

Approximation to Multivariate Normal Integral and Its Application in Time-Dependent Reliability Analysis

Xinpeng Wei

Graduate Research Assistant

Department of Mechanical and Aerospace Engineering
Missouri University of Science and Technology
400 West 13th Street, Rolla, MO 65409-0500, USA
E-mail: weixinp@mst.edu

Daoru Han

Assistant professor

Department of Mechanical and Aerospace Engineering
Missouri University of Science and Technology
400 West 13th Street, Rolla, MO 65409-0500, USA
E-mail: handao@mst.edu

Xiaoping Du

Professor

Department of Mechanical and Energy Engineering
Indiana University – Purdue University Indianapolis
723 W. Michigan Street, Indianapolis, IN 46202-5195, USA
E-mail: duxi@iu.edu

ABSTRACT

It is common to evaluate high-dimensional normal probabilities in many uncertainty-related applications such as system and time-dependent reliability analysis. An accurate method is proposed to evaluate high-dimensional normal probabilities, especially when they reside in tail areas. The normal probability is at first converted into the cumulative distribution function of the extreme value of the involved normal variables. Then the series expansion method is employed to approximate the extreme value with respect to a smaller number of mutually independent standard normal variables. The moment generating

1

This is the author's manuscript of the work published in final form as:

Wei, Xinpeng, Daoru Han, and Xiaoping Du. 2021. "Approximation to Multivariate Normal Integral and Its Application in Time-Dependent Reliability Analysis." *Structural Safety* 88 (C). <https://doi.org/10.1016/j.strusafe.2020.102008>

function of the extreme value is obtained using the Gauss-Hermite quadrature method. The saddlepoint approximation method is finally used to estimate the cumulative distribution function of the extreme value, thereby the desired normal probability. The proposed method is then applied to time-dependent reliability analysis where a large number of dependent normal variables are involved with the use of the First Order Reliability Method. Examples show that the proposed method is generally more accurate and robust than the widely used randomized quasi Monte Carlo method and equivalent component method.

Keywords: Multivariate normal distribution, Extreme value distribution, Dimension reduction, Saddlepoint approximation, Gauss-Hermite quadrature, Reliability

1. Introduction

Many uncertainty-related applications require the evaluation of multivariate normal probabilities, for instance, the system reliability analysis [1-3] and time-dependent reliability analysis [4-26]. Both analyses predict the reliability by integrating a multivariate normal density in the safe region if the First Order Reliability Method (FORM) [27] is employed. Other areas requiring a multivariate normal probability include the extreme value distribution [28], multivariate probit model [29], multiple comparisons [30], and multiple ordinal response models [31].

No methods exist for the exact computation of the multivariate normal probability, and many numerical and sampling methods have been developed to produce approximations [32]. The existing methods can be roughly grouped into two categories: random methods and deterministic methods.

Random methods generate a large number of samples of the involved random variables and then calculate the probability based on the statistical information of the samples. The most straightforward method is the crude Monte Carlo simulation (MCS) [33]. Other random methods are more or less based on the crude MCS. They include the quasi MCS [34, 35], the importance sampling method [36-38], the subset simulation method [39], and the Bayesian MCS [40]. The random methods are generally robust, easy to use, and accurate if the sample size is large enough. But they also have some shortcomings. First, samples are usually generated randomly and hence the result is not deterministic, resulting in unrepeatable results when different seed numbers, software, or computer platforms are

used. Second, they are inefficient to estimate a small probability. This makes reliability analysis difficult since engineering applications usually require a low probability of failure or high reliability. Note that some advanced random methods, such as importance sampling method [36-38] and the subset simulation method [39], can get over this shortcoming to some extent.

Deterministic methods do not need random sampling. The equivalent component methods [41-43] are widely used. They sequentially compound two components, i.e., two of the involved normal variables, into an equivalent one, and the multivariate normal probability is eventually estimated by a univariate normal probability. The methods differ from one another mainly in the way of evaluating the correlation coefficients between the equivalent component and the other components. The correlation coefficients are determined by the sensitivity equivalency and the finite difference method [41, 43]. The finite difference method is replaced by an analytical approach [42], resulting in better accuracy and efficiency. The correlation coefficients can also be evaluated by conditional probabilities [3]. Generally, the equivalent component methods are efficient, even for high-dimensional problems. They may not be accurate when solving high-dimensional problems with small probabilities. One reason is that the equivalent component is not necessarily a normal variable and the error accumulates with the increase of the dimensionality. In addition, other deterministic methods are also available, including the first-order methods [44, 45], the product of conditional marginal probabilities [46, 47], and conditioning approximation methods [32, 48]. Their accuracy still needs to be improved when solving high-dimensional problems with small probabilities.

Overall, evaluating a multivariate normal probability is challenging in terms of accuracy and efficiency when the dimension is large and the probability is small. The objective of this work is to develop a new method to improve both accuracy and efficiency. The proposed method involves the integration of dimension reduction, the expansion optimal linear estimation (EOLE) [49], the Gauss-Hermite quadrature method [50], and the saddlepoint approximation (SPA) [51, 52]. The proposed method is then applied to and evaluated by the time-dependent reliability analysis with a large number of dependent normal variables and small probabilities.

The remaining parts of the paper are organized as follows. Section 2 gives the problem statement. Section 3 reviews the existing methods. An overview of the proposed method is given in Section 4, followed by the detailed formulations in Section 5. Section 6 gives the application to time-dependent reliability analysis. Four examples are given in Section 7 to demonstrate the proposed method. Section 8 provides conclusions.

2. Problem statement

Suppose \mathbf{Y} is a vector of N normal random variables with the mean vector $\boldsymbol{\mu}$ and the correlation matrix $\boldsymbol{\Sigma}$. The joint probability density function (PDF) $f_{\mathbf{Y}}(\mathbf{y})$ of \mathbf{Y} is given by

$$f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right) \quad (1)$$

The cumulative distribution function (CDF) $F_{\mathbf{Y}}(\hat{\mathbf{y}}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ of \mathbf{Y} is given by

$$F_{\mathbf{Y}}(\hat{\mathbf{y}}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \int_{-\infty}^{\hat{\mathbf{y}}} f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) d\mathbf{y} \quad (2)$$

Note that Eq. (2) shows an N -dimensional integral.

Without losing generality, we assume that $\hat{\mathbf{y}} = \mathbf{0}$. We also assume that all components of \mathbf{Y} have a variance of 1. Then we only focus on calculating the following integral

$$F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) = \int_{-\infty}^{\mathbf{0}} f_Y(\mathbf{y}; \boldsymbol{\mu}, \mathbf{C}) d\mathbf{y} \quad (3)$$

where \mathbf{C} is the correlation coefficient matrix of \mathbf{Y} . A general problem can be solved by Eq. (3) using the following transformation

$$F_Y(\hat{\mathbf{y}}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = F_Y(\mathbf{0}; (\boldsymbol{\mu} - \hat{\mathbf{y}}) ./ \boldsymbol{\sigma}, \mathbf{C}) \quad (4)$$

where $\boldsymbol{\sigma}$ is the standard deviation vector of \mathbf{Y} , and the operator $./$ represents the elementwise division. $F_Y(\mathbf{0}; (\boldsymbol{\mu} - \hat{\mathbf{y}}) ./ \boldsymbol{\sigma}, \mathbf{C})$ shares the same format with $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$.

Introducing the indicator function $I(\cdot)$, Eq. (4) is written as

$$F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) = \int_{-\infty}^{+\infty} I(\mathbf{y} < \mathbf{0}) f_Y(\mathbf{y}; \boldsymbol{\mu}, \mathbf{C}) d\mathbf{y} \quad (5)$$

where $I(\mathbf{y} < \mathbf{0}) = \begin{cases} 1, \mathbf{y} < \mathbf{0} \\ 0, \text{otherwise} \end{cases}$.

In practical applications, high dimensions are commonly encountered. For example, in system reliability analysis, the dimensionality may be dozens or hundreds. Many existing methods require \mathbf{C} to be full-rank. However, a non-full-rank \mathbf{C} is also frequently encountered in engineering problems. The objective of the study is to calculate the high-dimensional normal probabilities with a \mathbf{C} being full-rank or not.

3. Review of existing methods

In this section, we briefly review four commonly used methods: the crude MCS, the sequential conditioned importance sampling method (SCIS) [38], the randomized quasi

MCS method [35], and the equivalent component method [42]. The first three are random methods while the last one is a deterministic method.

3.1. Crude MCS

Crude MCS is the origin of other random methods. It first randomly generates n_s samples of \mathbf{Y} using $f_Y(\mathbf{y}; \boldsymbol{\mu}, \mathbf{C})$ and then approximates Eq. (5) by

$$F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) \approx \tilde{F} = \frac{1}{n_s} \sum_{k=1}^{n_s} I(\mathbf{y}^k < \mathbf{0}) \quad (6)$$

where \tilde{F} represents the approximation, and \mathbf{y}^k is the k^{th} sample of \mathbf{Y} . \tilde{F} itself is a random variable. Therefore, different runs of crudes MCS lead to different realizations of \tilde{F} . This is known as random error. The variation coefficient V_{MCS} of \tilde{F} is used to measure the random error and is given by

$$V_{\text{MCS}} = \sqrt{\frac{1 - \tilde{F}}{n_s \tilde{F}}} \quad (7)$$

It shows that the convergence rate of crude MCS is $O(1/\sqrt{n_s})$ [35], which is independent of N . With this feature, crude MCS does not suffer from the curse of dimensionality. The convergence rate, however, is thought to be low. For example, if the exact value of $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$ is 10^{-5} and V_{MCS} is required to be no more than 10^{-2} , then according to Eq. (7), the sample size n_s must be at least about 10^9 .

Despite its low convergence rate, MCS is widely used and is especially treated as a benchmark method for an accuracy comparison study when an exact solution is not available.

3.2. Sequential conditioned importance sampling (SCIS) method

SCIS is based on the importance sampling method and makes use of the property that conditioned on given values of arbitrary components of \mathbf{Y} , the remaining components also follow (univariate or multivariate) normal distribution [38].

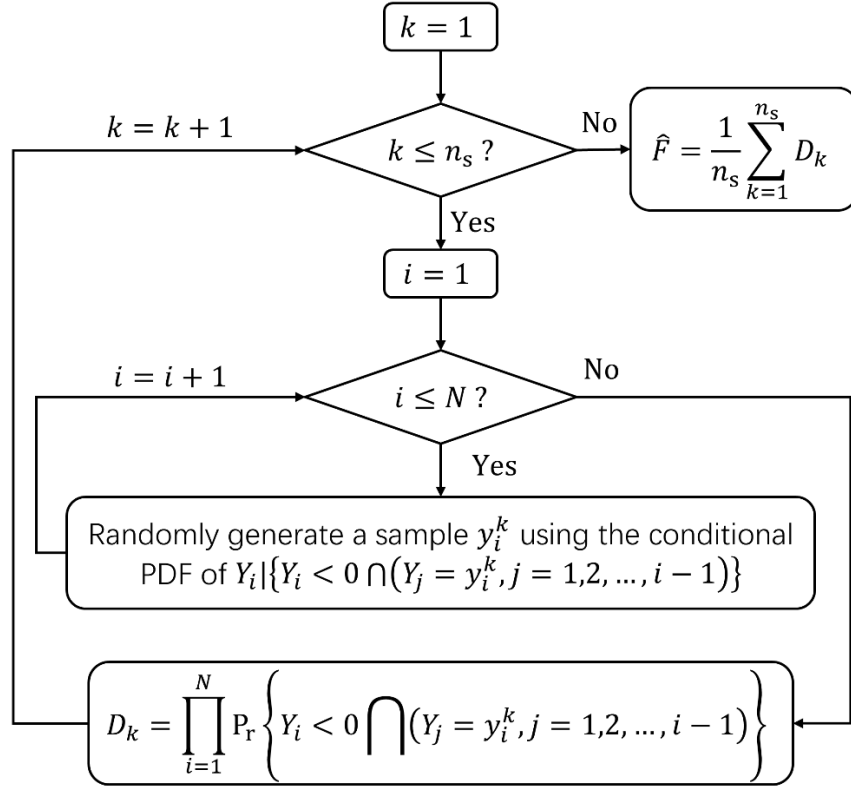


Fig. 1 Flowchart of SCIS

The flowchart of SCIS is shown in Fig. 1, where $P_r\{\cdot\}$ represents probability. Because of the property of multivariate normal variables, derivations of the conditional PDF and of D_k (in Fig. 1) are obtained easily. More details are given in Ref. [38]. Similar to crude MCS, the approximation \tilde{F} calculated by SCIS is also a random variable, with its variation coefficient V_{SCIS} given by

$$V_{\text{SCIS}} = \frac{1}{n_s \tilde{F}} \sqrt{\sum_{k=1}^{n_s} (D_k - \tilde{F})^2} \quad (8)$$

Compared to Eq. (7), Eq. (8) shows that the convergence rate of SCIS is significantly better than that of crude MCS.

3.3. Randomized quasi MCS

An effective way to improve the convergence rate of MCS is to replace the randomly generated samples by carefully selected, deterministic sequences of samples [35]. This approach is known as quasi MCS, and those samples are called low-discrepancy samples.

Fig. 2 shows 10^3 random samples and 10^3 low-discrepancy samples of \mathbf{Y} , given $\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The low-discrepancy samples are generated by Halton sequences [53].

The low discrepancy samples are regularly even while the random samples have irregular clusters. The evenness improves the convergence rate of the quasi MCS.

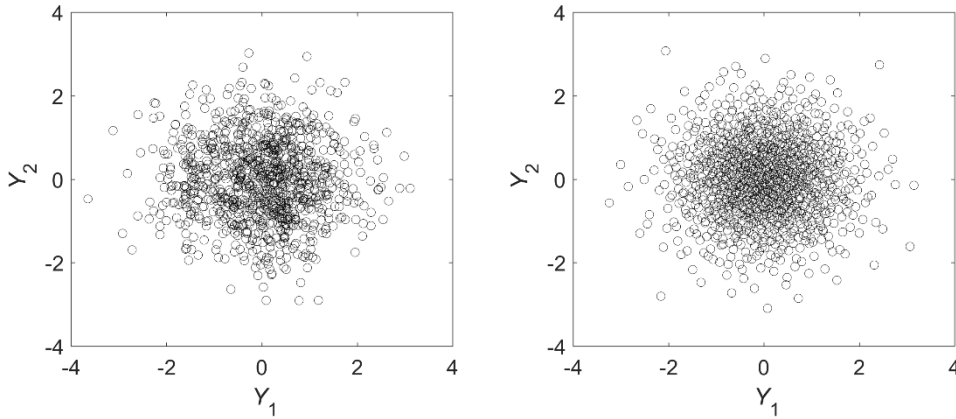


Fig. 2 Random samples (left) and low-discrepancy samples (right)

A drawback of quasi MCS is that it is hard to estimate the error. In order to estimate the error with the way similar to crude MCS, the deterministic low-discrepancy samples are randomized and the randomized quasi MCS was developed [35, 54]. It is worth mentioning that the quasi MCS developed by Genz and Bretz [35] is commonly used. This method has been coded into *pmvnorm()*, an R [55] function in R package *mvtnorm*.

3.4. Equivalent component methods

The equivalent component methods compound a pair of component normal variables into an equivalent normal variable sequentially so that the multivariate normal probability is finally estimated by a univariate normal probability. Fig. 3 shows the compounding procedure. Y_{12}^e is the equivalent component obtained by compounding Y_1 and Y_2 . Then Y_{12}^e and Y_3 are compounded into Y_{123}^e . This process continues until N normal variables have been compounded into one equivalent normal variable $Y_{123\dots N}^e$. Eq. (3) is then approximated by

$$F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) = \int_{-\infty}^0 f_e(y; \mu_e, \sigma_e^2) dy = \Phi\left(-\frac{\mu_e}{\sigma_e}\right) \quad (9)$$

where $f_e(y; \mu_e, \sigma_e^2)$, μ_e and σ_e^2 are the PDF, mean, and variance of $Y_{123\dots N}^e$, respectively, and $\Phi(\cdot)$ is the CDF of the standard normal variable.

The latest equivalent component method [42] has been proven to be effective for many problems. Assuming all the equivalent components to be normal variables, however, may introduce an unmeasurable error and hence compromise the accuracy of the method.

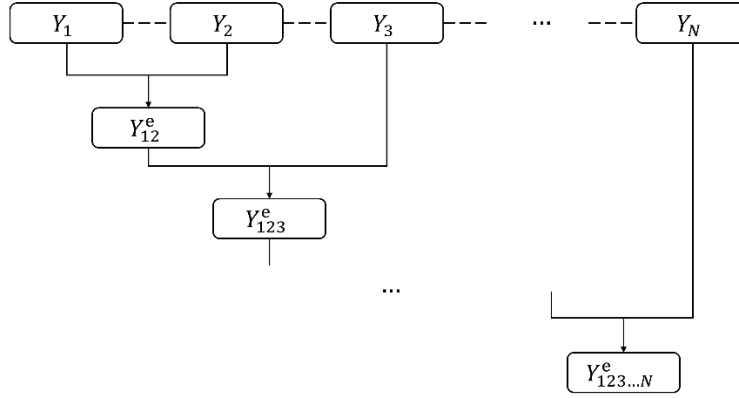


Fig. 3 Compounding procedures of the equivalent component method

4. Overview of the proposed method

The main idea of the proposed method is to convert the multidimensional probability into an equivalent extreme value probability. Eq. (3) is equivalent to

$$F_{\mathbf{Y}}(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) = \Pr \left\{ \bigcap_{i=1}^N Y_i < 0 \right\} = \Pr \{ \max(\mathbf{Y}) < 0 \} = \Pr \{ Z < 0 \} = F_Z(0) \quad (10)$$

where $Z = \max(\mathbf{Y})$ is the maximum value of \mathbf{Y} . Note that Z itself is a random variable, and we denote it by $Z(\mathbf{Y})$ since it is a function of \mathbf{Y} .

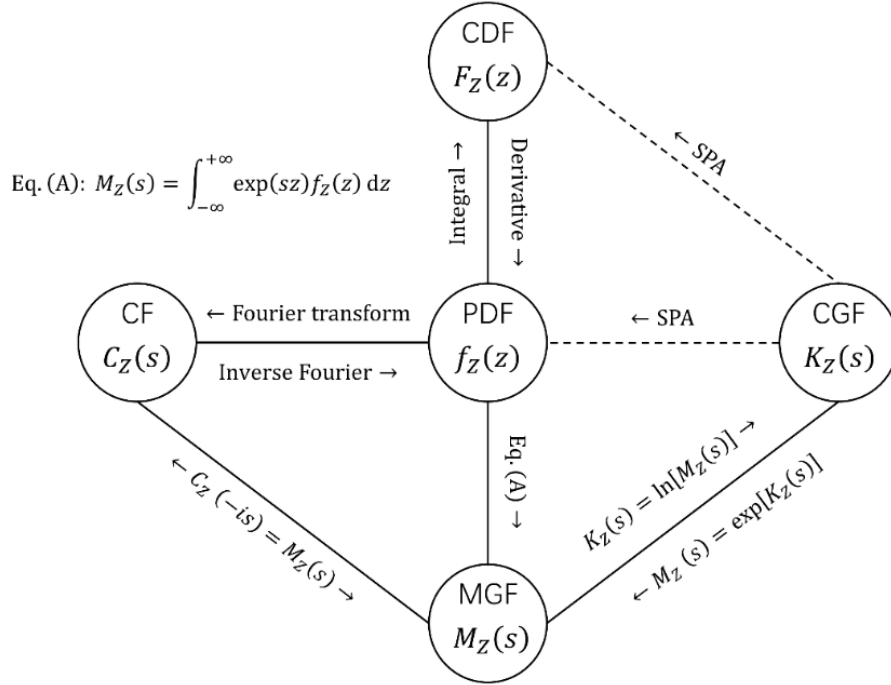


Fig. 4 Functions that fully define the distribution of Z

The distribution of Z can be determined from its PDF $f_Z(z)$, CDF $F_Z(z)$, moment generating function (MGF) $M_Z(s)$, cumulant generating function (CGF) $K_Z(s)$, or characteristic function (CF) $C_Z(s)$. The relationships among those functions are shown in Fig. 4. A solid line means a theoretically rigorous transformation between the two functions connected by the line, while a dotted line means an approximate transformation. Theoretically, once any of the five functions is obtained, the other four can also be obtained using the transformation indicated above or below a line.

The easiest starting point is the MGF given by

$$M_Z(s) = \int_{-\infty}^{+\infty} \exp(sz) f_Z(z) dz = \int_{-\infty}^{+\infty} \exp[sz(\mathbf{y})] f_Y(\mathbf{y}; \boldsymbol{\mu}, \mathbf{C}) d\mathbf{y} \quad (11)$$

Although Eq. (11) is also a high-dimensional integral similar to Eq. (5), it is much easier to calculate. The reason is that the integrand $\exp[sz(\mathbf{y})]$ is generally a continuous function,

which can be calculated effectively using quadrature methods, while the integrand $I(\mathbf{y} < \mathbf{0})$ in Eq. (5) is not. This is also the reason why we convert the multidimensional probability in Eq. (5) into the extreme value probability in Eq. (10).

Once $M_Z(s)$ is obtained, there are at least two ways to get $F_Z(z)$ or Eq. (10). As shown in Fig. 4, the first way is $M_Z(s) \rightarrow C_Z(s) \rightarrow f_Z(z) \rightarrow F_Z(z)$ and the second way is $M_Z(s) \rightarrow K_Z(s) \rightarrow F_Z(z)$. The first way, however, is not practical, and there are two reasons. First, $M_Z(s)$ calculated by the quadrature method is not exact, and neither is $C_Z(s)$, which generally has complex output values and may suffer from large errors. Second, currently there are no robust and effective methods to transform $C_Z(s)$ into $f_Z(z)$ using the inverse Fourier transform, especially when $C_Z(s)$ is not exact. In contrast, the second way is effective. The reason is that a simple logarithm is needed to obtain $K_Z(s)$ from $M_Z(s)$, and SPA is an accurate and efficient method to approximate $F_Z(z)$ from $K_Z(s)$, especially at the tails of $F_Z(z)$. Therefore, we use the latter way to calculate $F_Z(z)$.

Calculating Eq. (11), however, needs a heavy computational effort, since it may be a high-dimensional integral. To solve this problem, we propose two approaches to reduce the dimension of the integral. The first approach is to screen the random variables in \mathbf{Y} and remove the ones that barely contribute to the desired $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$. The second approach is to transform the integral from the \mathbf{Y} space, or physical space, into the eigenspace, using truncated series expansion of \mathbf{Y} . With the latter approach, we can further reduce the dimension of the integral. This approach can usually reduce the dimension significantly because \mathbf{C} is a low-rank matrix in many engineering problems.

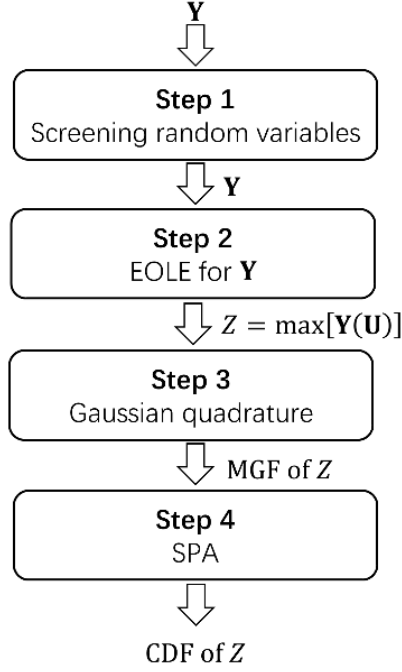


Fig. 5 Abstract flowchart of the proposed method

An abstract flowchart of the proposed method is given in Fig. 5. Step 1 screens random variables in \mathbf{Y} . (See Subsection 5.1.) Note that after the screening step, we still use \mathbf{Y} for the remaining random variables for convenience of presentation. In Step 2, we use EOLE to expand \mathbf{Y} and then truncate the expansion to N' orders. This step transforms the integral in Eq. (11) from the \mathbf{Y} space into the \mathbf{U} space (the eigenspace). (See Subsection 5.2.) In Step 3, the Gauss-Hermite quadrature is used to calculate the MGF of Z in the eigenspace. (See Subsection 5.3.) In Step 4, SPA is employed to transform the MGF into CDF of Z , and finally the desired $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$ is obtained through Eq. (18). (See Subsection 5.4)

There are four advantages of the proposed method. First, it can calculate multidimensional normal probabilities with arbitrary dimension N , as long as the dimension N' of the truncated eigenspace is not large. This is practical for dealing with engineering problems where the number of basic random variables is not large. Second,

the method is accurate even when calculating very small probabilities because SPA is able to recover CDF from MGF with sufficient accuracy, especially at tails of CDF. Third, it is generally more efficient than the abovementioned random methods, when $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$ is small, such as 10^{-5} and smaller. The reason is that random methods need large sample size to guarantee the accuracy when calculating small probabilities. Fourth, the result calculated by the proposed method is deterministic, instead of random by a random method.

5. Formulation of the proposed method

In this section we give all details involved in the steps shown in Fig. 5.

5.1. Step1: Screening random variables

The screening procedure is aimed at reducing the dimension by removing those components of \mathbf{Y} that are not important to $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$. If $\Pr\{Y_i < 0\}$ is almost equal to 1, or equivalently if $\Pr\{Y_i > 0\}$ is sufficiently small, then removing Y_i will not significantly affect the accuracy.

Since $\Pr\{Y_i > 0\}$ measures the importance of Y_i to $F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C})$, the most important component Y_* is determined by

$$Y_* = \arg \max_i \Pr\{Y_i > 0\} \quad (12)$$

and $\Pr\{Y_* > 0\}$ is used as a benchmark to screen the other components of \mathbf{Y} . The screening criterion is given by

$$\Pr\{Y_i > 0\} < c\Pr\{Y_* > 0\} \quad (13)$$

where c is a hyperparameter taking a small value, such as 10^{-4} . Theoretically, the smaller is c , the higher accuracy will we have. However, if c is too small, the screening step will not effectively screen out unimportant normal random variables. If Y_i meets the criterion, it will be removed. Since \mathbf{Y} are normal variables, Eq. (13) is equivalent to

$$\Phi(\mu_i) < c\Phi(\mu_*) \quad (14)$$

where μ_* is the mean value of Y_* . Note that we have assumed in Section 2 that all components of \mathbf{Y} have variance of 1, so Eq. (14) does not involve the variance of \mathbf{Y} . Fig. 6 shows an example of the screening of \mathbf{Y} . Initially, there are $N = 300$ components in \mathbf{Y} . Only 68 components shown by the circles, however, do not satisfy the criterion in Eq. (13). Therefore, only 68 components are kept and the other 232 ones are removed, reducing the dimensionality from $N = 300$ to $N = 68$. Note that after the screening step, we also use \mathbf{Y} to denote the remaining random variables for convenience.

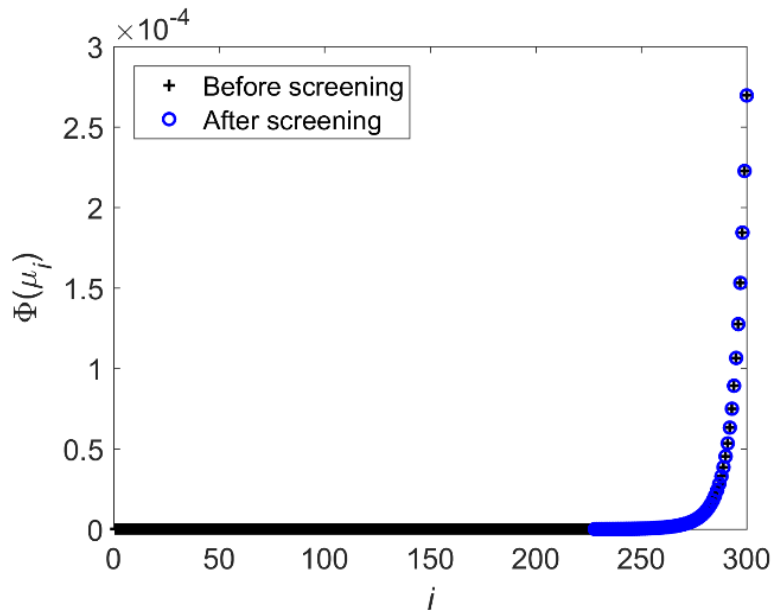


Fig. 6 An example of the screening step

5.2. Step 2: Series expansion with EOLE

The purpose of EOLE [49] is to transform the integral in Eq. (11) from the \mathbf{Y} space into the eigenspace. A key step of EOLE is the eigendecomposition [56] of \mathbf{C} . In linear algebra, eigendecomposition, or spectral decomposition, is the factorization of a matrix into a canonical form. With the decomposition, a square matrix \mathbf{C} is represented in terms of its eigenvalues and eigenvectors. A (non-zero) vector \mathbf{V} is an eigenvector of \mathbf{C} if it satisfies the linear equation

$$\mathbf{C}\mathbf{V} = \lambda\mathbf{V} \quad (15)$$

where λ is the eigenvalue corresponding to \mathbf{V} . The eigenvalues are obtained through solving the following equation

$$\det(\mathbf{C} - \lambda\mathbf{I}) = 0 \quad (16)$$

where $\det(\cdot)$ represents determinant, and \mathbf{I} is an identity matrix with the same size as \mathbf{C} . The number of eigenvalues obtained by solving Eq. (16) is N_{rank} , the rank of \mathbf{C} . Once an eigenvalue is obtained, we can calculate its corresponding eigenvector by substituting it into Eq. (15).

With the eigendecomposition, we obtain N_{rank} eigenvalues λ and N_{rank} eigenvectors $\mathbf{V}_j, j = 1, 2, \dots, N_{\text{rank}}$. Note that the eigenvalues are sorted from the largest to the smallest. Then the EOLE expansion of \mathbf{Y} is given by

$$Y_i(\mathbf{U}) = \mu_i + \sum_{j=1}^{N_{\text{rank}}} \frac{U_j}{\sqrt{\lambda_j}} \mathbf{V}_j^T \mathbf{C}(:, j), i = 1, 2, \dots, N \quad (17)$$

where $\mathbf{U} = [U_1, U_2, \dots, U_j, \dots, U_{N_{\text{rank}}}]$ are N_{rank} mutually independent standard normal variables, λ_j is the j^{th} eigenvalue, and $\mathbf{C}(:, j)$ is the j^{th} column of \mathbf{C} . The j^{th} eigenvalue λ_j measures how sensitive \mathbf{Y} is to U_j .

For a full-rank \mathbf{C} , $N_{\text{rank}} = N$, and hence there are $N + 1$ terms in the expansion. For a non-full-rank \mathbf{C} , with $N_{\text{rank}} < N$, there are only N_{rank} non-zero eigenvalues, and therefore there are less than $N + 1$ terms in the expansion. In practical engineering, however, not all the N_{rank} eigenvalues are at the same level of magnitude. Excluding the μ_i term, we only keep the first N' terms that have large eigenvalues, because they contribute most to the expansion. Hereafter, we let \mathbf{U} denote $[U_1, U_2, \dots, U_j, \dots, U_{N'}]$. The uncertainty of \mathbf{Y} is mainly propagated from the uncertainty of \mathbf{U} , and hence we call \mathbf{U} significant basic random variables.

Specifically, N' is determined as the smallest integer that meets the criterion as follows

$$\left(\sum_{j=1}^{N'} \lambda_j \right) / \left(\sum_{j=1}^{N_{\text{rank}}} \lambda_j \right) \geq \eta \quad (18)$$

where η is a hyperparameter determining the accuracy of the expansion. It takes a value close to, but not larger than, 1. The smaller is η , the less accurate is the expansion. If $\eta = 1$, the expansion is exact. Typically, η is set to 0.9999. When N' has been determined by Eq. (18), the truncated EOLE expansion is given by

$$Y_i(\mathbf{U}) = \mu_i + \sum_{j=1}^{N'} \frac{U_j}{\sqrt{\lambda_j}} \mathbf{v}_j^T \mathbf{C}(:, j), i = 1, 2, \dots, N \quad (19)$$

With the truncated EOLE expansion, each Y_i is a function of \mathbf{U} , and hence $Z(\mathbf{Y}) = \max(\mathbf{Y})$ is also a function of \mathbf{U} . Then Eq. (11) is converted into

$$M_Z(s) = \int_{-\infty}^{+\infty} \exp[sz(\mathbf{u})]f_U(\mathbf{u})d\mathbf{u} \quad (20)$$

where $f_U(\mathbf{u})$ is the PDF of \mathbf{U} , i.e., the PDF of N' -dimensional mutually independent standard normal variables.

Eq. (20) shows an N' -dimensional integral. Compared to Eq. (11) for an N -dimensional integral, Eq. (20) is more efficient because of the dimension reduction. With the dimension reduction, the efficiency of the proposed method mainly depends on N' instead of N . Intuitively, a larger N will lead to a larger N' . However, there is no direct relationship between N' and N . It is the number of significant eigenvalues of \mathbf{C} that directly determines N' . A \mathbf{C} with a dimension of 1,000 by 1,000 may have only two significant eigenvalues and hence $N' = 2$, while another \mathbf{C} with a dimension of 5 by 5 may have up to 5 significant eigenvalues and hence $N' = 5$.

5.3. Step 3: Calculate MGF with Gauss-Hermite quadrature

The purpose of this step is to calculate the multidimensional integral in Eq. (20) efficiently. Gauss-Hermite quadrature is a form of Gaussian quadrature for approximating the integrals with the following format

$$I = \int_{-\infty}^{+\infty} g(u)\exp(-u^2)du \quad (21)$$

where I is the integral result, $g(u)$ is a smooth and continuous function of u , and $\exp(-u^2)$ is called a weight function. With the Gauss-Hermite quadrature, Eq. (21) is approximated by

$$I = \sum_{q=1}^Q w^{(q)} g[u^{(q)}] \quad (22)$$

where Q , the quadrature order, is the number of quadrature points used, $w^{(q)}$ is the q^{th} weight, and $u^{(q)}$ is the q^{th} quadrature point. Table 1 shows the quadrature points and weights for some quadrature orders.

When the weight function is the PDF of the standard normal variable, i.e., $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$, instead of $\exp(-u^2)$, the quadrature weights and points should be modified accordingly. The modification rule is simply multiplying the weights by $\frac{1}{\sqrt{\pi}}$, and the points by $\sqrt{2}$. For example, the weights and points in Table 1 are modified to that in Table 2.

Table 1 Gauss-Hermite quadrature points and weights

Quadrature order Q	Quadrature point $u^{(q)}$	Quadrature weight w_q
1	0	1.772453
2	± 0.707107	0.886227
3	0	1.81635
	± 1.22474	0.295409
4	± 0.524648	0.804914
	± 1.65068	0.081312

Table 2 Modified Gauss-Hermite quadrature weights and points

Quadrature order Q	Quadrature point $u^{(q)}$	Quadrature weight w_q
1	0	$1.772453/\sqrt{\pi}$
2	$\pm 0.707107\sqrt{2}$	$0.886227/\sqrt{\pi}$
3	0	$1.81635/\sqrt{\pi}$
	$\pm 1.22474\sqrt{2}$	$0.295409/\sqrt{\pi}$
4	$\pm 0.524648\sqrt{2}$	$0.804914/\sqrt{\pi}$
	$\pm 1.65068\sqrt{2}$	$0.081312/\sqrt{\pi}$

The integral in Eq. (20) is N' -dimensional, and the unidimensional formulation in Eq. (22) is extended to its multidimensional counterpart using the tensor product rule. The N' -dimensional Gauss-Hermite quadrature formulation is given by

$$I = \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \dots \sum_{q_{N'}=1}^{Q_{N'}} w_1^{(q_1)} w_2^{(q_2)} \dots w_{N'}^{(q_{N'})} g \left[u_1^{(q_1)}, u_2^{(q_2)}, \dots, u_{N'}^{(q_{N'})} \right] \quad (23)$$

where Q_j is the quadrature order in the j^{th} dimension. Therefore, Eq. (20) is approximated by

$$M_Z(s) = \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \dots \sum_{q_{N'}=1}^{Q_{N'}} w_1^{(q_1)} w_2^{(q_2)} \dots w_{N'}^{(q_{N'})} \exp \left\{ sZ \left[u_1^{(q_1)}, u_2^{(q_2)}, \dots, u_{N'}^{(q_{N'})} \right] \right\} \quad (24)$$

Note that the weight function $f_U(\mathbf{u})$ in Eq. (20) is the PDF of N' mutually independent standard normal variables, and Eq. (24) should use the modified quadrature weights and points. The total number N_Q of quadrature points is equal to $\prod_{j=1}^{N'} Q_j$.

Generally, the higher are the quadrature orders $Q_j, j = 1, 2, \dots, N'$, the higher is the accuracy. Higher quadrature orders, however, mean lower efficiency. Therefore, a good tradeoff is needed. Since the j^{th} eigenvalue λ_j of \mathbf{C} measures how sensitive \mathbf{Y} is to U_j , as mentioned in Subsection 5.2, λ_j also measures how sensitive Z is to U_j . Hence, we assign values to $Q_j, j = 1, 2, \dots, N'$, according to the corresponding eigenvalues.

To determine $Q_j, j = 1, 2, \dots, N'$, we need the maximum and minimum allowable values Q_{\max} and Q_{\min} . Since λ_1 is the largest eigenvalue, we set Q_1 to Q_{\max} . $Q_j, j = 2, 3, \dots, N'$, are determined by

$$Q_j = \max \left\{ \text{round} \left(\frac{\lambda_j}{\lambda_1} Q_1 \right), Q_{\min} \right\} \quad (25)$$

where $\text{round}(\cdot)$ rounds its input value to the nearest integer. Eq. (25) shows that the larger is $\frac{\lambda_j}{\lambda_1}$, the larger is Q_j , but Q_j cannot be smaller than Q_{\min} . The specific values of the two hyperparameters Q_{\max} and Q_{\min} are dependent on the requirement of the calculation accuracy and efficiency.

5.4. Step 4: Transform MGF to CDF using SPA

SPA is a powerful tool to transform MGF to CDF as well as to PDF. Although the theory behind SPA is complicated, its implementation is straightforward.

First, the MGF $M_Z(s)$ is transformed to CGF $K_Z(s)$ through

$$K_Z(s) = \ln[M_Z(s)] \quad (26)$$

Then the first derivative $\dot{K}_Z(s)$ of $K_Z(s)$ is given by

$$\dot{K}_Z(s) = \frac{\dot{M}_Z(s)}{M_Z(s)} \quad (27)$$

where $\dot{M}_Z(s)$ is the first derivative of $M_Z(s)$ and is given by

$$\begin{aligned} \dot{M}_Z(s) = & \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \dots \sum_{q_{N'}=1}^{Q_{N'}} w_1^{(q_1)} w_2^{(q_2)} \dots w_{N'}^{(q_{N'})} Z \left[u_1^{(q_1)}, u_2^{(q_2)}, \dots, u_{N'}^{(q_{N'})} \right] \\ & * \exp \left\{ sZ \left[u_1^{(q_1)}, u_2^{(q_2)}, \dots, u_{N'}^{(q_{N'})} \right] \right\} \end{aligned} \quad (28)$$

The second derivative $\ddot{K}_Z(s)$ of $K_Z(s)$ is given by

$$\ddot{K}_Z(s) = -\frac{\dot{M}_Z^2(s)}{M_Z^2(s)} + \frac{\ddot{M}_Z(s)}{M_Z(s)} \quad (29)$$

where $\ddot{M}_Z(s)$ is the second derivative of $M_Z(s)$ and is given by

$$\begin{aligned} \ddot{M}_Z(s) = & \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \dots \sum_{q_{N'}=1}^{Q_{N'}} w_1^{(q_1)} w_2^{(q_2)} \dots w_{N'}^{(q_{N'})} z^2 \left[u_1^{(q_1)}, u_2^{(q_2)}, \dots, u_{N'}^{(q_{N'})} \right] \\ & * \exp \left\{ s z \left[u_1^{(q_1)}, u_2^{(q_2)}, \dots, u_{N'}^{(q_{N'})} \right] \right\} \end{aligned} \quad (30)$$

Daniels [57] derived the SPA to the PDF $f_Z(z)$ of Z as

$$f_Z(z) = \left[\frac{1}{2\pi \dot{K}_Z(s^*)} \right]^{\frac{1}{2}} \exp[K_Z(s^*) - s^* z] \quad (31)$$

where s^* , known as the saddlepoint, is the solution to the equation given by

$$\dot{K}(s) = z \quad (32)$$

The bisection method [58] is employed to solve Eq. (32). Apart from $f_Z(z)$, the CDF $F_Z(z)$ is given by

$$F_Z(z) = \Phi[w(z)] + \phi[w(z)] \left(\frac{1}{w(z)} - \frac{1}{v} \right) \quad (33)$$

where $\phi(\cdot)$ is the PDF of the standard normal variable,

$$w(z) = \text{sign}(s^*) \{2[s^* z - K_Z(s^*)]\}^{\frac{1}{2}} \quad (34)$$

and

$$v = s^* [\dot{K}_Z(s^*)]^{\frac{1}{2}} \quad (35)$$

Since we only need to calculate $F_Z(0)$, we can simply set $z = 0$ in Eqs. (32), (33) and (34).

Once $F_Z(0)$ is obtained, we also obtain the desired multivariate normal probability

$$F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) = F_Z(0).$$

An important property of SPA is that it is able to convert MGF to CDF with sufficient accuracy, especially at the tails of CDF [51, 57]. Some studies showed that in some cases,

SPA has tail exactness [59]. This property makes the proposed method able to calculate very small probabilities with high accuracy.

6. Application in time-dependent reliability analysis

Time-dependent reliability measures the probability that a component or system does not fail within a given period of time. With different theories, existing methods to time-dependent reliability analysis are roughly grouped into simulation methods [16, 17, 21, 33], surrogate model methods [6, 11, 12, 18-20], extreme value methods [13, 22, 23, 25], outcrossing rate methods [4, 7, 10, 15], and equivalent Gaussian process methods [5, 8, 14], etc.

Simulation methods are straightforward. A large number of samples of Y are generated first, whose statistic information is then used to estimate the reliability or the probability of failure. This group of methods are generally accurate as long as the sample size is sufficiently large. Generating a large number of samples, however, is usually expensive or even unaffordable, especially when the limit-state function is an expensive black-box function. To deal with this problem, surrogate model methods train a computationally cheap surrogate model to replace the original expensive limit-state function. Once the surrogate model is well trained, the time-dependent reliability may be estimated accurately and efficiently. This group of methods, however, introduce some additional issues, such as the design of experiment, training scheme, learning function, and convergence criteria, etc.

Extreme value methods convert the time-dependent problems into static ones by calculating the extreme values of the limit-state function with respect to time. Generally,

the calculation of extreme values needs a global optimization with respect to time. It limits the application of this group of methods since global optimization may not be efficient.

Outcrossing rate methods are traditional methods for time-dependent reliability analysis and are widely used. The methods are efficient if they are used with FORM. Their accuracy may not be good for problems with low reliability because the dependence among crossing events is neglected. On the contrary, the autocorrelation of the limit-state function is considered in equivalent Gaussian process methods, and hence more accurate results can be obtained. The procedures of equivalent Gaussian process methods are straightforward. FORM is first employed to convert the limit-state function into a Gaussian process whose discretization is a vector of correlated normal variables, and then a high-dimensional normal integral is used to calculate the reliability.

The existing equivalent Gaussian process methods mainly differ in the way the high-dimensional normal integral is estimated. Hu and Du [5] employed the crude MCS. Jiang et al. [14] employed the quasi MCS [35]. Gong and Frangopol [8] employed the equivalent component method. In this study, we apply the proposed method to improve the accuracy of equivalent Gaussian process methods without a random sampling method.

The reliability is predicted by a limit-state function given by

$$Y = G(\mathbf{X}, \mathbf{P}(t), t) \quad (36)$$

where \mathbf{X} are the basic input random variables, $\mathbf{P}(t)$ are the input random processes, and t is time. Generally, Y is a random process. The time-dependent reliability R over the time interval $[\underline{t}, \bar{t}]$ is given by

$$R = \Pr\{Y < 0, \forall t \in [\underline{t}, \bar{t}]\} \quad (37)$$

To calculate R numerically, we need to discrete $[t, \bar{t}]$ into N points $t_i, i = 1, 2, \dots, N$, where $t_1 = \underline{t}$ and $t_N = \bar{t}$. Then the random process Y is discretized into N random variables $Y_i = G(\mathbf{X}, \mathbf{P}(t_i), t_i), i = 1, 2, \dots, N$. With the discretization, Eq. (37) is rewritten as

$$R = \Pr \left\{ \bigcap_{i=1}^N Y_i < 0 \right\} \quad (38)$$

Although Y_i is in general not a normal variable, we can use FORM to transform it into an equivalent normal variable with a unit variance [5]. Therefore, we always assume that Y_i is normally distributed with a unit variance without losing generality. Then Eq. (38) is equivalent to

$$R = F_Y(\mathbf{0}; \boldsymbol{\mu}, \mathbf{C}) \quad (39)$$

The details of how to calculate \mathbf{C} using FORM is given in [5]. The time dependent probability of failure $P_f = 1 - R$.

For general time-dependent reliability problems, N can be hundreds. Although N is large, the number of significant basic random variables, i.e., N' , is not necessarily large. If there are no random processes in Eq. (36), N_{rank} will be exactly equal to the dimension of \mathbf{X} , i.e., the number of basic random variables. N' is no larger than N_{rank} . $N' = N_{\text{rank}}$ only if Y is sensitive to all the basic random variables. $N' < N_{\text{rank}}$ when Y is not sensitive to at least one basic random variable. If there are input random processes, N_{rank} is dependent on not only the number of basic random variables and random processes, but also the autocorrelation functions of the input random processes.

From response Y , N' is generally determined by the correlation length l_Y of Y and the length $l_t = (\bar{t} - \underline{t})$ of time interval $[\underline{t}, \bar{t}]$. More specifically, the larger $\frac{l_Y}{l_t}$ is, the smaller N' will we have. For problems with small $\frac{l_Y}{l_t}$, N' is large and hence the proposed method may not be efficient or may even fail.

7. Numerical examples

In this section, we demonstrate the effectiveness of the proposed method using four time-dependent reliability analysis examples. The first example has the exact solution and hence we can easily test the accuracy of the proposed method. In the second example, the limit-state function is given as a Gaussian random process. The third example involves a mechanism whose inputs only contain several random variables without a random process. The last example has an implicit limit-state function, which is a black-box model evaluated by the finite element method (FEM) [60]. Exact solutions are not available for the last three examples, and hence we employ the crude MCS, using a sufficiently large sample size, to obtain accurate results, which are treated as benchmarks. In all the examples, the hyperparameters c , η , Q_{\max} , and Q_{\max} are set to 10^{-4} , 0.9999, 35, and 5, respectively. Note that there are no criteria for selecting specific values for those hyperparameters. We set those values based on our experience from many experiments. Note that all the reported results and error are about the calculation of the multinormal probabilities, so the error due to FORM approximation is not included.

The proposed method is also compared with two widely used methods. The first one is the latest version of the equivalent component method [42], which is a deterministic

method. For convenience, we denote this method by IECA (improved equivalent component method). The second one is the randomized quasi MCS developed by Genz and Bretz [35], which has been implemented in the *R* programming language and has been widely used to calculate the high-dimensional normal probabilities. We can simply call the *R* function *pmvnorm()* to calculate the desired probability. Since it is a random method whose result is dependent on the seed of the random number generator, we will run this method three times to see the differences. For convenience, we denote the three solutions from the method by RQ1, RQ2, and RQ3.

7.1. Example 1: A math example with exact solution

The limit-state function $Y(t)$ is a stationary Gaussian process with mean value $\mu(t) = b$ and standard deviation $\sigma(t) = 1$. Its autocorrelation coefficient function $\rho(t_1, t_2)$ is given by

$$\rho(t_1, t_2) = \cos(t_1 - t_2) \quad (40)$$

The time interval $[\underline{t}, \bar{t}] = [0, 2\pi]$ s. $Y(t)$ is a function of $\mathbf{U} = [U_1, U_2]$ given by

$$Y(t) = b + U_1 \cos(t) + U_2 \sin(t) = b + \sqrt{U_1^2 + U_2^2} \sin \left[t + \tan^{-1} \left(\frac{U_1}{U_2} \right) \right] \quad (41)$$

Therefore, the maximum value Z of $Y(t)$ is given by

$$Z = b + \sqrt{U_1^2 + U_2^2} \quad (42)$$

Since $U_1^2 + U_2^2$ is a chi-square variable with freedom 2, the exact R is given by

$$R = \Pr\{Z < 0\} = \text{Prob}\{U_1^2 + U_2^2 < b^2\} = \Psi(b^2, 2) \quad (43)$$

where $\Psi(\cdot, 2)$ represents the chi-square CDF with the degree of freedom being 2.

$[\underline{t}, \bar{t}] = [0, 2\pi]$ is evenly discretized into $N = 500$ points, hence a 500-dimensional normal probability is to be calculated. With Eq. (40), we get the correlation coefficient matrix \mathbf{C} whose dimension is 500×500 . Since $Y(t)$ is a stationary Gaussian process, after discretization, $\mathbf{Y} = (Y_1, Y_2, \dots, Y_{500})$ share the same mean value b and standard deviation 1. As a result, no components in \mathbf{Y} are removed during the variable screening procedure.

Since there are only two input random variables in Eq. (41), $N_{\text{rank}} = 2$. The corresponding two eigenvalues of \mathbf{C} are 250.5 and 249.5, both of which are significant, and therefore there are $N' = 2$ significant basic random variables. Since $Q_{\text{max}} = 35$ and $Q_{\text{min}} = 5$, we use $Q_1 = 35$ and $Q_2 = 35$ quadrature points for U_1 and U_2 , respectively, and hence there are in total $N_Q = Q_1 Q_2 = 1225$ quadrature points. To test how the proposed method performs at different levels of P_f , we vary b . The values of P_f calculated by the proposed method, IECA, and RQ are given in Table 3. Note that the values in the parentheses under P_f are relative errors with respect to the accurate solutions and that the values in the square brackets are the estimated absolute errors (EAE) given by the RQ method.

When b takes $-2, -4, -6,$ and -8 , all the relative errors of the proposed method are less than 1%. It shows that the proposed method is accurate even when we calculate an extremely small P_f , such as 1.27×10^{-14} . The reason for the high accuracy is that there are only two significant basic random variables, and hence the Gauss-Hermite quadrature can obtain accurate MGF using Eq. (24). SPA can also produce an accurate CDF, and hence accurate P_f . In addition, this example shows that although $N = 500$, N' is only 2.

IECA is less accurate. When b is $-2, -4, -6,$ and $-8,$ the errors of IECA are $57.3\%, 34.7\%, 15.4\%,$ and $4.4\%,$ respectively. When $b = -2,$ RQ gets stable and accurate results. However, when calculating small probabilities ($b = -4, -6,$ or $-8,$), RQ1, RQ2 and RQ3 produce different results, showing instability. It is a typical feature of a random method.

Table 3 Results for Example 1

$b \rightarrow$ Methods \downarrow	-2	-4	-6	-8
Proposed	1.35×10^{-1} (0.0%)	3.34×10^{-4} (-0.5%)	1.52×10^{-8} (-0.2%)	1.27×10^{-14} (0.0%)
IECA	2.13×10^{-1} (57.3%)	4.52×10^{-4} (34.7%)	1.76×10^{-8} (15.4%)	1.32×10^{-14} (4.4%)
RQ1	1.35×10^{-1} (0.0%) [7.80×10^{-6}]	4.31×10^{-4} (28.5%) [2.92×10^{-4}]	1.60×10^{-8} (5.1%) [1.84×10^{-8}]	1.48×10^{-14} (16.7%) [1.18×10^{-14}]
RQ2	1.35×10^{-1} (0.0%) [6.77×10^{-6}]	3.39×10^{-4} (1.1%) [1.93×10^{-4}]	8.78×10^{-9} (-42.4%) [3.48×10^{-9}]	9.66×10^{-15} (-23.7%) [3.83×10^{-15}]
RQ3	1.35×10^{-1} (0.0%) [5.40×10^{-6}]	2.84×10^{-4} (-15.3%) [5.29×10^{-5}]	7.77×10^{-9} (-49.0%) [9.98×10^{-10}]	9.66×10^{-15} (-23.7%) [1.94×10^{-15}]
Exact	1.35×10^{-1}	3.35×10^{-4}	1.52×10^{-8}	1.27×10^{-14}

7.2. Example 2: A math example without an exact solution

The limit-state function $Y(t)$ is a nonstationary Gaussian process. The standard deviation is $\sigma(t) = 1$ and the mean $\mu(t)$ is given by

$$\mu(t) = -6 - t \cos(t) \quad (44)$$

where $t \in [\underline{t}, \bar{t}] = [0, 5]$ s. We consider three different correlation coefficient functions, given by Eq. (45), Eq. (46), and Eq. (47).

$$\text{Case 1: } \rho(t_1, t_2) = \sin(\pi|t_1 - t_2|)/(\pi|t_1 - t_2|) \quad (45)$$

$$\text{Case 2: } \rho(t_1, t_2) = \exp[-0.25(t_1 - t_2)^2] \quad (46)$$

$$\text{Case 3: } \rho(t_1, t_2) = \exp(-0.25|t_1 - t_2|)(1 + 0.25|t_1 - t_2|) \quad (47)$$

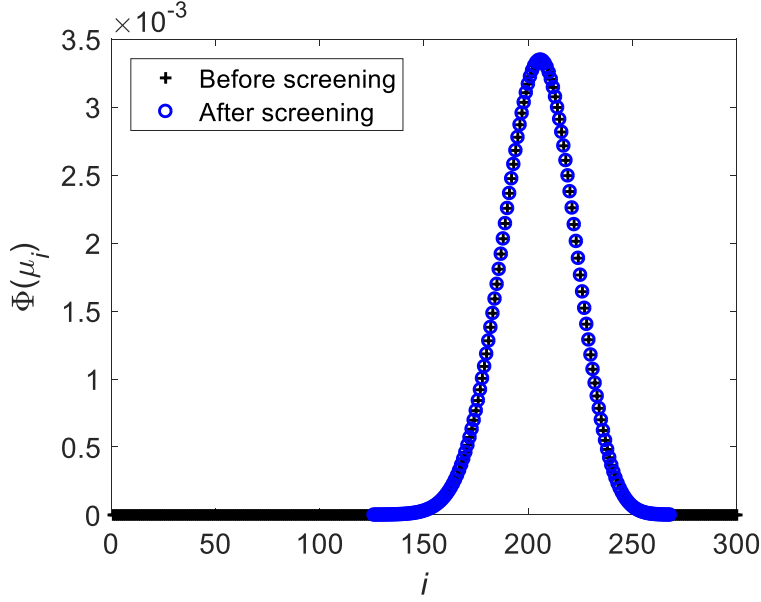


Fig. 7 Variable screening for Example 2

For numerical calculation, $[\underline{t}, \bar{t}]$ is evenly discretized into $N = 300$ points, and hence the dimension is 300. Fig. 7 shows the variable screening. 176 points among the 300 points do not contribute to P_f significantly and hence are removed. N is updated to 124. Note that the physical meaning of $\Phi(\mu_i)$ in Eq. (14) is the instantaneous probability of failure, and the variable screening procedure in fact removes those time points with low instantaneous probabilities of failure.

In Case 1, there are $N' = 5$ significant basic random variables. The numbers of quadrature points for them are 35, 31, 15, 5, and 5, and hence there are in total 406,875 quadrature points. The results are given in Table 4, where ε represents the relative error with respect to MCS. The sample size of MCS is 8×10^6 .

P_f calculated by the proposed method is 6.42×10^{-3} with a relative error of -0.1% , while IECA yields a P_f value of 6.93×10^{-3} with a relative error of 7.9% . The proposed method is significantly more accurate than IECA. RQ is more accurate than IECA, but not stable due to randomness.

Table 4 Results for Case 1 of Example 2

Methods	Proposed	IECA	RQ1	RQ2	RQ3	MCS
$P_f(\times 10^{-3})$	6.42	6.93	6.76	5.94	6.54	6.42
$\varepsilon(\%)$	-0.1	7.9	5.3	-7.5	1.8	-
EAE	-	-	6.37×10^{-4}	5.90×10^{-4}	3.40×10^{-4}	-

In Case 2, there are $N' = 4$ significant basic random variables. The numbers of quadrature points for them are 35, 7, 5, and 5, respectively, and hence there are in total 6125 quadrature points. The results are given in Table 5. The sample size of MCS is 1.2×10^7 . Again, the proposed method is more accurate than both IECA and RQ.

Table 5 Results for Case 2 of Example 2

Methods	Proposed	IECA	RQ1	RQ2	RQ3	MCS
$P_f(\times 10^{-3})$	3.96	3.60	3.76	4.17	4.17	3.99
$\varepsilon(\%)$	-0.8	-9.7	-5.8	4.6	4.5	-
EAE	-	-	4.49×10^{-4}	5.81×10^{-4}	4.69×10^{-4}	-

In Case 3, there are $N' = 4$ significant basic random variables. The numbers of quadrature points for them are 35, 5, 5, and 5, and hence there are in total 4,375 quadrature points. The results are given in Table 6. The sample size of MCS is 1.2×10^7 . All the three methods are accurate, and the proposed method is slightly more accurate.

Table 6 Results for Case 3 of Example 2

Methods	Proposed	IECA	RQ1	RQ2	RQ3	MCS
$P_f(\times 10^{-3})$	3.42	3.35	3.48	3.35	3.48	3.43

$\varepsilon(\%)$	-0.2	-2.3	1.6	-2.3	1.7	-
EAE	-	-	2.31×10^{-4}	8.54×10^{-7}	2.34×10^{-4}	-

7.3. Example 3: A slider-crank mechanism

Shown in Fig. 8 is a slider-crank mechanism. The link with lengths R_1 and R_3 rotates with an angular velocity of $\omega = \pi$ rad/s. The motion output is the difference between the displacements of two sliders A and B . The mechanism is supposed to work with small motion errors during time period $[\underline{t}, \bar{t}] = [0, 2]$ seconds. The motion error is defined as the difference between the desired motion output and the actual motion output. A failure occurs when the motion error is larger than 0.94 mm. The actual motion output Δs_{actual} is given by

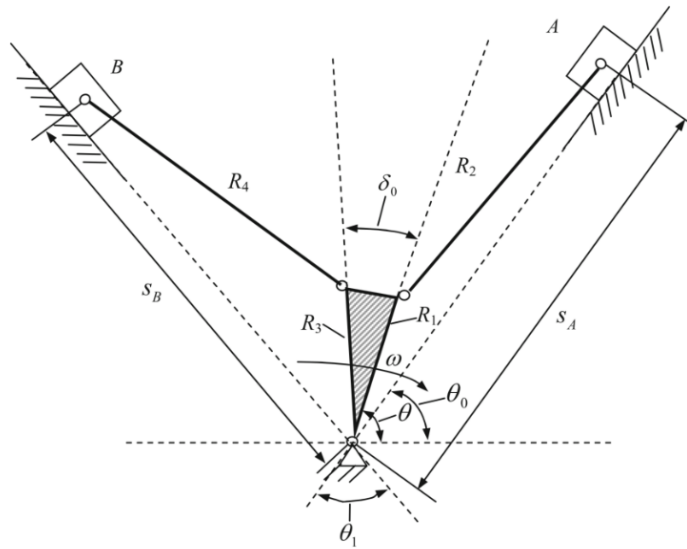


Fig. 8 A slider-crank mechanism

$$\begin{aligned}
\Delta s_{\text{actual}} = & R_1 \cos(\theta - \theta_0) + \sqrt{R_2^2 - R_1^2 \sin^2(\theta - \theta_0)} \\
& - R_3 \cos(\theta_1 + \theta_0 - \theta - \delta_0) \\
& - \sqrt{R_4^2 - R_3^2 \sin^2(\theta_1 + \theta_0 - \theta - \delta_0)}
\end{aligned} \tag{48}$$

where $\theta = \omega t$. The desired motion output $\Delta s_{\text{desired}}$ is given by

$$\begin{aligned}
\Delta s_{\text{desired}} = & 108 \cos(\theta - \theta_0) + \sqrt{211^2 - 108^2 \sin^2(\theta - \theta_0)} \\
& - 100 \cos(\theta_1 + \theta_0 - \theta - \delta_0) \\
& - \sqrt{213^2 - 100^2 \sin^2(\theta_1 + \theta_0 - \theta - \delta_0)}
\end{aligned} \tag{49}$$

Then the limit-state function $Y(t)$ is given by

$$Y(t) = (\Delta s_{\text{desired}} - \Delta s_{\text{actual}}) - 0.94 \tag{50}$$

Table 7 shows the random variables and other parameters.

Table 7 Variables and parameters of Example 3

Variable	Mean	Standard deviation	Distribution
R_1	108 mm	0.05 mm	Gaussian
R_2	211 mm	0.2 mm	Gaussian
R_3	100 mm	0.05 mm	Gaussian
R_4	213 mm	0.2 mm	Gaussian
θ_0	45°	0	Deterministic
θ_1	60°	0	Deterministic
δ_0	10°	0	Deterministic
ω	π rad/s	0	Deterministic

The time interval $[\underline{t}, \bar{t}]$ is evenly discretized into $N = 300$ points. Since $Y(t)$ is not a Gaussian random process, we need to transform it into an equivalent Gaussian process by applying FORM at each time point. After that we need to calculate a 300-dimensional normal probability to obtain P_f . Fig. 9 shows the variable screening step. No points among

the 300 points are removed because the instantaneous probabilities of failure at all the 300 points contribute to P_f significantly.

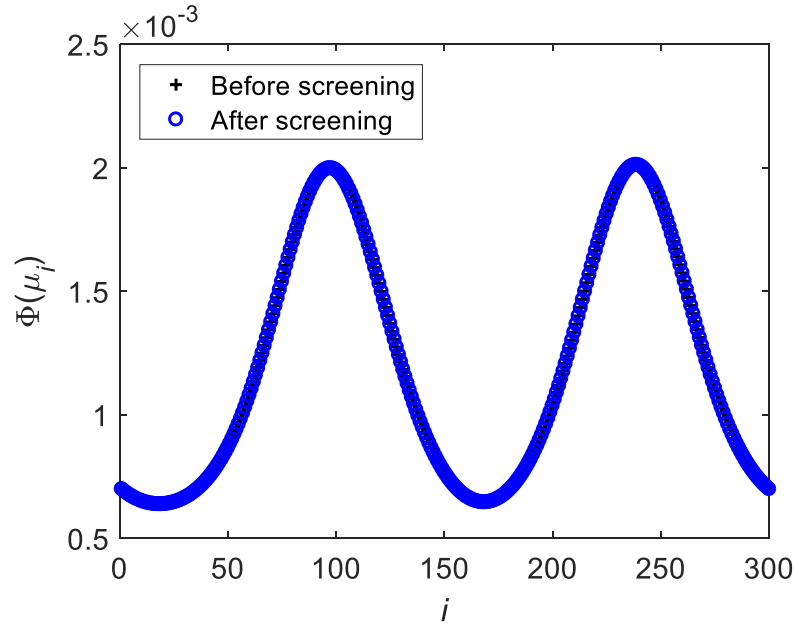


Fig. 9 Variable screening for Example 3

There are four significant basic random variables in \mathbf{U} after the dimension reduction is performed. The numbers of quadrature points for \mathbf{U} are 35, 5, 5, and 5, and hence there are in total 4,375 quadrature points. The results are given in Table 8. The sample size of MCS is 1.8×10^7 .

Table 8 Results of Example 3

Methods	Proposed	IECA	RQ1	RQ2	RQ3	MCS
$P_f (\times 10^{-3})$	2.38	2.11	2.48	2.39	2.48	2.38
$\varepsilon (\%)$	0.1	-11.4	4.1	0.5	4.1	-
EAE	-	-	3.24×10^{-4}	3.82×10^{-4}	4.62×10^{-4}	-

P_f calculated by the proposed method is 2.38×10^{-3} with a relative error of 0.1%, while P_f calculated by IECA is 2.11×10^{-3} with a relative error of -11.4% . RQ is more accurate than IECA but less accurate than the proposed method.

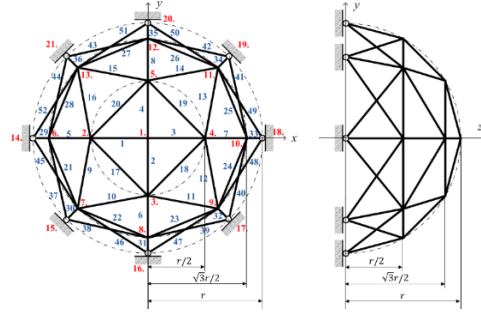
Note that there is no input random process in this example and hence the number N' of significant basic random variables is at most the number of input random variables. If $Y(t)$ was not sensitive to some input random variables, N' would be less than the number of input random variables.

7.4. Example 4: A 52-bar space truss

This example is modified from an example in [61]. Shown in Fig. 13 is a 52-bar space truss with 21 nodes. All the nodes are located on the surface of an imaginary hemisphere whose radius is $r = 240$ in. The cross-sectional areas of Bars 1~8 and 29~36 are 2 in^2 . The cross-sectional areas of Bars 9~16 and other bars are 1.2 in^2 and 0.6 in^2 , respectively. The Young's modulus of all bars is E . To distinguish the node numbers and the bar numbers, we add a decimal point after all node numbers in Fig. 13. Nodes 1~13 are subjected to external loads $F_1 \sim F_{13}$, all in the $-z$ direction. F_1 is a stationary Gaussian process whose autocorrelation coefficient function is given by

$$\rho(t_1, t_2) = \exp[-0.25(t_1 - t_2)^2] \quad (51)$$

E and $F_2 \sim F_{13}$ are random variables, and their distributions are given in Table 9.



(a) Top view

(b) Left view

Fig. 13 A 52-bar space truss

Table 9 Variables and parameters of Example 4

Variable	Mean	Standard deviation	Distribution	Autocorrelation
E	2.5×10^4 ksi	2.5×10^2 ksi	Gaussian	N/A
$F_1(t)$	40 kip	4 kip	Nonstationary Gaussian process	Eq. (51)
$F_2 \sim F_5$	50 kip	5 kip	Lognormal	N/A
$F_6 \sim F_{13}$	60 kip	6 kip	Lognormal	N/A

A failure occurs when the displacement δ of Node 1 along $-z$ direction exceeds the threshold $\delta_0 = 1.3$ in at any instant of time in the period $[\underline{t}, \bar{t}] = [0, 5]$ years. The limit-state function is given by

$$Y(t) = \delta_0 - \delta(E, \mathbf{F}) \quad (52)$$

where $\mathbf{F} = [F_1(t), F_2, F_3, \dots, F_{13}]$ is the vector all the loads. $\delta(E, \mathbf{F})$ is calculated by FEM. The linear bar element is used.

The time interval $[\underline{t}, \bar{t}]$ is evenly discretized into $N = 500$ points. Since $Y(t)$ is not a Gaussian random process, we need to transform it into an equivalent Gaussian process by applying FORM at each time point. After that we need to calculate a 500-dimensional normal probability to obtain P_f . Since $Y(t)$ becomes a stationary Gaussian process after

the transformation, $\mathbf{Y} = (Y_1, Y_2, \dots, Y_{500})$ share the same mean value and standard deviation. As a result, no components in \mathbf{Y} are removed during the variable screening procedure.

There are only $N' = 7$ significant basic random variables after the dimension reduction. The numbers of quadrature points for them are 35, 18, 6, 5, 5, 5, and 5, and hence there are in total 2,362,500 quadrature points. The sample size of MCS is 1.2×10^8 . The results are given in Table 10. The proposed method is significantly more accurate than both RQ and IECA.

Table 10 Results for Example 4

Methods	Proposed	IECA	RQ1	RQ2	RQ3	MCS
$P_f(\times 10^{-4})$	3.35	4.07	4.11	4.25	2.72	3.36
$\varepsilon(\%)$	-0.6	21.0	22.3	26.4	-19.1	-
EAE	-	-	2.51×10^{-4}	4.72×10^{-4}	2.13×10^{-4}	-

The four examples have demonstrated the high accuracy and robustness of the proposed method. IECA is accurate for some examples but less accurate for the others, and RQ is not robust for some problems because of large randomness in the solutions with different sampling seeds. The proposed method works particularly well for a time-dependent reliability analysis for which the limit-state function has been approximated by a Gaussian process.

8. Conclusions

Evaluating a multivariate normal probability is widely encountered in many engineering problems. It is a challenging task when the dimension is high and the probability is low. The proposed method addresses the problem by using the extreme value

of all the normal variables. Its moment generating function (MGF) is obtained by the Gauss-Hermite quadrature method, and the dimension is also reduced by screening out variables in both the physical space and the eigenspace. The saddlepoint approximation is used to recover the multivariate normal probability from MGF.

The main computational effort is the calculation of MGF by a multidimensional quadrature method. The efficiency depends on the dimension of the integral, or the reduced dimension. Therefore, the efficiency of the proposed method mainly depends on the number of the significant basic random variables after the dimension reduction, instead of the dimension of the original normal variables. This is a good feature for many engineering problems where the dimension can be reduced significantly because not all normal variables contribute significantly to the multivariate normal probability and the multivariate normal probability is not sensitive to all coordinates of the eigenspace.

Another advantage of the proposed method is its ability to calculate extremely small probabilities. The accuracy is achieved by the accurate generation of MGF, as well as saddlepoint approximation with its well-known accuracy for small probabilities. This feature makes the proposed method suitable for reliability applications where the probability of failure is inevitably small. The proposed method is also numerically stable, and the result is repeatable.

The method, however, may not work well if the reduced dimension is still high. For example, in time-dependent reliability problems, if the correlation length of the limit-state function is short and/or the time interval of interest is long, the reduced dimension will be high and the proposed method may not work well or may even fail. Our future work will focus on accommodating larger dimension in the reduced space.

Acknowledgements

We would like to acknowledge the support from the National Science Foundation under Grant No 1923799 (formerly 1727329).

Reference

- [1] Rausand M, Høyland A. System reliability theory: models, statistical methods, and applications. Hoboken: John Wiley & Sons; 2003.
- [2] Du X. System reliability analysis with saddlepoint approximation. *Structural and Multidisciplinary Optimization*. 2010;42:193-208.
- [3] Kang W-H, Song J. Evaluation of multivariate normal integrals for general systems by sequential compounding. *Structural Safety*. 2010;32:35-41.
- [4] Hu Z, Du X. Time-dependent reliability analysis with joint upcrossing rates. *Structural and Multidisciplinary Optimization*. 2013;48:893-907.
- [5] Hu Z, Du X. First order reliability method for time-variant problems using series expansions. *Structural and Multidisciplinary Optimization*. 2015;51:1-21.
- [6] Hu Z, Mahadevan S. A single-loop kriging surrogate modeling for time-dependent reliability analysis. *Journal of Mechanical Design*. 2016;138:061406.
- [7] Jiang C, Wei XP, Huang ZL, Liu J. An outcrossing rate model and its efficient calculation for time-dependent system reliability analysis. *Journal of Mechanical Design*. 2017;139:041402.
- [8] Gong C, Frangopol DM. An efficient time-dependent reliability method. *Structural Safety*. 2019;81:101864.
- [9] Wei X, Du X. Uncertainty Analysis for Time-and Space-Dependent Responses With Random Variables. *Journal of Mechanical Design*. 2019;141:021402.
- [10] Andrieu-Renaud C, Sudret B, Lemaire M. The PHI2 method: a way to compute time-variant reliability. *Reliability Engineering & System Safety*. 2004;84:75-86.
- [11] Wang Z, Chen W. Time-variant reliability assessment through equivalent stochastic process transformation. *Reliability Engineering & System Safety*. 2016;152:166-75.
- [12] Hu Z, Du X. Mixed efficient global optimization for time-dependent reliability analysis. *Journal of Mechanical Design*. 2015;137:051401.
- [13] Hu Z, Du X. A sampling approach to extreme value distribution for time-dependent reliability analysis. *Journal of Mechanical Design*. 2013;135:071003.
- [14] Jiang C, Wei XP, Wu B, Huang ZL. An improved TRPD method for time-variant reliability analysis. *Structural and Multidisciplinary Optimization*. 2018.

- [15] Sudret B. Analytical derivation of the outcrossing rate in time-variant reliability problems. *Structure and Infrastructure Engineering*. 2008;4:353-62.
- [16] Wang Z, Mourelatos ZP, Li J, Baseski I, Singh A. Time-dependent reliability of dynamic systems using subset simulation with splitting over a series of correlated time intervals. *Journal of Mechanical Design*. 2014;136:061008.
- [17] Singh A, Mourelatos Z, Nikolaidis E. Time-dependent reliability of random dynamic systems using time-series modeling and importance sampling. *SAE International Journal of Materials and Manufacturing*. 2011;4:929-46.
- [18] Wang Z, Wang P. A Nested Extreme Response Surface Approach for Time-Dependent Reliability-Based Design Optimization. *Journal of Mechanical Design*. 2012;134:121007-14.
- [19] Shi Y, Lu Z, Xu L, Chen S. An adaptive multiple-Kriging-surrogate method for time-dependent reliability analysis. *Applied Mathematical Modelling*. 2019;70:545-71.
- [20] Wang D, Jiang C, Qiu H, Zhang J, Gao L. Time-dependent reliability analysis through projection outline-based adaptive Kriging. *Structural and Multidisciplinary Optimization*. 2020:1-20.
- [21] Du W, Luo Y, Wang Y. Time-variant reliability analysis using the parallel subset simulation. *Reliability Engineering & System Safety*. 2019;182:250-7.
- [22] Shi Y, Lu Z, Cheng K, Zhou Y. Temporal and spatial multi-parameter dynamic reliability and global reliability sensitivity analysis based on the extreme value moments. *Structural and Multidisciplinary Optimization*. 2017;56:117-29.
- [23] Chen J-B, Li J. The extreme value distribution and dynamic reliability analysis of nonlinear structures with uncertain parameters. *Structural Safety*. 2007;29:77-93.
- [24] Yu S, Wang Z. A general decoupling approach for time-and space-variant system reliability-based design optimization. *Computer Methods in Applied Mechanics and Engineering*. 2019;357:112608.
- [25] Yu S, Wang Z, Meng D. Time-variant reliability assessment for multiple failure modes and temporal parameters. *Structural and Multidisciplinary Optimization*. 2018;58:1705-17.
- [26] Zhang D, Han X, Jiang C, Liu J, Li Q. Time-dependent reliability analysis through response surface method. *Journal of Mechanical Design*. 2017;139.
- [27] Madsen HO, Krenk S, Lind NC. *Methods of structural safety*. Englewood Cliffs: Prentice-Hall; 2003.
- [28] De Haan L, Ferreira A. *Extreme value theory: an introduction*. New York: Springer Science & Business Media; 2007.
- [29] Ochi Y, Prentice RL. Likelihood inference in a correlated probit regression model. *Biometrika*. 1984;71:531-43.
- [30] Dunnett CW. A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*. 1955;50:1096-121.
- [31] Anderson JA, Pemberton J. The grouped continuous model for multivariate ordered categorical variables and covariate adjustment. *Biometrics*. 1985:875-85.
- [32] Trinh G, Genz A. Bivariate conditioning approximations for multivariate normal probabilities. *Statistics and Computing*. 2015;25:989-96.
- [33] Mooney CZ. *Monte carlo simulation*. Thousand Oaks: Sage Publications; 1997.
- [34] Sobol' I. Quasi-monte carlo methods. *Progress in Nuclear Energy*. 1990;24:55-61.

- [35] Genz A, Bretz F. Computation of multivariate normal and t probabilities. Heidelberg: Springer Science & Business Media; 2009.
- [36] Melchers R. Importance sampling in structural systems. *Structural Safety*. 1989;6:3-10.
- [37] Phinikettos I, Gandy A. Fast computation of high-dimensional multivariate normal probabilities. *Computational Statistics & Data Analysis*. 2011;55:1521-9.
- [38] Ambartzumian R, Der Kiureghian A, Ohanian V, Sukiasiana H. Multinormal probability by sequential conditioned importance sampling: theory and application. *Probabilistic Engineering Mechanics*. 1998;13:299-308.
- [39] Au S-K, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic engineering mechanics*. 2001;16:263-77.
- [40] Cadini F, Gioietta A. A Bayesian Monte Carlo-based algorithm for the estimation of small failure probabilities of systems affected by uncertainties. *Reliability Engineering & System Safety*. 2016;153:15-27.
- [41] Roscoe K, Diermanse F, Vrouwenvelder T. System reliability with correlated components: Accuracy of the Equivalent Planes method. *Structural Safety*. 2015;57:53-64.
- [42] Gong C, Zhou W. Improvement of equivalent component approach for reliability analyses of series systems. *Structural Safety*. 2017;68:65-72.
- [43] Gollwitzer S, Rackwitz R. Equivalent components in first-order system reliability. *Reliability Engineering*. 1983;5:99-115.
- [44] Hohenbichler M, Rackwitz R. First-order concepts in system reliability. *Structural Safety*. 1982;1:177-88.
- [45] Tang L, Melchers R. Improved approximation for multinormal integral. *Structural Safety*. 1986;4:81-93.
- [46] Pandey M. An effective approximation to evaluate multinormal integrals. *Structural Safety*. 1998;20:51-67.
- [47] Yuan X-X, Pandey M. Analysis of approximations for multinormal integration in system reliability computation. *Structural Safety*. 2006;28:361-77.
- [48] Mendell NR, Elston R. Multifactorial qualitative traits: genetic analysis and prediction of recurrence risks. *Biometrics*. 1974:41-57.
- [49] Li C-C, Der Kiureghian A. Optimal discretization of random fields. *Journal of Engineering Mechanics*. 1993;119:1136-54.
- [50] Liu Q, Pierce DA. A note on Gauss—Hermite quadrature. *Biometrika*. 1994;81:624-9.
- [51] Butler RW. Saddlepoint approximations with applications. Cambridge: Cambridge University Press; 2007.
- [52] Hu Z, Du X. Saddlepoint approximation reliability method for quadratic functions in normal variables. *Structural Safety*. 2018;71:24-32.
- [53] Chi H, Mascagni M, Warnock T. On the optimal Halton sequence. *Mathematics and computers in simulation*. 2005;70:9-21.
- [54] Tuffin B. Randomization of quasi-monte carlo methods for error estimation: Survey and normal approximation. *Monte Carlo Methods and Applications*. 2004;10:617-28.
- [55] Ihaka R, Gentleman R. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*. 1996;5:299-314.

- [56] Abdi H, statistics. The eigen-decomposition: Eigenvalues and eigenvectors. Encyclopedia of measurement and statistics. 2007:304-8.
- [57] Daniels HE. Saddlepoint Approximations in Statistics. The Annals of Mathematical Statistics. 1954;25:631-50.
- [58] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes 3rd edition: The art of scientific computing. Cambridge: Cambridge university press; 2007.
- [59] Barndorff - Nielsen OE, Kluppelberg CJSjos. Tail exactness of multivariate saddlepoint approximations. 1999;26:253-64.
- [60] Zienkiewicz OC, Taylor RL, Nithiarasu P, Zhu J. The finite element method. London: McGraw-hill; 1977.
- [61] Zhang Z, Jiang C, Han X, Ruan X. A high-precision probabilistic uncertainty propagation method for problems involving multimodal distributions. Mechanical Systems & Signal Processing. 2019;126:21-41.