An Efficient Algorithm for Structural Reliability Based on Dichotomy Method

Zhengliang Xiang
Ph.D Candidate, School of Civil Engineering, Harbin Institute of Technology, Harbin, China

Yuequan Bao
Professor, School of Civil Engineering, Harbin Institute of Technology, Harbin, China

ABSTRACT: Monte Carlo Simulation (MCS) method is obviously a feasible and easy method for structural reliability evaluation, by which the multiple integral is replaced by sampling statistics. However, MCS is time-consuming because of its large number of simulations. To reduce the number of simulations, a structural reliability method based on dimensionality reduction and dichotomy has been presented, in the proposed method the dimensionality reduction technique is employed in grouping samples and the dichotomy method is applied to determining the partitioned limit state function (LSF). First, samples of direct MCS generated in original space are mapped to the independent standard Gaussian space and bi-dimensional space successively. Then the samples are divided into many groups according to the value of horizontal axis in the bi-dimensional space. Finally, the critical samples of each group are located by dichotomy method, and the partitioned LSF are approximated by the critical samples. With this method, the failure samples can be distinguished from whole samples by a relative little number of simulations. By several examples, the efficiency and robustness of the proposed algorithm were demonstrated, and the optimal number of the samples and the groups were respectively studied.

1. INTRODUCTION

Reliability analysis has been increasingly applied to structural design and structural assessment due the uncertainties involved with material, load and geometric properties. Reliability analysis aims to obtain the probability of failure of an event that is defined as Eq.(1)

$$P_f = \int_{g(X)<0} f(X)dX \tag{1}$$

where P_f is the probability of failure, $X = [X_1, X_2, ..., X_n]^T$ represents the vector of random variables and $f(X)$ stands for the joint probability density function(JPDF) of the vector of random variables. $g(X)$ is the response function, by which the stochastic domain is divided into safety and failure regions, i.e. $g(X) \le 0$ indicates X is located in the failure region.

However, it is often impossible to determine the probability of failure by the calculation of the integration of Eq.(1), because of the complicated response function and JPDF of random variables involved in the multiple integral. Instead, the probability of failure is commonly evaluated by approximate method or simulation method. First Order Reliability Method(FORM) is the most widely used reliability calculation method so far. FORM is defined in the standard Gaussian space, and the LSF is approximated at a reasonable point by linear polynomial expansion. Particularly, when the design point is selected as the expansion point, the reliability calculation in FORM is equal to finding the design point lies on the LSF which has the minimum distance from the origin of standard normal coordinate system, and the reliability index β defined by Hasofer−Lind is equal to the minimum distance Melchers (1987). Unfortunately, the accuracy of the evaluation results of FORM is unacceptable in case of limit state functions with large nonlinearity Valdebenito, Pradlwarter and Schu¨eller (2010). Accordingly, Second Order Reliability

Method(SORM) which approximates the performance function by a quadratic hypersurface was proposed to improve the accuracy of FORM Zhao and Ono (1999) Zhao (1999). But the accuracy of SORM is still insufficient when the performance function is highly nonlinear.

With the development of computer technology, simulation method is no doubt a feasible way to conduct reliability analysis. Conventional MCS can obtain a high accuracy in the evaluation of the probability of failure by increasing the number of samples Rubinstein (2008) Robert and Casella (2009). In order to reduce coefficient of variation of the estimate to lower than 0.1, the number of samples needed in conventional MCS should be select by Eq.(2).

$$N > 100\frac{1-\hat{P}_{f,FORM}}{\hat{P}_{f,FORM}} \tag{2}$$

where N is the number of samples, and $\hat{P}_{f,FORM}$ is the probability of failure estimated by FORM. Obviously, conventional MCS is inefficient in large-scale practical engineering problems, because it is computationally expensive and time consuming to simulate such large-scale practical engineering problems for millions of times. Additionally, conventional MCS may face "curse of dimensionality" problem in high dimensional cases, thus the estimated results are difficult to converge.

Many methods have been proposed over the past decades, aiming to reduce the number of samples of conventional MCS. Above all, Importance Sampling is the most widely employed method to reduce the number of samples in MCS Engelund and Rackwitz (1993) Olsson, Sandberg and Dahlblom (2003) Kurtz and Song (2013) Papaioannou, Papadimitriou and Straub (2016) Shayanfar, Barkhordari, Barkhori M, et al (2018). In IS, an optimal importance density is selected to make more samples located in the failure region, as shown in Eq. (3)

$$\begin{aligned} P_f &= \int_{R^d} I(X) f(X) dX \\ &= \int_{R^d} \frac{I(X) f(X)}{h(X)} h(X) dX = E_{\mathrm{f}}\left[\frac{I(X) f(X)}{h(X)}\right] \end{aligned} \tag{3}$$

where $I(X)$ is the indicator function of f, $h(X)$ is the optimal importance density function, and $E_{\mathrm{f}}[\cdot]$ denotes expectation function. But the optimal importance density is difficult or impractical to determine in many cases, e.g. cases with multiple failure regions which are not well separated Au and Beck (2001,2003).

Subset simulation changes the calculation of small probability of failure event into the calculation of the product of lager and conditional probabilities as follows Wang (2017)

$$P_f = P(F_m) = P(F_1)\prod_{i=1}^{m-1} P(F_{i+1}/F_i) \tag{4}$$

Because the required numbers of samples for the calculation of the conditional probabilities are away far smaller the calculation of small probability of failure, the total number of samples of subset simulation is smaller than conventional MCS. Through the concept of Subset simulation is logical, it is difficult to be carried out sometimes due to the appropriate proposal density function involved.

Besides, the existing methods to reduce the number of samples of MCS include Directional Simulation Nie and Ellingwood (2000), Line Sampling Pradlwarter, Schu¨ller and Koutsourelakis, et al (2007), Subdomain Sampling Methods Juang, Gong and Martin (2017) and Hierarchical Failure Clustering Yin and Kareem (2016). However, the applicability of these method is constrained by some drawbacks including the inefficiency in high dimensional problems.

To deal with high dimensional problem, an effective algorithm, based on the design point, for reducing the dimensionality of a structural reliability problem was proposed Hurtado (2012), then it was proposed to be useful for reducing the number of samples in MCS of the failure probability Hurtado, and Alvarez (2013). In this paper, dichotomy method which is common used to search the zero point of mathematical equations is combined with the aforementioned dimensionality reduction, and an algorithm aiming to save the computational expense of MCS

of the failure probability. In the proposed method, the dimensionality reduction is employed in sample clustering, and the dichotomy method is utilized to search the critical samples which can distinguish the failure samples from the save samples. Section 2 briefly introduces the algorithm for reducing the dimensionality of a structural reliability problem. In Section 3 is devoted to a detailed exposition of the proposed method for the calculation of probability of failure based on dimensionality reduction and dichotomy. Finally, the efficiency and robustness of the proposed method is assessed by examples in Section 4.

2. DIMENSIONALITY REDUCTION

Without loss of generality, consider a bi-dimensional case, as shown in Figure 1.

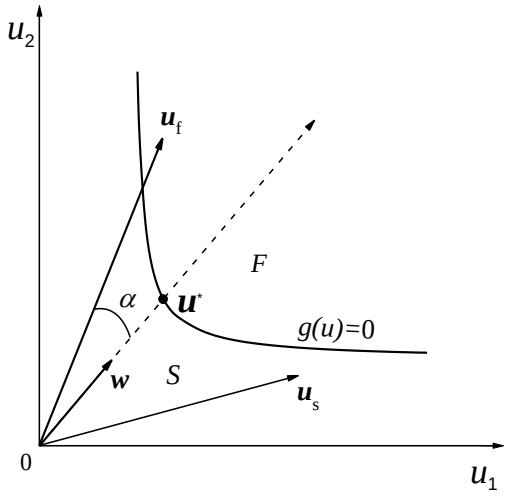

Figure 1: FORM and the polar features of Gaussian samples.

Firstly, the reliability problem is transformed to the standard Gaussian space with independent variables. This process is easy to achieve with the contribution of the appropriate transformations such as Nataf or Rosenblatt transformations Melchers (1987). Then the design point u^* , the Hasofer–Lind reliability index β ,and the corresponding failure probability $\hat{P}_{f,FORM}$ are determined. The unit vector of the design point is defined as

$$w = \frac{u^*}{\|u^*\|} \tag{5}$$

where $\|\cdot\|$ denotes 2-norm. Similarly, any sample can define a vector, as shown in Figure 1. The distance between the origin and the samples and the cosine of the angle between the vector of samples and the vector of design point are given by

$$v_1 = R = \|u\| \tag{6}$$

$$v_2 = \cos\alpha = \cos\angle(u, w) \tag{7}$$

By employing the defined variables v_1 and v_2 as the value of horizontal axis and vertical axis respectively, the samples in d-dimensional standard Gaussian space are mapped into a new independent bi-dimensional space. Moreover, it is proposed that the plot of samples observes a standard form in which the failure samples are accommodate in its upper-right sector, as show in Figure 6.

3. THE PROPOSED APPROACH

3.1. A Sample grouping and sorting

The bi-dimensional case is considered here again to introduce the proposed method for sample grouping, as shown in Figure 2. Figure 2 displays the target LSF and the LSF of FORM and SORM in standard Gaussian space. It is observed that FORM and SORM may fail to approximate the target LSF when a strong roughness exists in the target LSF, thus the samples around the approximate LSF will be classified falsely. To approximate LSF and distinguish the failure and save samples with a higher accuracy, hyperspheres of different radius R which is equal to v_1 introduced in Section 2 are applied to divide the standard Gaussian space into many annular regions, namely the region with radius ranges from R_{i-1} to R_i is defined as ith group G_i . When an appropriate number of groups are selected, the radius of the same group can be considered uniform. Besides, the failure samples in the same group often have bigger v_2 than the save samples. Consequently, the samples in each group are

descending sorted in term of v_2, and they are consumed to be classable by a critical sample S_{icr} which is a failure sample closest to the target LSF. In this way, the LSF in each regions of different radius are approximated separately.

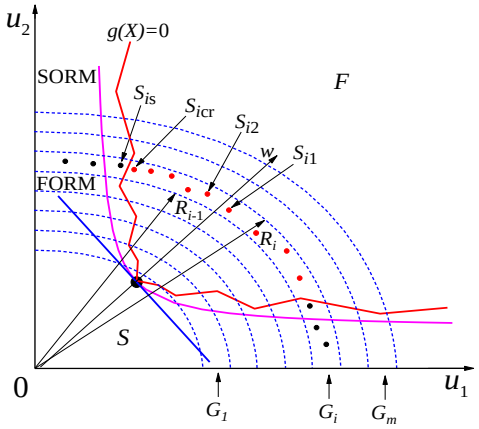

Figure 2: The groups of sample.

3.2. Critical samples seeking

By sample grouping and sorting, the problem of sample classification is changed into seeking the critical samples in each sample group. Dichotomy method is employed to located the critical samples.

Suppose there are *Ni* samples in *i*th group, and they are descending sorted in term of v_2, as show in Figure 3. The detailed process of dichotomy is as follows:

1. Set a=1 and b=Ni, then calculate $G(S_a)$ and $G(S_b)$;
2. Set bottom integral function $c=\lfloor (a+b)/2 \rfloor$, then calculate the LSF of $G(S_c)$;
3. If $G(S_c)$=0, S_c is the critical sample, and stop the algorithm;
4. If $G(S)_c \times G(S)$<0, set $b=(a+b)/2$, and if $G(S)_c \times G(S)$>0, set $a=(a+b)/2$;
5. Repeat steps (2)-(4) until $b-a$=1;
6. If $G(S_b)<0$, S_b is the critical sample, or S_a is the critical sample.

As show in Figure 3, sometimes a zero point S_k can be found to be the critical sample. However, in most cases, there is no zero point in the sample group, and two samples (S_{k-1},S_{k+1} or S_{k-2}, S_{k-1},) will be found at the end of the algorithm. Therefore, we need to check which one is the critical sample. After the critical samples are determined, the number of failure samples in each group is equal to the order number of the critical sample. For example, if S_k is the critical sample, there are k failure samples in ith group. With the dichotomy method, the critical samples of each group can be found without evaluating all limit state function of the samples, so a lot of computational cost can be saved.

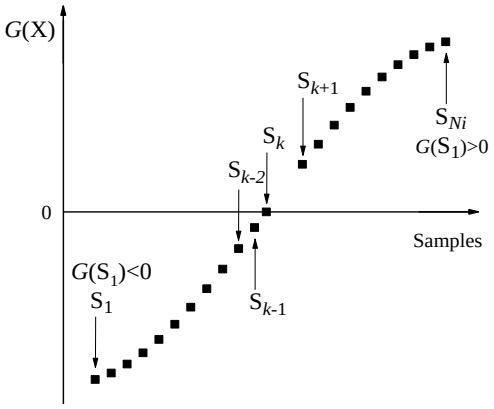

Figure 3 : The critical samples.

3.3. Calculation of failure probability

The proposed numerical procedure for the structural reliability problem in Eq.(1) is as follows:

1. Determine the design point u^*, the Hasofer–Lind reliability index β ,and the corresponding failure probability $\hat{P}_{f,FORM}$;
2. Generate N samples ordering to the distribution of X, where N is determined in term of Eq.(2);
3. Mapping the samples into the bi-dimensional space refer to the method introduced in Section 2;
4. Divide the samples into m groups by v_1, as described in Section 3.1, and sort the samples in each group. Generally, we can determine the upper and lower bounds of v_1, then divide the region $[v_{low}, v_{up}]$ uniformly by v_1.

5. Simulate the first several samples of each group (3 samples were selected in this paper). If the first several samples of a group are all save samples, all the samples in the group will be considered as save samples, otherwise, seek the critical samples in the group with the dichotomy method in Section 3.2, then determine the number of failure samples N_{fi} in the group.
6. Finally, the probability of failure can be evaluated by

$$P_f = \frac{\sum_{i=1}^{m} N_{fi}}{N} \tag{8}$$

4. VALIDATION OF THE METHOD

To demonstrate the efficiency and robustness of the proposed method for probability analysis, several widely used cases and a high dimensional nonlinear case are presented. The examples were firstly calculated by direct MCS, then the same samples were applied again to evaluate the failure probability by the proposed method. In order to compare the failure samples detected by direct MCS and proposed method, the results of two methods were both presented in bi-dimensional space introduced in Section 2. Additionally, failure probability of FORM was also obtained when the design point was determined, and it was compared with the results of the proposed method.

4.1. Example 1

Consider the case defined in Eq.(9)

$$g(x) = 2.5 - \frac{(x_1 + x_2)}{\sqrt{2}} + 0.1(x_1 - x_2)^2 \tag{9}$$

where x_1 and x_2 are independent standard normal random variables, and the probability information are $x_1 : N(0,1)$ and $x_2 : N(0,1)$. Figure 4 presents the samples simulated by direct MCS of 25000 samples, in which the failure samples are displayed in red. The number of groups was set as 40, and the samples simulated by proposed method are shown in Figure 5. It was observed that the failure samples of proposed method are almost the same as those of direct MCS. The results of failure probability can be seen in Table 1, which shows that the proposed method can save large computational expense with a high accuracy.

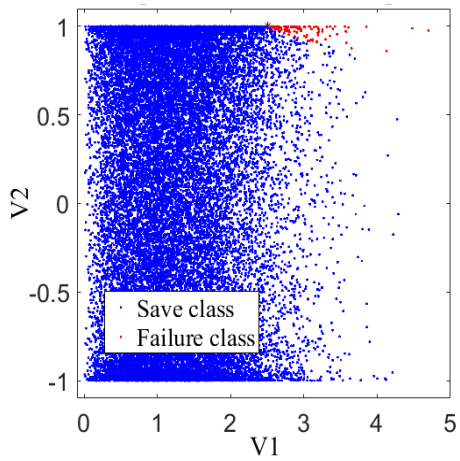

Figure 4: Results of direct MCS.

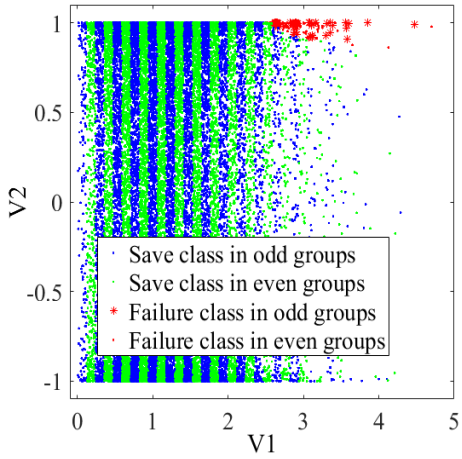

Figure 5: Results of proposed method.

Table 1: Results of Example 1.

Methods	*NS*	P_f $(\times 10^{-3})$	*Error (%)*
Direct MCS	*25000*	*4.280*	——
Proposed method	*155*	*4.120*	*3.7*
FORM	——	*6.210*	*45.1*

4.2. Example 2

In this section, a high dimensional nonlinear case is taken into consider. The response function is defined by Eq.(10)

$$g(\mathrm{x}) = \sum_{i=1}^{5} x_i^{\,2} - 2\sum_{j=1}^{5} x_j^{\,3} + 3\sum_{k=1}^{5} x_k^{\,4} + 20 \quad (10)$$

where x_1 and x_2 are independent standard normal random variables, and the probability information are $x_i : N(6, 0.6)$, $x_j : N(3, 0.3)$ and $x_k : N(2, 0.2)$. Figure 6 presents the samples simulated by direct MCS of 70000 samples, which are located in a round region because of high dimension. Set the number of groups as 60, and the samples simulated by proposed method can be seen in Figure 7. It was observed that the failure samples of proposed method are very similar to those of direct MCS. The results of failure probability can be seen in Table 2, which demonstrate that the proposed method can save large computational expense with a high accuracy in high dimensional nonlinear problems.

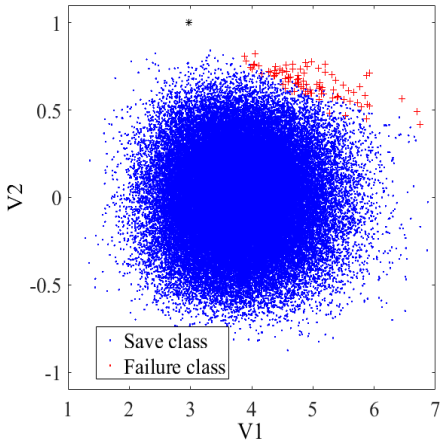

Figure 6: Results of direct MCS.

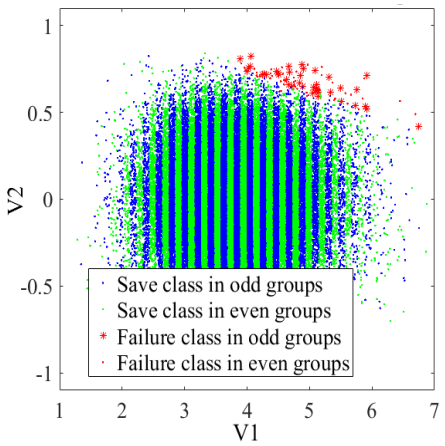

Figure 7: Results of proposed method.

Table 2: Results of Example 2.

Methods	*NS*	P_f *(×10^{-3})*	*Error (%)*
Direct MCS	*70000*	*1.286*	——
Proposed method	*326*	*1.257*	*2.2*
FORM	——	*1.484*	*15.4*

4.3. The influence of group number and sample number on the results

For the purpose of investigating the influence of group number and sample number on the results, the Example 1 in Section 4.1 is considered again in this section. By sequentially changing the sample number and the group number, respectively, simulations various sample numbers and group numbers were performed, then the results of which were compared as follows:

4.3.1. The influence of sample numbers

To evaluate the influence of sample numbers on results, the number of samples (*NS*) which ranged from 25000 to 45000 with an increment of 5000 were set, and the group number was set as 60 uniformly. Table 3 presents the results of simulation, including number of simulation of direct MCS (*NM*), number of simulation of proposed method (*NP*), failure probability of direct MCS (P_f-*M*) and failure probability of proposed method (P_f-*P*).

Table 3: Results of various sample numbers.

NS (×10^3)	*25*	*30*	*35*	*40*	*45*
NM (×10^3)	*25*	*30*	*35*	*40*	*45*
NP	*208*	*224*	*227*	*240*	*217*
P_f*-M (×10^{-3})*	*4.440*	*4.200*	*4.200*	*4.225*	*4.356*
P_f *-P (×10^{-3})*	*4.560*	*4.333*	*4.200*	*4.025*	*4.378*

The results are also displayed in Figure 8. It shows that the numbers of simulation of direct MCS are equal to the sample numbers, while the

numbers of simulation of proposed method which range around a relative small value are insensitive to the sample numbers. That is the reason why the proposed method can reduce large computational expense. Furthermore, the results of failure probability of each method achieve high accuracy when the sample numbers are larger than the recommended value of Eq.(2).

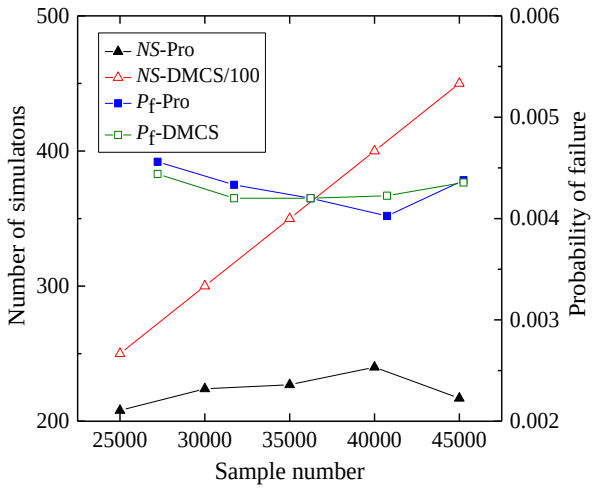

Figure 8 : Results of various sample numbers.

4.3.2. The influence of group numbers

Set the number of samples as 30000, and change the number of groups(*NG*) from 20 to 100, then the influence of group numbers on results was investigated. The results of simulation can be seen in Table 4, which are also displayed in Figure 9.

Table 4 : Results of various group numbers.

NG	*20*	*40*	*60*	*80*	*100*
NM ($\times 10^4$)	*3*	*3*	*3*	*3*	*3*
NP	*92*	*161*	*224*	*264*	*348*
P_f-*M* ($\times 10^{-3}$)	*3.267*	*3.733*	*4.200*	*4.433*	*4.067*
P_f -*P* ($\times 10^{-3}$)	*3.533*	*3.967*	*4.333*	*4.467*	*4.067*

As shown in Figure 9, the number of simulations of proposed method increases with the increase of the number of groups. But it is absolutely a small value compared with the number of simulations of direct MCS. In addition, the error between the results of two methods decrease gradually with the increase of the number of groups. As a conclusion, we can appropriately increase the number of groups to reduce the error of proposed method.

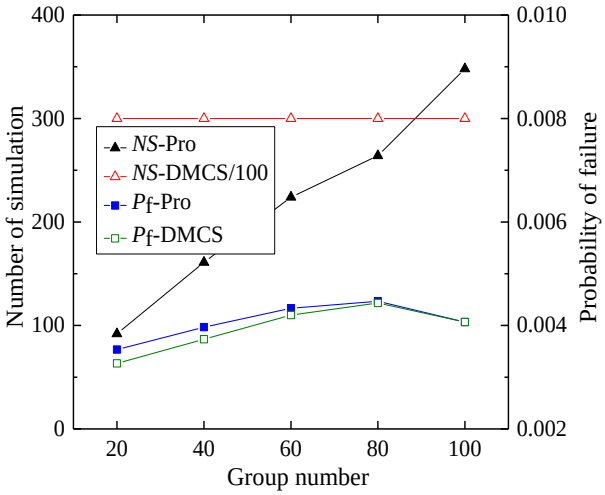

Figure 9: Results of various group numbers.

5. CONCLUSIONS

A structural reliability method based on dimensionality reduction and dichotomy has been presented, in which the dimensionality reduction technique is employed in grouping sample and the dichotomy method is applied to determining the partitioned LSF. Firstly, samples of direct MCS generated in original space are mapped to the independent standard Gaussian space and bi-dimensional space successively. Then the samples are divided into many groups according to the value of horizontal axis in the bi-dimensional space. Finally, the critical samples of each group are located by dichotomy method, and the partitioned LSFs are approximated by the critical samples. With this method, the failure samples can be distinguished from whole samples with a relative low computational cost, and the failure probability can be evaluated by Monte Carlo method. The efficiency and robustness of the proposed algorithm in high dimensional nonlinear problems are demonstrated by examples. According to results, an appropriate large number of samples and groups are recommended to

improve the accuracy of proposed method at the expense of a relatively little computational cost.

6. ACKNOWLEDGEMENTS

The authors express their acknowledgments to the authors of references and the anonymous reviewers.

7. REFERENCES

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

Au S K, Beck J L. (2003). Subset Simulation and its Application to Seismic Risk Based on Dynamic Analysis. *Journal of Engineering Mechanics*, 129(8):901-917.

Engelund S, Rackwitz R. (1993). A benchmark study on importance sampling techniques in structural reliability. *Structural Safety*, 12(4):255-276.

Hurtado J E. (2012). Dimensionality reduction and visualization of structural reliability problems using polar features. *Probabilistic Engineering Mechanics*, 29(7):16-31.

Hurtado, J. E., & Alvarez, D. A. (2013). A method for enhancing computational efficiency in Monte Carlo calculation of failure probabilities by exploiting form results. *Computers & Structures*, 117(3), 95-104.

Juang C H, Gong W, Martin J R. (2017). Subdomain sampling methods – Efficient algorithm for estimating failure probability. *Structural Safety*, 66:62-73.

Kurtz N, Song J. (2013). Cross-entropy-based adaptive importance sampling using Gaussian mixture. *Structural Safety*, 42(4):35-44.

Melchers, R. E. (1987). Structural reliability analysis and prediction. Wiley.

Nie J, Ellingwood B R. (2000). Directional methods for structural reliability analysis. *Structural Safety*, 22(3):233-249.

Olsson A, Sandberg G, Dahlblom O. (2003). On Latin hypercube sampling for structural reliability analysis. *Structural Safety*, 25(1):47-68.

Pradlwarter H J, Schu¨ller G I, Koutsourelakis P S, et al. (2007). Application of line sampling simulation method to reliability benchmark problems. *Structural Safety*, 29(3):208-221.

Papaioannou I, Papadimitriou C, Straub D. (2016). Sequential importance sampling for structural reliability analysis. *Structural Safety*, 62:66-75.

Robert C P, Casella G. (2009). Monte Carlo Statistical Methods. *Publications of the American Statistical Association*, 96(453):339-355.

Schu¨ller, G. I., Pradlwarter, H. J., & Koutsourelakis, P. S. (2004). A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics*, 19(4), 463-474.

Shayanfar M A, Barkhordari M A, Barkhori M, et al. (2018). An adaptive directional importance sampling method for structural reliability analysis. *Structural Safety*, 70:14-20.

Valdebenito M A, Pradlwarter H J, Schu¨ller G I. (2010). The role of the design point for calculating failure probabilities in view of dimensionality and structural nonlinearities. *Structural Safety*, 32(2):101-111.

Wang F, Li H. (2017). Subset simulation for non-Gaussian dependent random variables given incomplete probability information. *Structural Safety*, 67:105-115.

Yin C, Kareem A. (2016). Computation of failure probability via hierarchical clustering. *Structural Safety*, 61:67-77.

Zhao Y G, Ono T. (1999). New approximations for SORM: part I. *Journal of Engineering Mechanics*, 125(1):86-93.

Zhao Y G. (1999). New approximations for SORM: Part 2. *Journal of Engineering Mechanics*, 125(1):86-93.