots, 6)$, are considered as (truncated) Gaussian variables. The corresponding distribution information of these 24 structural random variables is given in Table 5.

Stochastic ground acceleration $\ddot{a}_g(t)$ which is modeled as Gaussian white noise $W(t)$ is applied to this the structure. When the peak interstory drift ratio over any of the stories of the structure exceeds a threshold level b, structural failure occurs. The damage level b=1.5% (Life-Safety) is considered herein. A duration of T =10s and time interval of $\Delta t = 0.05\mathrm{s}$ are assumed. The discrete approximation for $W(t)$ is applied at time instants $t_k = k\Delta t(k = 1, 2, \ldots, n_t)$, i.e., $W(t_k) = Z(t_k)\sqrt{2\pi S/\Delta t}$ where $S = 0.05\mathrm{m}^2/\mathrm{s}^3$ is the spectral intensity and $Z_k = Z(t_k)$ are independent identical distributed Gaussian random variables, thus there are $n_k = k/\Delta t = 200$ input random variables.

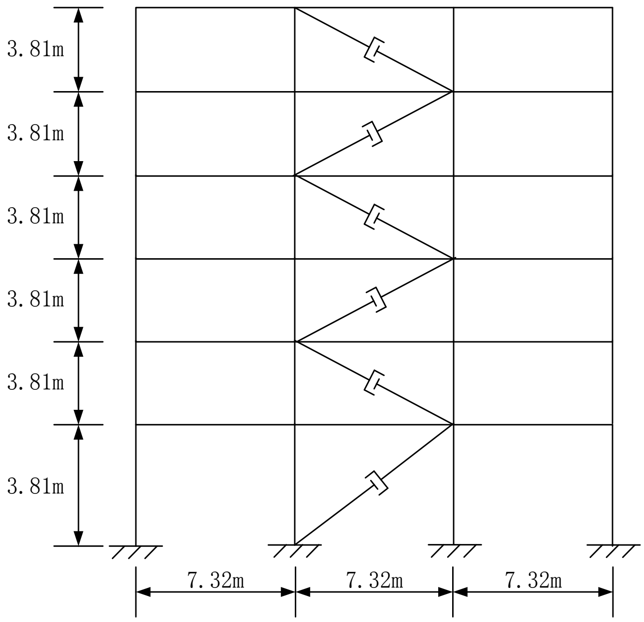

Fig. 7. Six-story steel frame structure with passive dampers

Here, the mean values of the first story stiffness is taken as the design parameter, i.e., $\theta = \mu_{k_1}$, and the design interval for θ is [100, 200] (kN/mm). The reason for considering the mean value of the first story is that in practical situations, it dominates the calculation of the failure probability (see, e.g. [34]).

Table 5. Distribution information of structural random variables

Variable	Mean	C.o.v.
k_1 (kN/mm)	$\theta = \mu_{k_1} \in [100, 200]$	0.1
k_2 (kN/mm)	367	0.1
k_3 (kN/mm)	246	0.1
k_4 (kN/mm)	246	0.1
k_5 (kN/mm)	175	0.1
k_6 (kN/mm)	175	0.1
m_1 (ton)	283	0.1
m_2 (ton)	263	0.1
m_3 (ton)	256	0.1
m_4 (ton)	255	0.1
m_5 (ton)	247	0.1
m_6 (ton)	215	0.1
C_{di} $(i = 1, \ldots, 6)$ (10^6Ns/m)	2	0.1
ζ_i $(i = 1, \ldots, 6)$ (%)	0.05	0.1

In the following, different approaches are applied to obtain the FPF of this frame structure with respect to θ. The proposed procedure with MCS, IS and SS is carried out in augmented space, i.e., ASI-MCS, ASI-IS and ASI-SS, respectively. And also, the weighted approach with MCS, IS, and SS is applied for comparison, i.e., WMCS, WIS, WSS. And it is assumed that the steel frame structure remains linear when the case for b =1.5% is considered in this example.

4.3.2 Results and discussion

Table 6 lists the computational cost (number of the samples used) by each approach. For both the proposed ASI-procedure and weighted approach, their implementations involve the same number of samples, i.e., MCS comprises 10^4 samples, IS comprises 1000 samples, and SS comprises 1000 for each level. These number of samples are selected as they provide a reasonable C.o.v. of FPF. Note that they are carried out in different spaces, i.e., the weighted approaches are carried out in the original random variable space whereas the proposed ASI procedure is implemented in the augmented space. Different number of failure samples are obtained. Note that for both ASI-IS and WIS, the importance sampling function is constructed based on the approach proposed in [33][35], and in this context, each of the generated sample will cause failure of system.

Fig. 8 shows the FPF results obtained by different methods. In addition, the direct MCS is carried out in ten independent simulation runs to generate point wise failure probability estimates, which are regarded as the exact reference results. Mostly, all the results by different approaches are consistent with the exact results. However, for this problem, the FPF results by the weighted approach possess remarkable error in both sides of the design domain i.e., when $\theta \in [100, 110]$ and $\theta \in [190, 200]$. The reason is similar as that illustrated in Example 1: the weighted approach may just concentrate in an important failure region corresponding to $\theta = \theta_0$, while the proposed approach can explore all the important failure regions. At this point, the proposed ASI-approaches perform better than weighted approach in accuracy, as they produce more accurate results.

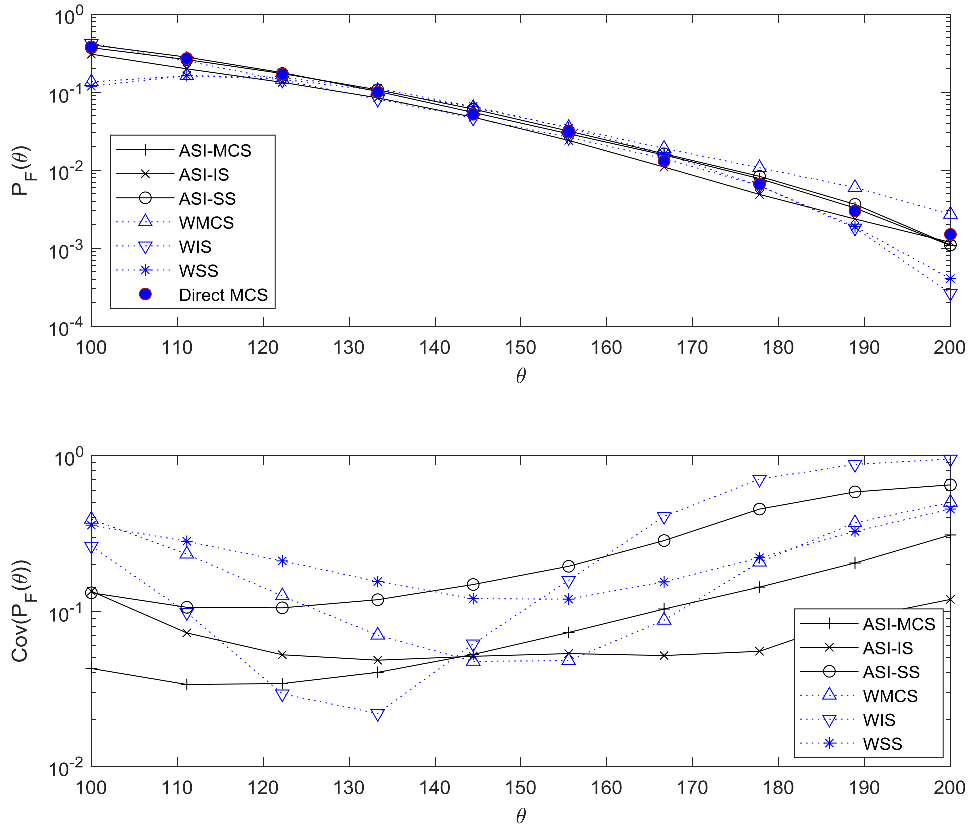

Fig. 8. The results of FPF obtained by the proposed approaches for Example 3

Table 6. Comparison of different methods for Example 3

Methods	No. of samples	No. of failure samples
ASI-MCS	10^4	927
ASI-IS	1000	1000
ASI-SS	1000×2	932
WMCS	10^4	442
WIS	1000	1000
WSS	1000×2	483
Direct MCS	10^5×10	-

5. Conclusions

In this contribution, an efficient procedure has been presented for the estimation of the structural failure probability function (FPF). It utilizes the augmented idea to handle the problem, which transforms the FPF into an expression involving three terms. An algorithm is proposed to efficiently obtain the posterior distribution term, which is the key point of the estimation of the FPF. The proposed procedure comprises three different practical implementations to estimate the FPF, i.e., ASI-MCS, ASI-IS and ASI-SS.

Numerical examples have been presented to show the advantages of the proposed approaches. The following conclusions summarize the most salient features of the proposed.

(1) It demands a single reliability analysis in augmented space, and repeated evaluations of reliability are avoided.

(2) There is no need to estimate the posterior distribution by using density fitting methods. In fact, in the proposed procedure, this posterior distribution is expressed as an integral, which allows it to be estimated directly through the failure samples.

(3) There is no need to predict (or select beforehand) the shape of the posterior probability density function.

While the results presented are encouraging, it should be note that the proposed approach also possesses some limitations. Specifically, the number of the design parameters that can be handled effectively cannot be that large, e.g. not beyond 10. This is due to the fact that estimating probability densities (as required in the proposed approach) becomes challenging in high dimensions, as documented in [25] and [36].

Future research will involve application of the proposed procedure to other fields. For example, the proposed procedure can be easily applied for solving of reliability-based optimization problems combined with a decoupling strategy. Also, it can be used in reliability sensitivity analysis, which comprises the derivative of FPF and also in imprecise reliability problems [36], which involve the estimation of the extreme value of failure probability function over a certain region.

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

The authors would like to acknowledge financial support from NSAF (Grant No. U1530122), the Aeronautical Science Foundation of China (Grant No. ASFC-20170968002), the Fundamental

Research Funds for the Central Universities of China (XMU-20720180072) and ANID (National Agency for Research and Development, Chile) under its program FONDECYT, grant number 1180271.

Appendix A

This appendix further derives the expression for $\Delta(\boldsymbol{x})=\int_{\underline{\boldsymbol{\theta}}}^{\bar{\boldsymbol{\theta}}} f(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$. Note that since all variables $\boldsymbol{x}$ are assumed independent, calculating this integral is straightforward. For example, it can be calculated using numerical algorithms. One alternative way is expressing $\Delta(\boldsymbol{x})$ as:

$$\Delta(\boldsymbol{x}) = \int_{\underline{\boldsymbol{\theta}}}^{\bar{\boldsymbol{\theta}}} f(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = E_{\boldsymbol{\theta}}\left[\frac{f(\boldsymbol{x}|\boldsymbol{\theta})}{\varphi(\boldsymbol{\theta})}\right] \tag{34}$$

where $E_{\boldsymbol{\theta}}[\cdot]$ is the expectation under $\varphi(\boldsymbol{\theta})$. This means that it can be solved through sampling.

For the particular case where $\boldsymbol{\theta}$ corresponds to the mean values of Gaussian random variables, $\Delta(\boldsymbol{x})$ can be derived further. Suppose $x_i \sim N(\theta_i, \sigma_i^2)$, and $\theta_i \sim U[\underline{\theta}_i,\ \bar{\theta}_i]$ then $\Delta(\boldsymbol{x})$ can be obtained as:

$$\Delta(\boldsymbol{x}) = \prod\nolimits_{i=1}^{n_\theta}\left[\Phi\left(\frac{\bar{\theta}_i - x_i}{\sigma_i}\right) - \Phi\left(\frac{\underline{\theta}_i - x_i}{\sigma_i}\right)\right] \tag{35}$$

where $\Phi(\cdot)$ is the cumulative probability function associated with a standard Gaussian distribution.

Appendix B

This appendix derives the C.o.v. of the estimator $\hat{P}_F(\boldsymbol{\theta})$ in Eq. (27) calculated by the proposed procedure with subset simulation.

In the following and for simplicity in notation, let $P_i = P(F_i|F_{i-1})$, $\hat{P}_i = \hat{P}(F_i|F_{i-1}), i = 1, \dots, m-1$, (where $F_0 = \Omega$), $P_m = P(F)$ and $I_{jk}^{(i)} = I_{F_i}(\boldsymbol{x}_{jk}^{(i-1)})$ where $\boldsymbol{x}_{jk}^{(i-1)}$ denotes the k-th sample in the j-th Markov chain at simulation level i.

1) C.o.v. of $\hat{P}_1$

$$\delta_1 = Cov(\hat{P}_1) = \sqrt{\frac{1-P_1}{P_1 N}} \approx \sqrt{\frac{1-\hat{P}_1}{\hat{P}_1 N}} \tag{36}$$

2) C.o.v. of $\hat{P}_i (2 \le i \le m-1)$

At $(i-1)$ th level, suppose the number of Markov chain is N_C and N / N_C samples are generated for each of these chains. It is assumed that the samples generated by different chains are uncorrelated.

The variance of $\hat{P}_i$ $(i = 2, \dots, m-1)$ is given by [11]:

$$\begin{aligned} Var(\hat{P}_i) = E\left[\hat{P}_i - P_i\right]^2 &= E\left[\frac{1}{N}\sum_{j=1}^{N_C}\sum_{k=1}^{N/N_C}\left(I_{jk}^{(i)} - P_i\right)\right]^2 \\ &= \frac{1}{N^2}\sum_{j=1}^{N_C} E\left[\sum_{1\Leftarrow 1}^{N/N_C}\left(I_{jk}^{(i)} - P_i\right)\right]^2 \end{aligned} \tag{37}$$

For the j-th Markov chain

$$E\left[\sum_{1\Leftarrow 1}^{N/N_C}\left(I_{jk}^{(i)}-P_i\right)\right]^2=\frac{N}{N_C}\left[R_i(0)+2\sum_{k=1}^{N/N_C-1}\left(1-\frac{kN_C}{N}\right)R_i(k)\right] \tag{38}$$

Substituting Eq. (38) into (37) yields:

$$Var\left(\hat{P}_i\right)=\frac{R_i(0)}{N}\left[1+2\sum_{k=1}^{N/N_C-1}\left(1-\frac{kN_C}{N}\right)\frac{R_i(k)}{R_i(0)}\right] \tag{39}$$

Based on the Markov chain samples $\{(\boldsymbol{x},\boldsymbol{\theta})_{jk}^{(i-1)}: j=1,\dots,N_C; k=1,\dots,N\,/\,N_C\}$ at the $(i-1)$th conditional level, the covariance $R_i(k)(k=0,\dots,N/N_C-1)$ can be estimated as:

$$\hat{R}_i(k)=\left(\frac{1}{N-kN_C}\sum_{j=1}^{N_C}\sum_{l=1}^{N/N_C-k}I_{jl}^{(i)}I_{j,l+k}^{(i)}\right)-\hat{P}_i^2 \tag{40}$$

3) C.o.v. of $\hat{P}_m(\boldsymbol{\theta})$

For the last level and for simplicity in notation, let $V_{jk}^{(m)}=\frac{I_F(\boldsymbol{x}^{(j)})f\left(\boldsymbol{x}^{(j)}|\boldsymbol{\theta}\right)}{\Delta(\boldsymbol{x}^{(j)})}$ and $\hat{P}_m=\hat{P}_m(\boldsymbol{\theta})$. Then the variance of $\hat{P}_m$ is given by:

$$\begin{aligned}Var\left(\hat{P}_m\right)&=E\left[\hat{P}_m-P_m\right]^2=E\left[\frac{1}{N}\sum_{j=1}^{N_C}\sum_{k=1}^{N/N_C}\left(V_{jk}^{(m)}-P_m\right)\right]^2\\&=\frac{1}{N^2}\sum_{j=1}^{N_C}E\left[\sum_{k=1}^{N/N_C}\left(V_{jk}^{(m)}-P_m\right)\right]^2\end{aligned} \tag{41}$$

For the *j*-th Markov chain,

$$E\left[\sum_{k=1}^{N/N_C}\left(V_{jk}^{(m)}-P_m\right)\right]^2=\frac{N}{N_C}\left[R_m(0)+2\sum_{k=1}^{N/N_C-1}\left(1-\frac{kN_C}{N}\right)R_m(k)\right] \tag{42}$$

Substituting Eq. (42) into (41) yields:

$$Var\left(\hat{P}_m\right)=\frac{R_m(0)}{N}\left[1+2\sum_{k=1}^{N/N_C-1}\left(1-\frac{kN_C}{N}\right)\frac{R_m(k)}{R_m(0)}\right] \tag{43}$$

Based on the Markov chain samples $\{(\boldsymbol{x},\boldsymbol{\theta})_{jk}^{(m-1)}: j=1,\dots,N_C; k=1,\dots,N\,/\,N_C\}$ at the $(m-1)$th conditional level, the covariance $R_m(k)(k=0,\dots,N/N_C-1)$ is estimated as:

$$R_m(k)\approx\hat{R}_m(k)=\left(\frac{1}{N-kN_C}\sum_{j=1}^{N_C}\sum_{l=1}^{N/N_C-k}V_{jl}^{(m)}V_{j,l+k}^{(m)}\right)-\hat{P}_m^2 \tag{44}$$

4) C.o.v. of $\hat{P}_F(\boldsymbol{\theta})$

At last, suppose all $\hat{P}_i$ $(i=1,\dots,m)$ are uncorrelated [11], then the C.o.v. of $\hat{P}_F(\boldsymbol{\theta})$ is given by:

$$Cov\left[\hat{P}_F(\boldsymbol{\theta})\right]=\sqrt{\sum_{i=1}^{m}\delta_i^2}=\sqrt{\sum_{i=1}^{m}\frac{Var(\hat{P}_i)}{P_i^2}}\approx\sqrt{\sum_{i=1}^{m}\frac{Var(\hat{P}_i)}{\hat{P}_i^2}} \tag{45}$$

where $Var(\hat{P}_i)$ can be calculated according to Eqs. (36), (39) and (43).

References

[1] Enevoldsen I, Sørensen JD. Reliability-based optimization in structural engineering. Structural Safety 1994; 15(3):169-96.

[2] Valdebenito MA, Schuëller GI. A survey on approaches for reliability-based optimization. Structural and Multidisciplinary Optimization 2010; 42(5):645-663.

[3] Valdebenito MA, Jensen HA, Hernandez HB, Mehrez L. Sensitivity estimation of failure probability applying line sampling. Reliability Engineering and System Safety 2018; 171:99-111.

[4] Papaioannou I, Breitung K, Straub D. Reliability sensitivity estimation with sequential importance sampling. Structural Safety 2018; 75: 24-34

[5] Zou T, Mahadevan S. A direct decoupling approach for efficient reliability-based design optimization. Structural and Multidisciplinary Optimization 2006; 31(3):190-200.

[6] Yuan XK, Lu ZZ. Efficient approach for reliability-based optimization based on weighted importance sampling approach. Reliability Engineering and System Safety 2014; 132: 107-114.

[7] Rackwitz R, Fiessler B. Structural reliability under combined random load sequences. Comput. Struct 1978; 9 (5):489–494.

[8] Breitung K. Asymptotic approximations for probability integrals. Probabilistic Engineering Mechanics 1989; 4 (4): 187–190.

[9] Metropolis N, Ulam S. The Monte Carlo method. J Am Stat Assoc 1949; 44(247):335–41.

[10] R.E. Melchers. Importance sampling in structural systems. Struct Safety 1989; 6(1):3–10.

[11] Au SK, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics 2001; 16(4): 263–77.

[12] Koutsourelakis PS, Pradlwarter HJ, Schuëller GI. Reliability of structures in high dimensions, Part I: algorithms and application. Probabilistic Engineering Mechanics 2004; 19: 409-417.

[13] Gasser M, Schuëller GI. Reliability-based optimization of structural systems. Mathematical Methods of Operations Research 1997; 46(3):287-307.

[14] Jensen HA. Structural optimization of linear dynamical systems under stochastic excitation: A moving reliability database approach. Computer Methods in Applied Mechanics and Engineering 2005; 194(12-16):1757-78.

[15] Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments. Stat Sci 1989; 4:409–423.

[16] Jones DR, Schonlau M, Welch WJ Efficient global optimization of expensive black-box functions. J Glob Optim 1998, 13(4):455–492.

[17] Cortes C, Vapnik VN. Support vector networks. Machine Learning, 1995, 20(3):273–297.

[18] Vapnik VN. An overview of statistical learning theory. IEEE Transaction on Neural Networks 1999, 10(5):988–998.

[19] Echard B., Gayton N., Lemaire M. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo Simulation. Struct. Saf. 2011, 33:145-154.

[20] Li HS, Lu ZZ, Yue ZF. Support vector machine for structural reliability analysis. Applied Mathematics and Mechanics 2006, 27 (10):1295-1303.

[21] Wei PF, Song JW, Bi SF, Broggi M, Beer M. Non-intrusive stochastic analysis with parameterized imprecise probability models: I. performance estimation, Mech. Syst. Signal Process 2019, 124:349–368.

[22] Wei PF, Song JW, Bi SF, Broggi M, Beer M. Non-intrusive stochastic analysis with parameterized imprecise probability models: II. Reliability and rare events analysis, Mech. Syst. Signal Process 2019, 126:227–247.

[23] Yuan XK. Local estimation of failure probability function by weighted approach, Probabilistic Engineering Mechanics 2013; 34: 1-11.

[24] Yuan XK, Zheng ZX, Zhang BQ. Augmented line sampling for approximation of failure probability function in reliability-based analysis. Applied Mathematical Modelling 2020, 80: 895-910.

[25] Au SK. Reliability-based design sensitivity by efficient simulation. Computers and Structures 2005; 83(14):1048-61.

[26] Ching J, Hsieh YH. Local estimation of failure probability function and its confidence interval with maximum entropy principle. Probabilistic Engineering Mechanics 2007; 22(1):39-49.

[27] Ching J, Hsieh YH. Approximate reliability-based optimization using a three-step approach based on subset simulation. Journal of Engineering Mechanics-ASCE 2007; 133(4):481-493.

[28] Casella G, Berger R.L. Statistical Inference, 2nd edn(M). Duxbury Press, 2001.

[29] Au SK, Beck JL. A new adaptive importance sampling scheme for reliability calculations. Structural Safety 1999, 21(2):135-158.

[30] Au SK, Beck JL, Zuev KM, Katafygiotis LS. Discussion of paper by F. Miao and M. Ghosn “Modified subset simulation method for reliability analysis of structural systems”, Structural Safety, 33:251–260, 2011. Structural Safety 2011; 34:379-380.

[31] Cai ZY. Precision design of roll-forging die and its application in the forming of automobile front axles. Journal of Materials Processing Technology 2005, 168(1):95-101.

[32] Xiao SN, Lu ZZ. Structural Reliability Analysis Using Combined Space Partition Technique and Unscented Transformation. Journal of Structural Engineering 2016, 142(11):04016089.

[33] Au SK, Beck JL. First excursion probabilities for linear systems by very efficient importance sampling. Probabilistic Engineering Mechanics 2001, 16 (3):193-207.

[34] Valdebenito, M.; Misraji, M.; Jensen, H. & Mayorga, C. Sensitivity Estimation of First Excursion Probabilities of Linear Structures Subject to Stochastic Gaussian Loading. Computers & Structures, 2021, 248: 106482.

[35] Yuan XK, Gu J, Wu MY, Zhang F. Efficient reliability-based optimization of linear dynamic system with random structural parameters, preprint submitted to elsevier (2020).

[36] Yuan XK, Faes MGR, Liu SL, Valdebenito MA, Beer M. Efficient imprecise reliability analysis using the Augmented Space Integral, Reliability Engineering & System Safety 2021, 210: 107477.