UNCECOMP 2015
1st ECCOMAS Thematic Conference on
International Conference on Uncertainty Quantification in
Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Crete Island, Greece, 25–27 May 2015

SUBSET SIMULATION FOR ASSESSING STRUCTURAL RELIABILITY OF MULTIPLE LIMIT STATE FUNCTIONS

Hong-Shuang Li [1], Yuan-Zhuo Ma [1], and Zijun Cao[2]

[1]Key Laboratory of Fundamental Science for National Defense-Advanced Design Technology of Flight Vehicles, Nanjing University of Aeronautics and Astronautics
Nanjing 210016, China
e-mail: hongshuangli@nuaa.edu.cn

[2]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

Keywords: Subset Simulation, multiple limit state functions, reliability analysis.

Abstract. *It remains a challenging task to calculate failure probabilities of multiple limit state functions (LSFs) using a single run of Subset Simulation in structural reliability analysis. To address this issue, this article presents a variant of standard Subset Simulation (SS), in which a unified intermediate event is designed to drive the simulation procedure progressively approaching multiple failure regions defined by all LSFs simultaneously. All failure probabilities of multiple LSFs are obtained by a single run of SS, which bypasses the sorting difficulty arising from the multiple LSFs. A representative example is used to demonstrate the efficiency, accuracy and robustness of the presented SS method.*

1 INTRODUCTION

During the past several decades, estimating the failure probability of a single limit state function (LSF) has been paid much attention [1-9]. However, relatively few efforts have been focused on the capacity of estimating the failure probabilities of multiple LSFs in a single run, which is a common issue in reliability-based design optimization. The well-known crude Monte Carlo Simulation (MCS) is capable of estimating the failure probabilities of multiple LSFs simultaneously, however, it still suffers from lack of efficiency at small failure probability levels. On the other hand, Subset Simulation (SS) improves significantly the computational efficiency for small probability levels, but it remains a difficulty to estimate all failure probabilities of multiple LSFs simultaneously. An improved strategy, Parallel Subset Simulation (PSS), in which a principle variable that is correlated with all LSFs of interest is defined to drive the simulation to gradually approach the multiple failure regions, has then been proposed for multiple LSFs case. However, the determination of a proper principle variable in the PSS is a not trivial task. Furthermore, the involved correlation hypothesis is only verified by numerical examples, i.e., empirically, rather than by a theoretical way [10].

This study is aiming to present a variant of standard SS, which can inherit the excellent properties of the standard SS, e.g., robustness to dimension, high efficiency for rare event simulation, and independence to model complexity and etc. The most attractive characteristic of this variant is that it can be applicable for estimating simultaneously all failure probabilities of multiple LSFs. In particular, a unified intermediate event is constructed in it to resolve the sorting difficulty arising in the standard SS. The determination of intermediate events for each LSF in the presented method is identical with that in the standard SS, while the unified intermediate events drive the simulation to, progressively and simultaneously, approach the multiple failure regions in a single run of simulation.

2 STANDARD SUBSET SIMULATION

The basic principle of SS is to decompose a small failure probability into a product of a sequence of relatively large conditional probabilities by introducing the intermediate events adaptively [3]. The target failure event F can be defined as $F=\{g(\mathrm{X}) \leq b\}$, where b is the desired response threshold for a performance quantity and $g(\mathrm{X}) \leq b$ indicates failure. Let $F= F_m \subset \ldots \subset F_2 \subset F_1$ denote a sequence of nested intermediate events. Hence, the target failure probability is given by

$$P_F = \mathrm{P}(F) = \mathrm{P}(F_m) = \mathrm{P}(F_m | F_{m-1})\mathrm{P}(F_{m-1}) = \cdots = \mathrm{P}(F_1)\prod_{i=2}^{m}\mathrm{P}(F_i | F_{i-1}) \quad (1)$$

Note that expressions of the intermediate events are similar to that of the target failure event F, i.e., $F_i = \{g(\mathrm{X}) \leq b_i, i=1, \ldots, m\}$ ($b = b_m < \ldots b_2 < b_1$), where m is the total number of intermediate events. Provided that the conditional probability is set as a constant value p_0, the intermediate events can be then adaptively determined [3]. It is very clear that generating conditional samples is vital for implementing the standard SS. Then, a modified Metropolis-Hasting algorithm has been developed and applied to generate the conditional samples in standard SS. More Details of the standard SS and the modified Metropolis-Hasting algorithm are referred to Ref. [3].

3 SUBSET SIMULATION FOR MULTIPLE LSFS

The failure probabilities of multiple LSFs of interest corresponding to multiple failure modes may also be estimated by the standard SS. However, the standard SS should be performed repetitively for each LSF. Inevitably, a large amount of computational cost is still required for multiple LSFs case. To be specific, it is not a trivial task to estimate the failure probabilities of multiple LSFs simultaneously using a single run of SS on account of the sorting difficulty. In the standard SS for a single LSF, a group of conditional samples is selected to provide "seeds" for generating samples in the next simulation level. The procedure involves sorting the samples according to their LSF values in each simulation level. However, it cannot determine a unique sequence of samples in the current simulation level for multiple LSFs case. Thus, the standard SS cannot handle the sorting difficulty arising in multiple LSFs. To resolve this the sorting difficulty, a unified intermediate event, which is defined as the union of the intermediate events for all LSFs concerned, is then constructed. Note that the union can be viewed as a single driving event, driving the simulation procedure to estimate all the failure probabilities of the multiple LSFs simultaneously using a single run of SS.

To provide a deep insight of the presented method, an illustrated problem with two LSFs and two uncertain parameters is firstly considered, as shown in Figure 1. Support that we have obtained two intermediate events $F_1^{(1)}=\{g^{(1)} \leq b_1^{(1)}\}$ and $F_1^{(2)}=\{g^{(2)} \leq b_1^{(2)}\}$ for two LSFs from the crude MCS step, respectively. Superscripts "(1)" and "(2)" denote these two LSFs, respectively, and the subscript "1" denotes the first intermediate event level. Then, a unified intermediate event is defined as the union of $F_1^{(1)}$ and $F_1^{(2)}$

$$F_1 = F_1^{(1)} \cup F_1^{(2)} = \left\{g^{(1)} \leq b_1^{(1)}\right\} \cup \left\{g^{(2)} \leq b_1^{(2)}\right\} \tag{2}$$

The unified intermediate event F_1 is employed to generate conditional samples for the next simulation level in the presented method. As discussed in the previous section, the determination of $F_1^{(1)}$ and $F_1^{(2)}$ in the presented method is similar to that in the standard SS. Consider three potential relationships between $F_1^{(1)}$and $F_1^{(2)}$, namely intersection, inclusion, and disjoin. It is obvious that the inclusion case can be simplified as a problem with a single LSF and natural considered by the principle of the standard SS. In accordance with Eq.(2), the i-th unified intermediate event can be expressed as

$$F_i = F_i^{(1)} \cup F_i^{(2)} = \left\{g^{(1)} \leq b_i^{(1)}\right\} \cup \left\{g^{(2)} \leq b_i^{(2)}\right\} \tag{3}$$

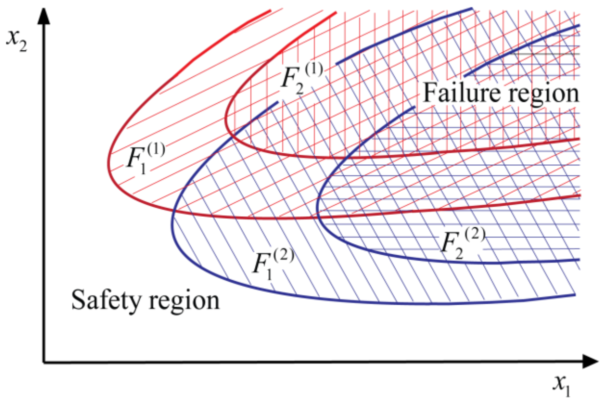

Figure 1: An illustrative example with two LSFs

It can be easily proved that $\{F_i\}(i=1,2,...)$ is also a nested sequence because both $\{F_i^{(1)}\}(i=1,2,...)$ and $\{F_i^{(2)}\}(i=1,2,...)$ are nested sequences. Without loss of generality, support that the failure probability $\mathrm{P}(F^{(1)})$ of the first target event $F^{(1)}=\{g^{(1)} \leq b^{(1)}\}$ is greater than that of the second target event $F^{(2)}=\{g^{(2)} \leq b^{(2)}\}$. The failure probability of $F^{(1)}$ is

$$\mathrm{P}_F^{(1)} = \mathrm{P}(F^{(1)} | F_{u-1})\mathrm{P}(F_{u-1} | F_{u-2}) \cdots \mathrm{P}(F_2 | F_1)\mathrm{P}(F_1) \tag{4}$$

where u is the number of intermediate events exactly when $b_u^{(1)} \leq b^{(1)}$ occurs. In practical implementation, the response thresholds satisfy $b^{(1)} = b_u^{(1)} < ... < b_2^{(1)} < b_1^{(1)}$. Let k denote the number of intermediate events required to reach $F^{(2)}$. The failure probability of $F^{(2)}$ is then given by

$$P_F^{(2)} = P\left(F^{(2)} \middle| F_{k-1}^{(2)}\right) \cdots P(F_u^{(2)} \middle| F_{u-1}) P(F_{u-1} \middle| F_{u-2}) \cdots P(F_2 \middle| F_1) P(F_1) \tag{5}$$

where $k \geqslant u$. According to the term $P(F_u^{(2)}| F_{u-1})$ $P(F_{u-1}| F_{u-2})$ in Eq.(5), no further consideration is placed on the target failure event $F^{(1)}$ further once it has been visited.

Different from the standard SS for a single LSF, the conditional probability $P(F_i| F_{i-1})$ is not a constant value in the presented method. It is expressed as

$$P(F_i \middle| F_{i-1}) = P(F_i^{(1)} \cup F_i^{(2)} \middle| F_{i-1}) = P(F_i^{(1)} \middle| F_{i-1}) + P(F_i^{(2)} \middle| F_{i-1}) - P(F_i^{(1)} F_i^{(2)} \middle| F_{i-1}) \tag{6}$$

Specifically, the first two terms (i.e., $P(F_i^{(1)}|F_{i-1})$ and $P(F_i^{(2)}|F_{i-1})$) on the right-hand side of Eq.(6) are still equal to q_0 since the same operation in standard SS is adopted. However, the third term $P(F_i^{(1)} F_i^{(2)}| F_{i-1})$ does not always remain constant. There are two extreme cases, i.e., the inclusion and disjoint cases. The third term reduces to $P(F_i^{(2)}| F_{i-1})$ in an inclusion case, while it is equal to zero in a disjoint case. As a result, the value of $P(F_i^{(1)} F_i^{(2)}| F_{i-1})$ indicates the correlation between the two events to some extent.

For a problem with two LSFs, on basis of Eq. (6), it can be reasoned that

$$p_0 \leq P(F_i \middle| F_{i-1}) \leq 2 p_0 \tag{7}$$

Eqs. (3)-(6) can be readily extended to a problem with M stochastic responses ($M \geqslant 3$), Then,the interval of the conditional probability is expressed as

$$p_0 \leq P(F_i \middle| F_{i-1}) \leq \min\{M p_0, 1\} \tag{8}$$

The upper boundaries of $P(F_i|F_{i-1})$ in Eq.(8) are generally larger than the proper value of $P(F_i|F_{i-1})$ suggested in the standard SS, i.e. $p_0 \in [0.1,0.3]$[8]. If p_0 increases, the number of simulation levels required to reach the target failure regions will increase and the efficiency of SS might decrease. Since the LSFs from a structure might generally share the common input random variables and the common numerical model, they are inevitably mutually dependent. Therefore, the upper bound values of conditional probability in Eq. (8) are rarely reached. Even for the worst case occurs, i.e., the unit upper boundary, the presented method will degrade to the crude MCS.

The main implementation procedure of the presented SS is similar to that of the standard SS except the determining the probabilities of the unified intermediate events and the number of conditional samples replenished to a simulation level. After the crude MCS stage, the first response threshold values $b_1^{(j)}$ (j=1,2,...,M) for the M LSFs of interest are adaptively determined Subsequently, the union of all the first intermediate events $F_1^{(j)}=\{g^{(j)} \leq b_1^{(j)}\}$ are regarded as the first unified intermediate event F_1. Note that the samples falling within F_1 shall only be counted once. Then, the probability of $P(F_1)$ can be estimated by

$$P(F_1) \approx \tilde{P}(F_1) = \frac{N_{1_1}}{N} \tag{9}$$

where N_1 is the number of samples in F_1.In accordance with the standard SS, another (N- N_1) samples are replenished into F_1 so that the number of samples in F_1 maintains N. New conditional samples are generated by the modified Metropolis-Hastings algorithm. The number of samples falling within F_2 is counted and is denoted as N_2. Thus, the conditional probability $P(F_2|F_1)$ is then calculated as

$$P\left(F_2|F_1\right) \approx \tilde{P}\left(F_2|F_1\right) = \frac{N_2}{N} \tag{10}$$

Repeat the above stage until the response threshold $b_u^{(j)}$ (j=1,2,...,M) for the j-th stochastic response satisfies $b_u^{(j)} \leq b^{(j)}$. Then, the failure probability of the j-th LSF is obtained and it is excluded from the unified intermediate event in the subsequent simulation levels. The simulation procedure shall not be terminated until all the target failure regions defined by all LSFs concerned are reached.

Note that the number of seed samples may vary in each simulation level because of that it is controlled by the correlation between all the intermediate events $F_i^{(j)}=\{g^{(j)} \leq b_i^{(j)}\}$. Furthermore, two properties of the presented method are attracting. Firstly, one interesting phenomenon is shown in Figure 2. That is, a seed sample selected by the intermediate event of a LSF might generate samples belongs to the intermediate events of the other LSFs in the next simulation level. The transitions in a Markov chain during simulation are denoted as the arrows. Note that the "jumping" property provides larger opportunity to completely explore the whole target failure regions.

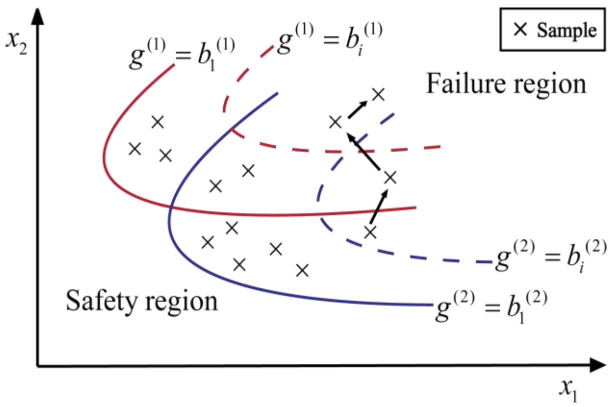

Figure 2 Illustration of the "jumping" property

Secondly, the statistical properties of estimators $P(F_1)$ and $P(F_i|F_{i-1})$ are identical with those in the standard SS because the simulation procedure is still driven by a unique event [11].

4 NUMERICAL EXAMPLE

A representative example modified from Re.[12] for designing a speed reducer is used to demonstrate the efficiency, accuracy and robustness of the presented method, as shown inFigure 3. During the implementation for both the standard SS and the presented SS, N and p_0 are set as 500 and 0.1. Beside, 30 independent runs are performed to estimate the coefficient of variation (COV), mean value of failure probabilities, and the total number of sample N_T required in the simulation, respectively.

A random vector $\mathbf{X}=(X_1, ..., X_7)$ consisting of seven normal variables is considered as input random parameters. Their statistical properties involve the corresponding mean values of (3.58, 0.70, 17.0, 7.43, 8.24, 3.37 5.31) and the same standard deviation of 0.005.

Totally, 11 LSFs are governed by the following equtions

$$
\begin{aligned}
&g^{(1)}(\mathbf{X})=\frac{27}{X_1X_2^2X_3}-1,\quad g^{(2)}(\mathbf{X})=\frac{397.5}{X_1X_2^2X_3^2}-1,\\
&g^{(3)}(\mathbf{X})=\frac{1.93X_4^3}{X_2X_3X_6^4}-1,\quad g^{(4)}(\mathbf{X})=\frac{1.93X_5^3}{X_2X_3X_7^4}-1,\\
&g^{(5)}(\mathbf{X})=\frac{\sqrt{\left(745X_4/\left(X_2X_3\right)\right)^2+16.9\times10^6}}{0.1X_6^3}-1100,\\
&g^{(6)}(\mathbf{X})=\frac{\sqrt{\left(745X_5/\left(X_2X_3\right)\right)^2+157.5\times10^6}}{0.1X_7^3}-850\\
&g^{(7)}(\mathbf{X})=X_2X_3-40,\quad g^{(8)}(\mathbf{X})=5-\frac{X_1}{X_2},\\
&g^{(9)}(\mathbf{X})=\frac{X_1}{X_2}-12,\quad g^{(10)}(\mathbf{X})=\frac{1.5X_6+1.9}{X_4}-1,\\
&g^{(11)}(\mathbf{X})=\frac{1.1X_7+1.9}{X_5}-1
\end{aligned}
\tag{11}
$$

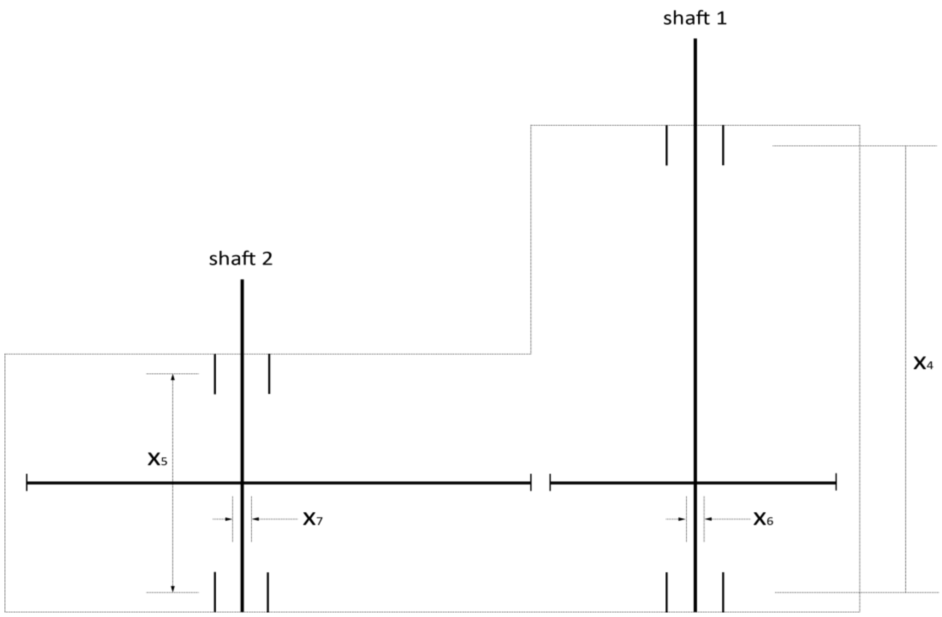

Figure 3: A speed reducer configuration

The problem is solved by the presented method, the standard SS and the crude MCS. Their results are listed in Table 1. Note that the maximum number of levels of the presented method and the standard SS are set as 30 and 10, respectively. For the crude MCS, 1×10^6 samples are simulated and the failure probability obtained from the crude MCS is considered as the exact failure probability. COV is estimated as $\left((1-\tilde{p}_F)/(\tilde{p}_F N_T)\right)^{1/2}$. The average number of samples in the presented method is 4032, which is much less than that of 47425 samples in the standard SS. It should be pointed out that each value in rows of failure probability is the mean value of failure probabilities obtained in 30 independent runs, so that failure probabilities of $g^{(5)}$ and $g^{(6)}$ are in the order of 10^{-8} and 10^{-7}.

In addition, COVs of the failure probabilities estimated from the standard SS and the presented method are listed in Table 1, where each “--” in rows of COV signifies that the corresponding COV cannot be calculated. Compared with the standard SS upon the mean estimated value of probabilities, total number of samples, errors and COVs, the presented method improves the computational efficiency significantly without much loss of accuracy and robustness.

LSF	Estimator	MCS	The standard SS	The presented SS
$g^{(1)}$	Failure probability	0	0	0
	Number of samples	1×10^{6}	4550	--
	COV	--	--	--
$g^{(2)}$	Failure probability	0	0	0
	Number of samples	1×10^{6}	4550	--
	COV	--	--	--
$g^{(3)}$	Failure probability	0	0	0
	Number of samples	1×10^{6}	4550	--
	COV	--	--	--
$g^{(4)}$	Failure probability	0	0	0
	Number of samples	1×10^{6}	4550	--
	COV	--	--	--
$g^{(5)}$	Failure probability	6.67×10^{-8}	0	0
	Number of samples	1×10^{6}	4550	--
	COV	3.87	--	--
$g^{(6)}$	Failure probability	5.33×10^{-7}	0	2.27×10^{-7}
	Number of samples	1×10^{6}	4550	--
	COV	1.34	--	33.08
$g^{(7)}$	Failure probability	0	0	0
	Number of samples	1×10^{6}	4550	--
	COV	--	--	--
$g^{(8)}$	Failure probability	8.57×10^{-4}	6.84×10^{-4}	7.54×10^{-4}
	Number of samples	1×10^{6}	1925	--

	COV	0.034	0.87	0.57
	Failure probability	0	0	0
$g^{(9)}$	Number of samples	1×10^6	4550	--
	COV	--	--	--
	Failure probability	0	0	0
$g^{(10)}$	Number of samples	1×10^6	4550	--
	COV	--	--	--
	Failure probability	0	0	0
$g^{(11)}$	Number of samples	1×10^6	4550	--
	COV	--	--	--
N_T		1×10^6	47425	4032

Table 1 Summary of computational results

5 DISCUSSIONS

It should be pointed out that the efficiency of the presented method is problem-dependent to some extent, but it is generally better than the strategy of repeatedly performing multiple SS for all LSFs. Given that p_0 and the sample size of N are fixed in each simulation level, the shared area for M=2, volume for M=3 or hypervolumes for M>3 of the unified intermediate events will affect the conditional probabilities $P(F_i|F_{i-1})$ and hence the efficiency of the presented method. If the intersections are relatively small, $P(F_i|F_{i-1})$ values will be relatively large, and the number of simulation levels required to reach the target failure regions will increase. As a result, the total number of samples might increase. In contrast, if the intersections are large, $P(F_i|F_{i-1})$ values will be relatively small. Then, more conditional samples are required to maintain the sample size N in each simulation level. In a word, there is a trade-off between the number of samples required in each simulation level and the number of simulation levels required to reach the target failure events.

Meanwhile, the performance of the presented method depends on the degree of correlation among the intermediate events defined by different LSFs. The efficiency of the presented method will be better than the standard SS when LSFs and intermediate events are highly correlated, while might deteriorate as this correlation decreases.

6 CONCLUSIONS

In the present study, a Subset Simulation method is developed to estimates failure probabilities of multiple LSFs simultaneously using a single simulation run. It shows significantly improvement in computational efficiency than repetitively implementing the original SS for all LSFs. However, this improvement is somehow problem-dependent, controlled by the shared area (M=2), volume (M=3) or hypervolume (M>3) of the unified intermediate events. However, even in the worst case, the presented method degenerates into the crude Monte Carlo simulation.

ACKNOWLEDGEMENTS

The authors are grateful for the supports by the Natural Science Foundation (Grant No. 11102084) and the Fundamental Research Funds for the Central Universities (Grant No. NS2015007).

REFERENCES

[1] R. Rackwitz and B. Flessler, *Structural reliability under combined random load sequences.* Computers & Structures, **9**(5), 489-494, 1978.

[2] O. Ditlevsen and H.O. Madsen, *Structural reliability methods*. *Vol. 178.* Wiley New York, 1996.

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

[4] R. Rackwitz, *Reliability analysis—a review and some perspectives.* Structural Safety, **23**(4), 365-395, 2001.

[5] S.K. Au and J.L. Beck, *Important sampling in high dimensions.* Structural Safety, **25**(2),139-163, 2003..

[6] S.K. Au, J. Ching, and J.L. Beck, *Application of subset simulation methods to reliability benchmark problems.* Structural Safety, **29**(3), 183-193, 2007.

[7] G.I. Schuëller and H. Pradlwarter, *Benchmark study on reliability estimation in higher dimensions of structural systems–an overview.* Structural Safety, **29**(3), 167-182, 2007.

[8] K.M. Zuev, et al., *Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions.* Computers & Structures, **92**, 283-296, 2012.

[9] A. Hasofer, N. Lind, *An exact and invariant first-order reliability format*. University of Waterloo, Solid Mechanics Division, 1973.

[10] W. C. Hsu and J. Ching, *Evaluating small failure probabilities of multiple limit states by parallel subset simulation.* Probabilistic Engineering Mechanics, **25**(3), 291-304, 2010.

[11] S.K. Au, *On the solution of first excursion problems by simulation with applications to probabilistic seismic performance assessment.* 2001.

[12] S. Rao, *Engineering Optimization: Theory and Practice*. New Age International, 1996.